



# Exclusion and Reintegration in a Social Dilemma

Alice Solda, Marie Claire Villeval

## ► To cite this version:

Alice Solda, Marie Claire Villeval. Exclusion and Reintegration in a Social Dilemma. 2017. halshs-01579216v2

**HAL Id: halshs-01579216**

**<https://shs.hal.science/halshs-01579216v2>**

Preprint submitted on 5 Sep 2017 (v2), last revised 1 Oct 2018 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

WP 1720 – August 2017

## Exclusion and Reintegration in a Social Dilemma

Alice Solda, Marie Claire Villeval

### Abstract:

The existing literature on ostracism in social dilemma games has focused on the impact of the threat of exclusion on cooperation within groups but so far, little attention has been paid to the behavior of the excluded members after their reintegration. This paper studies the effect of exclusion by peers followed by reintegration on cooperation once excluded individuals are readmitted in their group. Using a negatively framed public good game, we manipulate the length of exclusion and whether this length is imposed ex-ogenously or results from a vote. We show that people are willing to exclude the least cooperators although it is not an equilibrium strategy. Exclusion has a positive impact on cooperation and compliance to the group norm of withdrawal after reintegration when exclusion is followed by a quick rather than a slow reintegration and that the length of exclusion is chosen by the group. In this environment, a quicker reintegration also limits retaliation. Post-exclusion cooperation and forgiveness depend not only on the length of exclusion but also on the perceived intentions of others when they punish.

### Keywords:

Ostracism, exclusion, reintegration, social dilemma, cooperation, experiment

### JEL codes:

C92, H41, D23

# Exclusion and Reintegration in a Social Dilemma

Alice Solda and Marie Claire Villeval

August 31, 2017

## Abstract

The existing literature on ostracism in social dilemma games has focused on the impact of the threat of exclusion on cooperation within groups but so far, little attention has been paid to the behavior of the excluded members after their reintegration. This paper studies the effect of exclusion by peers followed by reintegration on cooperation once excluded individuals are readmitted in their group. Using a negatively framed public good game, we manipulate the length of exclusion and whether this length is imposed exogenously or results from a vote. We show that people are willing to exclude the least cooperators although it is not an equilibrium strategy. Exclusion has a positive impact on cooperation and compliance to the group norm of withdrawal after reintegration when exclusion is followed by a quick rather than a slow reintegration and that the length of exclusion is chosen by the group. In this environment, a quicker reintegration also limits retaliation. Post-exclusion cooperation and forgiveness depend not only on the length of exclusion but also on the perceived intentions of others when they punish.

**JEL-Code:** C92, H41, D23

**Keywords:** Ostracism, exclusion, reintegration, social dilemma, cooperation, experiment

## Contact Information:

Alice Solda, Univ Lyon, CNRS, GATE UMR5824, 93 Chemin des Mouilles, F-69130, Ecully, France; Queensland University of Technology, Brisbane, Australia. E-mail: solda@qut.edu.au

Marie Claire Villeval, Univ Lyon, CNRS, GATE UMR5824, 93 Chemin des Mouilles, F-69130, Ecully, France; IZA, Bonn, Germany. E-mail: villeval@gate.cnrs.fr

**Acknowledgments:** We are grateful to R. Slonim, J.R. Tyran, R. Sheremata and participants at the GATE-Lab seminar in Lyon, the 11th Annual Australia New Zealand Workshop on Experimental Economics, the QUT brown bag seminar, the 2017 Asia-Pacific ESA Conference and the Experimental Public Choice Workshop

in Lille for valuable comments on this paper. We also address special thanks to Q. Thévenet for his help in programming this experiment and C. Rimbaud for valuable research assistance. This research has been supported by a grant from the French National Research Agency (ANR, HEIDI grant 11-EMCO-011 01) and QuBE (Queensland Behavioral Economics Group) and was performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program “Investissements d’Avenir” (ANR-11-IDEX-007) operated by the French National Research Agency (ANR).

## 1 Introduction

In the presence of social dilemmas, *i.e.* when private and social interests are at odds, the temptation of free-riding is strong. Several mechanisms, including communication, rewards, sanctions, or taxation, have been designed to encourage group members to cooperate durably. In particular, the effectiveness of punishment mechanisms is now well established, although their positive impact on efficiency may develop mainly the long-run (Gächter *et al.*, 2008). An extreme form of punishment is the exclusion of group members who do not comply with the norm of the community. Various forms of exclusion can be found in societies. Informal and decentralized exclusion (ostracism) sanctions individuals with undesirable behavior who are isolated from the community for a certain amount of time, usually determined by customs.<sup>1</sup> Formal and centralized exclusion sanctions by imprisonment people who commit crimes, according to judicial rules. In both cases, the goal of this extremely severe punishment is both punitive and corrective. The intention is to make wrongdoers comply *ex ante* with the rules of the community if the threat is sufficient, or *ex-post* if the threat was not sufficient, once they are rehabilitated. However, we cannot ignore that exclusion has also drawbacks, not only on the individual well-being of excluded people, but also on the risk of anti-social behavior and recidivism after reintegration.<sup>2</sup>

While the deterrence theory of Becker (1968) predicts that increasing the penalty should lead to a reduction in misbehavior, there is still little conclusive empirical evidence of the effect of the length of exclusion on post-release compliance and re-

---

<sup>1</sup>The word ‘ostracism’ comes from the pieces of pottery that were used in the Athenian democracy to vote once a year about the banishment for ten years of non-compliant citizens. The anthropological literature is full of examples of ostracism and social exclusion in tribes and villages. Ethological studies have shown that scape-goating and shunning exist in social animal species. Forms of ostracism evolve over time, with cyberostracism as a modern expression of the refusal to communicate with someone (Wesselmann and Williams, 2013).

<sup>2</sup>For example, psychologists have found that ostracism may entail devastating psychological damages if exclusion is long, such as depression (Pressman and Cohen, 2005). In neurosciences, William (2002) has found that exclusion leads to the activation of the anterior cingulate cortex, an area of the brain that responds to pain, and the right ventral pre-frontal cortex, an area that is involved in coping with pain (see also Eisenberger *et al.*, 2003; Masten *et al.*, 2009; Bolling *et al.*, 2011). Exclusion generates anger and lowers self-esteem (Zadro *et al.* 2004; van Beest and Williams, 2006). In contrast, forgiveness coincides with activation in regions implicated in “theory of mind” and perspective-taking (Will *et al.* 2015).

cidivism. For example, studies on the length of prison sentences find little evidence that this length affects either recidivism (Martinson, 1974; Blumstein *et al.*, 1978), employment or income (Needels, 1996; Kling, 2006) after rehabilitation.<sup>3</sup> The absence of conclusive evidence is partly due to the fact that the empirical analysis based on data from real settings is often plagued with endogeneity and selection issues (a higher likelihood of recidivism for an inmate with a long prison sentence may be due to his own characteristics that led him to jail, as well as to peer effects in prison and to stigma on the labor market).<sup>4</sup> Moreover, we are not aware of any study measuring the impact of the length of informal exclusion on deterrence and compliance. Thus, overall, we still know surprisingly little about the impact of the length of exclusion on behavior after reintegration.

Our research objective is studying how the possibility to exclude group members through a vote for a given amount of time affects the behavior of wrongdoers after they are readmitted in their group, in the context of a social dilemma game. Our aim is, first, to compare how different lengths of exclusion influence behavior after reintegration, and, second, whether the length of exclusion has a different impact when voters had a choice between different possible lengths or when the length of exclusion is imposed exogenously by an external authority. We consider two dimensions of behavior after reintegration: the pro-sociality of the decisions made in the social dilemma game, and the forgiveness *vs.* the willingness of excluded people to seek revenge by voting to exclude other group members. Addressing these questions is highly policy-relevant because of the monetary, psychological and social costs of exclusion: if a longer length of exclusion generates a decreasing marginal benefit in terms of post-reintegration compliance or even net costs because of antisocial behavior, it would be important to moderate the severity of sanctions.

It is difficult to address these research questions with data from real settings because of the endogeneity issues mentioned above. A laboratory experiment controls the environment and keeps it constant (*e.g.*, it guarantees the same reference group and the same level of information before, during and after exclusion), which helps in identifying the direct effect of the length of exclusion and its mode of determination on behavior. It also allows us to close two channels that play a role in real settings: the loss of experience while excluded for a long period, which reduces the opportunities on the labor market after reintegration and affects deterrence negatively,

---

<sup>3</sup>Smith, Gendreau and Goggin (2002) have found that more punishment was associated with either a slight increase in recidivism or with no effect. In contrast, Lott (1992) found a large effect of the prison sentence length, notably on embezzlement offenders.

<sup>4</sup>To address these issues, Drago *et al.*, 2009 used a natural quasi-experiment in Italy where a law implemented an immediate three-year reduction in detention that led to the release of 40 per cent of the population of inmates, and a random manipulation of the length of the remaining sentence at the date of the pardon. They found that longer sentences *ex ante* have a higher deterrent effect on recidivism, but the inmates who experienced longer time in prison responded less to the variation in incentives. In contrast, Maurin and Ouss (2009) who exploited the discontinuity in the relationship between the time served in prison and the date of release created by a collective reduction of sentence in France for the Bastille Day showed a resulting increase on recidivism. Kuziemko (2007) also found that longer prison sentences decrease recidivism in the U.S.

and the channel of peer effects during exclusion that permit social learning in the behavior that led to exclusion.

Our experiment is based on a taking game. We formed groups of four subjects who have to decide simultaneously how many units to withdraw from a public good in each of 20 periods. In this context, the dominant selfish strategy is to withdraw as much as possible from the public good, whereas reaching the social optimum requires that no one takes any unit. After being informed about the withdrawal decision of each partner, individuals can vote to exclude some members. The decision of excluding someone requires at least two votes to become effective. Excluded subjects remain inactive, cannot communicate, and receive a low inactivity payoff. Since the initial size of the public good depends directly on the number of members in the group, excluding a member is potentially costly. In one treatment, the length of exclusion is fixed at one period; in another treatment it is fixed at three periods, and in a third treatment, individuals have to vote also about the length of exclusion, from one to three periods. We compare behavior in these treatments with a baseline in which exclusion is not possible. This design allows us to investigate the effect of exclusion followed by reintegration on excluded members' behavior once they have been readmitted in their group, depending on how long they have been excluded and on the procedure of exclusion.

Our study makes three contributions to the previous literature. The first contribution is comparing the impact of different lengths of exclusion on the cooperative behavior of the excluded members after their reintegration in the group. Earlier studies have been more interested in measuring the effect of the threat of exclusion on contribution behavior in public good games than in the implementation of the threat (Masclet, 2003; Cinyabuguma *et al.*, 2005; Kerr *et al.*, 2009; Sheremeta *et al.* 2009; Maier-Rigaud *et al.*, 2010). Most of the time, the threat of exclusion is sufficient to discourage free-riding. In real settings, however, we observe exclusion, and the question of how the length of exclusion determines behavior after reintegration is important.

In the previous experiments, we observe three types of length of exclusion. In some studies, exclusion is irreversible (Cinyabuguma *et al.*, 2005; Maier-Rigaud *et al.*, 2010). A subject is excluded from the game until the end of the session. In other studies, exclusion from the group is irreversible but subjects can move immediately to another group (Charness and Yang, 2014). Finally, in other studies exclusion is only for one period (Masclet, 2003; Kerr *et al.*, 2009; Sheremeta *et al.*, 2009; Koike *et al.* 2015). In these studies there is no comparison between different lengths of exclusion, except in Neuhofel and Kittel (2015) who compare in a between-subject experiment a treatment with one-period exclusion and another treatment with irreversible exclusion, showing that the latter has a stronger impact. The lack of variation of the length of exclusion in the literature is surprising because exclusion is the only punishment mechanism that has been studied in a binary way: someone is or not excluded, and if punished, each excluded member gets the same punishment regardless of his action. Yet in real life, the severity of exclusion depends on

the gravity of the harm that has been caused. In contrast, in our study, we consider only reversible exclusion and we manipulate the possible length of exclusion.<sup>5</sup>

Overall, the previous literature has found that exclusion is an effective mechanism to enforce social norms and limit free-riding in social dilemmas (see also Bowles and Gintis 2004).<sup>6</sup> However, irreversible exclusion decreases efficiency because of the loss of the potential input of the excluded members. Moreover, it is not clear that a more severe punishment will always lead to more compliance. In his theory of reintegrative shaming in criminology, Braithwaite (1989) underlines the importance of the perception of a fair punishment to trigger positive responses after rehabilitation (see Coricelli *et al.*, 2014, for an experimental test). A punishment that is perceived as unfair because too severe as regards previous actions may trigger negative responses.

Our second contribution is studying whether the choice of the length of exclusion through a voting procedure modulates the impact of exclusion on future cooperative behavior. In the previous literature, the decision to exclude is usually implemented via voting but the length of exclusion is fixed exogenously. In contrast, we are interested in examining whether choosing a less severe sanction when a more severe one was available is perceived by the excluded individuals as a warning that signals a willingness to forgive after reintegration. Choosing instead a tougher punishment in this context may be seen as aggressive and limit the willingness to cooperate after reintegration. Using both a within- and a between-subject design, we are able to test whether the choice of the length of exclusion affects the cooperative behavior of the excluded group members after reintegration differently from when the length is fixed exogenously.

Our third contribution is studying the impact of exclusion on retaliation behavior, depending on the length of exclusion and on the decision rule. Previous literature has shown that the introduction of punishment in social dilemma games may lead to counter-punishment (Nikiforakis, 2008; Fatas and Mateu 2015) and sometimes can even generate feuds (Nikiforakis and Engelmann, 2011), especially in the presence of normative conflicts (Nikiforakis *et al.*, 2011). In contrast, in our study punishment is not individual; it results from a group decision, which may signal more clearly what the norm of the group is and lead to less retaliation. We examine whether such a collective punishment resulting from a vote influences the frequency of counter-punishment and whether letting group members choose the length of exclusion has a different impact from when it is fixed exogenously. Am-

---

<sup>5</sup>Davis and Johnson (2015) also manipulate the length of exclusion by using a two-phase mechanism: in each period, participants vote to exclude group members and to reintegrate those who have previously been excluded. But excluded members are only excluded from a chat and they are still allowed to contribute to the public good. Being excluded from cheap talk communication is not monetary costly for the excluded subjects.

<sup>6</sup>Studies in psychology have also shown that ostracized individuals tend to conform more when they are re-admitted in the group (*e.g.*, William *et al.*, 2000; Feinberg *et al.*, 2014), they are likely to work harder in group tasks (William and Sommer, 1997) and they become more sensitive to social cues (Gardner *et al.*, 2005).

brus and Greiner (2015) have shown that using democratic punishment instead of individual punishment decreases anti-social punishment, but we ignore whether this applies also in our context.

Our results confirm that, compared to the sanction-free baseline, the threat of exclusion and its implementation decrease the mean level of withdrawals from the group account. Group members vote for excluding the least cooperators, confirming the previous literature in a negatively framed public good game. We also show that when the length of exclusion is exogenous, a longer exclusion sanctions larger deviations from the norm of the group but it has only a marginal positive effect on norm compliance after reintegration compared to a shorter exclusion.

When the length of exclusion results from a vote, the violation of the norm of the group determines the decision to exclude a group member but not the length of exclusion. The length of exclusion depends partly on recidivism: a long exclusion is more likely when a group member has already been excluded whereas a short exclusion is more likely for members who have not been excluded yet. In this environment, a shorter exclusion improves cooperation after reintegration whereas a longer exclusion does not. Moreover, in this context the group members excluded for a short period of time are less likely to seek revenge than participants excluded for a longer duration. When it is possible to adjust the length of exclusion to the level of disapproval, the voluntary choice of a more severe sanction by the group members may generate stronger negative emotions that make forgiveness by the reintegrated members more difficult. The stronger risk of retaliation in this condition may also result from the screening of less cooperative subjects who have less reluctance to retaliate compared to more pro-social individuals. In contrast, choosing a short length of exclusion when one could have chosen a more severe sanction leads to more forgiveness and reduces the risk of revenge. These results are strong arguments that not only the severity of the sanction matters on behavior after reintegration, but also the intentions of group members that can be inferred from it.

The remaining sections are organized as follows. Section 2 details the experimental design and the procedures. In Section 3, we derive our theoretical predictions and behavioral conjectures. Section 4 displays our main results. Section 5 discusses these results and concludes.

## 2 Experimental Design and Procedures

We first present our design before detailing the procedures.

### 2.1 Design

The experiment is based on a taking game, *i.e.* a negatively framed public good game, that is played for 20 periods with a fixed matching protocol. It consists of four treatments that are played between subjects. In all treatments we form groups



of four subjects who receive a random identifying number that they keep throughout the 20 periods. A group account is given to the group at the beginning of each period and its value depends on the number of active (non-excluded) members in the group. In period 1, the group account always consists of 80 Experimental Currency Units (ECU), *i.e.* 20 ECU per group member. Each of the four group members can withdraw  $w_i \in [0, 20]$  ECU from the group account that they can put in their private account. Each ECU left in the group account yields 0.4 ECU to each member. Each ECU extracted by a member and placed in her private account pays 1 ECU to this member.

The payoff function for subject  $i$  in period  $t$  is the following:

$$\pi_{i,t} = w_{i,t} + \alpha(n * 20 - \sum_{j=1}^n w_{j,t}) \quad (1)$$

with  $\alpha$  the marginal per capita return of the group account, with  $\alpha = 0.04$ , and  $n$  the number of active members in the group.  $n$  represents the number of active members in the group.

#### *Baseline treatment*

In this treatment, the number of active members in the group is fixed throughout the game, as there is no possible exclusion. The value of the group account is equal to 80 ECU at the beginning of each period. In each period, each group member decides simultaneously how many ECU to withdraw from the group account, between 0 and 20 ECU. At the end of each period, subjects receive a feedback with their own withdrawal and earnings for the period, and the withdrawal and earnings of each other group member with their id. We rewrite the payoff function as follows:

$$\pi_{i,t} = w_{i,t} + 0.4 * (80 - \sum_{j=1}^4 w_{j,t}) \quad (2)$$

The following treatments introduce voting and exclusion.

#### *Exo1P treatment*

Each period of the Exo1P treatment (for “Exogenous one Period”)<sup>7</sup> consists now in two stages. The first stage is similar to the Baseline: subjects decide simultaneously on the number of ECU they want to withdraw from the group account, between 0 and 20 ECU included. At the beginning of the period, they are informed about the number of active members in the group and the amount of the group account. At the end of the first stage, all group members are informed of the withdrawal decisions and earnings of each group member with their id.

---

<sup>7</sup>Note that in these treatments exclusion is endogenous since it results from a vote in the group. The name of the treatment refers to the fact that the length of exclusion is exogenously imposed.

At the beginning of the second stage, subjects have the opportunity to exclude one or more members of their group for one period. Each member of the group makes an exclusion decision for each of the other members of her group. All group members make their decisions simultaneously. If someone is excluded in period  $t$ , she will remain inactive in period  $t+1$ . This means that in  $t+1$  she will not be able to participate in the taking stage, nor to vote in the punishment stage, and she will receive a fixed payoff of 10 ECU. Thus, being excluded is costly because the associated payoff is lower than when being in the group even if others free ride.<sup>8</sup> During the exclusion period, an excluded member receives feedback on the withdrawal decisions and earnings of her group members, and on the exclusion of other group members in  $t+1$ .

There is no direct cost for excluding a group member. Exclusion is indirectly costly since the value of the group account at the beginning of a period depends on the number of active group members and decreases with the number of excluded members. By excluding a group member, the remaining members lose the return from the ECU that the excluded members could have left in the group account if they were not excluded. Indirect costs are null only in the case where the subjects correctly anticipate that the excluded subject would have withdrawn 20 ECU in the next period. Therefore, the payoff function of player  $i$  for a period  $t$  is:

$$\pi_{i,t} \begin{cases} 10 & \text{if } i \text{ is excluded} \\ w_{i,t} + 0.4(n * 20 - \sum_{j=1}^n w_{j,t}) & \text{otherwise} \end{cases} \quad (3)$$

with  $n$  the number of active members in the group.

Excluding a group member for one period requires that at least two members vote to exclude this member. Since a member cannot vote to exclude herself, this corresponds to a majority rule when  $N=4$  and a unanimity rule when  $N=3$ . When two or less members remain in a period, the voting stage is withdrawn and it is not possible to exclude anyone.<sup>9</sup> If  $N=0$ , each subject receives the exclusion payoff of 10 ECU in  $t+1$  and period  $t+2$  starts.

After casting their votes, participants receive a feedback on the id of each group member that has been excluded (including themselves), if any. When  $N=4$ , subjects are not informed on the number of votes cast to exclude a group member (between 2 or 3). A member who is not excluded never knows whether another member voted for her exclusion or not.

### *Exo3P treatment*

---

<sup>8</sup>The exclusion payoff is only higher than the payoff of a cooperator who withdraws less than 4 ECU from the public account whereas others free-ride in full.

<sup>9</sup>Imposing either the unanimity or the majority voting rule regardless of the group size would have been a concern because it would have imposed different requirements to exclude someone according to the group size.

The Exo3P (for “Exogenous three Periods”) treatment is similar to Exo1P except that the length of exclusion is exogenously set at three periods instead of one. If a group member is excluded in period  $t$ , she will remain inactive in periods  $t + 1$ ,  $t + 2$  and  $t + 3$ . During each of these three periods, she receives the 10 ECU exclusion payoff and she is informed on the withdrawal decisions and earnings of her group members and on the exclusion of other group members. In period 18, since there are only two periods left before the end of the game, if a group member is excluded from the group, she will be excluded for two periods only (from  $t$  until the end of the game). On their screen, participants are reminded that if they choose to exclude another group member, this group member will only be excluded for two periods. In the second-to-last period, an excluded member is excluded for one period only and in the last period it is not possible to exclude any group member from the group. This change in procedure for the last three periods is explained to the subjects in the instructions.<sup>10</sup>

At the end of the second stage, subjects receive a feedback on the id of each player that has been excluded (including themselves), if any, and the length of exclusion. They are also given the number of remaining periods of exclusion of the group members that have been excluded earlier.

#### *Endo treatment*

The Endo (for “Endogenous”) treatment differs from the other treatments with possible exclusion in that the length of exclusion is endogenous. This has two implications. First, the number of periods of exclusion is no longer unique, but it is restricted to the interval  $[1, 3]$ . Second, the length of exclusion results from a vote from the group members. If a group member votes to exclude another group member, she also has to decide for how many periods she wants to exclude this member.<sup>11</sup>

The actual length of exclusion of a subject is determined by the median of the length chosen by all the voters in the group. It is legitimate to assume that a subject who agrees to exclude someone for  $t$  periods tacitly agrees to exclude him for any shorter amount of time. In each period, when  $N=4$  each subset of three members vote about the exclusion of the fourth one. If one voter chooses not to exclude, the underlying value of the length is 0. In this case, the median is the lowest number between the choices of the two remaining voters. If less than two voters choose to exclude another group member, this group member will not be excluded, therefore

---

<sup>10</sup>We chose this procedure to avoid having no risk of sanction in three periods whereas severe sanctions were available in the previous periods.

<sup>11</sup>We acknowledge that it would be better to introduce a single change with the Exo treatments. A possibility would have been to add a treatment with three possible lengths of exclusion, but with a random assignment of each length of exclusion. This would have allowed us to isolate the pure effect of the diversity of lengths of exclusion on behavior after reintegration. However, this would have certainly modified the decision of group members to exclude. Therefore, comparability would have been limited. Instead of a random assignment we could have imposed thresholds for triggering each length, based on the empirical thresholds identified in our current Endo treatment. However, it would have been difficult to set a threshold in the case of retaliatory behavior.

no length is computed. Subjects are told that if all the voters agree on a given length, the program applies the chosen amount of time. If the voters chose different amounts of time and if three members are voting in favor of the exclusion of another member, the program implements the intermediate number. If only two members vote in favor of the exclusion of another member but disagree on the length, the program implements the shortest length. We provide examples in the instructions and subjects have to answer a quiz to show their understanding.

Like in the other treatments with voting on exclusion, at the end of the second stage group members receive a feedback on each member who has been excluded (including themselves) and - if any - the number of periods of exclusion. They are also reminded of the number of remaining periods of exclusion of the group members who have been excluded in earlier periods.

#### *Individual characteristics*

Before playing this taking game, in a preliminary part we elicited the subjects' social preferences regarding cooperation.<sup>12</sup> We adapted the method developed by Fischbacher *et al.* (2001) to our taking game environment. This allows us to classify our subjects into five categories (free-riders, conditional cooperators, unconditional cooperators, subjects with U-shaped preferences, and others). In this preliminary part, we formed groups of four subjects randomly and assigned an id number to each group member. Subjects had to take two different types of decisions. The first decision is unconditional and the other decisions are conditional. For the unconditional decision, subjects had to indicate the number of ECU they were willing to withdraw from a group account in a one-shot taking game. For the conditional decisions, they had to make a withdrawal decision for each possible average number of ECU taken by three other group members. Therefore, each subject had 21 conditional decisions to make. The payment mechanism is the same as in Fischbacher *et al.* (2001): one member among the four group members is randomly selected by the program. The selected member is paid according to her conditional withdrawal decision while the three other members are paid according to their unconditional decisions. The payoff of the selected member is therefore determined by the conditional decision that matches the average amount of ECU taken from the group account by the three other members. The earnings are calculated according to the payoff function described at the beginning of this section.

Subjects did not receive any feedback regarding their payoff in this preliminary part before the end of the session. They were also informed at the beginning of the following part that the groups were rematched and that the id numbers identifying each group member were reassigned randomly.

---

<sup>12</sup>In the instructions this preliminary part is called part 1 and the main game is presented in part 2 (see Appendix 1). Subjects were informed from the beginning of the session that the session consisted of two parts but they received the instructions for the second part only after completing the first part.

At the end of the session, we also collected standard information on the subjects' gender, age, student status and discipline.

## 2.2 Procedures

The experiment has been run at GATE-Lab, Lyon, France. A total of 208 participants were recruited via HRoot (Bock *et al.*, 2014) mainly among students from local engineering, business and medical schools. No subject participated in more than one session. We ran 12 sessions that involved an average of 20 subjects per session. We collected observations for 12 groups of subjects in both the Endo and Exo1P treatments, 13 groups in the Exo3P treatment and 15 groups in the Baseline treatment (see Table A1 in Appendix 2). The experiment was programmed using z-Tree (Fischbacher, 2007).

Upon arrival, subjects randomly drew a ticket from an opaque bag indicating their terminal number. The instructions for the preliminary part were distributed at the beginning of the session and those for the main taking game were distributed after completion of the preliminary part (see Appendix 1). In the instructions, we used a neutral language, avoiding the terms “exclusion” and “excluded members” in favor of more neutral terms, as “remove” and “players who do not belong in the group”. After the experimenter finished reading the instructions aloud, subjects were asked to read again the instructions and they had to complete a quiz to check their understanding before starting each part. Responses were checked individually.

Sessions lasted about 1.5 hour. The average earnings were 18.34 Euros (s.e. = 2.60). Subjects were paid the sum of their payoffs from all 20 periods in addition to their payoff for the first part and a €5 show-up fee. They received their payment in cash and in private in a separate payment room.

## 3 Theoretical Predictions and Behavioral Hypotheses

In the Baseline treatment, it is straightforward that if subjects are selfish and payoff-maximizers, the dominant strategy is for each player to take the maximum of 20 ECU from the group account in every period, which gives them a payoff of 20 ECU if every group member plays the same strategy. On the opposite, the Pareto optimum is attained if players fully cooperate and do not withdraw anything from the group account. In this case, each player would earn 32 ECU.

In the treatments with possible exclusion, using backward induction and considering a finite number of periods  $T$ , it is easy to see that if participants are rational, they will all defect in the last period since there is no more interest in cooperation at this stage and no risk of exclusion. No one will exclude in  $T - 1$  if exclusion is costly. Therefore, every one also defects in  $T - 1$  and this goes back to the first period. Thus, the only sub-game perfect Nash equilibrium is that no one excludes

and everyone free rides in all periods by withdrawing the maximum amount from the group account.

If there is no direct cost of exclusion, everyone still defects in  $T$ . In  $T - 1$ , players are weakly willing to exclude knowing that the payoff in  $T$  will not depend on the number of players in the group. However, if the possibility of someone not free riding in full exists, exclusion becomes costly. Indeed, each remaining group member will lose the marginal return of the potential amount of ECU that the excluded group member would have left in the group account if he would have not been excluded. The 'trembling hand' refinement of the Nash equilibrium by Selten (1975) considers the possibility of mistakes in the withdrawal decisions, which makes exclusion irrational in this case since there is a positive probability that a group member will leave a positive amount of ECU in the group account (see also Maier-Rigaud *et al.*, 2010). If there is any chance that a player does not withdraw 20 ECU in the next period, nobody should exclude in  $T - 1$ . The same reasoning as before should apply. To sum up, there is no cooperation under standard assumptions but if for whatever reason, there is a possibility that someone leaves a positive number of ECU in the group account, there will also be no exclusion.

**Hypothesis 1.a:** *Each group member extracts the maximum amount of ECU from the group account.*

**Hypothesis 2.a:** *In treatments with possible exclusion, no one is excluded.*

However, Hirshleifer and Rasmusen (1989) have shown that costless ostracism can lead to a cooperative equilibrium in some contexts. They consider a  $n$ -player prisoner's dilemma game in which people can vote to exclude other group members for one period. They show that ostracizing defectors is a sub-game-perfect equilibrium when players follow the 'banishment strategy' according to which everyone cooperates and excludes defectors until the last round where everyone defects. In their settings, a player is excluded for one period if at least one vote is cast against her but since the threat is sufficient to enforce cooperation throughout the game, neither the length of exclusion nor the severity of the punishment affect the equilibrium. However, these predictions as well as the previous ones may not hold if players have social preferences. Thus, we also consider models that take into account behavioral dimensions.

Previous works on Voluntary Contribution Mechanism games have shown that it is possible to reach full cooperation by using punishment mechanisms (Yamagishi, 1986 ; Fehr and Gächter, 2000 ; Gächter *et al.*, 2008). Fehr and Schmidt's (1999) inequality aversion theory predicts that punishment - even when costly - can be used by players in equilibrium: if individuals care about payoffs equity, they return cooperation in order to avoid favorable unequal payoffs and they are also willing to punish free-riders. Moreover, we expect a different effect of each length of exclusion on cooperation before any exclusion occurs, because this variable captures the severity of potential punishment. Indeed, previous literature has

shown that the severity of punishment affects cooperation in public goods games (e.g., Nikiforakis and Normann, 2012). Questioning the prediction of Hirshleifer and Rasmusen (1989), Neuhofer and Kittel (2015) have found that the threat of irreversible exclusion leads to a higher level of cooperation than the threat of a one-period exclusion. These models and findings suggest that we should observe that not all subjects withdraw 20 ECU from the group account in every treatment with and without exclusion. We should expect a higher level of compliance to the norm of the group when there is a threat of longer rather than shorter length of exclusion. Since non-selfish players usually coexist with selfish ones, such heterogeneity may lead to Nash equilibria depending on the composition of the groups. This leads to the following hypothesis that is alternative to Hypothesis 1.a.

**Hypothesis 1.b:** *The level of cooperation is positive and increases in the length of potential exclusion.*

Previous experiments on ostracism in public good games have shown that participants are willing to exclude even when the decision is costly and they generally exclude the least cooperator (Cinyabuguma *et al.*, 2005; Maier-Rigaud *et al.*, 2010; Davis and Johnson, 2015; Neuhofer and Kittel, 2015). We also expect to observe exclusion, and potentially even more in the negatively framed context of our taking game than in the previous experiments. Indeed, in our game the selfish action is an active decision (taking from the group account), whereas in the standard VCM game the selfish action is remaining passive by not contributing.<sup>13</sup> This leads to the following hypothesis that is alternative to Hypothesis 2.a.

**Hypothesis 2.b:** *In treatments with possible exclusion, subjects exclude the group members that deviate the most from the norm of the group and withdraw the most from the group account.*

After excluded subjects have been reintegrated in their group, the length of exclusion may matter in two different ways: first, it may affect the level of withdrawal compared to the level of withdrawal that triggered exclusion, and second, it may influence the willingness to seek revenge for past exclusion. The impact on these two dimensions may depend on how the length of exclusion is determined.

Let us consider first the case in which the length of exclusion is determined exogenously. Regarding withdrawal behavior, we expect that subjects who have been excluded should be more willing to conform to the withdrawal norm of the group after reintegration than before being excluded. Because of the higher cost of exclusion, a more severe punishment should have a stronger impact on withdrawal after reintegration than a less severe one. However, a more severe punishment may

---

<sup>13</sup>This idea is in line with the findings of Dickinson *et al.* (2015) who compared punishment in a taking game and in a contribution game and observed that the socially desirable behavior was less frequent and subjects punished more in the negative framing (the taking game) compared to the positive one (the contribution game).

also be more likely to lead to counter-punishment after reintegration: subjects who have been excluded longer may be more willing to seek revenge by voting for the exclusion of others, compared to those excluded for one period. This leads to the following two hypotheses.

**Hypothesis 3:** *When the length of exclusion is exogenous, a longer exclusion decreases more the level of withdrawal after reintegration than a shorter exclusion.*

**Hypothesis 4:** *When the length of exclusion is exogenous, subjects excluded for three periods are more likely to seek revenge than those excluded for a shorter duration.*

Consider now the impact of endogeneity in the choice of the exclusion length on behavior after reintegration. Choosing the length of exclusion by voting conveys a richer signal on the group's disapproval and forgiveness compared to the settings in which the length of exclusion is fixed exogenously. Excluding for one period when a longer exclusion was possible can act as a warning on free-riders. Because of positive reciprocity (Falk *et al.*, 2003; Falk and Fischbacher, 2006), subjects who have been excluded for one period in the Endo treatment may adapt their behavior more positively than people who have been out for one period in the Exo1P treatment.

Conversely, by coordinating on a three-period exclusion when it was possible to exclude only for one period, group members signal even more clearly their discontent towards the excluded member, compared to both the Exo3P treatment and the one-period exclusion in Endo. Indeed, group members not only need to coordinate on the decision to exclude but also on the choice of this length (whereas a one-period exclusion may result from votes for different lengths in the group). If voting for a three period exclusion is interpreted as a less forgiving action, excluded members may be willing to retaliate after being reintegrated by increasing their level of withdrawal and/or by casting votes for excluding the other group members. In contrast, if a one-period exclusion is perceived as a fair and acceptable punishment, retaliation is less likely. This prediction is based on the theory of 'reintegrative shaming' that predicts that negative responses are more likely to occur if a sanction is perceived as being unfair by wrongdoers. Thus, we expect that other things equal, participants that are excluded for a shorter period of time in these settings will cooperate more than others who have been excluded for more periods when they reintegrate their group. This leads to our last hypotheses.

**Hypothesis 5:** *When the length of exclusion results from a vote, (a) choosing a shorter exclusion decreases more the target's level of withdrawal after reintegration compared to a longer exclusion, and (b) this effect is also greater than in the Exo1P treatment.*

**Hypothesis 6:** *When the length of exclusion results from a vote, (a) choosing a longer exclusion increases the risk that excluded subjects seek revenge compared*



*to a shorter exclusion, and (b) this effect is greater than in the Exo3P treatment.*

## 4 Results

We start by examining the treatment effects on cooperation and efficiency. We next investigate exclusion and its determinants. Then, we analyze the effect of exclusion followed by reintegration on cooperation and compliance to the withdrawal norm in the group. Finally, we study retaliation behavior.

Throughout the section we refer to Table 1 that displays summary statistics (means, standard deviations and non-parametric tests) for the following variables, by treatment: level of withdrawal per period, as measured by the number of ECU taken from the group account, withdrawal of non-excluded group members, withdrawal of excluded group members in the period of exclusion, number of votes in favor of exclusion, number of active group members that cast at least one vote to exclude another group member (considering only the cases where there are more than two active members in the group), number of group members excluded in a period (considering only the cases where a new subject been excluded), length of exclusion conditional on being excluded, number of votes cast by excluded members against other group members once they have been readmitted in their group, payoffs of all group members and of non-excluded members separately.

Table 1: Summary statistics

	Baseline	Exo1P	Exo3P	Endo
Average withdrawal (All members)	16.94 (5.40)	7.81*** (8.20)	10.65*** (7.56)	6.43*** (8.60)
Average withdrawal of non-excluded members	16.94 (5.40)	8.39*** (8.01)	5.74*** (7.30)	6.50*** (8.28)
Average withdrawal of excluded members	-	16.43 (5.39)	12.31*** (7.24)	15.86 (5.65)
Average number of votes in favor of exclusion (with a max. of 12)	-	3.06 (3.19)	1.58*** (2.14)	1.68*** (2.32)
Average number of active group members in favor of exclusion (with a max. of 4)	-	0.29 (0.14)	0.20*** (0.15)	0.21*** (0.19)
Average number of excluded subjects (with a max. of 4)	-	0.88 (1.30)	0.33*** (1.07)	0.43*** (1.25)
Average length of exclusion (nb periods)	-	1 (0.00)	3*** (0.00)	2.08*** (0.83)
Average number of exclusion votes cast by excluded members after their reintegration	-	1.64 (0.92)	1.29*** (0.95)	1.35 (0.81)
Average payoff (in ECU)	21.93 (4.68)	21.58* (8.16)	22.55 (8.62)	22.51 (8.69)
Average payoff of non-excluded members (in ECU)	21.93 (4.68)	24.82*** (6.09)	26.23*** (6.02)	26.15*** (6.22)
Number of observations	1200	960	1040	960

*Notes:* Standard deviations are in parentheses. The Table also reports the results of Mann-Whitney two-tailed tests comparing each treatment with the Baseline, or comparing Exo1P and Exo3P to Endo when the Baseline is omitted. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ .

## 4.1 Cooperation and Efficiency

We introduce our first result:

**Result 1:** *The existence of an exclusion mechanism helps sustain cooperation and does not harm social welfare.* This supports Hypothesis 1.b and rejects Hypothesis 1.a.

*Support for Result 1.* Figure 1 displays the evolution of withdrawal over time by treatment, as measured by the average number of ECU taken from the group account by each active group member (we ignore the excluded members, by definition). We can easily see that subjects take more ECU in the Baseline than in any other treatment and they do so right from the first period and then, throughout the game. As also visible in Table 1, the level of withdrawal by active members is lower on average in treatments in which exclusion is possible (Exo1P: 10.65 ECU; Endo: 7.81; Exo3P: 6.43) compared to the Baseline (16.94 ECU) (Mann-Whitney tests - MW, hereafter: Baseline *vs.* Endo:  $p < 0.001$ ; Baseline *vs.* Exo1P:  $p < 0.001$ ; Baseline *vs.* Exo3P:  $p < 0.001$ )<sup>14</sup>, which is in line with the literature on punishment. Considering period 1 in isolation, the mean level of withdrawal is also lower in the treatments with possible exclusion (Exo1P: 8.16 ECU; Endo: 7.56; Exo3P: 7.08) compared to the Baseline (12.14 ECU). These differences are significant (MW tests: Baseline *vs.* Exo1P:  $p < 0.001$ ; Baseline *vs.* Endo:  $p < 0.001$ ; Baseline *vs.* Exo3P:  $p < 0.001$ ), confirming previous findings that the threat of exclusion itself plays a role in enhancing cooperation even before getting any information about others' behavior in the group.

As suggested by Figure 1, we find significant differences in the level of withdrawal depending on the length of exclusion when comparing the Exo1P and the Exo3P treatments (MW test:  $p < 0.001$ ), which supports our hypothesis that the length of exclusion matters. However, we find no significant difference between the other treatments with possible exclusion (MW tests: Exo1P *vs.* Endo:  $p = 0.100$ ; Exo3P *vs.* Endo:  $p = 0.320$ ). Interestingly, in the first period there is no significant difference between Exo1P (12.27 ECU), Exo3P (13.19 ECU) and Endo (12.44 ECU) (MW tests: Exo1P *vs.* Endo:  $p = 0.860$ ; Exo1P *vs.* Exo3P:  $p = 0.346$ ; Endo *vs.* Exo 3P:  $p = 0.439$ ). The threat of a harsher punishment in some treatments does not affect the initial withdrawal decisions. This indicates that the effect of possible exclusion in terms of cooperation at the beginning of the game depends on whether there is a punishment mechanism available but not on the severity of this punishment, which differs from Neuhofer and Kittel (2015). This discrepancy could be explained by the fact that the difference in the threat between a three-period exclusion and a one-period exclusion is much smaller than the difference between irreversible exclusion and a one-period exclusion. Overall, our results highlight the role of the risk of exclusion in sustaining cooperation and the higher cooperativeness of an environment with a three-period exclusion compared to an environment with

---

<sup>14</sup>The level of independent observation for the Mann-Whitney tests is the mean number of ECU withdrawn in all periods per group. All tests are two-tailed.

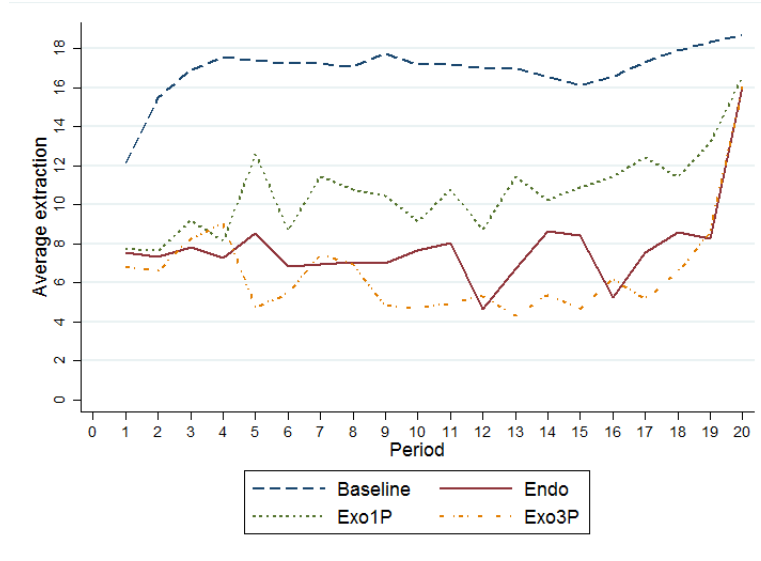


Figure 1: Evolution of average withdrawal over time, by treatment (including all the subjects).

one-period exclusion when the length of exclusion is exogenous.

Regarding efficiency, as indicated in Table 1, the average payoffs in the Endo (22.51 ECU) and Exo3P (22.55 ECU) treatments are higher than in the baseline (21.93 ECU) but these differences are not statistically significant (MW tests: Endo *vs.* Baseline:  $p=0.276$ ; Exo3P *vs.* Baseline:  $p=0.608$ ; Exo3P *vs.* Endo:  $p=0.825$ ). Interestingly, participants in the Exo1P treatment are worse off on average than in the absence of punishment mechanism (21.58 ECU) and the difference is significant (MW test: Baseline *vs.* Exo1P:  $p=0.048$ ). However, exclusion significantly increases the payoffs of the non-excluded members (21.93 ECU in the Baseline, 26.15 ECU in Endo, MW test:  $p < 0.001$ ; 24.82 ECU in Exo1P,  $p < 0.001$ ; and 26.23 ECU in Exo3P,  $p < 0.001$ ). Hence, exclusion does not systematically harm efficiency when it is followed by reintegration. This result is consistent with previous findings (Cinyabuguma *et al.*, 2005; Sheremeta *et al.*, 2011; Neuhofer and Kittel, 2015). Moreover, the possibility to exclude group members rebalances the distribution of earnings in favor of the more cooperative players since, as shown in the next subsection, cooperative players are less likely to be excluded.

## 4.2 Determinants of Exclusion

We next analyze the exclusion patterns. We first consider the decision to exclude before examining the choice of the length of exclusion in Endo. Our second result is the following:

**Result 2:** *Individuals are willing to punish and generally they exclude the least cooperators.* This supports Hypothesis 2.b.

*Support for Result 2.* There are more people voting in favor of exclusion in Exo 1P (0.29) than in Endo (0.20) and Exo 3P (0.21) and the differences are significant (MW tests: Exo1P *vs.* Endo:  $p=0.007$ ; Exo1P *vs.* Exo3P:  $p=0.002$ ). The difference between Endo and Exo3P is not significant (MW test:  $p=0.956$ ). This corresponds to an overall percentage of group members in favor of exclusion of 26.16% in Exo1P, 14.20% in Exo3P and 15.46% in Endo. These results remain the same when excluding the last three periods where long exclusion is no longer possible in Exo3P and in Endo (MW tests: Exo1P *vs.* Endo:  $p=0.001$ ; Endo *vs.* Exo3P:  $p=0.912$ ; Exo1P *vs.* Exo3P:  $p=0.006$ ). As a result, there are more people excluded in Exo 1P (0.88) than Endo (0.43) and Exo 3P (0.33) and the differences are significant (MW tests: Exo1P *vs.* Endo:  $p<0.001$ ; Exo1P *vs.* Exo3P:  $p<0.001$ ). The difference between Endo and Exo3P is not significant (MW test:  $p=0.150$ ). This corresponds to an overall percentage of excluded subjects of 21.29% in Exo1P, 22.69% in Exo3P and 22.5% in Endo. These results remain the same when excluding the last three periods where long exclusion is no longer possible in Exo3P and in Endo (MW tests: Exo1P *vs.* Endo:  $p<0.001$ ; Endo *vs.* Exo3P:  $p=0.149$ ; Exo1P *vs.* Exo3P:  $p<0.001$ ). These statistics support the general findings from the previous literature that individuals are willing to exclude others even when it is costly to them. A novel finding is that when subjects cannot adjust the severity of sanctions, this severity does not impact the willingness to exclude others.

Subjects excluded for one period withdrew on average 16.43 ECU in Exo1P and 15.44 in Endo; those excluded for three periods withdrew 12.31 ECU in Exo3P and 14.9 in Endo; and those excluded for two periods in Endo withdrew 17.5 ECU. Figure 2 displays the mean rate of exclusion of individuals depending on the difference between their withdrawal and the mean withdrawal in their group (excluding themselves) that is usually considered in the literature as a measure of the norm prevailing in the group. Each bar represents the proportion of members who, among those subjects who extracted  $x$  to  $y$  ECU more or less than their group average, have been excluded in each treatment. For example, the bars on the far right of Figure 2 show the proportion of excluded members among those who withdrew 10 to 20 ECU more than their group average. The numbers above the bars indicate the number of observations. We can see that group members who extract on average 10 to 20 ECU more than the average of their group members are more likely to be excluded than group members who extract on average less than 5 ECU more than their group average. This is observed in all treatments. The left-hand side of the figure shows evidence of anti-social exclusion: some subjects are excluded although they withdrew less than their group average. This is all the more surprising in our context as exclusion requires at least two votes from other group members. This anti-social behavior will be analyzed in details in section 4.4.

To study the determinants of exclusion, we proceed to a regression analysis in which the dependent variable is the voting decision of subject  $i$  to exclude subject  $j$  in period  $t$ . Table 2 reports the marginal effects of the estimates of the vote for exclusion of a group member, using random-effects logit regressions with standard

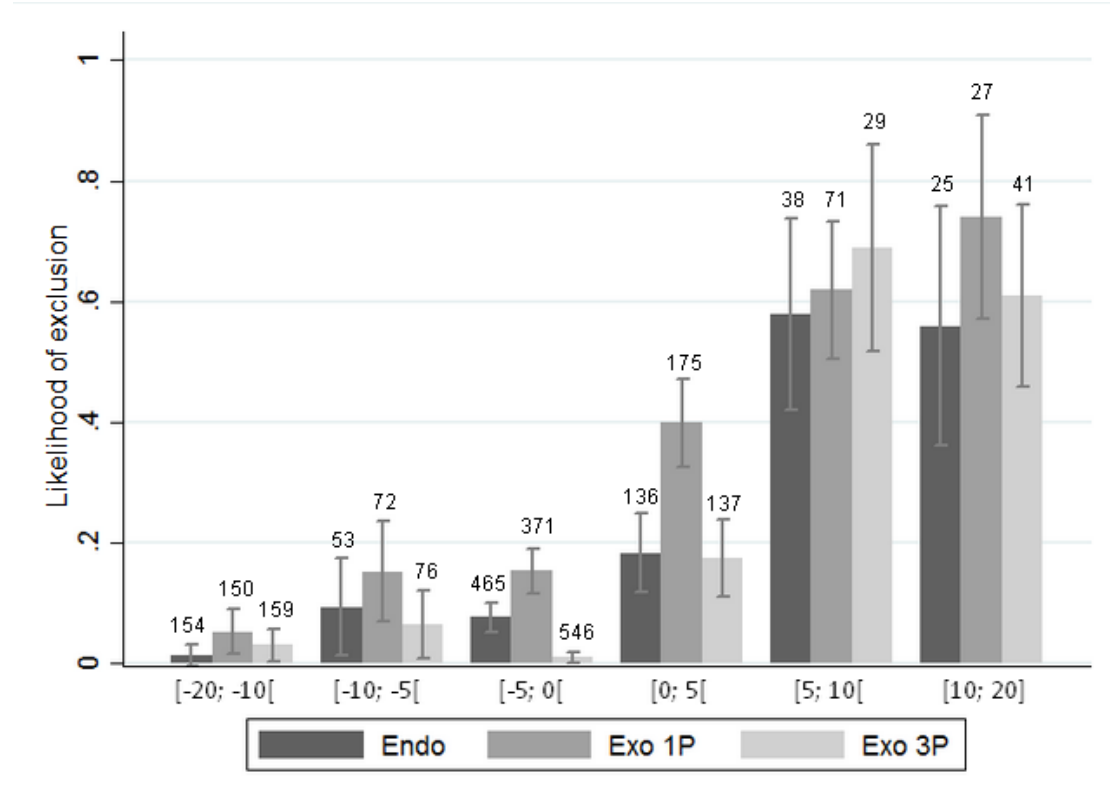


Figure 2: Exclusion rate according to the withdrawal deviation to the group average, by treatment.

errors clustered at the group level. We only consider the observations in which exclusion is possible: we exclude the cases in which  $i$  or  $j$  are already being excluded and the observations corresponding to groups smaller than three active members. We only consider periods 1 to 17 because long exclusion is no longer possible in Endo and Exo3P after period 17.

In addition to treatment dummies, the independent variables include the negative deviation of  $j$ 's withdrawal to the average withdrawal of the group (excluding  $j$ ) during period  $t$  ( $\max(0; w_{j,t} - \sum_{-j=1, -j \neq j}^{n-1} w_{-j,t})$ ) and the positive deviation to the average withdrawal of the group during period  $t$  ( $\max(0; \sum_{-j=1, -j \neq j}^{n-1} w_{-j,t} - w_{j,t})$ ). Therefore, the negative deviation variable takes a positive value equal to the difference between the number of ECU taken by  $j$  and the mean number of ECU withdrawn by the remaining members of her group when  $j$  takes more ECU than her group average, and 0 otherwise. Conversely, the positive deviation variables takes a positive value equal to the difference between the mean number of ECU taken by the remaining group members and the number of ECU withdrawn by  $j$  when  $j$  takes less ECU than her group average, and 0 otherwise. The independent variables also include a control for whether  $j$  has already been excluded in the past to identify possible harassment strategies, controlling for withdrawal decisions. They include a reintegration dummy that equals 1 if  $i$  is reintegrated in  $t$  and 0 otherwise. In model (2) we add interaction terms between this reintegration dummy

and the treatment dummies. The variable “ $i$  is reintegrated in  $t$  after  $x$ P exclusion (Y)” takes value 1 if  $i$  has been reintegrated in  $t$  after  $x$  period(s) of exclusion, in the Y treatment, and 0 otherwise. This allows us to test whether subjects who have been recently punished are more severe or more forgiving than non excluded members. We also control for the group size, for time, and for individual characteristics (gender, age, student status and cooperation type of  $i$ )<sup>15</sup>.

Table 2 indicates that the more ECU  $j$  takes over her group average, the more likely subject  $i$  is to vote in favor of her exclusion. The probability to vote to exclude a group member is the same across treatments, regardless of the severity of punishment. In these models we assume that the positive and negative deviations in withdrawal from the group average has the same impact across treatments. We relax this assumption in additional models reported in Appendix Table A2, in which the independent variables include not only the treatments and the positive and negative deviations in withdrawal but also interaction terms between the treatments and these deviations. The only significant interaction effect is the positive deviation in the Endo treatment: subjects who withdraw less than the average are less likely to be excluded in Endo than in Exo1P.

Table 2 also shows that  $i$  is more likely to vote to exclude another group member in  $t$  when she has been herself reintegrated in  $t$ . This suggests that there is some retaliation going on. Model (2) indicates that  $i$  is less likely to vote to exclude another group member when she has been excluded for one period in Endo compared to when she has been excluded in Exo1P. Therefore, when it results from a vote a short exclusion affects differently the decision to exclude after reintegration.

In the Endo treatment, subjects who vote for the exclusion of a group member have also to choose the length of exclusion. 30.77% of the excluded members were excluded for one period, 30.77% for two periods and 38.46% for three periods (the average length of exclusion is 2.08 periods, see Table 1). We investigate the determinants of the choice between one, two or three periods of exclusion by estimating random-effects ordered logit regressions. Our dependent variable is a categorical variable that equals 3 if individual  $i$  votes to exclude  $j$  for three periods, 2 if  $i$  votes

---

<sup>15</sup>Following the methodology of Fischbacher *et al.* (2001), we assign a cooperation type to each subject, depending on her conditional decisions made in part 1 of the session. We created five dummies, one for each cooperation type: unconditional cooperators, conditional cooperators, unconditional free-riders, individuals with U-shaped preferences, and individuals who cannot be classified in any of the previous categories. A subject is classified as an unconditional free-rider if she chose to withdraw the maximum possible amount of ECU regardless of her group members’ average withdrawal decision. A subject is classified as an unconditional cooperator if she chose to take 0 ECU regardless of her group members’ average withdrawal decision. A subject is classified as a conditional cooperator if her withdrawal decision increases when the average withdrawal decision of her group members also increases, and vice-versa. A subject with U-shaped preferences is a conditional cooperator above a certain average level of withdrawal by group members but who free-rides while others are on average more cooperative below this threshold. Subjects whose withdrawal profile does not fit in any of the previous categories is classified in the last category. None of these categories were ever significant in any regression, and thus we do not comment on them in the text. Most of the time it is similar with the other individual characteristics.

Table 2: Determinants of an exclusion vote

Dep. variable: Decision to exclude $j$ in $t$	All treatments (1)	with exclusion (2)
Exo1P	<i>Ref.</i>	<i>Ref.</i>
Exo 3P	-0.052 (0.061)	-0.045 (0.060)
Endo	-0.060 (0.063)	-0.074 (0.060)
Max $(0; w_{j,t} - \sum_{-j=1, j \neq -j}^{n-1} w_{-j,t})$	0.044*** (0.010)	0.043** (0.010)
Max $(0; \sum_{-j=1, j \neq -j}^{n-1} w_{-j,t} - w_{j,t})$	0.003 (0.003)	0.004 (0.003)
Exclusion of $j$ prior to $t$	0.001 (0.020)	0.002 (0.020)
$i$ is reintegrated in $t$	0.113*** (0.038)	0.094** (0.043)
$i$ is reintegrated in $t$ after 1P exclusion (Exo1P)	-	Ref.
$i$ is reintegrated in $t$ after 1P exclusion (Endo)	-	-0.088*** (0.030)
$i$ is reintegrated in $t$ after 2P exclusion (Endo)	-	0.267 (0.212)
$i$ is reintegrated in $t$ after 3P exclusion (Endo)	-	0.301 (0.236)
$i$ is reintegrated in $t$ after 3P exclusion (Exo3P)	-	-0.038 (0.029)
Group size	-0.016 (0.019)	-0.016 (0.019)
Period	0.002 (0.001)	0.001 (0.002)
Individual controls	Yes	Yes
Observations	4198	4198
Log pseudo-likelihood	-1325.23	-1307.26

*Notes:* The Table reports the marginal effects from random-effects logit regressions with robust standard errors clustered at the group level in parentheses. We consider only: *i*) observations from periods 1 to 17; *ii*) groups with more than two active members, as it is not possible to exclude when the group has less than three active members; *iii*) cases where  $i$  and  $j$  are not excluded (as  $i$  cannot vote if she is excluded and  $i$  cannot vote to exclude  $j$  if  $j$  is already excluded). \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ .

to exclude  $j$  for two periods, and 1 if  $i$  votes to exclude  $j$  for one period. Our



independent variables are the same as in Table 2.<sup>16</sup> Table 3 reports the marginal effects of the independent variables on the vote to exclude for one period in models (1) and (2), for two periods in models (3) and (4), and for three periods in models (5) and (6).

Table 3 shows that the vote for a specific length of exclusion is not directly affected by the number of ECU taken by  $j$  relative to the average withdrawal in the group. This may be partly driven by the fact that a fraction of the exclusion votes are motivated by retaliation; in that case even votes for a long exclusion may be directed against the more cooperative subjects. There is a marginally significant effect of the reputation of the individual who is punished. When  $j$  has already been excluded in the past,  $i$  is less likely to choose to exclude her for one period (column (1)) and more likely to choose to exclude her for three periods (column (5)). This suggests that the severity of the punishment depends partly on the adjustment of behavior after a previous exclusion. However, the significance of this variable is lost when we control for the length of exclusion of the punisher (columns (2) and (6)). These models also indicate that  $i$  is more likely to vote to exclude another group member for three periods and less likely to exclude for one or two periods when herself was reintegrated in period  $t$  after two or three periods of exclusion compared to a reintegration in the same period but after only one period of exclusion, expressing the willingness to retaliate severely in reaction to a non-minimal length of exclusion for oneself.

To summarize, these results indicate that if the decision to exclude is generally motivated by the willingness to punish someone who has extracted more than the norm of the group, individuals in the Endo treatment vote for a more severe exclusion to punish recidivism but also to take revenge after a more severe punishment.

### 4.3 Reintegration

The analysis of reintegration patterns delivers two main findings.

**Result 3:** *Exclusion for three periods improves cooperation in the reintegration period compared to exclusion for one-period in Exo treatments but not in the Endo treatment.* This supports Hypothesis 3 and Hypothesis 5(a).

**Result 4:** *The positive effect of exclusion for one period on compliance after reintegration is larger when the length of exclusion results from a vote by the group compared to when the length of exclusion cannot be modulated and is fixed exogenously.* This supports Hypothesis 5(b).

*Support for Result 3.* Figure 3 displays the difference between the number of ECU

---

<sup>16</sup>We estimated the same model using random-effects GLS after we applied a 2-step Heckman correction, in order to control for a possible selection bias and for unobserved heterogeneity between group members who vote for exclusion and those who do not vote for exclusion. Since the inverse Mills ratio was not significant, indicating that there is no need to control for a selection bias, we do not report these estimates.

Table 3: Determinants of the length of exclusion in the Endo treatment

Dep. variable: Choice of $i$ regarding the length of exclusion of $j$	1 period (1)	1 period (2)	Exclusion for 2 periods (3)	2 periods (4)	3 periods (5)	3 periods (6)
Max $(0; w_{j,t} - \sum_{-j=1, j \neq -j}^{n-1} w_{-j,t})$	-0.005 (0.005)	-0.007 (0.006)	-0.002 (0.002)	-0.002 (0.002)	0.007 (0.007)	0.009 (0.007)
Max $(0; \sum_{-j=1, j \neq -j}^{n-1} w_{-j,t} - w_{j,t})$	-0.002 (0.011)	-0.002 (0.012)	-0.001 (0.004)	-0.001 (0.004)	0.002 (0.015)	0.003 (0.016)
Exclusion of $j$ prior to $t$	-0.097* (0.050)	-0.087 (0.055)	-0.032 (0.028)	-0.025 (0.029)	0.130* (0.075)	0.112 (0.081)
$i$ is reintegrated in $t$	-0.127* (0.074)	0.110 (0.099)	-0.041*** (0.015)	0.031 (0.037)	0.169** (0.080)	-0.142 (0.131)
$i$ is reintegrated in $t$ after 2P exclusion (Endo)	-	-0.349** (0.178)	-	-0.099 (0.062)	-	0.451** (0.194)
$i$ is reintegrated in $t$ after 3P exclusion (Endo)	-	-0.296** (0.152)	-	-0.083* (0.047)	-	0.382** (0.158)
Group size	-0.047 (0.088)	-0.042 (0.089)	-0.015 (0.030)	-0.012 (0.027)	0.063 (0.118)	0.054 (0.116)
Period	-0.002 (0.002)	-0.001 (0.002)	-0.001 (0.001)	-0.001 (0.001)	0.002 (0.002)	0.002 (0.002)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	335	335	335	335	335	335
Log pseudo-likelihood	-267.20	-263.40	-267.20	-263.40	-267.20	-263.40

*Notes:* Table 3 reports random-effects ordered logit regressions with robust standard errors clustered at the group level in parentheses. We consider only: *i*) periods from 1 to 17; *ii*) observations from the Endo treatment in which  $i$  chose to exclude  $j$ ; *iii*) groups with more than two members; *iv*) cases where  $j$  and  $i$  are not excluded (as  $i$  cannot vote if she is excluded and  $i$  cannot vote to exclude  $j$  if  $j$  is already excluded). \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ .

taken by the excluded members and the average withdrawal of their group both in the period of exclusion (left bars) and in the period of reintegration (right bars), by treatment.<sup>17</sup> It shows that reintegrated subjects after three periods of exclusion continue to withdraw more than their group average in the reintegration period in both Endo (+1.83) and Exo3P (+2.48), whereas subjects reintegrated after one period of exclusion comply to the norm of the group in Exo1P (+0.36) and withdraw less than the group average in Endo (-0.92) but the variance is large<sup>18</sup>.

<sup>17</sup>Although not illustrated on Figure 3 to keep the comparison symmetric, group members excluded for two periods in Endo take on average 1.56 ECU more than their group members in the period of exclusion and 0.31 ECU more than their group members in the reintegration period.

<sup>18</sup>Note that the difference cannot be due to less information in the case of a longer exclusion since excluded subjects receive also feedback on decisions and earnings of the group members during each period of exclusion.

The less-than-group-average withdrawal by reintegrated subjects after a one period exclusion in Endo is particularly striking, as in the period of exclusion the deviation to the group average was higher compared to Exo1P, Exo3P and long exclusion in Endo. Precisely, in the period when they are reintegrated, the members excluded for one period in Endo take 8.73 ECU while the group average withdrawal is 9.65 ECU, whereas in Exo1P they take 15.07 ECU while their group average withdrawal is 14.71 ECU. Conversely, group members excluded for three periods in Endo take on average 14.53 ECU when they are reintegrated while the group average is 12.70 ECU, and in Exo3P they take 9.20 ECU while their group average is 6.72 ECU. Finally, when reintegrated, group members excluded for one period in Endo seem to take on average less ECU than group members excluded for three periods in the same treatment, while group members excluded in Exo1P take on average more ECU than group members excluded in Exo3P; the difference is marginally significant (MW test:  $p=0.097$ ). The differences between Endo and the other treatments are not significant (MW test: Endo *vs.* Exo1P:  $p=0.179$ ; Endo *vs.* Exo3P:  $p=0.447$ ).<sup>19</sup>

Using regressions allows us to compare the effect of exclusion on behavior after reintegration, controlling for the average withdrawal in the group in the period prior to reintegration. Table 4 reports the marginal effects from random-effects GLS regressions with robust standard errors clustered at the group level. The dependent variable is the number of ECU withdrawn by the excluded members when they reintegrate their group in the treatments with possible exclusion. The independent variables include dummy variables for each possible length of exclusion interacted with the treatment. The only difference between models (1) and (2) is the reference category: one-period exclusion in Exo1P in model (1) and one-period exclusion in Endo in model (2). We control for the average withdrawal of the other group members in  $t-1$  and for the group size in  $t^{20}$ . Finally, we include a time trend and the same individual characteristics as in the previous Tables. The data only include the observations of reintegrated subjects from periods 3 to 20 but we do not include the observations corresponding to subjects excluded after period 17 (since the length of exclusion is censored in Endo and this could bias the comparison).

Model (1) in Table 4 shows almost no significant difference between the effect of any length of exclusion on withdrawal after reintegration compared to one-period exclusion in Exo1P. The only marginally significant difference is between Exo3P and Exo1P: reintegrated subjects withdraw marginally less after three periods of exclusion than after one, controlling for the average withdrawal in the group in the period preceding immediately the reintegration, which gives some support to Hypothesis 3. A similar regression performed on the sub-sample of the two Exo treatments indicates that the variable is significant at the 5 percent level ( $N=179$ , marginal effect = 3.972, standard error = 2.001). Choosing one-period exclusion in

---

<sup>19</sup>We cannot test the significance of the differences between members excluded for one period and members excluded for three periods in Endo because observations are not independent.

<sup>20</sup>We consider the group average withdrawal in  $t-1$  and not in  $t$  because when subjects make their withdrawal decision in  $t$  they are only aware of their group members' decisions in  $t-1$ .

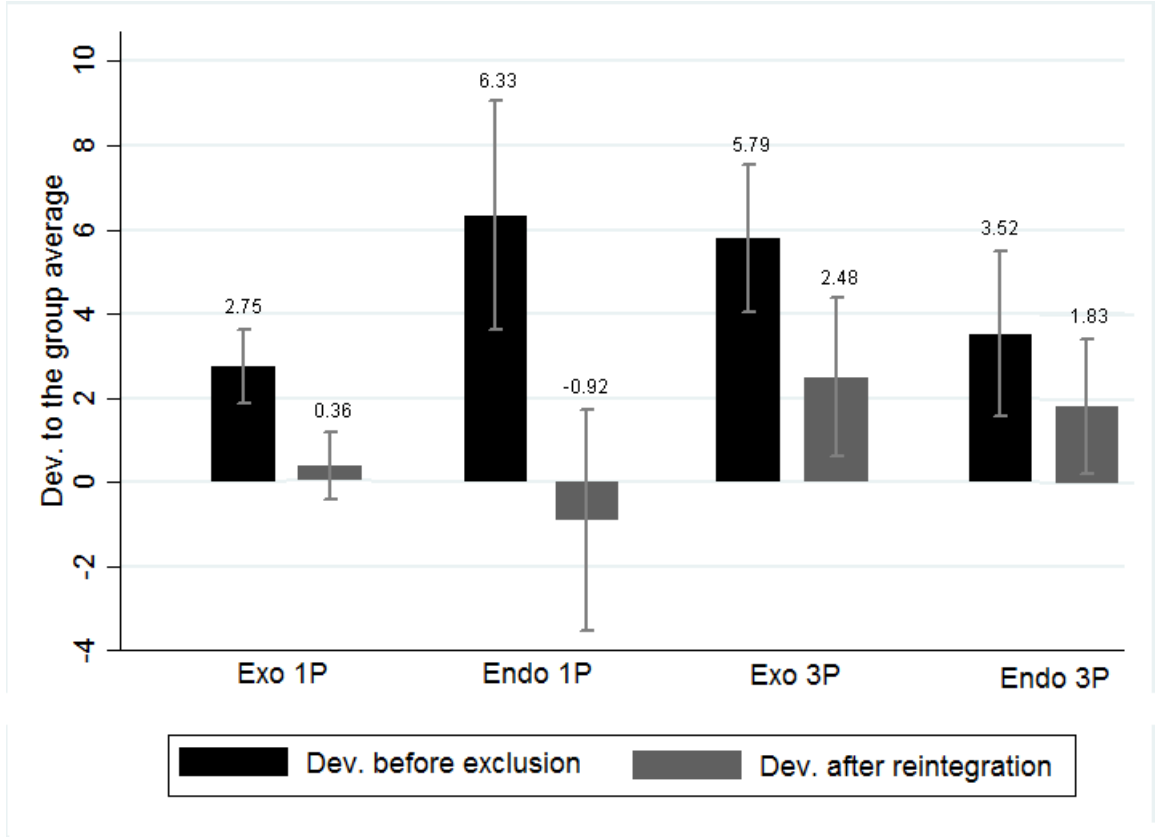


Figure 3: Difference between the number of ECU withdrawn by excluded members and by their fellow group members in the exclusion period (left bars) and in the reintegration period (right bars), by treatment.

Endo as the reference category instead of Exo1P in model (2) indicates that being excluded for two or three periods instead of one in Endo increases significantly the level of withdrawal during the reintegration period. There is no significant difference in the effect of being excluded for two or for three periods on withdrawal after reintegration ( $t$ -test:  $p=0.200$ ). This indicates that the length of exclusion does not have the same effect on withdrawal after reintegration depending on whether or not the severity of sanction can be modulated.

*Support for Result 4.* A limitation of the previous regression analysis is that we did not control for the subject's withdrawal behavior before exclusion, although it contributes to determine the conditions of exclusion. To better characterize the impact of exclusion on the evolution of withdrawal behavior, we now estimate random-effects GLS regressions with robust standard errors clustered at the group level. The dependent variable is the difference in the withdrawal level between  $t$  and  $t-2$  in model (1) or between  $t$  and  $t-4$  in model (2), for both reintegrated and non-excluded members<sup>21</sup>. If a subject has been excluded for one period (for three periods, respectively) and reintegrated in  $t$ , this indicates the difference in with-

<sup>21</sup>These regressions exclude subjects that cannot be observed in both  $t$  and  $t-2$  or  $t-4$ .

Table 4: Effect of the length of exclusion on the withdrawal (in ECU) of excluded subjects in the reintegration period

Dep. variable: number of ECU withdrawn	All treatments (1)	with exclusion (2)
1P exclusion in Exo1P	<i>Ref.</i>	3.017 (2.294)
1P exclusion in Endo	-3.017 (2.294)	<i>Ref.</i>
2P exclusion in Endo	0.492 (1.591)	3.509** (1.540)
3P exclusion in Endo	0.944 (1.734)	3.960** (1.714)
3P exclusion in Exo3P	-3.973** (1.917)	-0.956 (2.461)
Group average withdrawal excluding $i$ in $t-1$	0.286*** (0.100)	0.286*** (0.100)
Group size	-1.394 (0.944)	-1.394 (0.944)
Period	0.107 (0.091)	0.107 (0.091)
Individual controls	Yes	Yes
Constant	7.037 (4.406)	4.020 (4.436)
Observations	244	244
R-squared	0.30	0.30

*Notes:* The model estimates a random-effect GLS with robust standard errors clustered at the group level in parentheses. The data include the observations from subjects reintegrated in periods 3 to 20, excluding subjects who were excluded after period 17. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ .

drawal levels between  $t$ , the period of reintegration, and  $t-2$  ( $t-4$ , respectively), the period of exclusion. The independent variables include a dummy variable equal to 1 if the subject has been reintegrated in  $t$  after being excluded in  $t-2$  in Endo and Exo1P or in  $t-4$  in Endo and Exo3P, and 0 if not excluded. We add an interaction term between being reintegrated and the Endo treatment to measure whether the impact of exclusion differs when the length of exclusion results from a vote. We also include the group average withdrawal (excluding the individual) in  $t-1$  and the evolution of the group average withdrawal between  $t-1$  and  $t-2$  or  $t-4$ , according to the model. Indeed, even when subjects have been excluded they can observe the evolution of cooperation within their group in each period and if there is a decay of cooperation over time they can observe it. We add a time trend and controls for the same individual characteristics as before. Except for the fact that we include variables to capture the impact of reintegration after exclusion, these models are standard in the analysis of the evolution of behavior after being punished. Table

5 reports these estimates. As in Table 4, the data include the observations from subjects reintegrated in periods 3 to 20, excluding subjects who were excluded after period 17.

Table 5: Determinants of the evolution of the level of withdrawal over time

Dep. variable: Evolution of number of ECU withdrawn	$(wi_t - wi_{t-2})$ All 1P (1)	$(wi_t - wi_{t-4})$ All 3P (2)
Reintegrated after exclusion	-3.807*** (1.20)	-2.913** (1.307)
Reintegrated after exclusion *Endo	-4.116** (1.662)	-1.070 (1.679)
Group average in $t-1$	0.004 (0.030)	-0.089 (0.064)
$GA_{t-1} - GA_{t-2}$	0.082 (0.056)	-
$GA_{t-1} - GA_{t-4}$	-	0.273*** (0.106)
Period	0.077 (0.050)	0.223*** (0.062)
Individual controls	Yes	Yes
Constant	-1.047 (0.920)	-0.314 (1.346)
Observations	1339	1217
R-squared	0.03	0.07

*Notes:* Table 5 reports random-effect GLS regressions with robust standard errors clustered at the group level in parentheses. GA for group average. The data include the observations from subjects reintegrated in periods 3 to 20, excluding subjects who were excluded after period 17. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ .

Model (1) shows that controlling for the level and the evolution of average withdrawals at the group level, being reintegrated after being excluded for one period decreases the extraction level between  $t$  and  $t - 2$  and that the effect is even significantly stronger in the Endo treatment. Model (2) shows also a significant decrease in the evolution of withdrawal between  $t$  and  $t-4$  for the individuals reintegrated after being excluded for three periods compared to those who have not been excluded. However, there is no significant difference between the Exo3P and Endo treatment. Again, these results confirm that being excluded for one period has a positive impact on compliance to the group norm of withdrawal when it is possible to modulate the length of exclusion. In contrast, being excluded for three periods has the same impact on cooperation independently of how this length of exclusion

has been determined.

## 4.4 Retaliation

In section 4.2 we observed that group members who are reintegrated in  $t$  are more likely to vote to exclude another group member (and for a higher number of periods in Endo) especially if they have been excluded for three periods themselves. This suggests that voting for exclusion is not only motivated by the willingness to enforce norm compliance but is also driven by the willingness to retaliate<sup>22</sup>. Hence, we now focus on retaliation behavior by examining more directly the voting behavior of the excluded members after their reintegration, depending on how long they have been excluded and by which procedure. Our last results are the following.

**Result 5:** *When the length of exclusion is endogenous, subjects who have been excluded for a shorter duration are less likely to seek revenge when they reintegrate their group than subjects excluded for a longer duration; no such effect is found when the length of exclusion is imposed exogenously.* This rejects Hypothesis 4 and supports Hypothesis 6(a).

**Result 6:** *Subjects excluded for three periods (one period) seek only marginally (significantly at the 1 percent level, resp.) more (less, resp.) revenge when the length of exclusion is endogenous rather than exogenous.* This supports only weakly Hypothesis 6(b).

*Support for Results 5 and 6.* The left panel of Figure 4 displays the average number of exclusion votes cast by excluded members when they reintegrate their group, by treatment. They can cast between 0 and 3 exclusion votes. Subjects excluded for one period in Endo cast on average 0.58 exclusion vote, whereas those excluded for three periods cast on average 1.89 vote. In contrast, subjects excluded in Exo1P cast on average 1.64 exclusion vote and those excluded in Exo3P cast on average 1.29 exclusion vote. The differences between Exo3P and the other treatment is significant (MW tests: Exo3P *vs.* Exo 1P:  $p=0.003$ ; Exo3P *vs.* Endo:  $p=0.076$ ). The difference between Endo and Exo1P is not significant (MW test:  $p=0.202$ ).

The right panel of Figure 4 displays the share of participants who were previously excluded and who exert total revenge when reintegrating their group. "Total revenge" is defined as the fact for a previously excluded subject to vote to exclude all the other active group members after reintegration. Such behavior is blind except in the case when  $N=3$  since in that case the subject's exclusion necessarily results from the vote of the two other group members. In Exo1P 39.05% of the reintegrated

---

<sup>22</sup>Moreover, in a post-experimental questionnaire we asked subjects to comment on their motivation if they voted at least once to exclude a fellow group member. Most responded that they voted in favor of exclusion to punish free-riders. However, 25% of them reported that they were willing to retaliate from previous exclusion.

subjects vote for excluding all the other group members. This percentage is only 12.5% for reintegrated participants excluded for one period in Endo. In contrast, it is 21.18% in Exo 3P but it rises to 42.5% for those excluded for three periods in Endo. The percentage in Exo1P is significantly higher than in Endo (30.77%, regardless of the length of exclusion) (two-tailed two-sample test of proportions:  $p < 0.001$ ) and in Exo3P ( $p < 0.001$ ). The difference between Exo3P and Endo is also significant ( $p = 0.022$ ). As before, we cannot test the comparisons involving the split samples of Endo since the observations are not independent<sup>23</sup>. Both panels show that one-period exclusion limits counter-punishment in Endo, but it is not the case when the length of exclusion is imposed exogenously.

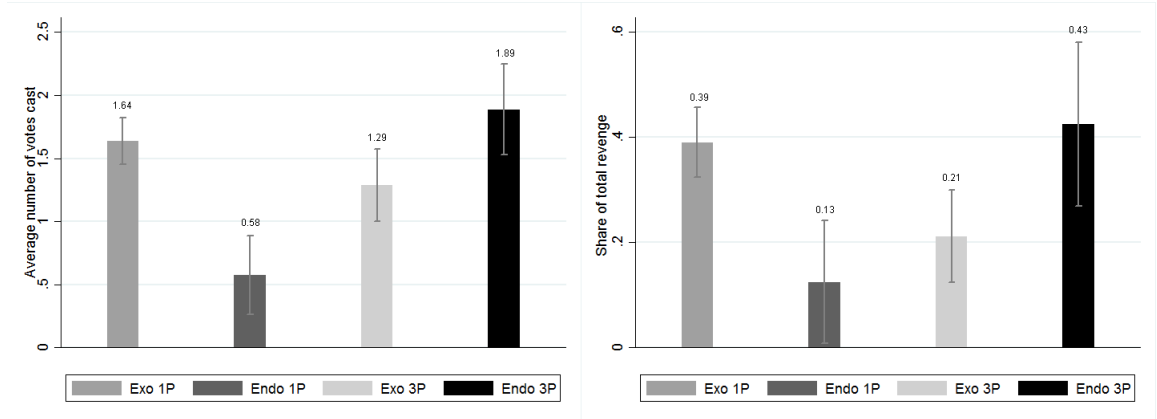


Figure 4: Average number of exclusion votes cast (left panel) and share of total revenge (right panel) by excluded members in the reintegration period, by treatment

Using regressions allows us to identify the effect of exclusion for one *vs.* three periods in Endo on retaliation, and to control for the average withdrawal level in the group and for the size of the group that also potentially influence the number of exclusion votes.

The first model reported in Table 6 estimates a random-effects Poisson regression with standard errors clustered at the group level. The sample pools the data from all the treatments with exclusion but includes only the observations from subjects who have been previously excluded and reintegrate their group in the current period and who can vote to exclude others as long as they like; therefore, we only consider periods 3 to 17 and groups where exclusion is possible ( $N > 2$ ). The dependent variable is a categorical variable equal to 3 if the subject casts three exclusion votes when she reintegrates her group, 2 if she cast two votes, 1 if she casts only one vote, and 0 if the subject does not want to exclude any group member. The independent variables include dummy variables for each possible length of exclusion of subject  $i$  interacted with each treatment, setting Exo1P as the reference category. We control for the amount withdrawn by the subject, the average amount

<sup>23</sup>While not illustrated in Figure 4 to keep the comparison symmetric, group members excluded for two periods in Endo cast on average 1.37 exclusion votes and 34.38% of them perform total revenge after reintegration.



withdrawn by the other group members in period  $t$  (since when they vote subjects know the current level of withdrawal in their group), the group size and the usual individual characteristics.

Table 6: Effect of the length of exclusion on the number of exclusion votes and on the likelihood of total revenge by reintegrated subjects

Dep. variables:	Nb of exclusion votes (1)	Total revenge (2)
1P exclusion in Exo1P	<i>Ref.</i>	<i>Ref.</i>
1P exclusion in Endo	-0.646*** (0.190)	-0.289*** (0.071)
2P exclusion in Endo	-0.064 (0.195)	-0.007 (0.165)
3P exclusion in Endo	0.205** (0.113)	0.074 (0.141)
3P exclusion in Exo3P	0.079 (0.205)	-0.046 (0.152)
Amount withdrawn by $i$ in $t$	0.010 (0.007)	0.016** (0.007)
Group average withdrawal in $t$ excluding $i$	0.030** (0.013)	0.020* (0.011)
Group size	0.456*** (0.123)	0.009 (0.105)
Period	-0.007 (0.009)	0.010 (0.008)
Individual controls	Yes	Yes
Constant	-1.582*** (0.612)	-0.090 (0.166)
Observations	323	323
Log pseudo-likelihood	-479.72	-169.23

*Note:* Model (1) reports marginal effects from a random-effects Poisson regression of the number of exclusion votes cast by excluded members when they reintegrate their group. Model (2) reports marginal effects from a random-effects logit regression of the probability of excluded members when reintegrated in their group to vote for total revenge (for the exclusion of all other group members). We retain only the observations for the groups that contain at least three members and the observations from periods 3 to 17. In all models, standard errors (in parentheses) are clustered at the group level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ .

Model (1) shows that group members reintegrated after one period of exclusion in Endo cast significantly less exclusion votes than group members excluded in Exo1P. They are less likely to retaliate when group members chose a short exclusion when they could have chosen a longer duration. Group members reintegrated

after three periods in Endo cast more exclusion votes than members excluded in Exo1P. Since there is no significant difference between Exo1P and Exo3P, this indicates that individuals excluded for a longer time are marginally more likely to seek revenge when group members chose the length than when it was imposed to them, supporting Result 6. A longer exclusion increases significantly the number of exclusion votes cast by reintegrated members only in Endo<sup>24</sup>.

To explore the effect of the length of exclusion on the likelihood of voting for total revenge, model (2) reported in Table 6 estimates a random-effects logit regression with standard errors clustered at the group level. The dependent variable is a dummy variable equal to 1 if the excluded member votes for the exclusion of all the other active group members when she reintegrates her group, and 0 otherwise. The independent variables are the same as in model (1).

Controlling for the individual and the group average withdrawals, model (2) shows that group members excluded for one period in Endo are significantly less likely to vote for the exclusion of all the other group members when reintegrating the group compared to similar subjects in Exo1P, with no other significant treatment differences. Consistent with model (1), this indicates that group members excluded for one period are less likely to retaliate systematically when the length of exclusion has been chosen than when it is fixed exogenously. There is no significant difference in the likelihood of total revenge between group members excluded for one period or for three periods when the length of exclusion is fixed exogenously. In contrast, subjects excluded for two or three periods are significantly more likely to use total revenge than group members excluded for one period when the length of exclusion resulted from a vote, supporting Result 5<sup>25</sup>. To conclude, subjects excluded for a longer time are more likely to seek revenge than people excluded for one period but only when the length of exclusion results from a vote. When it is fixed exogenously, the length of exclusion does not affect the likelihood of counter-punishment. Thus, it is not the length of exclusion that matters so much, but the intention of the punishers to harm more severely behind the choice of a longer exclusion.

## 5 Discussion and Conclusion

In the context of social dilemmas, understanding how punishment institutions work to limit free-riding is important but challenging. In particular, if the threat of exclusion is not sufficient to deter free-riding, it is important to know whether actual exclusion does more good than bad. How do people behave when they are reintegrated in their group after being excluded from that group by their fellow members,

---

<sup>24</sup>Consistently, a similar regression conducted only on the data of the Endo treatment indicates that the marginal effect of a two-period exclusion is 0.750\*\*\* (0.287) and the marginal effect of a three-period exclusion is 1.024\*\*\* (0.267), compared to a one-period exclusion.

<sup>25</sup>Consistently, a similar regression conducted only on the data of the Endo treatment indicates that the marginal effect of a two-period exclusion is 0.673\*\*\* (0.203) and the marginal effect of a three-period exclusion is 0.685\*\*\* (0.154), compared to a one-period exclusion.

both in terms of cooperation and retaliation? Do behavior adjustments depend on the length of exclusion and its mode of determination? How are these potential effects taken into account by group members who are willing to use punishment to increase cooperation and norm compliance? To address these questions we have designed an experiment based on a negatively framed public goods game in which the length of possible exclusion is either fixed exogenously or chosen by group members through a vote among three options.

Our findings first replicate the previous literature on ostracism showing that people are willing to exclude other members from their group even if it is costly to them. People tend to exclude the group members who deviate the most from the average withdrawal from the public goods. Then, we find some evidence that the severity of the sanction is influenced by the reputation of a group member. People adopt a second-chance strategy: they are more likely to forgive when a peer defects for the first time and to choose a one-period exclusion in that case, but they punish more harshly subsequent defections.

Contrary to previous experiments on punishment, we found that when exclusion is followed by reintegration, cooperation is sustainable without damaging social welfare even in the short run. When the length of exclusion cannot be modulated, members excluded for three periods cooperate more when they return compared to one-period exclusion. By endogeneizing the length of exclusion, we have shown that when the length of exclusion results from a vote and can be adjusted to the importance of the deviation to the group average withdrawal, a shorter length has a positive effect on excluded participants' behavior when they reintegrate their group. Members excluded for one period comply more to their group average after reintegration and they are less likely to seek revenge. On the opposite, subjects excluded for the maximum length are less likely to comply with the group norm when they return and they are more likely to adopt anti-social behavior by trying to exclude those who excluded them.

Our results highlight the joint importance of adaptive punishment and a quick reintegration of the individuals who have been excluded to sustain cooperation and avoid anti-social punishment strategies. Short exclusion has a positive effect when it can be interpreted as forgiving, which is less the case when there is only one possible length of exclusion. Hence, second-chance policy interventions may limit counter-punishment and help reintegration. This echoes the results of Coricelli *et al.* (2014) who tested experimentally the 'reintegrative shaming' theory (Braithwaite, 1989) and showed that the pardon of tax evaders reduces future tax evasion compared to the case in which there is no pardon for offenders who, once their reputation is lost, have less motivation to comply with the rule.

We acknowledge, however, a number of limitations of our study. The positive effect of short exclusion on cooperation in our Endo treatment may be driven by two effects. First, subjects have been excluded for a short duration and they know that they could have been punished more severely, which may relieve them. Second,

the length of exclusion results from a vote of group fellows, and this “democratic” expression of disapproval may facilitate its acceptance. Our design does not allow us to disentangle these two possible effects. Similarly, the counter-effects of a long exclusion may result from two effects. First, some people can find that the punishment was too harsh, considering that a shorter sentence was available. This may have raised negative emotions against their group members who could have signaled their disapproval with a milder sanction. Second, the possibility to select different lengths of exclusion help screen the free-riders. Those excluded for a longer time are more deviant and less compliant. They are more likely to behave selfishly and even aggressively towards group members after reintegration. In that case, it is not so much the negative emotions triggered by exclusion, but the screening permitted by the choice of the length of exclusion that would drive our results. In our regressions we controlled for the type of social preferences of the subject and we did not find a significant effect of any category of social preferences on behavior. But this may result from the fact that severe punishment is directed both toward free-riders and toward cooperators. Therefore, it is difficult to weight each of these two possible explanations precisely. A random assignment of the length of exclusion and direct measures of emotions could help isolate and weight the first effect. We leave this to further investigation.

## References

- [1] AMBRUS, A., AND GREINER, B. Democratic punishment in public good games with perfect and imperfect observability, 2015.
- [2] BECKER, G. S. *Crime and punishment: An economic approach*. Springer, 1968, pp. 13–68.
- [3] BLUMSTEIN, A., COHEN, J., AND NAGIN, D. *Deterrence and incapacitation: Estimating the effects of criminal sanctions on crime rates*. National Academy of Sciences, Washington, DC, 1978.
- [4] BOCK, O., BAETGE, I., AND NICKLISCH, A. Hroot: Hamburg registration and organization online tool. *European Economic Review* 71 (2014), 117–120.
- [5] BOLLING, D., PITSKEL, N., AND DEEN, B. E. A. Dissociable brain mechanisms for processing social exclusion and rule violation. *Neuroimage* 54 (2011), 2462–2471.
- [6] BOWLES, S., AND GINTIS, H. The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theoretical Population Biology* 65, 1 (2004), 17–28.
- [7] BRAITHWAITE, J. *Crime, shame and reintegration*. Cambridge University Press, 1989.

- [8] CHARNES, G., AND YANG, C.-L. Starting small toward voluntary formation of efficient large groups in public goods provision. *Journal of Economic Behavior and Organization* 102 (2014), 119–132.
- [9] CINYABUGUMA, M., PAGE, T., AND PUTTERMAN, L. Cooperation under the threat of expulsion in a public goods experiment. *Journal of Public Economics* 89, 8 (2005), 1421–1435.
- [10] CORICELLI, G., RUSCONI, E., AND VILLEVAL, M. C. Tax evasion and emotions: An empirical test of re-integrative shaming theory. *Journal of Economic Psychology* 40 (2014), 49–61.
- [11] DAVIS, B. J., AND JOHNSON, D. B. Water cooler ostracism: Social exclusion as a punishment mechanism. *Eastern Economic Journal* 41, 1 (2015), 126–151.
- [12] DICKINSON, D. L., MASCLÉT, D., AND VILLEVAL, M. C. Norm enforcement in social dilemmas: An experiment with police commissioners. *Journal of Public Economics* 126 (2015), 74–85.
- [13] DRAGO, F., GALBIATI, R., AND VERTOVA, P. The deterrent effects of prison: Evidence from a natural experiment. *Journal of political Economy* 117, 2 (2009), 257–280.
- [14] EISENBERGER, N., LIEBERMAN, M., AND WILLIAMS, K. D. Does rejection hurt? an fmri study of social exclusion. *Science* 302 (2003), 290–292.
- [15] FALK, A., FEHR, E., AND FISCHBACHER, U. On the nature of fair behavior. *Economic Inquiry* 41, 1 (2003), 20–26.
- [16] FALK, A., AND FISCHBACHER, U. A theory of reciprocity. *Games and Economic Behavior* 54, 2 (2006), 293–315.
- [17] FATAS, E., AND MATEU, G. Antisocial punishment in two social dilemmas. *Frontiers in Behavioral Neuroscience* 9 (2015).
- [18] FEHR, E., AND GÄCHTER, S. Cooperation and punishment in public goods experiments. *American Economic Review* 90 (2000), 980–994.
- [19] FEHR, E., AND SCHMIDT, K. M. A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics* 114, 3 (1999), 817–868.
- [20] FEINBERG, M., WILLER, R., AND SCHULTZ, M. Gossip and ostracism promote cooperation in groups. *Psychological Science* 25, 3 (2014), 656–664.
- [21] FISCHBACHER, U. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10, 2 (2007), 171–178.
- [22] FISCHBACHER, U., GÄCHTER, S., AND FEHR, E. Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters* 71, 3 (2001), 397–404.

- [23] GARDNER, W. L., PICKETT, C. L., AND KNOWLES, M. Social snacking and shielding. *The social outcast: Ostracism, social exclusion, rejection, and bullying* (2005), 227–242.
- [24] GÄCHTER, S., RENNER, E., AND SEFTON, M. The long-run benefits of punishment. *Science* 322, 5907 (2008), 1510–1510.
- [25] HIRSHLEIFER, D., AND RASMUSEN, E. Cooperation in a repeated prisoners’ dilemma with ostracism. *Journal of Economic Behavior & Organization* 12, 1 (1989), 87–106.
- [26] KERR, N. L., RUMBLE, A. C., PARK, E. S., OUWERKERK, J. W., PARKS, C. D., GALLUCCI, M., AND VAN LANGE, P. A. “how many bad apples does it take to spoil the whole barrel?": Social exclusion and toleration for bad apples. *Journal of Experimental Social Psychology* 45, 4 (2009), 603–613.
- [27] KLING, J. R. Incarceration length, employment, and earnings. *American Economic Review* 96, 3 (2006), 863–876.
- [28] KOIKE, S., NAKAMARU, M., OTAKA, T., SHIMAO, H., SHIMOMURA, K.-I., AND YAMATO, T. Reciprocity and exclusion in informal financial institutions: An experimental study of rotating savings and credit associations. *Discussion Paper Tokyo Institute of Technology 2015-05* (2015).
- [29] KUZIEMKO, I. Going off parole: How the elimination of discretionary prison release affects the social cost of crime. Report, National Bureau of Economic Research, 2007.
- [30] LOTT, J. R. Do we punish high income criminals too heavily? *Economic Inquiry* 30, 4 (1992), 583.
- [31] MAIER-RIGAUD, F. P., MARTINSSON, P., AND STAFFIERO, G. Ostracism and the provision of a public good: experimental evidence. *Journal of Economic Behavior & Organization* 73, 3 (2010), 387–395.
- [32] MARTINSON, R. What works?-questions and answers about prison reform. *The Public Interest*, 35 (1974), 22.
- [33] MASCLET, D. Ostracism in work teams: a public good experiment. *International Journal of Manpower* 24, 7 (2003), 867–887.
- [34] MASTEN, C., EISENBERGER, N., AND BOROFKY, L. E. A. Neural correlates of social exclusion during adolescence: understanding the distress of peer rejection. *Social Cognitive and Affective Neuroscience* 4 (2009), 143–157.
- [35] MAURIN, E., AND OUSS, A. Sentence reductions and recidivism: Lessons from the bastille day quasi experiment, 2009.
- [36] NEEDELS, K. E. Go directly to jail and do not collect? a long-term study of recidivism, employment, and earnings patterns among prison releasees. *Journal of Research in Crime and Delinquency* 33, 4 (1996), 471–496.

- [37] NEUHOFFER, S., AND KITTEL, B. Long- and short-term exclusion in the public goods game: An experiment on ostracism. *Department of Economic Sociology, University of Vienna* (2015).
- [38] NIKIFORAKIS, N. Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics* 92 (2008), 91–112.
- [39] NIKIFORAKIS, N., AND ENGELMAN, D. Altruistic punishment and the threat of feuds. *Journal of Economic Behavior & Organization* 78 (2011), 319–332.
- [40] NIKIFORAKIS, N., AND NORMANN, H. T. A comparative statics analysis of punishment in public good experiments. *Experimental Economics* 11 (2012), 358–369.
- [41] NIKIFORAKIS, N., NOUSSAIR, C. N., AND WILKENING, T. Normative conflict and feuds: The limits of self-enforcement. *Journal of Public Economics* 96, 9-10 (2011), 797–807.
- [42] PRESSMAN, S. D., AND COHEN, S. Does positive affect influence health? *Psychological Bulletin* 131, 6 (2005), 925.
- [43] SELTEN, R. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory* 4, 1 (1975), 25–55.
- [44] SHEREMETA, R. M., TUCKER, S. J., AND ZHANG, J. Creating self-sustained social norms through communication and ostracism, 2009.
- [45] SMITH, P., GENDREAU, P., AND GOGGIN, C. *The effects of prison sentences and intermediate sanctions on recidivism: General effects and individual differences*. Solicitor General Canada Ottawa, ON, 2002.
- [46] VAN BEEST, I., AND WILLIAMS, K. D. When inclusion costs and ostracism pays, ostracism still hurts. *Journal of Personality and Social Psychology* 91 (2006), 918–928.
- [47] WESSELMANN, E. D., AND WILLIAMS, K. D. *Ostracism and Stages of Coping*. Oxford: Oxford University Press, 2013.
- [48] WILL, G.-J., CRONE, E. A., AND GUROGLU, B. Acting on social exclusion: neural correlates of punishment and forgiveness of excluders. *Social Cognitive and Affective Neuroscience* 10 (2015), 209–218.
- [49] WILLIAMS, K., CHEUNG, C. K.-T., AND CHOI, W. Cyberostracism: Effects of being ignored over the internet. *Journal of Personality and Social Psychology* 79 (2000), 748–762.
- [50] WILLIAMS, K., AND SOMMER, K. Social ostracism by coworkers: Does rejection lead to loafing or compensation? *Personality and Social Psychology Bulletin* 23 (1997), 693–706.

- [51] WILLIAMS, K. D. *Ostracism: The power of silence*. Guilford Press, 2002.
- [52] YAMAGISHI, T. The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology* 51, 1 (1986), 110.
- [53] ZADRO, L., WILLIAMS, K. D., AND RICHARDSON, R. How low can you go? ostracism by a computer is sufficient to lower self-reported levels of belonging, control, self-esteem and meaningful existence. *Journal of Experimental Social Psychology* 40 (2004), 560–567.

# Appendices

## Appendix 1. Instructions

All the instructions are translated from French.

### **Instructions for the introduction and part 1 (common to all treatments):**

We thank you for your participation in this experiment on decision making. Please be aware that you are not allowed to communicate with the other participants during the experiment. Your payoffs will depend on your own decisions as well as on the decisions of other participants during this experiment. There is no right or wrong answers.

This experiment is divided into two parts. At the beginning of each part, the software will randomly form groups of 4 participants. You will be matched with three other participants at the beginning of the first part. At the beginning of the second part, the software will form new groups. In the second part, you will be matched with three participants who will be different from the participants you were matched with in the first part. You will stay in the same group for the whole duration of each part.

You will never know the other group members' identity. In each group, each member will be identified by a number that will remain the same throughout the part.

Every decision you make and every answer you give will remain strictly anonymous.

All the transactions during the experiment are made in ECU (Experimental Currency Units). At the end of the session, your earnings in ECU in the two different parts will be added up and converted into Euros at the following exchange rate:



## 35 ECU = 1 Euro

Your earnings will be paid to you in cash and in private at the end of the experiment in a separate room. You will also earn 5 euros for having arrived on time for the experiment.

First, we will distribute the instructions for the first part. You will get the instructions for the second part at the end of the first part. Please press « OK » to move on to the next screen.

### PART 1

At the beginning of this part, a total endowment of 80 ECU will be assigned to your group of 4 participants. This endowment will be deposited on a group account. The group account yields a 40% return on its total amount to each member of the group.

Each member of the group can withdraw a certain number of ECU from this group account to put on his private account. Each member decides individually how much he is willing to withdraw from the group account, between 0 and 20 ECU inclusive.

Each ECU that you withdraw from the group account pays you 1 ECU. If you decide to keep this ECU in the group account instead, your payoff from the group account is 40% of 1 ECU = 0.4 ECU. The payoff for each other group member is also 0.4 ECU. This means that keeping ECU on the group account increases the payoff of the other members. Similarly, you earn 0.4 ECU for each ECU kept on the group account by the other group members. For each ECU kept on the group account by one of your group members, you get a 40% return of 1 ECU = 0.4 ECU.

Example:

Suppose that the total amount withdrawn is 20 ECU. The total amount left on the group account is 60 ECU ( $80-20=60$ ). In this example, each group member receives a revenue from the group account equal to 40% of  $60 = 24$  ECU. Now, suppose that the amount withdrawn is 65 ECU. The amount left on the group account is 15 ECU. Each member receives a revenue from the group account equal to 40% of  $15 = 6$  ECU.

Payoff calculation:

Your final payoff in this part will be:

- On the one hand, the amount you have withdrawn from the group account.

- On the other hand, the revenue from the group account: this revenue is equal to 40% of the amount left on the group account after each group member has made his decision (the amount left on the account is 80 minus the sum withdrawn by each member). Each ECU kept on the group account yields a return of 0.4 ECU.

As a consequence, your total payoff is computed as follows:

$$\text{The amount you have withdrawn from the group account} \\ + 40\% (80 \text{ ECU} - \text{total amount withdrawn by the group members})$$

Example:

Suppose that you withdraw 10 ECU from the group account and each of the other 3 group members withdraws 1 ECU from the group account. The total amount left on the group account is 80 ECU minus the 10 ECU that you have withdrawn, minus 3 ECU withdrawn by the other group members, *i.e.* 67 ECU. In this case, you get from the group account 40% of 67 ECU = 27 ECU. You also need to add the 10 ECU that you withdrew from the group account. In this example, your final payoff is 37 ECU (10+27).

\*\*\*

Before moving on to the rest of the instructions, please read these instructions again and fill out the comprehension questionnaire that was handed out to you. We will check your answers individually. If you have any question, please raise your hand or press the red button on the side of your desk and we will answer your question in private.

In this part, you are going to make two types of decisions. In the rest of these instructions, the first kind of decision will be called « unconditional decision », and the second one « conditional decision ».

**Unconditional decision:** Here is an example of the screen you will get for an unconditional decision:

Remaining Time (sec) 112

**UNCONDITIONAL DECISION**

**The value of the group account is 80 ECU.**

Please indicate the number of ECU you want to withdraw from the group account:

For this particular decision, you must chose the amount of ECU you wish to withdraw from the group account between 0 and 20 ECU. Once you have made your decision, press « OK ». Then, once each participant has made his decision, you will be redirected to the conditional decision screen.

### Conditional decision:

Here is a screenshot of the conditional decision table you will get:

Remaining Time (sec) 112

**CONDITIONALS DECISIONS TABLE**

The value of the group account is 80 ECU. You have to decide of the number of ECU that you want to withdraw from the group account for each possible average number of ECU withdrawn from the group account by each of the 3 other members in your group. For each of the 21 possible cases, we ask you to indicate the number between 0 and 20 ECU included that you want to withdraw.

0 ECU <input style="width: 50px;" type="text"/> 1 ECU <input style="width: 50px;" type="text"/> 2 ECU <input style="width: 50px;" type="text"/> 3 ECU <input style="width: 50px;" type="text"/> 4 ECU <input style="width: 50px;" type="text"/> 5 ECU <input style="width: 50px;" type="text"/> 6 ECU <input style="width: 50px;" type="text"/>	7 ECU <input style="width: 50px;" type="text"/> 8 ECU <input style="width: 50px;" type="text"/> 9 ECU <input style="width: 50px;" type="text"/> 10 ECU <input style="width: 50px;" type="text"/> 11 ECU <input style="width: 50px;" type="text"/> 12 ECU <input style="width: 50px;" type="text"/> 13 ECU <input style="width: 50px;" type="text"/>	14 ECU <input style="width: 50px;" type="text"/> 15 ECU <input style="width: 50px;" type="text"/> 16 ECU <input style="width: 50px;" type="text"/> 17 ECU <input style="width: 50px;" type="text"/> 18 ECU <input style="width: 50px;" type="text"/> 19 ECU <input style="width: 50px;" type="text"/> 20 ECU <input style="width: 50px;" type="text"/>
---	---	--

Here you will have to make 21 decisions. You will have to decide the amount of your withdrawal from the group account for each possible average withdrawal from each of the three other group members. For every 21 possible situations, please indicate the amount you wish to withdraw, between 0 and 20 ECU.

The number right next to each box is the average amount of ECU withdrawn by each of the three other members of your group. The task consists in indicating in each box the amount of ECU you wish to withdraw, depending on the average amount withdrawn by the three other members of the group. For example, you have to indicate how many ECU you want to withdraw if each of the three other members withdraw on average 1 ECU, 2 ECU, 3 ECU ... You will have to fill each box with an integer between 0 and 20. Once you have made your decision for each of the 21 possibilities, please press « OK » to submit your decisions.

### **Payoff computation:**

Once each participant has made his decision, the program will randomly select one of the four group members. For this specific participant, only the conditional decision table will be taken into account for payment. For the three other members, only the unconditional decision will be used. At the time of the decision, you do not know if you have been randomly selected or not. As a consequence, make your decisions carefully, as each of them can be the one taken into account for the payoff computation. You will know your payoff for this part at the end of the experiment only.

#### **Example 1:**

Suppose you have been randomly selected by the software. You will then be paid based on the conditional decision table. The three other members of your group will be paid according to their unconditional decision. Suppose their unconditional decision are 0, 2 and 4 ECU. The average withdrawal from the members of your group is 2 ECU in this case.

- If you have submitted in your conditional decision table that you wish to withdraw 1 ECU if the average amount withdrawn by the others is 2 ECU, the total amount withdrawn from the group account is  $0+2+4+1=7$  ECU. Each member of the group will get a payoff of  $0.4 \cdot (80-7) = 29.2$  ECU from the group account. Each member will get a payoff of 29.2 ECU from the group account for this part, plus the number of ECU that they have put on their private account.
- If instead you submitted that you wish to withdraw 19 ECU if the average amount withdrawn by the others is 2 ECU, the total amount withdrawn from the group account is  $0+2+4+19=25$  ECU. Each member of the group will get a payoff of  $0.4 \cdot (80-25) = 22$  ECU from the group account. Each member will get a payoff of 22 ECU from the group account for this part, plus the number of ECU that they have put on their private account.

### Example 2:

Suppose that you have not been selected by the program. This means that you and two other members of your group will be paid according to your unconditional decision. Suppose your unconditional decision is 16 ECU, and the others have an unconditional decision of 18 and 20 ECU. The average amount withdrawn by you and the two other members is 18 ECU in this case.

- If the selected member of your group submitted that he wanted to withdraw 1 ECU if the three other members average withdrawal was 18 ECU, the total amount withdrawn from the group account is  $16+18+20+1=55$  ECU. Each member of the group will get a payoff of  $0.4 \cdot (80-55) = 10$  ECU from the group account. Each member will get a payoff of 10 ECU from the group account for this part, plus the number of ECU that they have put on their private account.
- If instead the selected member of your group submitted that he wanted to withdraw 19 ECU if the three other members average withdrawal was 18 ECU, the total amount withdrawn from the group account is  $18+16+20+19=73$  ECU. Each member of the group will get a payoff of  $0.4 \cdot (80-73) = 2.8$  ECU from the group account. Each member will get a payoff of 2.8 ECU from the group account for this part, plus the number of ECU that they have put on their private account.

\*\*\*

If you have any question, please raise your hand or press the red button on the side of your desk and we will come to assist you in private.

## Instructions Part 2 - Baseline treatment

At the beginning of this part, the software will randomly form new groups of four members that will remain fixed throughout this part. These groups are different than those in the first part. As a consequence, you will be interacting with three other participants. This part consists of 20 periods.

Description of each period :

At the beginning of each period, a total amount of 80 ECU is assigned to your group and deposited on a group account.

Each group member chooses the amount of ECU he is willing to withdraw from the group account, between 0 and 20 ECU, to put on his private account. Once all decisions have been made, the number of ECU left on the group account is computed. The group account yields a 40% return to each group member.

Each ECU that you withdraw from the group account pays you 1 ECU. If instead you are willing to keep this ECU in the group account, your payoff from the group account will be 40% of 1 ECU = 0.4 ECU. This ECU will also pay 0.4 ECU to each other group member. This means that keeping ECU on the group account increases the payoff of the other members. Similarly, you earn 0.4 ECU for each ECU kept on the group account by the other members of your group.

Once you have made your decision, please submit it by pressing the « OK » button. At the end of this period, you will be informed on the total amount left on the group account, the number of ECU withdrawn by each group member, and the payoff of each group member.

Next periods:

Each new period will start automatically. At the beginning of a period, the group is given a new group account. You are matched with the same 3 group members for the 20 periods.

To sum up, in each period:

- At the beginning, the group is given a new group account.
- Each member decides on the amount he is willing to withdraw from the group account.
- You are informed on the amount withdrawn by each group member, and on his payoff.

### **Payoff calculation:**

Your final payoff in each period will be:

- On the one hand, the amount you have withdrawn from the group account.
- On the other hand, the revenue from the group account: this revenue is equal to 40% of the amount left on the group account. The amount left is the difference between the initial amount (80) and the sum of group members' withdrawals. Every ECU kept on the group account yields a return of 0.4 ECU.

Your total payoff is computed as follows:

**The amount you have withdrawn from the group account  
+ 40% (80 ECU – total amount withdrawn by the group members)**

The payoff is computed in the same way for each group member. This means that each member gets the same payoff from the group account.

\*\*\*

Please read these instructions again. If you have any question, raise your hand or press the red button on the side of your desk. We will answer to your questions in private.

**Instruction Part 2 - Exo1P treatment (*The only differences with the Exo3P treatment are indicated in italics*)**

At the beginning of this part, the software will randomly form new groups of four members that will remain fixed throughout this part. These groups are different from those in the first part. As a consequence, you will be interacting with three other participants. This part consists of 20 periods. From one period to the other, your group size may vary. As it will be explained right after, you cannot be assigned to another group.

Description of each period :

At the beginning of each period, a total amount is assigned to your group. This amount of ECU is deposited on a group account and it depends on the number of members in the group.

If there are 4 members in the group, the value of the group account is 80 ECU.  
If there are 3 members in the group, the value of the group account is 60 ECU.  
If there are 2 members in the group, the value of the group account is 40 ECU.  
If there is only 1 member in the group, the value of the group account is 20 ECU.

Each period is divided in two stages:

Stage 1:

First, each member of the group decides how many ECU he is willing to withdraw from the group account, between 0 and 20, to put on his private account. Once all decisions have been made, the number of ECU left on the group account is computed. The group account yields a 40% return to each group member.

Each ECU that you withdraw from the group account pays you 1 ECU. If instead you are willing to keep this ECU in the group account, your payoff from the group account will be 40% of 1 ECU = 0.4 ECU. This ECU will also pay 0.4 ECU to each other group member. This means that keeping ECU on the group account increases the payoff of the other members. Similarly, you earn 0.4 ECU for each

ECU kept on the group account by the other members of your group.

Once you have made your decision, please submit it by pressing the « OK » button. At the end of the period, you will be informed on the total amount left on the group account, the number of ECU withdrawn by each group member, and the payoff of each group member.

Stage 2:

#### *Decision rule*

In this stage, the group members have the opportunity to remove one or several other group members for the following period (*for the three following periods*). Removing a member from the group means that he will not be able to participate in any of the two stages in the following period (*of the three following periods*): he will not be able to withdraw ECU from the group account or decide to remove other members. Instead, he will get a fixed payoff of 10 ECU. If one or several members of the group are removed, the total amount of the group account will be adjusted in the next period (*the three following periods*) as explained previously (60 ECU if the group is composed of 3 members, 40 ECU if the group is composed of 2 members, 20 ECU if the group is composed of only 1 member).

Several members of the group can be removed during a period. If all the group members are removed during a period, each removed member will get a fixed payoff of 10 ECU and the next period will start.

Removing one or several members of the group for the next period (*the three following periods*) is a group decision based on a vote within the group. During this stage, each member of the group has to vote for deciding to remove or not each of the other members of the group.

Here is an example of the vote decision screen:



Period: 1 sur 4 <span style="float: right;">Remaining Time [sec] 112</span>	
<b>DECISIONS:</b>  You have withdrawn 5 ECU from the group account.  Player 1 withdrew 20 ECU.  Player 3 withdrew 0 ECU.  Player 4 withdrew 10 ECU.	<b>PAYOFFS:</b>  You earned 23.0 ECU in this period.  Player 1 earned 38 ECU.  Player 3 earned 18 ECU.  Player 4 earned 28 ECU.
If at least two of you vote to withdraw one of more member(s) from the group, these members will not be able to participate in the next period and they will not receive any payoffs from the group account in this period. Instead, these members will receive a fixed wage of 10 ECU.	
<div style="display: flex; justify-content: space-between;"> <div>           Do you want to withdraw player 1 from the group during the next period?             Do you want to withdraw player 3 from the group during the next period?             Do you want to withdraw player 4 from the group during the next period?         </div> <div style="text-align: right;"> <input type="button" value="No"/>  <input type="button" value="Yes"/>   <input type="button" value="No"/>  <input type="button" value="Yes"/>   <input type="button" value="No"/>  <input type="button" value="Yes"/> </div> </div>	
<input type="button" value="OK"/>	

In the top left corner, you can see how many ECU you and each member of the group have withdrawn from the group account. On the right, you can see your payoff for this period, as well as the payoff of each other group member.

In the bottom of the screen, you will be asked to make a decision for each of the member of the group (you excepted). You vote by pressing « Yes » if you are willing to remove this member from the group for the next period (*the three following periods*), or « No » if you want him to remain part of the group in the next period.

You cannot remove a member of the group during the last period. (*At the 18th period, you can only remove someone for a maximum of two periods, and at the 19th period, only for one period.*) The last period only consists in choosing how many ECU you are willing to withdraw from the group account.

### *Consequences of the vote*

A member will only be removed **if at least two group members have pressed « Yes » to remove him.**

The removed participants will not be able to participate in any of the stages of the next period (*the next three periods.*). They will get a fixed payoff of 10 ECU but they will not get any return from the group account.

### *Information*

Once each participant has made his decision for each other member of his group and submitted his choice by pressing « OK » at the bottom of the screen, everyone

will be informed of the result of the votes.

The results only indicate for each group member if he will participate in the next period or if he will be removed from the group for one period (*for three periods*).

The number of votes for or against the removal of a participant will never be displayed.

### **Next periods:**

Each new period will start automatically. At the beginning of a period, the group is given a new group account. The total amount deposited on this group account varies depending on the number of members in the group.

You will be reminded of the number of participants left in your group for this period on your screen. You are matched with the same participants for these 20 periods, even if some of them may not participate in each period.

To sum up, in each period:

- At the beginning, the group is given a new group account, which amount depends on the number of members in the group.
- Each member decides on the amount he is willing to withdraw from the group account.
- You are informed on the amount withdrawn by each group member and on his earnings.
- You vote to remove or not one or several members of your group for the next period (*the next three periods*).
- You are informed about the removal of other group members, if any.

### **Payoff calculation:**

**If you are part of the group,** your final payoff in each period will be:

- On the one hand, the amount you have withdrawn from the group account.
- On the other hand, the revenue from the group account: this revenue is equal to 40% of the amount left on the group account. The amount left is the difference between the initial amount (80, 60, 40 or 20 depending on the number of group members for the current period) and the sum of group members' withdrawals. Every ECU kept on the group account yields a return of 0.4 ECU.

Your total payoff is computed in the following way:

$$\begin{aligned} &\text{The amount you have withdrawn from the group account} \\ &+ 40\% (\text{value of the group account at the beginning of the period} \\ &\quad - \text{total amount withdrawn by the group members}) \end{aligned}$$

The payoff is computed in the same way for each group member. This means that each member gets the same payoff from the group account.

**If you are not part of the group**, your payoff is 10 ECU for each period you have been removed from the group.

\*\*\*

Please read these instructions again and fill out the comprehension questionnaire that was handed out to you. We will check your answers individually. If you have any question, please raise your hand or press the red button on the side of your desk. We will answer to your questions in private.

### **Instruction Part 2 - Endo treatment**

At the beginning of this part, the software will randomly form new groups of four members that will remain fixed throughout this part. These groups are different from those in the first part. As a consequence, you will be interacting with three other participants. This part consists of 20 periods. From one period to the other, your group size may vary. As it will be explained right after, you cannot be assigned to another group.

Description of each period :

At the beginning of each period, a total amount is assigned to your group. This amount of ECU is deposited on a group account and it depends on the number of members in the group.

If there are 4 members in the group, the value of the group account is 80 ECU.  
If there are 3 members in the group, the value of the group account is 60 ECU.  
If there are 2 members in the group, the value of the group account is 40 ECU.  
If there is only 1 member in the group, the value of the group account is 20 ECU.

Each period is divided in two stages:

Stage 1:

First, each member of the group decides how many ECU he is willing to withdraw from the group account, between 0 and 20, to put on his private account. Once all decisions have been made, the number of ECU left on the group account is computed. The group account yields a 40% return to each group member.

Each ECU that you withdraw from the group account pays you 1 ECU. If instead you are willing to keep this ECU in the group account, your payoff from the group account will be 40% of 1 ECU = 0.4 ECU. This ECU will also pay 0.4 ECU to each other group member. This means that keeping ECU on the group account increases the payoff of the other members. Similarly, you earn 0.4 ECU for each ECU kept on the group account by the other members of your group.

Once you have made your decision, please submit it by pressing the « OK » button. At the end of the period, you will be informed on the total amount left on the group account, the number of ECU withdrawn by each group member, and the payoff of each group member.

Stage 2:

#### *Decision rule*

In this stage, the group members have the opportunity to remove one or several other group members for one, two or three following periods. Removing a member from the group means that he will not be able to participate in any of the two stages for the number of periods he has been removed: he will not be able to withdraw ECU from the group account or decide to remove other members. Instead, he will get a fixed payoff of 10 ECU. If one or several members of the group are removed, the total amount of the group account will be adjusted in the next period as explained previously (60 ECU if the group is composed of 3 members, 40 ECU if the group is composed of 2 members, 20 ECU if the group is composed of only 1 member).

Several members of the group can be removed during a period. If all the group members are removed during a period, each removed member will get a fixed payoff of 10 ECU and the next period will start.

Removing one or several members of the group for a certain number of periods is a group decision based on a vote within the group. During this stage, each member of the group has to vote for deciding to remove or not each of the other members of the group. If you vote to remove one group member, you will then have to decide for how long you want this member to be removed: from 1 period to a maximum of 3 periods.

Here is an example of the vote decision screen:

Period: 1 on 4		Remaining Time [sec]: 112
<b>DECISIONS:</b>  You have withdrawn 5 ECU from the group account.  Player 1 withdrew 20 ECU.  Player 2 withdrew 15 ECU.  Player 4 withdrew 10 ECU.	<b>PAYOFFS:</b>  You earned 17.9 ECU in this period.  Player 1 earned 32 ECU.  Player 2 earned 27 ECU.  Player 4 earned 22 ECU.	
If at least two of you vote to withdraw one or more member(s) from the group, these members will not be able to participate in a number of periods determined by the rules in your instructions and they will not receive any payoffs from the group account in these periods. Instead, these members will receive a fixed wage of 10 ECU.		
<div style="display: flex; justify-content: space-between;"> <div> <p>Do you want to withdraw player 1 from the group during the next period?</p> <p>Indicate a number of period(s) between 1, 2 and 3: 1 2 3</p> <p>Do you want to withdraw player 2 from the group during the next period?</p> <p>Do you want to withdraw player 4 from the group during the next period?</p> </div> <div style="text-align: right;"> <div style="margin-bottom: 5px;"> <input type="button" value="No"/> <input style="background-color: red; color: white;" type="button" value="Yes"/> </div> <div style="margin-bottom: 5px;"> <input type="button" value="No"/> <input style="background-color: red; color: white;" type="button" value="Yes"/> </div> <div style="margin-bottom: 5px;"> <input type="button" value="No"/> <input style="background-color: red; color: white;" type="button" value="Yes"/> </div> <div style="margin-bottom: 5px;"> <input type="button" value="No"/> <input style="background-color: red; color: white;" type="button" value="Yes"/> </div> </div> </div>		
<input type="button" value="OK"/>		

In the top left corner, you can see how many ECU you and each member of the group have withdrawn from the group account. On the right, you can see your payoff for this period, as well as the payoff of each other group member.

In the bottom of the screen, you will be asked to make a decision for each of the member of the group (you excepted). You vote by pressing « Yes » if you are willing to remove this member from the group, or « No » if you want him to remain part of the group in the next period.

If you press « Yes », an additional line will appear, allowing you to choose the length of the removal for this group member. You can exclude a member for 1, 2, or 3 periods.

At the 18th period, you can only remove someone for a maximum of two periods, and at the 19th period, only for one period. The last period only consists in choosing how many ECU you are willing to withdraw from the group account.

### *Consequences of the vote*

A member will only be removed **if at least two group members have pressed « Yes » to remove him**. In this case, the duration of the removal will be determined as follows:

- If each other member has agreed on the removal and has chosen the same number of periods, the participant will be removed for this number of periods.
- If each other member has agreed on the removal but they have chosen a different number of periods, the program will select the median number of

periods (the median is the intermediate number of periods, or if two of the three members have chosen the same number of periods, then it is this number of periods.)

- If two members agreed on the removal but have chosen a different number of periods, then the program selects the smaller number of periods.

Example:

Suppose that three members vote to remove the fourth member. If the first member chooses one period and the others choose two periods, then the fourth member will be removed for two periods. Now suppose that two members vote to remove a group member, one for one period and the other for three periods. Then, the length of removal will be one period.

The removed participants will not be able to participate in any of the stages for the number of periods decided at the end of the vote. They will get a fixed payoff of 10 ECU but they will not get any return from the group account.

### *Information*

Once each participant has made his decision for each other member of his group and submitted his choice by pressing « OK » at the bottom of the screen, everyone will be informed of the vote results.

The results only indicate for each group member if he will participate in the next period or if he will be removed from the group. In the latter case, they will also show for how long the member will be removed.

The number of votes for or against the removal of a participant will never be displayed.

### **Next periods**

Each new period will start automatically. At the beginning of a period, the group is given a new group account. The total amount deposited on this group account varies depending on the number of members in the group.

You will be reminded of the number of participants left in your group for this period on your screen. You are matched with the same participants for these 20 periods, even if some of them may not participate in each period.

To sum up, in each period:

- At the beginning, the group is given a new group account, which amount depends on the number of members in the group.

- Each member decides on the amount he is willing to withdraw from the group account.
- You are informed on the amount withdrawn by each group member and on his earnings.
- You vote to remove or not one or several members of your group and you decide for how long.
- You are informed about the removal of other group members, if any, and for those removed, the length of their removal.

### **Payoff calculation:**

**If you are part of the group,** your final payoff in each period will be:

- On the one hand, the amount you have withdrawn from the group account.
- On the other hand, the revenue from the group account: this revenue is equal to 40% of the amount left on the group account. The amount left is the difference between the initial amount (80, 60, 40 or 20 depending on the number of group members for the current period) and the sum of group members' withdrawals. Every ECU kept on the group account yields a return of 0.4 ECU.

Your total payoff is computed in the following way:

<p style="text-align: center;"> <b>The amount you have withdrawn from the group account</b>  <b>+ 40% (value of the group account at the beginning of the period</b>  <b>– total amount withdrawn by the group members)</b> </p>
--

The payoff is computed in the same way for each group member. This means that each member gets the same payoff from the group account.

**If you are not part of the group,** your payoff is 10 ECU for each period you have been removed from the group.

\*\*\*

Please read these instructions again and fill out the comprehension questionnaire that was handed out to you. We will check your answers individually. If you have any question, please raise your hand or press the red button on the side of your desk. We will answer to your questions in private.

## Appendix 2. Tables

Table A.1 displays basic statistics on the subject pool, by treatment.

Table A.1: Individual characteristics, by treatment

	Number of individuals	Number of sessions	Number of groups	Males (Percentage)	Mean age	Students (Percentage)
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline	60	3	15	0.383	25.68	0.783
Exo 1P	48	3	12	0.479	24.73	0.792
Endo	48	3	12	0.50	28.60	0.667*
Exo 3P	52	3	13	0.539**	23.38	0.808

*Notes:* A binomial test shows that there are significantly more males in the Exo3P treatment than in the Baseline (two-sided binomial test:  $p=0.031$ ), while the difference between the Baseline and each other treatment is never significant (*vs.* Exo1P:  $p=0.103$ ; *vs.* Endo:  $p=0.183$ ). There is no significant difference in the distribution of participants in terms of age between the Baseline and each other treatment (Kolmogorov-Smirnov tests: *vs.* Exo1P:  $p=1.000$ ; *vs.* Exo3P:  $p=0.731$ ; *vs.* Endo:  $p=0.161$ ). The percentage of students is significantly lower in the Endo treatment compared to the Baseline (two-sided binomial test:  $p=0.055$ ), while the difference between the Baseline and the other treatment is never significant (*vs.* Exo1P:  $p=1.000$ ; *vs.* Exo3P:  $p=0.740$ ).

Table A.2 reports a regression analysis complementary to Table 2 in the main text.



Table A.2: Determinants of an exclusion vote in treatments with possible exclusion

Dep. variable: Decision to exclude $j$ in $t$	All treatments (1)	with exclusion (2)
Exo1P	<i>Ref.</i>	0.024 (0.091)
Exo 3P	-0.054 (0.067)	-0.032 (0.088)
Endo	-0.022 (0.081)	<i>Ref.</i>
Max $(0; w_{j,t} - \sum_{-j=1, j \neq -j}^{n-1} w_{-j,t})$	0.043*** (0.013)	0.040*** (0.008)
Max $(0; \sum_{-j=1, j \neq -j}^{n-1} w_{-j,t} - w_{j,t})$	0.012** (0.005)	-0.008* (0.005)
Max $(0; w_{j,t} - \sum_{-j=1, j \neq -j}^{n-1} w_{-j,t})$ * Exo1P	<i>Ref.</i>	0.003 (0.011)
Max $(0; w_{j,t} - \sum_{-j=1, j \neq -j}^{n-1} w_{-j,t})$ * Exo3P	0.009 (0.014)	0.012 (0.015)
Max $(0; w_{j,t} - \sum_{-j=1, j \neq -j}^{n-1} w_{-j,t})$ * Endo	-0.003 (0.011)	<i>Ref.</i>
Max $(0; \sum_{-j=1, j \neq -j}^{n-1} w_{-j,t} - w_{j,t})$ * Exo1P	<i>Ref.</i>	0.020*** (0.008)
Max $(0; \sum_{-j=1, j \neq -j}^{n-1} w_{-j,t} - w_{j,t})$ * Exo3P	-0.011 (0.008)	0.009 (0.008)
Max $(0; \sum_{-j=1, j \neq -j}^{n-1} w_{-j,t} - w_{j,t})$ Endo	-0.020*** (0.008)	<i>Ref.</i>
Exclusion of $j$ prior to $t$	-0.003 (0.019)	-0.003 (0.019)
$i$ is reintegrated in $t$	0.104*** (0.036)	0.104*** (0.036)
Group size	-0.011 (0.020)	-0.011 (0.020)
Period	0.002 (0.001)	0.002 (0.001)
Individual controls	Yes	Yes
Observations	4198	4198
Log pseudo-likelihood	-1311.17	-1311.17

*Notes:* The Table reports marginal effects from random-effects logit regressions with robust standard errors clustered at the group level in parentheses. We consider only: *i*) observations from periods 1 to 17; *ii*) groups with more than two active members, as it is not possible to exclude when the group has less than three active members; *iii*) cases where  $i$  and  $j$  are not excluded (as  $i$  cannot vote if she is excluded and  $i$  cannot vote to exclude  $j$  if  $j$  is already excluded). \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ .