GATE LYON SAINT-ÉTIENNE
UMR 5824

93, chemin des Mouilles
69130 Ecully - France

Maison de l'Université, Bâtiment B
10, rue Tréfilerie
42023 Saint-Etienne cedex 02 - France

http://www.gate.cnrs.fr
gate@gate.cnrs.fr

WP 1720 – August 2017, revised April 2018

Exclusion and Reintegration in a Social Dilemma

Alice Solda, Marie Claire Villeval

Abstract:

Using a social dilemma game, we study the cooperative behavior of individuals who reintegrate their group after being excluded by their peers. 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 and they punish more, and more severely, chronic defections. In return, a longer exclusion has a higher disciplining effect on cooperation after reintegration, but only when the length of exclusion is not chosen by group members. Its relative disciplining effect is smaller when the length of exclusion results from a vote. In this environment, a quicker reintegration also limits retaliation. The difference in the impact of long vs. short exclusion on retaliation is larger when the length of exclusion is chosen by group members than when it is exogenous. Post-reintegration 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, cooperation, experiment

JEL codes:
C92, H41, D23

Working Paper

Exclusion and Reintegration in a Social Dilemma

Alice Solda and Marie Claire Villeval

April 24, 2018

Abstract

Using a social dilemma game, we study the cooperative behavior of individuals who reintegrate their group after being excluded by their peers. 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 and they punish more, and more severely, chronic defections. In return, a longer exclusion has a higher disciplining effect on cooperation after reintegration, but only when the length of exclusion is not chosen by group members. Its relative disciplining effect is smaller when the length of exclusion results from a vote. In this environment, a quicker reintegration also limits retaliation. The difference in the impact of long *vs.* short exclusion on retaliation is larger when the length of exclusion is chosen by group members than when it is exogenous. Post-reintegration 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, 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 L. Balafoutas, T. Offerman, R. Slonim, R. Sheremata, R. Sloof, J.R. Tyran, and participants at the 11th Annual Australia New Zealand Workshop on Experimental Economics, the 2017 Asia-Pacific ESA Conference, the Experimental Public Choice Workshop in Lille, and at seminars at the University of Amsterdam, Remnin University in Beijing, and Zhejiang University of Finance and Economics in Hangzhou 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, but cooperation can be sustained thanks to several mechanisms, including communication, rewards, sanctions, or taxation. In particular, the effectiveness of punishment mechanisms is now well established. 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.[1] Formal and centralized exclusion mechanisms sanction, notably by imprisonment, people who do not comply with judicial rules. In both cases, the goal is both preventive, 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, exclusion may also increase the risk of anti-social behavior and recidivism after they are reintegrated in their group.[2]

Exclusion is an effective mechanism to enforce social norms and limit free-riding in social dilemmas (*e.g.*, Bowles and Gintis, 2004).[3] But while the deterrence theory (Becker, 1968) predicts that increasing a penalty should discourage more misbehavior, there is little conclusive empirical evidence of the effect of the *length* of exclusion on post-release compliance and recidivism. For example, studies on the length of prison sentences find little evidence that length affects either recidivism (Martinson, 1974; Blumstein *et al.*, 1978), employment or income (Needels, 1996; Kling, 2006) after rehabilitation.[4] Moreover, we are not aware of any study measuring the impact of the length of informal

[1]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 social exclusion in tribes. Ethological studies have shown that scapegoating and shunning exist in social animal species. The forms of exclusion evolve over time, with cyberostracism as a modern expression of the refusal to communicate with someone (Wesselmann and Williams, 2013).

[2]This risk exists, as exclusion affects emotionally those who are excluded. Exclusion generates anger, lowers self-esteem (Zadro *et al.* 2004; van Beest and Williams, 2006), and may lead to depression if it lasts (Pressman and Cohen, 2005). In neurosciences, William (2002) found that exclusion activates 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). In contrast, forgiveness coincides with activation in regions implicated in "theory of mind" and perspective-taking (Will *et al.* 2015).

[3]Studies in psychology show that ostracized individuals conform more after re-admission in the group (*e.g.*, William *et al.*, 2000; Feinberg *et al.*, 2014), they work harder in group tasks (William and Sommer, 1997) and they become more sensitive to social cues (Gardner *et al.*, 2005).

[4]Smith *et al.* (2002) found that more punishment leads to at most a slight increase in recidivism. In contrast, Lott (1992) found a large effect of the prison sentence length and Zapryanova (2014) estimated that one month in prison results in a 1.12 percentage point decrease in recidivism probability (see also Bhuller *et al.*, 2016). By examining the effects of a reform that increased the sentence length by one month, Landerso (2015) showed that longer incarceration spells result in lower unemployment rates and higher earnings.

exclusion on deterrence and compliance. The absence of conclusive evidence is partly due to the fact that the empirical analysis based on natural data is often plagued with endogeneity and selection issues.[5]

Our objective is studying how the length of exclusion affects the behavior of wrongdoers *after* they are readmitted in their group, in the context of a social dilemma game. We also study whether the length of exclusion has a different impact when voters had a choice between different lengths, compared to when it is imposed exogenously by an external authority. We consider two dimensions of behavior after reintegration: cooperation and forgiveness *vs.* retaliation. Addressing these questions is highly policy-relevant because the recommendations will differ if a longer exclusion is shown to generate higher net benefits in terms of post-reintegration compliance or if it triggers violent antisocial behavior. By keeping the environment constant (*e.g.*, the same reference group and the same level of information before, during and after exclusion), a laboratory experiment helps identify the direct effect of the length of exclusion and its mode of determination on behavior. It can also mute two channels that play a role in real settings: the variation in human capital associated with durable exclusion and peer effects that permit social learning during exclusion in the behavior that led to exclusion.

Our experiment uses a taking game that is played for 20 periods. We formed groups of four subjects who have to decide simultaneously how many units to withdraw from a preexisting public good. After being informed about the withdrawal decision of each partner, individuals can vote to exclude some members. The decision of excluding someone requires two votes to become effective. Excluded members remain inactive, cannot communicate, receive information about the evolution of cooperation in their group, and are paid a low inactivity payoff. Since the initial value of the public good depends on the group size, excluding a member is potentially costly. In one treatment, the length of exclusion is fixed at one period; in another it is fixed at three periods, and in a third treatment, individuals vote also about the length of exclusion, up to three periods. This environment does not allow group mobility: excluded members are always reintegrated in their previous group. We compare behavior in these treatments with a baseline in which exclusion is not allowed. Thus, we can investigate the effect of exclusion followed by reintegration on excluded members' behavior once they are readmitted in their original group, depending on how long they have been excluded and on the procedure of exclusion.

Our study makes three contributions to the literature. The first contribution is comparing the impact of various lengths of exclusion on the cooperative behavior of excluded

[5]For example, an inmate with a long prison sentence is more at risk of recidivism due to his own characteristics that led to his exclusion, to peer effects in prison and to stigma on the labor market. 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, giving a random manipulation of the length of the remaining sentence at the date of the pardon. They found that longer sentences had ex ante a higher deterrent effect on recidivism, but the inmates who stayed longer 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 sentences decrease recidivism. Using instead random judge assignments, Green and Winik (2010) found that randomly assigned variations in prison and probation times have no effect on recidivism, echoing the results of Berube and Green (2007). In contrast, Roach and Schanzenbach (2016) estimated that one-month extra prison sentence decreases recidivism by one percentage point.

members *after* they reintegrate their group. Earlier studies on exclusion in social dilemmas have focused more on the threat of exclusion than on reintegration and most did not manipulate the length of exclusion (*e.g.*, Masclet, 2003; Cinyabuguma *et al.*, 2005; Kerr *et al.*, 2009; Sheremeta *et al.* 2009; Maier-Rigaud *et al.*, 2010). For an exception, see Balafoutas *et al.* (2017) who conducted an experiment with inmates and found that a longer incarceration is associated with a lower level of pro-sociality. Our second contribution is studying whether endogeneizing the length of exclusion through a voting procedure modulates the impact of exclusion on future cooperation after reintegration in the same group. In the previous literature on exclusion in social dilemmas, the length of exclusion was fixed exogenously. In contrast, we examine whether choosing a less severe sanction when a more severe one was available is perceived as a warning, signaling a willingness to forgive after reintegration, compared to when the length is fixed exogenously.

Our third contribution is studying the impact of the length of exclusion and of its mode of determination on retaliation. Previous studies have highlighted the risk of revenge (*e.g.*, Nikiforakis, 2008) and feuds (Nikiforakis and Engelmann, 2011) induced by punishment in social dilemmas. In our study punishment results from a group decision, which may signal more strongly what the norm of the group is, especially when subjects choose also the length of exclusion. Ambrus and Greiner (2015) have shown that democratic punishment decreases anti-social behavior compared to individual punishment. In our case, we can compare retaliation behavior for a given length of exclusion according to the mode of choice of this length.

Our results show that exclusion and the choice of its duration are driven by the willingness to punish larger deviations from the withdrawal norm in the group and chronic defections. A longer exclusion has a larger effect on norm compliance after reintegration compared to a shorter exclusion, but only when the length of exclusion is exogenous. When the length of exclusion results from a vote, the effect of a longer exclusion on the withdrawal level after reintegration is smaller. Choosing a shorter exclusion when a more severe sanction is available leads to more forgiveness by the reintegrated members and limits retaliation compared to when people are excluded for three periods. When the length of exclusion is chosen by group members, the difference in the impact on retaliation of a long sentence compared to a short sentence is much larger than when the length of exclusion is fixed exogenously. Indeed, behavior after reintegration does not only depend on the severity of the sanction but also on the intentions of group members that can be inferred from it.

The remaining sections are organized as follows. Section 2 reviews the related litarature very briefly. Section 3 details the experimental design and the procedures. In Section 4, we present behavioral conjectures. Section 5 displays our main results. Section 6 concludes.

2 Related Literature

The previous literature has considered various exclusion mechanisms in different environments, in particular in principal-agent games when people can be fired or in gift-exchange games (*e.g.*, Brown *et al.*, 2004; Charness *et al.*, 2011; Berninghaus *et al.*, 2013; Falk *et*

al., 2015; Bernard *et al.*, 2017), in coordination games and bargaining games (Fischer and Güth, 2012). For example, Kopanyi-Peuker *et al.* (2018) study a weak-link game with noisy information in which a principal can replace some members of a team. The fear of being fired increases cooperation even when workers are imperfectly monitored; but to be effective, the threat has to be maintained permanently. Croson *et al.* (2015) compare the impact of excludability in three team production settings (weak-link, voluntary contribution, and a best-shot games) when the lowest contributor in the team is automatically excluded. They find that excludability increases efficiency dramatically, but less so in the weak-link game. In contrast, we study exclusion in a social dilemma environment without competition (between group members or between unemployed players); the decision to exclude someone is made by the group members themselves,[6] without any formal rule defining who should be excluded (also in contrast with Swope, 2002). Note that we also differ from these previous studies in that our design allows us to study whether the reintegrated players retaliate by trying to exclude their punishers.

We also differ from the previous literature in which group members can exclude each others, as we manipulate the length of exclusion. We observe three types of length of reversible exclusion. In some studies, exclusion is irreversible (Cinyabuguma *et al.*, 2005; Maier-Rigaud *et al.*, 2010; Akpalu and Martinsson, 2011). An excluded subject remains ostracized until the end of the session. In contrast, in Charness and Yang (2014) exclusion from the group is irreversible but subjects can move immediately to another group. Finally, in other studies exclusion is only for one period (Masclet, 2003; Kerr *et al.*, 2009; Sheremeta *et al.*, 2009; Koike *et al.* 2015). None of these studies compares different lengths of exclusion. Neuhofer and Kittel (2015) show that irreversible exclusion has a stronger impact than a one-period exclusion but this comparison is based on a between-subject design. The lack of within-subject variation of the length of exclusion in the previous literature is somewhat 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 usually depends on the harm caused.[7]

Our study is also related to the literature on endogenous group formation, as excluding group members enables players to choose their partners, to a certain extent. For example, Ahn *et al.* (2008) showed that restricted entry in an endogenous group formation mechanism in a public goods game leads to the formation of mid-sized highly cooperative groups. Restricted entry also increases average contributions and earnings in a congestible public good game (Ahn *et al.*, 2009). Riedl *et al.* (2016) found that letting players form networks in a weak-link game leads to fully efficient coordination, demonstrating that the possibility of choosing interaction partners is fundamental to achieve efficiency (see also Riedl and Ule, 2014, on exclusion in prisoners dilemma games). Sääksvuori (2014) shows that competitive pressure between groups decreases the likelihood of social exclusion. In our game, players can choose their partners if they are able to coordinate to exclude

[6] In our setting all group members have the same status; thus, we also differ from Güth *et al.* (2007) in which a leader decides whom to exclude in the group and from Fatas *et al.* (2010) where the subject excluded is selected randomly.

[7] Davis and Johnson (2015) manipulate the length of exclusion in a game in which participants vote to exclude group members and to reintegrate them. However, excluded members are only excluded from a chat and they can still contribute. Exclusion has no monetary consequences.

some of them. Once individuals have chosen the length of exclusion, they do not have to vote about on reintegration. Moreover, players reintegrate the same group after being excluded. This allows us to study to which extent reintegrated individuals are willing to cooperate with those who have previously excluded them.

3 Experimental Design and Procedures

We first present our design before detailing the procedures.

3.1 Design

The experiment consists of four treatments that are played between subjects. In each treatment, a taking game is played for 20 periods with a fixed matching protocol. 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.[8] 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:

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

with $\alpha = 0.4$ the marginal per capita return of the group account and n 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:

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

The following treatments introduce voting and exclusion.[9]

[8] The exchange rate is 35 ECU = 1 Euro.

[9] Note that a public good like national defense or street lighting is non-excludable, by nature. But as noted by Cinyabuguma *et al.* (2005), there are goods that are public to the members of a group, but not to those who do not belong to the group. Hirshleifer and Rasmusen (1989) also evoke the excludability

Exo1P Treatment

Each period of the Exo1P treatment (for "Exogenous one Period")[10] 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.

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 other group member. All group members make their decisions simultaneously. If someone is excluded in period t, he will remain inactive in period t+1. This means that in t+1 he will not be able to participate in the taking stage, nor to vote in the punishment stage, and he 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.[11] During the exclusion period, an excluded member receives feedback on the withdrawal decisions and earnings of each group member and on the exclusion of other members in t+1.

There is no direct cost for excluding a group member (like in *e.g.*, Hirshleifer and Rasmusen, 1989; Masclet, 2003; Maier-Rigaud *et al.*, 2010). Exclusion is indirectly costly for two reasons. First, 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 since the good is non-rival. Indirect costs are null only in the case where the excluded subject would have withdrawn 20 ECU in the next period if not excluded. Second, excluding more than one group member generates an additional indirect cost because it destroys the incentive to cooperate at the group level. Indeed, when there are more than two members in the group, the game involves a social dilemma: the Marginal Per Capita Return (MPCR, hereafter) is 0.4, which creates an incentive to free-ride (MPCR<1), but the social optimum corresponds to full cooperation (no withdrawal) since N*MPCR>1. However, when N<3, the structure of incentives changes: it is still an equilibrium to withdraw to the maximum extent but it is no longer socially inefficient to withdraw from the group account (N*MPCR<1), and the social dilemma no longer exists. This means that if more than one person is excluded, cooperation can no longer maximize the social surplus. When N<3, even cooperators have no more reason

of resources from non-members. Teams, associations, partnerships, clubs develop collective action with a risk of free-riding, and they can usually exclude members who do not comply with the rules of the group. Our setting is closer to club goods (*e.g.*, private parks, collective gardens owned by an association, ponds maintained by fishing associations) (for studies of excludability in club goods, see Swope, 2002, and Chakravarty and Fonseca, 2017).

[10]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.

[11]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.

to limit withdrawal since there is no social benefit from cooperation at the group level. This change can be seen as an indirect cost of exclusion if group members exclude more than one of them. We acknowledge that by using an exclusion mechanism that is not directly costly to the punishers, we probably observe an upper bound of exclusion decisions.

The payoff function of player i for a period t becomes:

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

with n the number of active members in the group.

Excluding a member for one period requires at least two votes. 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, it is not possible to exclude anyone.[12] 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 are informed on the ID of each excluded member, if any. They are never informed on the number of votes cast to exclude someone. A member who is not excluded never knows whether another member voted for his exclusion.

Exo3P Treatment

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, he will remain inactive in periods $t+1$, $t+2$ and $t+3$. During each of these three periods, he receives the 10 ECU exclusion payoff and he is informed on the withdrawal decisions and earnings of the active members and on the exclusion of other members. In period 18 (19, respectively), since there are only two (one, resp.) periods left before the end of the game, if a group member is excluded, he will be excluded for two (one, resp.) periods only. In the last period it is not possible to exclude anyone. This change in procedure for the last three periods is made common information.[13] At the end of the second stage, like in the Exo1P treatment, subjects receive a feedback on the ID of each player that has been excluded, if any. They are also given the number of remaining periods of exclusion of the group members that have been excluded earlier.

Endo Treatment

In this treatment, the length of exclusion is endogenous. This has two implications. First, the number of periods of exclusion is no longer unique; it is restricted to the interval

[12]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.

[13]We chose this procedure to avoid having no risk of sanction in three periods whereas severe sanctions were available in the previous periods. An alternative procedure would have been to introduce random ending instead of deterministic ending in the game.

[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, he also has to decide for how many periods to exclude this member.[14]

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 member, this group member will not be excluded, therefore 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. At the end of the second stage, group members are informed about the members who have been excluded and, if any, about the number of periods of exclusion. They are 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.[15] 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).[16]

In this preliminary part, we formed groups of four subjects randomly. Subjects had

[14]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. 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.

[15]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 that the session consisted of two parts but they received the instructions for part 2 only after completing part 1.

[16]A subject is classified as an unconditional free-rider if he withdraws the maximum possible amount of ECU regardless of his group members' average withdrawal. A subject is classified as an unconditional cooperator if he withdraws 0 ECU regardless of the group average withdrawal. A subject is classified as a conditional cooperator if his withdrawal increases when the group average withdrawal 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.

to make first an unconditional decision and then, conditional decisions. 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 withdrawn by three other group members. Therefore, each subject had 21 conditional decisions to make. For payment, one member among the four was randomly selected by the program and paid according to his conditional decision while the three other members were 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 withdrawn from the group account by the three other members. 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 were reassigned randomly. At the end of the session, we also collected standard information on the subjects' gender, age, student status and discipline.

3.2 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 with an average of 20 subjects per session. We collected observations for 12 groups in both the Endo and Exo1P treatments, 13 groups in the Exo3P treatment and 15 groups in the Baseline (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 taking game were distributed after completion of the preliminary part (see Appendix 1). We avoided the terms "exclusion" and "excluded members" in favor of more neutral terms, as "remove" and "players who do not participate in the group". After the instructions were read aloud, subjects had to complete a quiz to check their understanding. 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.

4 Behavioral Hypotheses

In the Baseline treatment the dominant strategy for a selfish and payoff-maximizing player is to withdraw the maximum from the group account in each period. On the opposite, the social optimum requires that players do not withdraw anything. In the treatments with possible exclusion, exclusion is indirectly costly if the excluded member does not free

ride in full or if the outcome of exclusion is that the structure of incentives is changed.[17] Indeed, each remaining group member will lose the marginal return of the amount left by that excluded member in the group account if not excluded. Thus, nobody should exclude.[18] 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 withdraws the maximum amount of ECU from the group account.*

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

These predictions may not hold if players have social preferences. Previous works on Voluntary Contribution Mechanism games have shown that in the presence of social dilemmas, it is possible to reach full cooperation by using punishment mechanisms (Yamagishi, 1986 ; Fehr and Gächter, 2000 ; Gächter *et al.*, 2008), and that the severity of punishment affects cooperation (*e.g.*, Nikiforakis and Normann, 2012; Neuhofer and Kittel, 2015). So, we anticipate that subjects may not be willing to withdraw all their units from the group acccount, especially when N>2. But when the number of members in the group is lower than 3, it is no longer socially inefficient to withdraw units from the group account since the social dilemma disappears. When N<3 only altruistic subjects should refrain from withdrawing since N*MPCR<1 (by not withdrawing, they increase the payoff of their group members but at a personal cost). This reasoning 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 when N>2. In smaller groups, most individuals withdraw at their full capacity.*

Subjects may be willing to exclude others if they believe that it is economically advantageous to exclude. This is the case if they believe that cooperation will improve in future periods, both because cooperators should no longer fear of being exploited by free-riders if the latter are excluded, and because reintegrated subjects should be more willing to comply with the norm of the group to prevent a new exclusion. Moreover, 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 that the willingness to exclude a group member depends on the size of the deviation between the withdrawal of the target and the norm of the group, as excluding a group member that deviates only by a little from the group average is certainly not profitable. The size of the deviation required to trigger an exclusion vote should be even higher in the case of longer exclusion. Finally, if they

[17]The 'trembling hand' refinement of the Nash equilibrium (Selten, 1975) could explain such mistakes in the withdrawal decisions (see also Maier-Rigaud *et al.*, 2010).

[18]Note that Hirshleifer and Rasmusen (1989) have shown that costless ostracism may lead to a cooperative equilibrium. Considering a n-player prisoner's dilemma game, ostracizing defectors is a subgame-perfect equilibrium when players follow the 'banishment strategy' according to which everyone cooperates and excludes defectors until the last round where everyone defects.

are willing to maximize the social surplus, cooperators should not be willing to exclude more than one group member when N=4 (and exclude no one when N=3), and they are expected to coordinate on the exclusion of the least cooperative player. This leads to the following hypothesis that is alternative to Hypothesis 2.a.

Hypothesis 2.b: *Some subjects are willing to exclude. If they care about the social surplus they should not exclude more than one member when N=4 and they should not exclude when N=3. If they are willing to exclude, they target the group members that withdraws more from the group account than the average of the group.*

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, and second, it may influence the willingness to seek revenge for past exclusion. These effects may depend on how the length of exclusion is determined.

Subjects who have been excluded should be more willing to conform to the withdrawal norm of the group after reintegration than before being excluded. A more severe punishment should lead to more compliance after reintegration than a milder one. On the other hand, the 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: *A longer exclusion decreases more withdrawals after reintegration than a shorter exclusion.*

Hypothesis 4: *A longer exclusion increases more the risk of retaliation than a shorter exclusion.*

From a standard economic perspective, a similar severity of sanction should have the same effect on behavior after reintegration in both Exo and Endo treatments. However, from a behavioral point of view, a better ability to perceive the punishers' intentions through the choice between several lengths of exclusion may influence the reaction of those who are punished. According to Braithwaite (1989), reintegrative shaming (compared to disintegrative shaming) takes place when disapproval is partnered with respect for the individual, forgiveness, and acceptance back into the group. The choice of the length of exclusion conveys a richer signal on the group's level of disapproval, forgiveness, and willingness to reintegrate. In contrast, it is more difficult to infer the punishers' forgiveness when a single level of severity is available. Being excluded for three periods when exclusion for one or two periods was possible may be interpreted as a harsher decision than being excluded for three periods when severity could not be modulated. As a result, the excluded members in this context may express their anger after being reintegrated by complying less and by casting more votes for excluding the punishers than when the latter did not choose the length of exclusion. In the same line, being excluded for one period in this context may be interpreted as a milder and more forgiving punishment, which may be rewarded by a higher willingness to comply to the norm of the group and by less retaliation after reintegration than when punishers do not choose the severity of punishment. This leads to our last hypotheses.

Hypothesis 5: *The effect of long (short, respectively) exclusion on cooperation after reintegration is smaller (larger, resp.) when the length of exclusion is chosen by the punishers than when it is exogenous.*

Hypothesis 6: *The difference in the frequency of retaliation between those excluded for a longer duration and those excluded for a short duration is larger when the length of exclusion is chosen by the punishers than when it is exogenous.*

5 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 withdrawal behavior. 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 main variables of interest (withdrawals, length of exclusion, and payoff), by treatment.

5.1 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 against Hypothesis 1.a.

Support for Result 1. Figure 1 displays the evolution over time of the average withdrawal from the group account by each active group member, by treatment, when N>2 (*i.e.*, when there is a social dilemma; see the same figure in Appendix 3 for all groups regardless of their size). 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 mean individual level of withdrawal is lower on average in the treatments in which exclusion is possible (for all groups: Exo1P: 10.65 ECU; Exo3P: 6.43; Endo: 7.81; for groups with N>2: Exo1P: 10.13 ECU; Exo3P: 4.97; Endo: 7.28) compared to the Baseline (16.94 ECU) (Mann-Whitney tests - MW, hereafter: for all groups: Exo1P: p=0.005; Endo: p=0.008; Exo3P: p <0.001; for groups with N>2: Exo1P: p=0.004; Endo: p=0.007; Exo3P: p <0.001).[19] Considering period 1 in isolation (N=4), the mean withdrawal is also lower in the treatments with possible exclusion (Exo1P: 7.73 ECU; Exo3P: 6.81; Endo: 7.56) compared to the Baseline (12.18 ECU) (MW tests: Exo1P: p=0.002; Endo: p=0.002; Exo3P: p <0.001). This confirms the previous literature showing that the threat of exclusion itself enhances cooperation even before getting any information about others' behavior.

[19]The values indicated in Table 1 are means of individual decisions. But for the non-parametric tests we consider only one observation per group. For the Mann-Whitney tests, the mean withdrawal over all periods for each group gives one independent observation, except when we consider only the first period, in which case the mean withdrawal of each group in this period gives one independent observation. All tests are two-tailed.

Table 1: Summary statistics

	Baseline	Exo1P	Exo3P	Endo
Average withdrawal	16.94	10.65***	6.43 ***	7.81***
(All groups and active members)	(5.40)	(8.20)	(7.56)	(8.60)
Average withdrawal	16.94	10.13***	4.97***	7.28***
(Groups with N>2)	(5.40)	(8.13)	(6.44)	(8.38)
Average withdrawal	-	16.76	15.32	14.84
(Groups with N<3)		(6.45)	(7.78)	(8.51)
Average withdrawal	16.94	7.38***	3.93***	5.76***
of non-excluded members (N>2)	(5.40)	(7.58)	(5.61)	(7.85)
Average withdrawal	-	16.43	12.31***	15.86
of excluded members (N>2)		(5.39)	(7.24)	(5.65)
Average number of votes	-	4.05	2.04*	2.09*
in favor of exclusion (N>2)		(3.19)	(2.42)	(2.52)
(with a max. of 12)				
Average number of	-	0.88	0.33***	0.43**
excluded subjects		(1.30)	(0.73)	(0.95)
(with a max. of 4)				
Average length of	-	1	3***	2.08***
exclusion (nb periods)		(0.00)	(0.00)	(0.83)
Average number of exclusion votes	-	1.64	1.29	1.35
cast by excluded members		(1.30)	(1.19)	(1.18)
after their reintegration (N>2)				
Average payoff	21.93	21.58	22.55	22.51
(in ECU, all players)	(4.68)	(8.16)	(8.62)	(8.69)
Average payoff of non-excluded	21.93	24.82**	26.23***	26.15*
members (in ECU)	(4.68)	(6.09)	(6.02)	(6.22)
Number of observations	1200	960	1040	960

Notes: The Table reports mean values, with standard deviations in parentheses. It 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. The withdrawals per period are measured by the number of ECU taken from the group account. The average withdrawal of excluded members is at the period of exclusion. The average number of active members in favor of exclusion is the mean number of subjects who cast at least one vote to exclude another member when N>2. The average number of excluded members considers only the cases where a new subject has been excluded. The average length of exclusion is conditional on being excluded.

Cooperativeness is higher in an environment with a more severe potential punishment (consistent with *e.g.*, Nikiforakis and Normann, 2008). As suggested by Figure 1, when N>2 we find significant differences in the level of withdrawal depending on the length of exclusion when comparing Exo1P and Exo3P (MW test: p=0.014), which supports our hypothesis that the length of exclusion matters. We find no significant difference between Endo and the two other treatments with exclusion (MW test: Exo1P: p=0.488; Exo3P: p=0.301). Not surprisingly, when N<3 the average level of withdrawal is close to the maximum, since there is no more social dilemma. It remains, however, significantly lower than 20 (binomial tests: p <0.001 for all treatments), possibly because cooperators do not want to make a decision that could be interpreted as a signal of future free-riding when the size of the group will increase again, or because of possible confusion. The percentages of groups with less than three members is 23.75% in Exo1P, 27.31% in Exo3P and 22.08% in Endo.

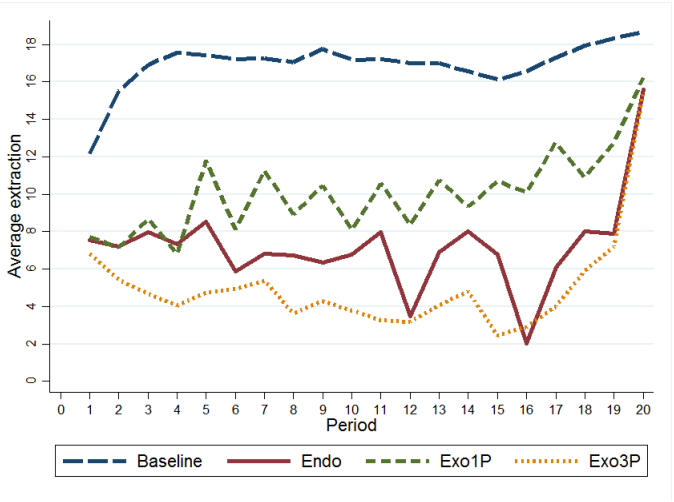


Figure 1: Evolution of average withdrawal over time, by treatment (includ ing all the active subjects when N>2).

Interestingly, in the first period there is no significant difference between Exo1P (7.73 ECU), Exo3P (6.81 ECU) and Endo (7.56 ECU) (MW tests: Exo1P *vs.* Endo: p=0.931; Exo1P *vs.* Exo3P: p=0.643; Endo *vs.* Exo 3P: p=0.703). 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 a punishment mechanism is available or not, but not on the severity of this punishment. This differs from Neuhofer and Kittel (2015), probably because the difference in the threat between a three-period and a one-period exclusion is much smaller than the difference between irreversible exclusion and a one-period exclusion. This suggests that individuals perceive the risk of being punished if they do not cooperate, but they are not able to adjust their risk perception to the severity of punishment when they do not know yet what the group norm is and when there is no clear compliance rule.[20] Individuals can probably better adjust their behavior to the severity of sanctions in the first period of a game when they are informed *ex ante* of the deterrence regime applied to

[20]Note that Nikiforakis and Normann (2008) also observed no difference in the first period between treatments implementing different punishment ratios in a public goods game.

a precise behavior (*e.g.*, De Angelo and Charness, 2012; Schildberg-Horisch and Strassmair, 2012).

Regarding efficiency, as indicated in Table 1, including all periods and groups, the average payoffs in Exo1P (21.58 ECU), Exo3P (22.55 ECU) and Endo (22.51 ECU) do not differ significantly from the payoffs in the Baseline (21.93 ECU) (MW tests: Exo1P *vs.* Baseline: p=0.329; Exo3P *vs.* Baseline: p=0.800; Endo *vs.* Baseline: p=0.591). However, an environment with possible exclusion significantly increases the payoffs of the non-excluded members (21.93 ECU in the Baseline, 24.82 ECU in Exo1P, p=0.028; 26.23 ECU in Exo3P, p <0.001; 26.15 ECU in Endo, MW test: p=0.088). 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.

5.2 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.

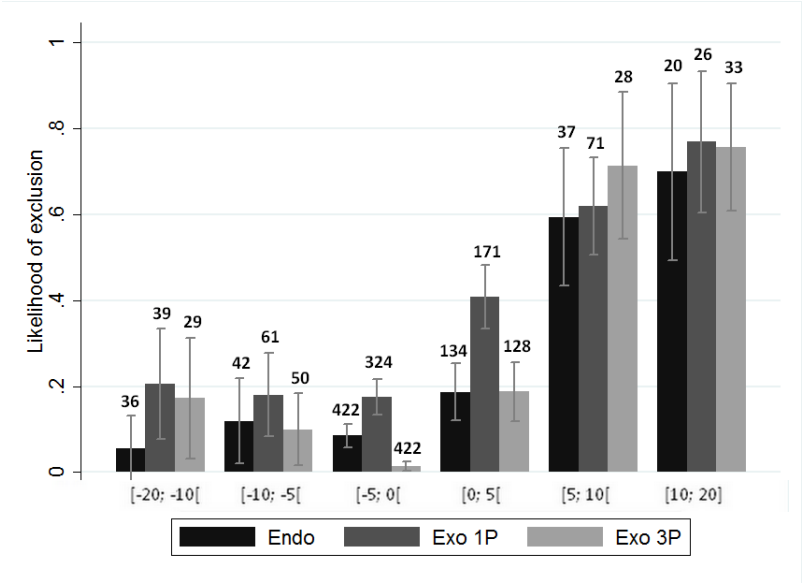
Result 2: *Individuals are willing to punish and generally they exclude the least cooperators.* This supports Hypothesis 2.b. against Hypothesis 2.a. *But contrary to Hypothesis 2.b., the size of the group has no effect. The choice of a length of exclusion depends on current and chronic free-riding.*

Support for Result 2. As visible in Table 1, the average number of individual votes in favor of the exclusion of another member in a period when groups have more than two subjects (out of a maximum of 12) is higher in Exo1P (4.05) than in Exo3P (2.04) and in Endo (2.09), but the difference is only marginally significant (MW tests: Exo1P *vs.* Exo3P: p=0.073; Exo1P *vs.* Endo: p=0.069; Endo *vs.* Exo3P: p=0.849). As a result, there are more people excluded in Exo1P (0.88) than in Endo (0.43) and in Exo3P (0.33) (out of a maximum of four) (MW tests: Exo1P *vs.* Exo3P: p=0.003; Exo1P *vs.* Endo: p=0.028). The difference between Endo and Exo3P is not significant (MW test: p=0.478).

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), taken as the group norm.[21] Each bar represents the proportion of members who, among those sub-

[21]We consider the group average as the social norm in the group, based on the previous literature on sanctioning in social dilemma games. To support this point, in a post experimental questionnaire subjects were asked to report one or more reasons explaining their decision to exclude another participant at least once during the experiment. The proposed reasons were: (i) the participant was taking more than me, (ii) the participant was taking more than the average of the group, (iii) I wanted to signal that I would stop cooperating if the participant did not change her behavior, (iv) I was removed from the

jects, extracted x to y ECU more or less than their group average, have been excluded in each treatment. We can see that group members who extract on average 10 to 20 ECU more than the group average (bars on the far right of Figure 2) are more likely to be excluded than those who extract on average less than 5 ECU more than the group average. This is observed in all treatments.



Notes: The numbers above the bars indicate the number of observations. A negative (positive) difference means that the subjet withdraw less (more) than the average of his group.

Figure 2: Exclusion rate according to the difference between the individual withdrawal and the average withdrawal in the group, by treatment (N>2)

A regression analysis in which the dependent variable is the voting decision of subject i to exclude subject j in period t shows that the willingness to exclude does not depend on the size of the group and confirms that it increases significantly (at the 1% level) when j withdraws more than the group average. Tables A.2 and A.3 in Appendix 2 report the estimates from random-effects logit regressions with standard errors clustered at the group level. They show that for a given difference in withdrawal between the target's withdrawal and the group average, the probability to vote to exclude a group member is the same across treatments, regardless of the severity of punishment (the only exception is that in Endo a lower extraction by the target is less likely to be punished than in the other treatments). In these Tables, chronic free-riding is captured by the ratio between the current number of periods prior to t in which j has withdrawn more than the average group withdrawal and the current number of periods prior to t in which j was active in the group (considering only the cases in which the group size was higher than two). A higher ratio increases significantly (at the 1% level) the likelihood of a vote for exclusion.

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

group myself and I wanted to take revenge, (v) other reasons. 33.33% of the subjects reported reason (i), 52.53% reported reason (ii), 51.52% reason (iii), and 25.25% reason (iv).)

of exclusion is 2.08 periods, see Table 1). To investigate the determinants of the choice of a length we estimated random-effects ordered logit regressions. Our dependent variable is a categorical variable that equals 1 if individual i votes to exclude j for one period, 2 if i votes to exclude j for two periods, and 3 if i votes to exclude j for three periods. Table 2 reports the marginal effects of the independent variables on the vote to exclude for one period in columns (1) and (2), two periods in columns (3) and (4), and three periods in columns (5) and (6). The independent variables include the difference between j's withdrawal and the group average withdrawal (excluding j) in period t when j withdraws more than the group (max $(0; w_{j,t} - \sum_{-j=1,-j\neq j}^{n-1} w_{-j,t})$) and the difference between the group average withdrawal and j's withdrawal in t when j withdraws less than his group (max $(0; \sum_{-j=1,-j\neq j}^{n-1} w_{-j,t} - w_{j,t})$). They include the ratio between the current number of periods prior to t in which j has withdrawn more than the average of his group and the current number of periods prior to t in which j was active in the group. They include a control for whether j has already been excluded in the past and a reintegration dummy that equals 1 if i is reintegrated in t and 0 otherwise. In model (2) reported in columns (2), (4) and (6), we add variables detailing the length of exclusion of i. The variable "i is reintegrated in t after xP exclusion" takes value 1 if i has been reintegrated in t after x period(s) of exclusion 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 include a time trend and its squared value to test whether exclusion votes are more likely at the beginning of the game when the expected benefits of punishment are potentially higher or to impose a norm as soon as possible. We control for group size and for individual characteristics (gender, age, student status and social preferences).

Table 2 shows that the vote for a specific length of exclusion is affected by the number of ECU taken by j relative to the average withdrawal in the group. There is also a significant effect of the reputation of the individual who is punished: a chronic positive deviation from the group average withdrawal leads to votes for a longer exclusion. When including this variable, the fact that j has already been excluded in the past does not affect the choice of a specific length of exclusion. 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. We analyze this behavior in sub-section 5.4. Note again that the group size has no significant effect.[22]

To summarize, both the probability to exclude and the severity of punishment increase not only in reaction to the importance of the deviation from the withdrawal norm of the group in the current period, but also in reaction to chronic free-riding over time. There is no evidence that exclusion votes are more frequent at the beginning of the game as the time trend is significant in no regression.[23] The choice of a longer exclusion depends also on the desire to take revenge after having suffered oneself from a more severe punishment.

[22]Note also that the category of social preferences of the subject is never significant in any regressions. Most of the time, the other individuals characteristics are not significant either.

[23]This is in contrast with standard experiments on monetary punishment in public goods game in which sanctions occur from the very beginning the game to impose a norm in the group. Maier-Rigaud *et al.* (2012) also observed that ostracism takes place mainly at the beginning of the game.

Table 2: Determinants of the length of exclusion in the Endo treatment

Dep. variable: Choice of i regarding the length of exclusion of j	Exclusion for 1 period (1)	1 period (2)	2 periods (3)	2 periods (4)	3 periods (5)	3 periods (6)
Max (0; $w_{j,t} - \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t}$)	-0.009*	-0.012*	-0.003*	-0.003*	0.013*	0.015**
	(0.005)	(0.006)	(0.002)	(0.002)	(0.007)	(0.007)
Max (0; $\sum_{-j=1,j\neq -j}^{n-1} w_{-j,t} - w_{j,t}$)	-0.006	-0.007	-0.002	-0.002	0.008	0.009
	(0.011)	(0.013)	(0.004)	(0.004)	(0.015)	(0.017)
Ratio current nb of pos. dev. of j prior to t/	-0.121*	-0.116*	-0.040**	-0.033*	0.162**	0.150*
Nb of periods with j in the group prior to t	(0.069)	(0.067)	(0.020)	(0.020)	(0.083)	(0.078)
Exclusion of j prior to t	-0.060	-0.046	-0.020	-0.013	0.080	0.060
(dummy)	(0.064)	(0.076)	(0.029)	(0.027)	(0.093)	(0.103)
i is reintegrated in t	-0.125*	0.148	-0.042***	0.042	0.167**	-0.192
(dummy)	(0.066)	(0.122)	(0.015)	(0.043)	(0.074)	(0.159)
i is reintegrated in t	-	-0.360**	-	-0.102*	-	0.466**
after 2P exclusion (Endo)		(0.179)		(0.061)		(0.208)
i is reintegrated in t	-	-0.362**	-	-0.103**	-	0.469**
after 3P exclusion (Endo)		(0.183)		(0.049)		(0.196)
Group size in t	-0.073	-0.070	-0.024	-0.020	0.098	0.090
	(0.072)	(0.074)	(0.025)	(0.022)	(0.095)	(0.094)
Period	-0.014	-0.017	-0.005	-0.005	0.019	0.022
	(0.012)	(0.013)	(0.004)	(0.004)	(0.016)	(0.016)
Squared Period	0.000	-0.000	-0.000	0.000	-0.000	-0.000
	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	337	337	337	337	337	337
Clusters	11	11	11	11	11	11
Log pseudo-likelihood	-263.15	-257.94	-263.15	-257.94	-263.15	-257.94

Notes: Table 2 reports the marginal effects from two random-effects ordered logit regressions with robust standard errors clustered at the group level in parentheses. The second model, corresponding to columns (2), (4), and (6), differs from the first model simply by controlling for the length of exclusion of player i. 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). Pos. dev. for Positive deviation ($w_{j,t} - \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t}$) >0). Individual controls include gender, age, student status, and category of social preferences. *** p <0.01, ** p <0.05, * p <0.10.

5.3 Reintegration

The analysis of reintegration patterns delivers two main findings.

Result 3: *Among punished individuals, a long exclusion has a higher disciplining effect on withdrawal levels in the reintegration period compared to a one-period exclusion, but this is observed only in the Exo treatments.* Hypothesis 3 is not supported in the Endo treatment.

Result 4: *Among punished individuals, the disciplining effect of a three-period exclusion*

(one-period exclusion, respectively) on the withdrawal level after reintegration is smaller (is similar) when the length of exclusion results from a vote compared to when the length of exclusion cannot be modulated and is fixed exogenously. Hypothesis 5 is not supported when exclusion is short.

Support for Results 3 and 4. Figure 3 displays the difference between the number of ECU withdrawn 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 when N>2 (*i.e.*, when the game involves a social dilemma; see a similar figure in Appendix 3 for all groups regardless of their size).[24] We can see that in every treatment, the average difference with the group average is lower during the reintegration period than during the exclusion period. Subjects reintegrated after three periods of exclusion continue to withdraw more than their group average in the reintegration period in both Endo (+1.89) and Exo3P (+2.25), whereas subjects reintegrated after one period of exclusion comply to the norm of the group in Exo1P (+0.35) and withdraw less than the group average in Endo (-0.92) but the variance is also large.[25]

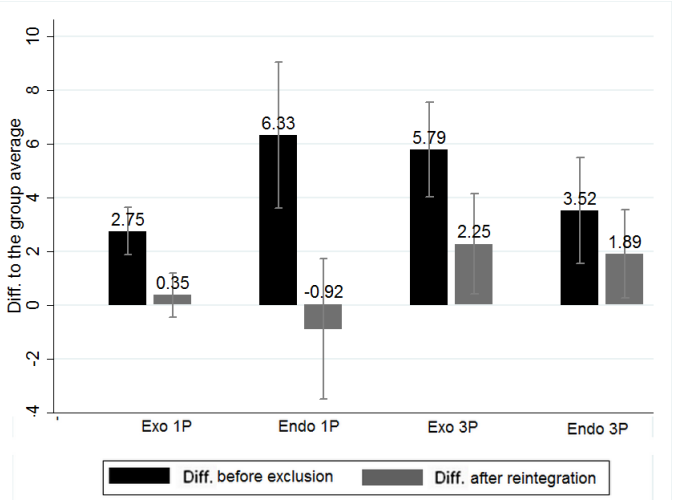


Figure 3: Difference between the number of ECU withdrawn by excluded subjects and by their group members in the exclusion period (left bars) and in the reintegration period (right bars), by treatment when N>2.

The fact that subjects who have been excluded for a longer time in Endo withdraw more after reintegration relative to subjects who were excluded for a shorter duration may result from three motives. First, there is a selection effect: these subjects are more likely to be strong free-riders. Second, they may be willing to signal their anger to their group members after such a harsh punishment. Third, subjects excluded for a longer duration may attempt at making up after reintegration for lost earnings during exclusion.

[24] Although not illustrated on Figure 3 to keep the comparison symmetric, group members excluded for two periods in Endo take on average 0.33 ECU more than their group members in the reintegration period.

[25] 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.

Table 3: Determinants of the evolution of the level of withdrawal over time

Dependent variable:	$(wi_t - wi_{t-2})$	$(wi_t - wi_{t-4})$
Evolution of number	All 1P	All 3P
of ECU withdrawn	(1)	(2)
Reintegrated after exclusion	-5.522***	-5.184***
	(1.276)	(1.047)
Reintegrated after exclusion	0.906	0.883
*Endo	(1.477)	(1.308)
Group average in t-1	0.099**	0.006
	(0.044)	(0.067)
$GA_{t-1} - GA_{t-2}$	0.019	-
	(0.075)	
$GA_{t-1} - GA_{t-4}$	-	0.212*
		(0.110)
Group size in t	1.172	1.591**
	(0.763)	(0.713)
Period	-0.826***	-1.396***
	(0.173)	(0.319)
Squared period	0.041***	0.069***
	(0.008)	(0.013)
Individual controls	Yes	Yes
Constant	-0.514	-0.252
	(3.654)	(3.709)
Observations	1097	1053
Clusters	24	25
R-squared	0.10	0.16

Notes: This Table 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 in groups with more than two members, excluding subjects who were excluded after period 17. Individual controls include gender, age, student status, and category of social preferences. *** p <0.01, ** p <0.05, * p <0.10.

Using regressions allows us to compare the effect of exclusion on the behavior of the excluded subjects after reintegration, controlling for the average withdrawal in the group in the period prior to reintegration and for the time of the decision. A regression analysis in which the dependent variable is the number of ECU withdrawn from the group account by an excluded subject in period t when he reintegrates his group is reported in Table A.4 in Appendix 2. It is based on the estimates of random-effects GLS with robust standard errors clustered at the group level. This analysis reveals a significant difference, at the 5% level, between Exo3P and Exo1P: reintegrated subjects withdraw less after three periods of exclusion than after one, controlling for the group average in the period preceding immediately reintegration (model (1)). In contrast, choosing one-period exclusion in 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 is associated with a significantly (at the 5% level) higher level of withdrawal during the reintegration period. This supports Result 3.

Model (1) also shows that the disciplining effect of a one-period exclusion is the same in the Exo and the Endo treatments, supporting Result 4 against Hypothesis 5.

To control for the subjects' withdrawal behavior before exclusion and to compare them with non-excluded members, Table 3 reports random-effects GLS regressions with robust standard errors clustered at the group level in which 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. These regressions exclude subjects that cannot be observed in both t and t-2 or t-4. If a subject has been excluded for one period (for three periods, respectively) and reintegrated in t, this indicates the difference in withdrawal levels between t, the period of reintegration, and t-2 (t-4, resp.), 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 the subject was 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 could be modulated by the group members. 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. We add a time trend and its squared value, and controls for individual characteristics. The data include the data from subjects reintegrated in periods 3 to 20 in groups with N>2, excluding subjects excluded after period 17.

The two models show that controlling for the level and the evolution of the group average withdrawals, being reintegrated after being excluded for one or for three periods decreases the extraction level between the period of exclusion and the period of reintegration, compared to individuals who have not been excluded. In both models there is no significant difference between the Exo and the Endo treatments. There is a non-linear effect of time, indicating an improvement of cooperation over time but at a declining rate.

5.4 Retaliation

In this section we focus on retaliation behavior by examining the voting behavior of the excluded members after their reintegration, conditional on how long they have been excluded and by which procedure. Our last results are the following.

Result 5: *Subjects who have been excluded for a shorter duration are less likely to retaliate when they reintegrate their group than subjects excluded for a longer duration. This is found only when the length of exclusion has been chosen by group members.* This supports Hypothesis 4 for the Endo treatment but not for the Exo treatments.

Result 6: *The difference in the impact of long vs. short exclusion on the willingness to retaliate is larger when the length of exclusion is chosen by group members than when it is exogenous.* This supports Hypothesis 6.

Support for Results 5 and 6. An extreme form of retaliation consists for a previously excluded subject of voting to exclude all the other active group members after reintegra-

tion ("Total revenge"). In Exo1P 43.39% of the reintegrated subjects vote for excluding all the other group members. This percentage is only 8.33% for reintegrated subjects excluded for one period in Endo. It is 26.56% in Exo 3P but it rises to 50% for those excluded for two periods in Endo and 45.71% for those excluded for three periods in Endo. The percentage in Exo1P is significantly higher than in Exo3P (two-tailed two-sample test of proportions: p=0.017). The difference between Endo and the other treatments with exclusion is not significant (35.80%, regardless of the length of exclusion) (two-tailed two-sample tests of proportions: Endo *vs.* Exo1P: p=0.246; Endo *vs.* Exo3P: p=0.235). [26]. One-period exclusion limits counter-punishment in Endo, but not when the length of exclusion is not chosen. Does the size of the group influence the willingness to vote for total revenge? When N=3 and the group size remains constant after reintegration, a subject who has been excluded knows for sure that the two other group members voted for his exclusion and he can target them directly. This is less precise when N=4 since people may have been excluded by two or by three group members. In fact, when N=4 total revenge occurs in 40.28% of the cases, whereas when N=3 it occurs in 27.45% of the cases. Thus, subjects vote for total revenge not only when they are sure of who in the group has made the decision to exclude.

We next investigate the determinants of the excluded subjects' willingness to cast 0, 1, 2 or 3 exclusion votes when they reintegrate their group, using a random-effects Poisson regression with standard errors clustered at the group level (model (1) in Table 4). The dependent variable is a categorical variable equal to the number of exclusion votes by the reintegrated subject. The independent variables include dummy variables for each possible length of exclusion of i interacted with each treatment, setting Exo1P as the reference category. We control for the amount withdrawn by i in t, the average amount withdrawn by the other group members in t, the group size at the time of reintegration,[27] a time trend and its squared value, and the usual individual characteristics. The sample pools the data from all the treatments with exclusion but includes only the excluded subjects who reintegrate their group in period t. We only consider periods 3 to 17 and groups where exclusion is possible (N > 2).

Model (1) shows that members reintegrated after one-period exclusion in Endo cast significantly less exclusion votes than those excluded in Exo1P. Members reintegrated after three periods in Endo cast more exclusion votes than those excluded in Exo1P. But there is no significant difference between Exo1P and Exo3P. A longer exclusion increases significantly the number of exclusion votes cast by reintegrated members only in Endo. A similar regression conducted only on the Endo data indicates that the marginal effect of a two-period exclusion is 0.789*** (0.236) and the effect of a three-period exclusion is 1.100*** (0.192), compared to a one-period exclusion. The impact of the length of exclusion on retaliation is thus larger in Endo than in the Exo treatments.

[26] 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.

[27] Alternatively we replaced the group size in t with the group size at the time of exclusion in these regressions and in the following ones. This variable was never significant.

Table 4: Effect of the length of exclusion on the number of exclusion votes and on the likelihood of total revenge by reintegrated subjects

Dependent variables:	Number of exclusion votes	Total revenge
	(1)	(2)
1P exclusion in Exo1P	*Ref.*	*Ref.*
1P exclusion in Endo	-0.630***	-0.292***
	(0.180)	(0.073)
2P exclusion in Endo	0.026	0.181
	(0.196)	(0.201)
3P exclusion in Endo	0.275**	0.232
	(0.113)	(0.176)
3P exclusion in Exo3P	0.131	0.043
	(0.208)	(0.203)
Amount withdrawn	0.009	0.016**
by i in t	(0.007)	(0.007)
Group average withdrawal	0.030***	0.018*
in t excluding i	(0.012)	(0.009)
Number of exclusion	0.039	0.081
prior to t	(0.057)	(0.064)
Group size in t	0.456***	0.002
	(0.116)	(0.113)
Period	-0.052	-0.035
	(0.053)	(0.052)
Squared period	0.001	0.001
	(0.002)	(0.002)
Individual controls	Yes	Yes
Constant	-1.400	
	(0.582)	
Observations	323	323
Clusters	35	35
Log pseudo-likelihood	-479.24	-167.33

Notes: Model (1) reports the marginal effects from a random-effects Poisson model where the dependent varable is the number of exclusion votes cast by excluded members when they reintegrate the group. Model (2) reports the marginal effects from a random-effects logit regression where the dependent varable is the probability of the reintegrated members to vote for total revenge (*i.e.*, for the exclusion of all the 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. Standard errors (in parentheses) are clustered at the group level. *** p <0.01, ** p <0.05, * p <0.10.

To explore the determinants of voting for total revenge, model (2) reports the estimates of 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 reintegrated member votes for the exclusion of all the other active group members, and 0 otherwise. The independent variables are the same as in model (1). Controlling for individual and group average withdrawals, members excluded for one period in Endo are significantly less likely

to vote for the exclusion of all the other members when reintegrating their group compared to similar subjects in Exo1P. There is no significant difference in the likelihood of total revenge between members excluded for one or for three periods when the length of exclusion is fixed exogenously. A similar regression conducted only on the data of the Endo treatment indicates that the marginal effect of a two-period exclusion is 0.845*** (0.121) and the effect of a three-period exclusion is 0.821*** (0.117), compared to a one-period exclusion. The impact of the length of exclusion on total revenge is thus also larger in Endo than in the Exo treatments.

To conclude, subjects excluded for a longer time are more likely to seek revenge than those 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 in itself that matters so much in triggering negative reciprocity, but the perceived intentions of the group. There is no visible time trend in these effects.

6 Conclusion

The exclusion of non-compliers is a strong form of punishment. One question is to know whether the threat of exclusion is sufficient to discipline members to follow the norm of the community after rehabilitation. Another question is to know the extent to which people who reintegrate their group after being excluded are willing to cooperate with those who previously excluded them. The answer to this under-studied question may depend on the length of exclusion and on how this length has been determined. To address this question we have conducted a controlled experiment based on a taking game that has the properties of a social dilemma game when the group size is not too small. In this game, group members can vote for excluding their peers for different durations and the severity of punishment is either imposed exogenously or chosen by the group members.

Our findings show that individuals are willing to ostracize others, in particular those who withdraw more: larger deviations from the group average withdrawal behavior determine both the vote for exclusion and the choice of the length of exclusion. People also punish more, and more severely, chronic defections. In return, a longer exclusion has a higher disciplining effect on cooperation after reintegration compared to a shorter exclusion, but only when the length of exclusion is not chosen by group members. The disciplining effect of a longer exclusion on the withdrawal level after reintegration is smaller when the length of exclusion results from a vote compared to when the length of exclusion cannot be modulated and is fixed exogenously. Another important dimension to be considered in the evaluation of the effects of exclusion is whether previously excluded individuals are willing to retaliate. We found that when group members choose the length of exclusion, individuals excluded for a shorter duration are less likely to retaliate when they reintegrate their group compared to people excluded for a longer duration. The difference in the impact of long *vs.* short exclusion on the willingness to retaliate is larger when the length of exclusion is chosen by group members than when it is exogenous.

Our results highlight the importance of quick reintegration of the individuals who

have been punished to sustain cooperation and limit anti-social behavior after reintegration when the community members choose the length of exclusion. 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. This is consistent with the 'reintegrative shaming' theory in criminology (Braithwaite, 1989) and the results of Coricelli *et al.* (2014) showing in a tax evasion context that pardon reduces future violations whereas in the absence of pardon, offenders have no more moral motivation to comply once their reputation is lost.

We acknowledge some limitations of our study. The differences in behavior between our Exo and Endo treatments may be driven by two differences in the design. First, in the Endo treatment the excluded individuals were able to compare the sentence they received with another, non-selected, sentence. In this context, learning that one receives the less severe sanction may help accept the sentence compared to a situation in which only one sentence is available. 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 effects. Similarly, the counter-effects of a long exclusion may result from two dimensions. A longer exclusion may trigger negative emotions against group members who could have signaled their disapproval with a milder sanction and it may encourage people to free-ride more on others to recover the lost earnings during exclusion. But the possibility to choose the length of exclusion helps also screen the free-riders. Those who are repeatedly less compliant may also be those who are more likely to behave aggressively towards others after reintegration. The fact that in our regressions we never found any effect of the type of social preferences on behavior speaks in favor of the first effects. But to weight each of these effects precisely, a random assignment of the length of exclusion and measures of emotions might be useful. Also, assigning the excluded members to different groups after reintegration would allow us to disentangle the effect of negative reciprocity from that of the willingness to recover a loss. We leave this to further investigation.

References

[1] Ahn, T.-K., Isaac, R., and Salmon, T. C. Endogenous group formation. *Journal of Public Economic Theory 10*, 2 (2008), 171–194.

[2] Ahn, T.-K., Isaac, R., and Salmon, T. C. Coming and going: experiments on endogenous group sizes for excludable public goods. *Journal of Public Economics 93* (2009), 336–351.

[3] Akpalu, W., and Martinsson, P. Ostracism and common pool resource management in a developing country: Young fishers in the laboratory. *Journal of African Economies 21*, 2 (2011), 266–306.

[4] Ambrus, A., and Greiner, B. Democratic punishment in public good games with perfect and imperfect observability. *Mimeo* (2015).

[5] Balafoutas, L., Garcia-Gallego, A., Georgantzis, N., Jaber-Lopez, T., and Mitrokostas, E. Rehabilitation and social behavior: Experiments in prison. *Mimeo* (2017).

[6] Becker, G. S. *Crime and punishment: An economic approach.* Springer, 1968.

[7] Bernard, M., Fanning, J., and Yuksel, S. Finding cooperators: Sorting through repeated interaction. *Mimeo* (2017).

[8] Berninghaus, S., Bleich, S., and Güth, W. The consequences of (de)regulation on employment duration and efficiency: An experimental study. *Economic Inquiry 51*, 1 (2013), 1050–1065.

[9] Berube, D., and Green, D. The effects of sentencing on recidivism: Results from a natural experiment. *Paper presented at the Second Annual Conference on Empirical Legal Studies, New York* (2007).

[10] Bhuller, M., Dahl, G. B., Loken, K. V., and Mogostad, M. Incarceration, recidivism and employment.

[11] 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.

[12] Bock, O., Baetge, I., and Nicklisch, A. Hroot: Hamburg registration and organization online tool. *European Economic Review 71* (2014), 117–120.

[13] Bolling, D., Pitskel, N., and Deen, B. e. a. Dissociable brain mechanisms for processing social exclusion and rule violation. *Neuroimage 54* (2011), 2462–2471.

[14] Bowles, S., and Gintis, H. The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theoretical Population Biology 65*, 1 (2004), 17–28.

[15] Braithwaite, J. *Crime, shame and reintegration.* Cambridge University Press, 1989.

[16] Brown, M., Falk, A., and Fehr, E. Relational contracts and the nature of market interactions. *Econometrica 72*, 3 (2004), 747–780.

[17] Chakravarty, S., and Fonseca, M. A. Discrimination via exclusion: An experiment on group identity and club goods. *Journal of Public Economic Theory 19*, 1 (2017), 244–263.

[18] Charness, G., and Yang, C.-L. Starting small toward voluntary formation of efficient large groups in public goods provision. *Journal of Economic Behavior & Organization 102* (2014), 119–132.

[19] 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.

[20] 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.

[21] Croson, R., Fatas, E., Neugebauer, T., and Morales, A. J. Excludability: A laboratory study on forced ranking in team production. *Journal of Economic Behavior & Organization 114* (2015), 13–26.

[22] Davis, B. J., and Johnson, D. B. Water cooler ostracism: Social exclusion as a punishment mechanism. *Eastern Economic Journal 41*, 1 (2015), 126–151.

[23] 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.

[24] Eisenberger, N., Lieberman, M., and Williams, K. D. Does rejection hurt? an fmri study of social exclusion. *Science 302* (2003), 290–292.

[25] Falk, A., Huffman, D., and Macleod, W. B. Institutions and contract enforcement. *Journal of Labor Economics 33*, 3 (2015), 571–590.

[26] Fatas, E., and Mateu, G. Antisocial punishment in two social dilemmas. *Frontiers in Behavioral Neuroscience 9* (2015).

[27] Fatas, E., Morales, A. J., and Ubeda, P. Blind justice. an experimental analysis of random punishment in team production. *Journal of Economic Psychology 31*, 3 (2010), 358–373.

[28] Fehr, E., and Gächter, S. Cooperation and punishment in public goods experiments. *American Economic Review 90* (2000), 980–994.

[29] Feinberg, M., Willer, R., and Schultz, M. Gossip and ostracism promote cooperation in groups. *Psychological Science 25*, 3 (2014), 656–664.

[30] Fischbacher, U. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*, 2 (2007), 171–178.

[31] Fischbacher, U., Gächter, S., and Fehr, E. Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters 71*, 3 (2001), 397–404.

[32] Fischer, S., and Güth, W. Effects of exclusion on acceptance in ultimatum games. *Journal of Economic Psychology 33* (2012), 1100–1114.

[33] Gächter, S., Renner, E., and Sefton, M. The long-run benefits of punishment. *Science 322*, 5907 (2008), 1510–1510.

[34] 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.

[35] Green, D. P., and Winik, D. Using random judge assignments to estimate the effects of incarceration and probation on recidivism among drug offenders. *Criminology 48*, 2 (2010), 357–387.

[36] Güth, W., I., L. M.-V., Sutter, M., and Van der Heijden, E. Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics 91* (2007), 1023–1042.

[37] Hirshleifer, D., and Rasmusen, E. Cooperation in a repeated prisoners' dilemma with ostracism. *Journal of Economic Behavior & Organization 12*, 1 (1989), 87–106.

[38] 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.

[39] Kling, J. R. Incarceration length, employment, and earnings. *American Economic Review 96*, 3 (2006), 863–876.

[40] 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).

[41] Kopanyi-Peuker, A., Offerman, T., and Sloof, R. Team production benefits from a permanent fear of exclusion. *European Economic Review 103* (2018), 125–149.

[42] Kuziemko, I. Going off parole: How the elimination of discretionary prison release affects the social cost of crime. *National Bureau of Economic Research* (2007).

[43] Landerso, R. Does incarceration length affect labor market outcomes? *The Journal of Law and Economics 58*, 1 (2015), 205–234.

[44] Lott, J. R. Do we punish high income criminals too heavily? *Economic Inquiry 30*, 4 (1992), 583.

[45] 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.

[46] Martinson, R. What works?-questions and answers about prison reform. *The Public Interest*, 35 (1974), 22.

[47] Masclet, D. Ostracism in work teams: a public good experiment. *International Journal of Manpower 24*, 7 (2003), 867–887.

[48] Masten, C., Eisenberger, N., and Borofsky, 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.

[49] Maurin, E., and Ouss, A. Sentence reductions and recidivism: Lessons from the bastille day quasi experiment. *IZA Discussion Paper n3990* (2009).

[50] 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.

[51] Neuhofer, 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).

[52] Nikiforakis, N. Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics 92* (2008), 91–112.

[53] Nikiforakis, N., and Engelman, D. Altruistic punishment and the threat of feuds. *Journal of Economic Behavior & Organization 78* (2011), 319–332.

[54] Nikiforakis, N., and Normann, H.-T. A comparative statics analysis of punishment in public good experiments. *Experimental Economics 11* (2008), 358–369.

[55] Pressman, S. D., and Cohen, S. Does positive affect influence health? *Psychological Bulletin 131*, 6 (2005), 925.

[56] Riedl, A., and Aljaz, U. Exclusion and cooperation in social network experiments. *Mimeo* (2014).

[57] Riedl, A., Rohde, I., and Strobel, M. Efficient coordination in weakest-link games. *Review of Economic Studies 83*, 2 (2016), 737–767.

[58] Roach, M., and Schanzenbach, M. The effect of prison sentence length on recidivism: Evidence from random judicial assignment, 2016.

[59] Schildberg-Hörisch, H., and Strassmair, C. An experimental test of the deterrence hypothesis. *Journal of Law, Economics, and Organization 28*, 3 (2012), 447–459.

[60] Selten, R. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory 4*, 1 (1975), 25–55.

[61] Sheremeta, R. M., J., T. S., and Jingjing, Z. Creating self-sustained social norms through communication and ostracism. *Mimeo* (2009).

[62] 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.

[63] Swope, K. J. An experimental investigation of excludable public goods. *Experimental Economics 5*, 3 (2002), 209–222.

[64] Sääksvuori, L. Intergroup conflict, ostracism, and the evolution of cooperation under free migration. *Behavioral Ecology and Sociobiology 68*, 8 (2014), 1311–1319.

[65] 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.

[66] Wesselmann, E. D., and Williams, K. D. *Ostracism and Stages of Coping*. Oxford: Oxford University Press, 2013.

[67] 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.

[68] 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.

[69] 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.

[70] Williams, K. D. *Ostracism: The power of silence*. Guilford Press, 2002.

[71] Yamagishi, T. The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology 51*, 1 (1986), 110.

[72] 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.

[73] Zapryanova, M. The effects of time in prison and time on parole on recidivism. *Mimeo* (2014).

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:

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

Example:

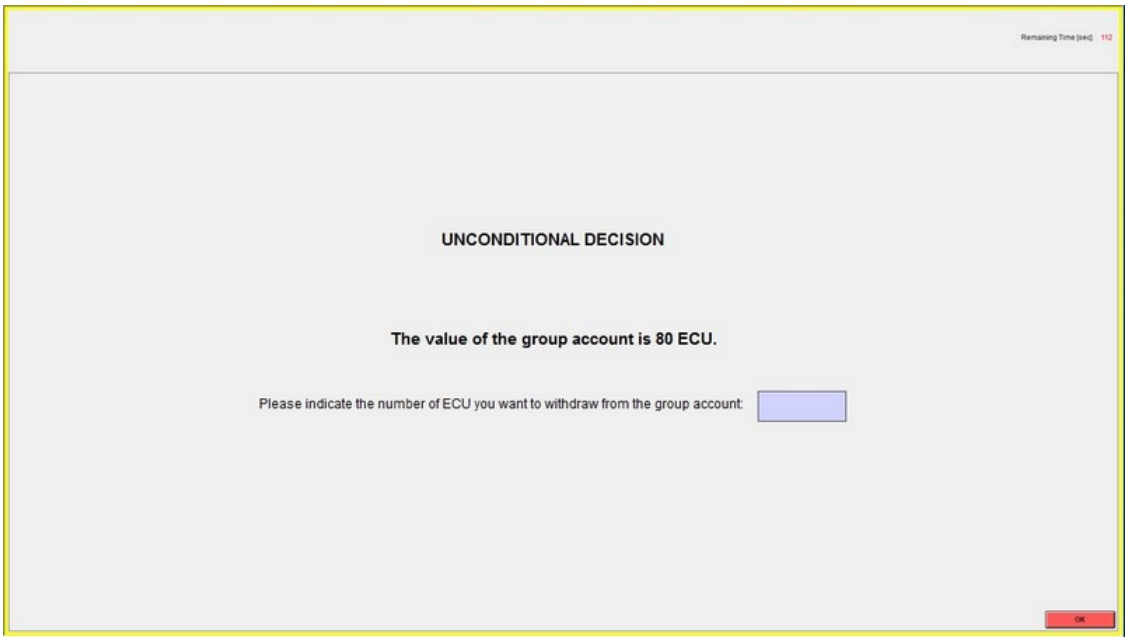
Suppose that you withdraw 10 ECU from the group account and each of the other 3 group members withdraws 1 ECU from the group account. The total amount left on the group account is 80 ECU minus the 10 ECU that you have withdrawn, minus 3 ECU withdrawn by the other group members, *i.e.* 67 ECU. In this 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 ».

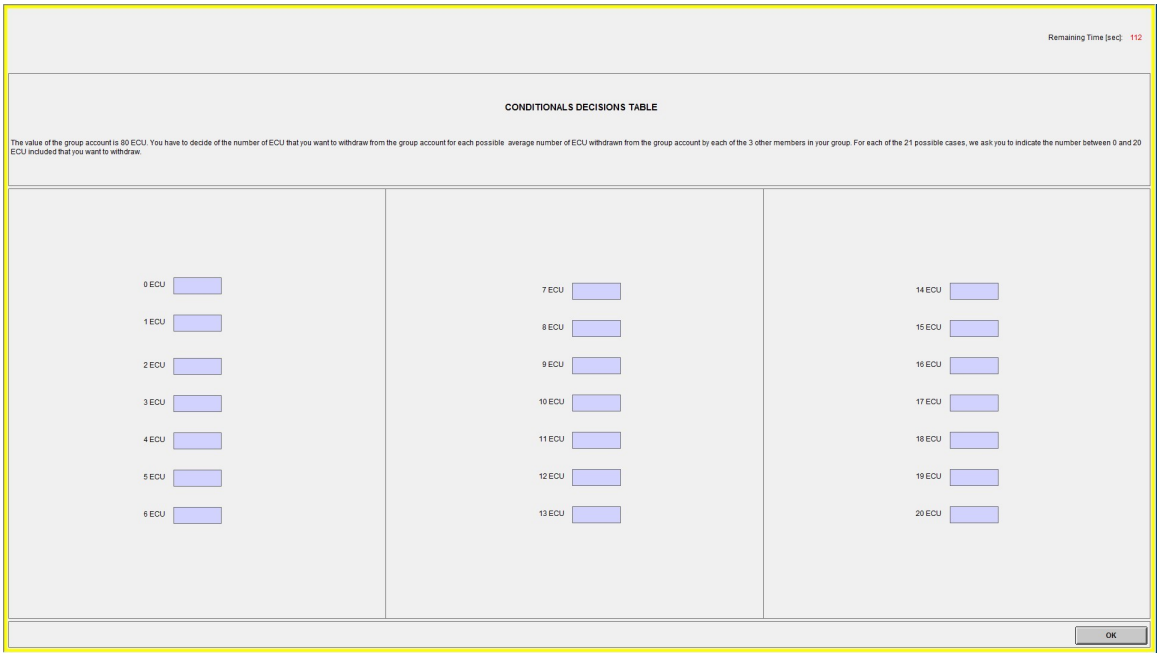
Unconditional decision: Here is an example of the screen you will get for an unconditional decision:



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:



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*(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*(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*(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*(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.

Instructions 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:

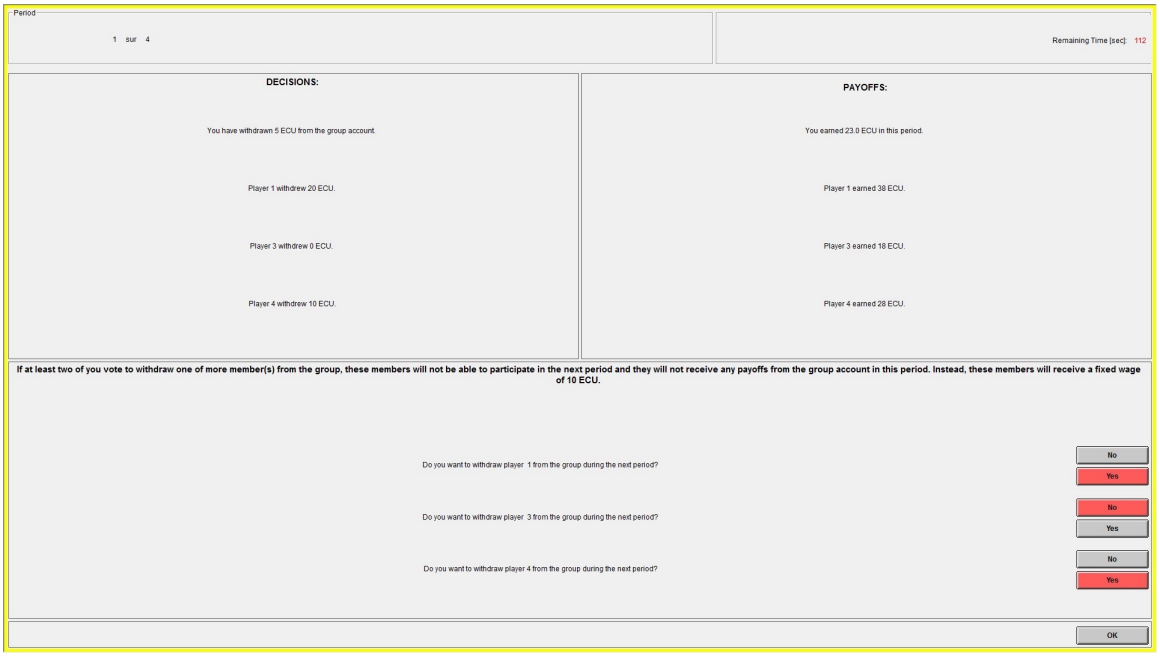
Decision rule

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:



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:

The amount you have withdrawn from the group account
+ 40% (value of the group account at the beginning of the period
− 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.

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.

Instructions 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:

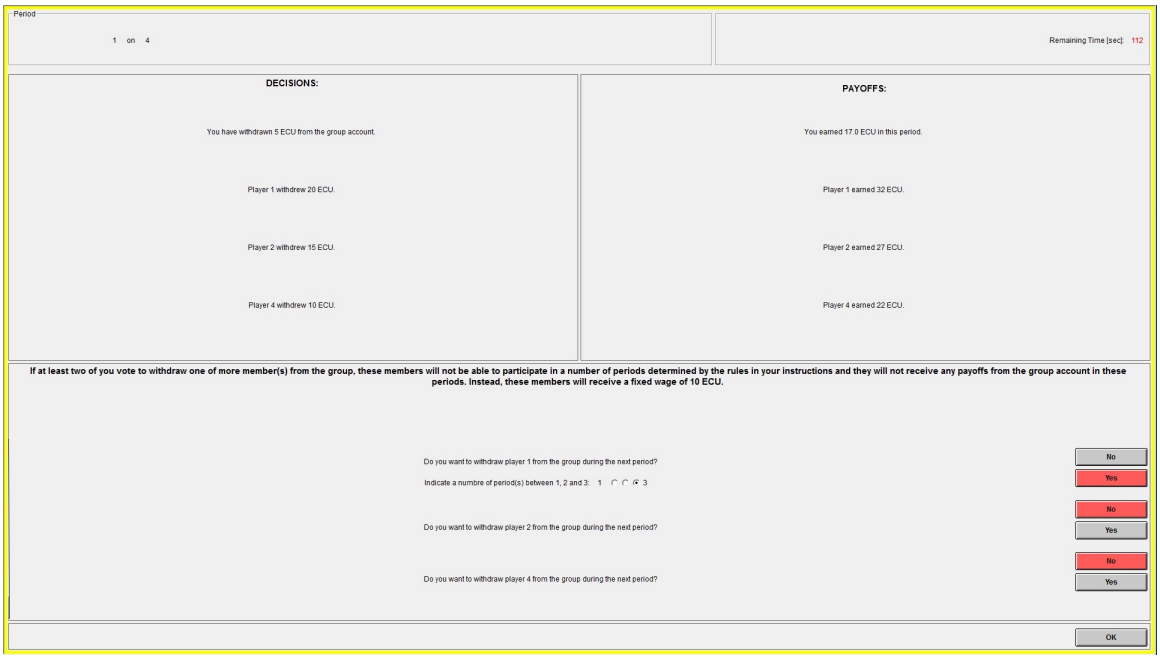
Decision rule
In this stage, the group members have the opportunity to remove one or several other group members for one, two or three following periods. Removing a member from the group means that he will not be able to participate in any of the two stages for the number of periods he has been removed: 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:



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:

The amount you have withdrawn from the group account
+ 40% (value of the group account at the beginning of the period
– 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.
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

Appendix 2.1 Subject pool

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 any 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 treatments is never significant (*vs.* Exo1P: p=1.000; *vs.* Exo3P: p=0.740).

Appendix 2.2 Determinants of exclusion

Table A.2 reports the marginal effects of the estimates of the vote for exclusion of a subject i, using random-effects logit regressions with standard 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 difference between j's withdrawal and the group average withdrawal (excluding j) in period t when j withdraws more than the group (max $(0; w_{j,t} - \sum_{-j=1,-j\neq j}^{n-1} w_{-j,t})$) and the difference between the group average withdrawal and j's withdrawal in t when j withdraws less than the group (max $(0; \sum_{-j=1,-j\neq j}^{n-1} w_{-j,t} - w_{j,t})$). The ratio between the current number of periods prior to t in which j has withdrawn more than the average group withdrawal ($w_{j,t} - \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t} > 0$) and the current number of periods prior to t in which j was active in the group (considering only the cases in which the group size was higher than two) captures the impact of chronic free-riding by j on i's willingness to vote for exclusion. The independent variables also include a control for whether j has already been excluded in the past and 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 xP exclusion (Y)" takes value 1 if i has been reintegrated in t after x period(s) of

Table A.2: Determinants of an exclusion vote

Dep. variable:	All treatments	with exclusion
Decision to exclude j in t	(1)	(2)
Exo1P	*Ref.*	*Ref.*
Exo 3P	-0.054	-0.047
	(0.059)	(0.058)
Endo	-0.062	-0.078
	(0.060)	(0.057)
Max (0; $w_{j,t} - \sum_{-j=1,j\neq-j}^{n-1} w_{-j,t}$)	0.044***	0.043***
	(0.010)	(0.010)
Max (0; $\sum_{-j=1,j\neq-j}^{n-1} w_{-j,t} - w_{j,t}$)	0.004	0.005*
	(0.003)	(0.003)
Ratio current nb of pos. dev. of j prior to t/	0.077***	0.079***
Nb of periods with j in the group prior to t	(0.022)	(0.021)
Exclusion of j prior to t	-0.027	-0.028
	(0.018)	(0.018)
i is reintegrated in t	0.114***	0.089**
	(0.041)	(0.044)
i is reintegrated in t	-	Ref.
after 1P exclusion (Exo1P)		
i is reintegrated in t	-	-0.087***
after 1P exclusion (Endo)		(0.031)
i is reintegrated in t	-	0.347
after 2P exclusion (Endo)		(0.242)
i is reintegrated in t	-	0.321
after 3P exclusion (Endo)		(0.245)
i is reintegrated in t	-	-0.037
after 3P exclusion (Exo3P)		(0.029)
Group size in t	-0.010	-0.010
	(0.018)	(0.018)
Period	0.002	0.003
	(0.003)	(0.003)
Squared period	0.000	0.000
	(0.000)	(0.000)
Individual controls	Yes	Yes
Observations	4197	4197
Clusters	32	32
Log pseudo-likelihood	-1321.18	-1300.54

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 he is excluded and i cannot vote to exclude j if j is already excluded). Pos. dev. for Positive deviation ($w_{j,t} - \sum_{-j=1,j\neq-j}^{n-1} w_{-j,t}$) >0). Individual controls include gender, age, student status, and category of social preferences. *** p <0.01, ** p <0.05, * p <0.10.

Table A.3: Determinants of an exclusion vote with controls for deviations by treatment

Dep. variable: Decision to exclude j in t	All treatments with exclusion
Exo1P	$Ref.$
Exo 3P	-0.060
	(0.064)
Endo	-0.024
	(0.078)
Max (0; $w_{j,t} - \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t}$)	0.042***
	(0.013)
Max (0; $\sum_{-j=1,j\neq -j}^{n-1} w_{-j,t} - w_{j,t}$)	0.012**
	(0.005)
Max (0; $w_{j,t} - \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t}$) * Exo1P	$Ref.$
Max (0; $w_{j,t} - \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t}$) * Exo3P	0.010
	(0.014)
Max (0; $w_{j,t} - \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t}$) * Endo	-0.004
	(0.011)
Max (0; $\sum_{-j=1,j\neq -j}^{n-1} w_{-j,t} - w_{j,t}$) * Exo1P	$Ref.$
Max (0; $\sum_{-j=1,j\neq -j}^{n-1} w_{-j,t} - w_{j,t}$) * Exo3P	-0.009
	(0.008)
Max (0; $\sum_{-j=1,j\neq -j}^{n-1} w_{-j,t} - w_{j,t}$) Endo	-0.019**
	(0.008)
Ratio current nb of pos. dev. of j prior to t/ Nb of periods with j in the group prior to t	0.070***
	(0.021)
Exclusion of j prior to t	-0.027
	(0.019)
i is reintegrated in t	0.106***
	(0.038)
Group size in t	-0.006
	(0.018)
Period	0.002
	(0.003)
Squared period	-0.000
	(0.000)
Individual controls	Yes
Observations	4197
Clusters	32
Log pseudo-likelihood	-1307.67

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). Pos. dev. for Positive deviation ($w_{j,t} - \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t}$) >0). Individual controls include gender, age, student status, and category of social preferences. *** p <0.01, ** p <0.05, * p <0.10.

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 include a time trend (period and period squared to control for a possible non linear effect of time). We also control for the group size and for individual characteristics (gender, age, student status and social preferences of i).

Table A.2 indicates that the more ECU j takes above his group average, the more likely subject i is to vote in favor of his exclusion. Chronic free-riding compared to the average withdrawal in the group increases the likelihood of i voting for j's exclusion. Controling for j's withdrawal behavior, 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 differences in withdrawal between the individual and his group average has the same impact across treatments. We relax this assumption in Table A.3, in which the independent variables include not only the treatments and the differences in withdrawal but also interaction terms between the treatments and these differences. The only significant interaction effect is the negative difference in Endo: subjects who withdraw less than the average are less likely to be excluded in Endo than in Exo1P. We find no effect of the size of the group: contrary to our expectation, subjects are not less likely to vote for excluding a group member when the size of the group is smaller.

Table A.2 also shows that i is more likely to vote to exclude another group member in t when he has been himself 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 he has been excluded for one period in Endo compared to when he has been excluded in Exo1P. Therefore, when it results from a vote, a short exclusion affects differently the willingness to exclude after reintegration. We find no effect of time.

Appendix 2.3 Reintegration

Table A.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 (considering instead the group size at the time of exclusion does not change the result). 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. Finally, we include a time trend and its squared value, and the same individual characteristics as in the previous Tables. The data only include the observations of reintegrated subjects from periods 3 to 20 in groups that contain at least three members, but we do not include the observations corresponding to subjects excluded after period 17 (the length of exclusion is censored in Endo and this could bias the comparison).

Model (1) in Table A.4 shows the difference between the effect of any length of ex-

clusion on withdrawal after reintegration compared to one-period exclusion in Exo1P. Reintegrated subjects withdraw significantly less (at the 5% level) after three periods of exclusion in Exo3P than after one in Exo1P, controlling for the average withdrawal in the group in the period preceding immediately reintegration, which supports Hypothesis 3. Choosing one-period exclusion in 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 is associated with a significantly higher 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 (chi^2 test: p=0.790). The length of exclusion does not have the same effect on withdrawal after reintegration depending on how it has been determined.

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

Dependent variable: Number of ECU withdrawn	All treatments with exclusion (1)	(2)
1P exclusion in Exo1P	*Ref.*	2.209
		(2.256)
1P exclusion in Endo	-2.209	*Ref.*
	(2.256)	
2P exclusion in Endo	0.820	3.029**
	(1.776)	(1.543)
3P exclusion in Endo	1.171	3.380**
	(1.776)	(1.543)
3P exclusion in Exo3P	-3.804**	-1.595
	(1.942)	(2.440)
Group average withdrawal	0.294***	0.294***
excluding i in t-1	(0.107)	(0.107)
Group size in t	-0.258	-0.258
	(0.827)	(0.827)
Period	-0.491	-0.491
	(0.455)	(0.455)
Squared period	0.027	0.027
	(0.019)	(0.019)
Individual controls	Yes	Yes
Constant	3.806	1.597
	(4.447)	(4.453)
Observations	237	237
Clusters	35	35
R-squared	0.32	0.31

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 when the group contains at least three members, excluding subjects who were excluded after period 17. Individual controls include gender, age, student status, and category of social preferences. *** p <0.01, ** p <0.05, * p <0.10.

Appendix 3. Figures

Appendix 3.1 Evolution of average withdrawal over time, by treatment (including all the subjects and all group sizes)

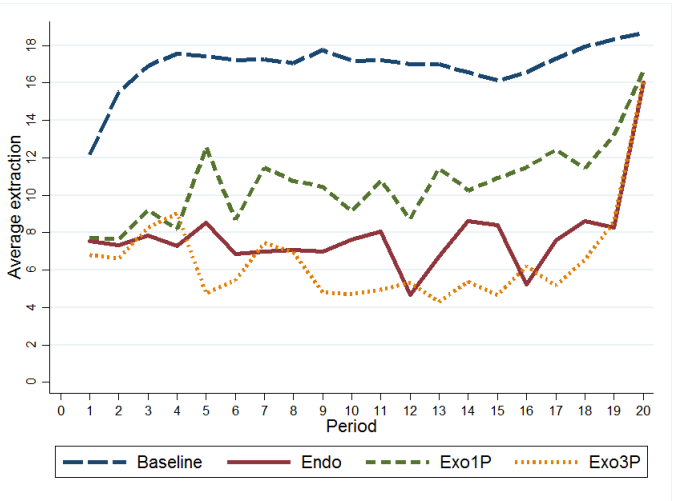


Figure 4: Evolution of average withdrawal over time, by treatment (including all the subjects and all group sizes).

Appendix 3.2 Difference between the number of ECU withdrawn by excluded subjects and by their group members in the exclusion period and in the reintegration period (including all group sizes)

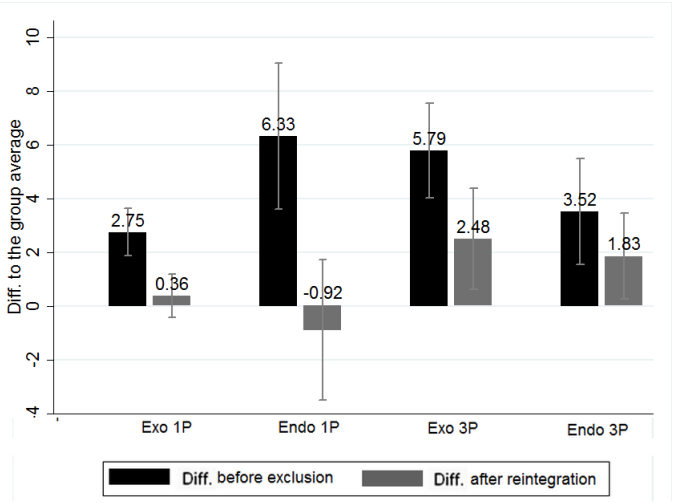


Figure 5: Difference between the number of ECU withdrawn by excluded subjects and by their group members in the exclusion period (left bars) and in the reintegration period (right bars), by treatment (including all group sizes).

Exclusion and Reintegration in a Social Dilemma

Alice Solda and Marie Claire Villeval

August 4, 2018

Abstract

Using a social dilemma game, we study the cooperative behavior of individuals who reintegrate their group after being excluded by their peers. 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 and they punish more, and more severely, chronic defections. In return, a longer exclusion has a higher disciplining effect on cooperation after reintegration, but only when the length of exclusion is not chosen by group members. Its relative disciplining effect on cooperation after reintegration is smaller when the length of exclusion results from a vote. In this environment, a quicker reintegration also limits retaliation. The difference in the impact of long *vs.* short exclusion on retaliation is larger when the length of exclusion is chosen by group members than when it is exogenous. Post-reintegration 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, 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 UMR 5824, 93 Chemin des Mouilles, F-69130, Ecully, France; IZA, Bonn, Germany. E-mail: villeval@gate.cnrs.fr

Acknowledgments: We are grateful to T. Offerman, R. Slonim, J.R. Tyran, R. Sheremata and participants at the 11th Annual Australia New Zealand Workshop on Experimental Economics, the 2017 Asia-Pacific ESA Conference, the Experimental Public Choice Workshop in Lille, and at seminars at the University of Amsterdam, Remnin University in Beijing, and Zhejiang University of Finance and Economics in Hangzhou 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, but cooperation can be sustained thanks to several mechanisms, including communication, rewards, sanctions, or taxation. In particular, the effectiveness of punishment mechanisms is now well established. 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.[1] Formal and centralized exclusion mechanisms sanction, notably by imprisonment, people who do not comply with judicial rules. In both cases, the goal is both preventive, 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, exclusion may also increase the risk of anti-social behavior and recidivism after individuals are reintegrated in their group.[2]

Exclusion is an effective mechanism to enforce social norms and limit free-riding in social dilemmas (*e.g.*, Bowles and Gintis, 2004).[3] But while the deterrence theory (Becker, 1968) predicts that increasing a penalty should discourage more misbehavior, there is little conclusive empirical evidence of the effect of the *length* of exclusion on post-release compliance and recidivism. For example, studies on the length of prison sentences find little evidence that length affects either recidivism (Martinson, 1974; Blumstein *et al.*, 1978), employment or income (Needels, 1996; Kling, 2006) after rehabilitation.[4] Moreover, we are not aware of any study measuring the impact of the length of informal exclusion on deterrence and compliance. The absence of conclusive evidence is partly due to the fact that the empirical analysis based on natural data is often plagued with endogeneity and selection issues.[5]

[1]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 social exclusion in tribes. Ethological studies have shown that scapegoating and shunning exist in social animal species. The forms of exclusion evolve over time, with cyberostracism as a modern expression of the refusal to communicate with someone (Wesselmann and Williams, 2013).

[2]This risk exists, as exclusion affects emotionally those who are excluded. Exclusion generates anger, lowers self-esteem (Zadro *et al.* 2004; van Beest and Williams, 2006), and may lead to depression if it lasts (Pressman and Cohen, 2005). In neurosciences, William (2002) found that exclusion activates 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). In contrast, forgiveness coincides with activation in regions implicated in "theory of mind" and perspective-taking (Will *et al.* 2015).

[3]Studies in psychology show that ostracized individuals conform more after re-admission in the group (*e.g.*, William *et al.*, 2000; Feinberg *et al.*, 2014), they work harder in group tasks (William and Sommer, 1997) and they become more sensitive to social cues (Gardner *et al.*, 2005).

[4]Smith *et al.* (2002) found that more punishment leads to at most a slight increase in recidivism. In contrast, Lott (1992) found a large effect of the prison sentence length and Zapryanova (2014) estimated that one month in prison results in a 1.12 percentage point decrease in recidivism probability (see also Bhuller *et al.*, 2016). By examining the effects of a reform that increased the sentence length by one month, Landerso (2015) showed that longer incarceration spells result in lower unemployment rates and higher earnings.

[5]For example, an inmate with a long prison sentence is more at risk of recidivism due to his own characteristics that led to his exclusion, to peer effects in prison and to stigma on the labor market. To

Our objective is studying how the length of exclusion affects the behavior of wrongdoers *after* they are readmitted in their group, in the context of a social dilemma game. We also study whether the length of exclusion has a different impact when voters had a choice between different lengths, compared to when it is imposed exogenously by an external authority. We consider two dimensions of behavior after reintegration: cooperation and forgiveness *vs.* retaliation. Addressing these questions is highly policy-relevant because the recommendations will differ if a longer exclusion is shown to generate higher net benefits in terms of post-reintegration compliance or if it triggers violent antisocial behavior. By keeping the environment constant (*e.g.*, the same reference group and the same level of information before, during and after exclusion), our laboratory experiment helps identify the direct effect of the length of exclusion and its mode of determination on behavior. It can also mute two channels that play a role in real settings: the variation in human capital associated with durable exclusion and peer effects that permit social learning during exclusion in the behavior that led to exclusion.

Our experiment uses a taking game that is played for 20 periods. We formed groups of four subjects who have to decide simultaneously how many units to withdraw from a preexisting public good. 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 members remain inactive, cannot communicate, receive information about the evolution of cooperation in their group, and are paid a low inactivity payoff. Since the initial value of the public good depends on the group size, excluding a member is potentially costly. In one treatment, the length of exclusion is fixed at one period; in another it is fixed at three periods, and in a third treatment, individuals vote also about the length of exclusion, up to three periods. Group mobility is not allowed: excluded members are always reintegrated in their previous group. We compare behavior in these treatments with a baseline in which exclusion is not allowed. Thus, we can investigate the effect of exclusion followed by reintegration on excluded members' behavior once they are readmitted in their original group, depending on how long they have been excluded and on the procedure of exclusion.

Our study makes three contributions to the literature. The first contribution is comparing the impact of various lengths of exclusion on the cooperative behavior of excluded members *after* they reintegrate their group. Earlier studies on exclusion in social dilemmas have focused more on the threat of exclusion than on reintegration and most did not manipulate the length of exclusion (*e.g.*, Masclet, 2003; Cinyabuguma *et al.*, 2005; Kerr *et al.*, 2009; Sheremeta *et al.* 2009; Maier-Rigaud *et al.*, 2010). For an exception, see Bal-

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, giving a random manipulation of the length of the remaining sentence at the date of the pardon. They found that longer sentences had ex ante a higher deterrent effect on recidivism, but the inmates who stayed longer 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 sentences decrease recidivism. Using instead random judge assignments, Green and Winik (2010) found that randomly assigned variations in prison and probation times have no effect on recidivism, echoing the results of Berube and Green (2007). In contrast, Roach and Schanzenbach (2016) estimated that one-month extra prison sentence decreases recidivism by one percentage point.

afoutas *et al.* (2017) who conducted an experiment with inmates and found that a longer incarceration is associated with a lower level of pro-sociality. Our second contribution is studying whether endogeneizing the length of exclusion through a voting procedure modulates the impact of exclusion on future cooperation after reintegration in the same group. In the previous literature, the length of exclusion was fixed exogenously. In contrast, we examine whether choosing a less severe sanction when a more severe one was available is perceived as a warning, signaling a willingness to forgive after reintegration, compared to when the length is fixed exogenously. Our third contribution is studying the impact of the length of exclusion and of its mode of determination on retaliation. Previous studies have highlighted the risk of revenge (*e.g.*, Nikiforakis, 2008) and feuds (Nikiforakis and Engelmann, 2011) induced by punishment in social dilemmas. In our study punishment results from a group decision, which may signal more strongly what the norm of the group is, especially when subjects choose also the length of exclusion. Ambrus and Greiner (2015) have shown that democratic punishment decreases anti-social behavior compared to individual punishment. In our case, we can compare retaliation behavior for a given length of exclusion according to the mode of choice of this length.

Our results show that exclusion and the choice of its duration are driven by the willingness to punish larger deviations from the withdrawal norm in the group and chronic defections. A longer exclusion has a larger effect on norm compliance after reintegration compared to a shorter exclusion, but only when the length of exclusion is exogenous. When the length of exclusion results from a vote, the effect of a longer exclusion on the withdrawal level after reintegration is smaller. Choosing a shorter exclusion when a more severe sanction is available leads to more forgiveness by the reintegrated members and limits retaliation compared to when people are excluded for three periods. When the length of exclusion is chosen by group members, the difference in the impact on retaliation of a long sentence compared to a short sentence is much larger than when the length of exclusion is fixed exogenously. Indeed, behavior after reintegration does not only depend on the severity of the sanction but also on the intentions that can be inferred from it.

The remaining sections are organized as follows. Section 2 reviews the literature briefly. Section 3 details the experimental design and the procedures. In Section 4, we present behavioral conjectures. Section 5 displays our main results. Section 6 concludes.

2 Related Literature

The previous literature has considered various exclusion mechanisms in different environments, in particular in principal-agent games when people can be fired or in gift-exchange games (*e.g.*, Brown *et al.*, 2004; Charness *et al.*, 2011; Berninghaus *et al.*, 2013; Falk *et al.*, 2015; Bernard *et al.*, 2017), in coordination games and bargaining games (Fischer and Güth, 2012). For example, Kopanyi-Peuker *et al.* (2017) study a weak-link game with noisy information in which a principal can replace some members of a team. The fear of being fired increases cooperation even when workers are imperfectly monitored; but to be effective, the threat has to be maintained permanently. Croson *et al.* (2015) compare the impact of excludability in three team production settings (weak-link, voluntary contribution, and a best-shot games) when the lowest contributor in the team is automatically excluded. They find that excludability increases efficiency dramatically,

but less so in the weak-link game. In contrast, we study exclusion in a social dilemma setting without competition (between group members or between unemployed players); the decision to exclude someone is made by the group members themselves,[6] without any formal rule defining who should be excluded (in contrast with Swope, 2002). We also differ in that our design allows us to study whether the reintegrated players retaliate by trying to exclude their punishers.

We also differ from the previous literature in which group members can exclude each others, as we manipulate the length of exclusion. We observe three types of length of reversible exclusion. In some studies, exclusion is irreversible (Cinyabuguma *et al.*, 2005; Maier-Rigaud *et al.*, 2010; Akpalu and Martinsson, 2011). An excluded subject remains ostracized until the end of the session. In contrast, in Charness and Yang (2014) exclusion from the group is irreversible but subjects can move immediately to another group. Finally, in other studies exclusion is only for one period (Masclet, 2003; Kerr *et al.*, 2009; Sheremeta *et al.*, 2009; Koike *et al.* 2015). None of these studies compares different lengths of exclusion. Neuhofer and Kittel (2015) show that irreversible exclusion has a stronger impact than a one-period exclusion but this comparison is based on a between-subject design. The lack of within-subject variation of the length of exclusion in the previous literature is somewhat 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 usually depends on the harm caused.[7]

Our study is also related to the literature on endogenous group formation, as excluding group members enables players to choose their partners, to a certain extent. For example, Ahn *et al.* (2008) showed that restricted entry in an endogenous group formation mechanism in a public goods game leads to the formation of mid-sized highly cooperative groups. Restricted entry also increases average contributions and earnings in a congestible public good game (Ahn *et al.*, 2009). Riedl *et al.* (2016) found that letting players form networks in a weak-link game leads to fully efficient coordination, demonstrating that the possibility of choosing interaction partners is fundamental to achieve efficiency (see also Riedl and Aljaz (2014), on exclusion in prisoners dilemma games). Sääksvuori (2014) shows that competitive pressure between groups decreases the likelihood of social exclusion. In our game, players can choose their partners if they are able to coordinate to exclude some of them. Once individuals have chosen the length of exclusion, they do not have to vote about on reintegration. Moreover, players reintegrate the same group after being excluded. This allows us to study to which extent reintegrated individuals are willing to cooperate with those who have previously excluded them.

[6] In our setting all group members have the same status; thus, we differ from Güth *et al.* (2007) in which a leader decides whom to exclude in the group and from Fatas *et al.* (2010) where the subject excluded is selected randomly.

[7] Davis and Johnson (2015) manipulate the length of exclusion in a game in which participants vote to exclude group members and to reintegrate them. However, excluded members are only excluded from a chat and they can still contribute. Exclusion has no monetary consequences.

3 Experimental Design and Procedures

We first present our design before detailing the procedures.

3.1 Design

The experiment consists of four treatments that are played between subjects. In each treatment, a taking game is played for 20 periods with a fixed matching protocol. 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.[8] 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:

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

with $\alpha = 0.4$ the marginal per capita return of the group account and n 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:

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

The following treatments introduce voting and exclusion.[9]

Exo1P Treatment

[8] The exchange rate is 35 ECU = 1 Euro.

[9] Note that a public good like national defense or street lighting is non-excludable, by nature. But as noted by Cinyabuguma *et al.* (2005), there are goods that are public to the members of a group, but not to those who do not belong to the group. Hirshleifer and Rasmusen (1989) also evoke the excludability of resources from non-members. Teams, associations, partnerships, clubs develop collective action with a risk of free-riding, and they can usually exclude members who do not comply with the rules of the group. Our setting is closer to club goods (*e.g.*, private parks, collective gardens owned by an association, ponds maintained by fishing associations) (for studies of excludability in club goods, see Swope, 2002, and Chakravarty and Fonseca, 2017).

Each period of the Exo1P treatment (for "Exogenous one Period")[10] 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.

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 other group member. All group members make their decisions simultaneously. If someone is excluded in period t, he will remain inactive in period t+1. This means that in t+1 he will not be able to participate in the taking stage, nor to vote in the punishment stage, and he 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.[11] During the exclusion period, an excluded member receives feedback on the withdrawal decisions and earnings of each group member and on the exclusion of other members in t+1.

There is no direct cost for excluding a group member (like in *e.g.*, Hirshleifer and Rasmusen, 1989; Masclet, 2003; Maier-Rigaud *et al.*, 2010). Exclusion is indirectly costly for two reasons. First, 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 since the good is non-rival. Indirect costs are null only in the case where the excluded subject would have withdrawn 20 ECU in the next period if not excluded. Second, excluding more than one group member generates an additional indirect cost because it destroys the incentive to cooperate at the group level. Indeed, when there are more than two members in the group, the game involves a social dilemma: the Marginal Per Capita Return (MPCR, hereafter) is 0.4, which creates an incentive to free-ride (MPCR<1), but the social optimum corresponds to full cooperation (no withdrawal) since N*MPCR>1. However, when N<3, the structure of incentives changes: it is still an equilibrium to withdraw to the maximum extent but it is no longer socially inefficient to withdraw from the group account (N*MPCR<1), and the social dilemma no longer exists. This means that if more than one person is excluded, cooperation can no longer maximize the social surplus. When N<3, even cooperators have no more reason to limit withdrawal since there is no social benefit from cooperation at the group level. This change can be seen as an indirect cost of exclusion if group members exclude more than one of them. We acknowledge that by using an exclusion mechanism that is not directly costly to the punishers, we probably observe an upper bound of exclusion decisions.

The payoff function of player i for a period t becomes:

[10]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.

[11]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.

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

with n the number of active members in the group.

Excluding a member for one period requires at least two votes. 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, it is not possible to exclude anyone.[12] 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 are informed on the ID of each excluded member, if any. They are never informed on the number of votes cast to exclude someone. A member who is not excluded never knows whether another member voted for his exclusion.

Exo3P Treatment

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, he will remain inactive in periods $t+1$, $t+2$ and $t+3$. During each of these three periods, he receives the 10 ECU exclusion payoff and he is informed on the withdrawal decisions and earnings of the active members and on the exclusion of other members. In period 18 (19, respectively), since there are only two (one, resp.) periods left before the end of the game, if a group member is excluded, he will be excluded for two (one, resp.) periods only. In the last period it is not possible to exclude anyone. This change in procedure for the last three periods is made common information.[13] At the end of the second stage, like in the Exo1P treatment, subjects receive a feedback on the ID of each player that has been excluded, if any. They are also given the number of remaining periods of exclusion of the group members that have been excluded earlier.

Endo Treatment

In this treatment, the length of exclusion is endogenous. This has two implications. First, the number of periods of exclusion is no longer unique; 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, he also has to decide for how many periods to exclude this member.[14]

[12]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.

[13]We chose this procedure to avoid having no risk of sanction in three periods whereas severe sanctions were available in the previous periods. An alternative procedure would have been to introduce random ending instead of deterministic ending in the game.

[14]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. 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

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 member, this group member will not be excluded, therefore 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. At the end of the second stage, group members are informed about the members who have been excluded and, if any, about the number of periods of exclusion. They are 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.[15] 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).[16]

In this preliminary part, we formed groups of four subjects randomly. Subjects had to make first an unconditional decision and then, conditional decisions. 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 withdrawn by three other group members. Therefore, each subject had 21 conditional decisions to make. For payment, one member among the four was randomly selected by the program

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.

[15] 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 that the session consisted of two parts but they received the instructions for part 2 only after completing part 1.

[16] A subject is classified as an unconditional free-rider if he withdraws the maximum possible amount of ECU regardless of his group members' average withdrawal. A subject is classified as an unconditional cooperator if he withdraws 0 ECU regardless of the group average withdrawal. A subject is classified as a conditional cooperator if his withdrawal increases when the group average withdrawal 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.

and paid according to his conditional decision while the three other members were 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 withdrawn from the group account by the three other members. 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 were reassigned randomly. At the end of the session, we also collected standard information on the subjects' gender, age, student status and discipline.

3.2 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 with an average of 20 subjects per session. We collected observations for 12 groups in both the Endo and Exo1P treatments, 13 groups in the Exo3P treatment and 15 groups in the Baseline (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 taking game were distributed after completion of the preliminary part (see Appendix 1). We avoided the terms "exclusion" and "excluded members" in favor of more neutral terms, as "remove" and "players who do not participate in the group". After the instructions were read aloud, subjects had to complete a quiz to check their understanding. 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.

4 Behavioral Hypotheses

Selfish payoff-maximizing players should neither cooperate, nor vote for exclusion. In the Baseline the dominant strategy for a player is to withdraw the maximum from the group account in each period. On the opposite, the social optimum requires that players do not withdraw anything. In the treatments with possible exclusion, exclusion is indirectly costly if excluded members do not free ride in full or if the outcome of exclusion changes the structure of incentives.[17] Indeed, each remaining member will lose the marginal return of the amount left by excluded members in the group account if not excluded.[18]

[17]The 'trembling hand' refinement of the Nash equilibrium (Selten, 1975) could explain such mistakes in the withdrawal decisions (see also Maier-Rigaud *et al.*, 2010).

[18]Note that Hirshleifer and Rasmusen (1989) have shown that costless ostracism may lead to a cooperative equilibrium. Considering a n-player prisoner's dilemma game, ostracizing defectors is a sub-

The previous literature has shown that these predictions do not hold when players have social preferences. Such players may not be willing to withdraw all their units from the group acccount when N>2. When N<3, the social dilemma disappears (since N*MPCR<1) and it is no longer socially inefficient to withdraw units from the group account; thus, only altruistic subjects should refrain from withdrawing, increasing the payoff of others at a personal cost. Subjects may be willing to exclude others if it is economically advantageous to exclude. This is the case if they believe that cooperation will improve in future periods, both because cooperators should no longer fear of being exploited by free-riders if the latter are excluded, and because reintegrated subjects should be more willing to comply with the norm of the group to prevent a new exclusion. Moreover, previous experiments on ostracism have shown that participants are willing to exclude even when the decision is costly (Cinyabuguma *et al.*, 2005; Maier-Rigaud *et al.*, 2010; Davis and Johnson, 2015; Neuhofer and Kittel, 2015). In our setting, cooperators who care about the maximization of social surplus should not be willing to exclude more than one group member when N=4 and no one when N=3 to avoid forming groups with N<3 (by design exclusion is impossible when N=2). If some players are willing to exclude, we also anticipate that individuals will reduce their level of withdrawal, especially when the severity of sanction can be increased by the choice of a longer length of exclusion. We thus retain the following hypothesis.

Hypothesis 1: *The level of cooperation increases in the length of potential exclusion when N>2. Some subjects are willing to exclude others.*

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, and second, it may influence the willingness to seek revenge for past exclusion. These effects may depend on how the length of exclusion is determined. Subjects who have been excluded should be more willing to conform to the withdrawal norm of the group after reintegration than before being excluded. A more severe punishment should lead to more compliance after reintegration than a milder one. On the other hand, the 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 2: *A longer exclusion decreases more withdrawals after reintegration than a shorter exclusion.*

Hypothesis 3: *A longer exclusion increases more the risk of retaliation than a shorter exclusion.*

From a standard economic perspective, a similar severity of sanction should have the same effect on behavior after reintegration in both Exo and Endo treatments. However, from a behavioral point of view, a better ability to perceive the punishers' intentions through the choice between several lengths of exclusion may influence the reaction of those who are punished. According to Braithwaite (1989), reintegrative shaming (compared to disintegrative shaming) takes place when disapproval is partnered with forgiveness and

game-perfect equilibrium when players follow the 'banishment strategy' according to which everyone cooperates and excludes defectors until the last round where everyone defects.

acceptance back into the group. The choice of the length of exclusion conveys a richer signal on the group's level of disapproval, forgiveness, and willingness to reintegrate. In contrast, it is more difficult to infer the punishers' forgiveness when a single level of severity is available. Being excluded for three periods when a shorter exclusion was possible may be interpreted as a harsher decision than being excluded for three periods when severity could not be modulated. As a result, the excluded members in this context may express their anger after being reintegrated by complying less and by casting more votes for excluding the punishers than when the latter did not choose the length of exclusion. Similarly, being excluded for one period in this context may be interpreted as a more forgiving punishment, which may be rewarded by a higher willingness to comply to the group norm and by less retaliation after reintegration than when punishers do not choose the severity of punishment. This leads to our last hypotheses.

Hypothesis 4: *The positive effect of long (short, respectively) exclusion on cooperation after reintegration is smaller (larger, resp.) when the length of exclusion is chosen by the punishers than when it is exogenous.*

Hypothesis 5: *The difference in the frequency of retaliation between those excluded for a longer duration and those excluded for a short duration is larger when the length of exclusion is chosen by the punishers than when it is fixed exogenously.*

5 Results

We start by presenting general findings about treatment effects on cooperation, efficiency and exclusion, as although not central to our research question, these findings serve as a background for our main analyses. We then focus on the effect of exclusion followed by reintegration on withdrawal behavior. 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 main variables of interest (withdrawals, length of exclusion, and payoff), by treatment.

5.1 General Findings

We introduce our first result:

Result 1: *The existence of an exclusion mechanism helps sustain cooperation and does not harm social welfare. Individuals are willing to punish and generally they exclude the least cooperators.* This supports Hypothesis 1.

Support for Result 1. Figure 1 displays the evolution over time of the average withdrawal from the group account by each active group member, by treatment, when N>2 (*i.e.*, when there is a social dilemma; see the same figure in Appendix 3.1 for all groups regardless of their size). As also visible in Table 1, 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.[19],[20] Considering period 1 in isolation, the mean withdrawal is also lower in the treatments with possible exclusion.[21] This confirms the previous literature showing that the threat of exclusion itself enhances cooperation even before getting any information about others' behavior.

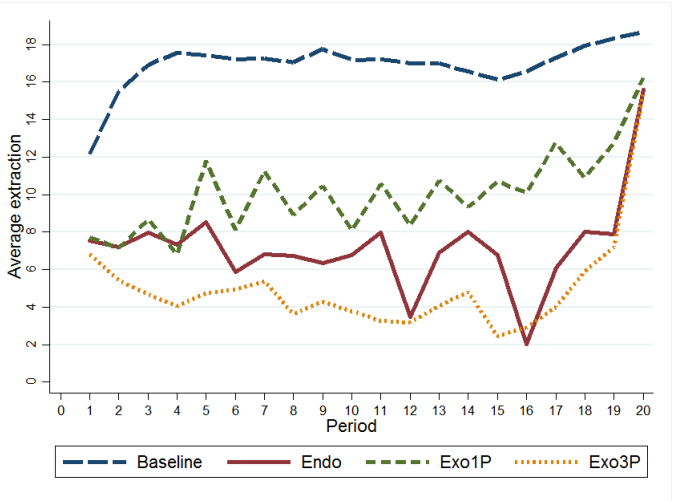


Figure 1: Evolution of average withdrawal over time, by treatment (includ ing all the active subjects when N>2).

We find significant differences in the level of withdrawal depending on the length of possible exclusion when comparing Exo1P and Exo3P (MW test: p=0.014). This supports our hypothesis that cooperativeness is higher in an environment with a more severe potential punishment (consistent with *e.g.*, Nikiforakis and Normann, 2008). We find no significant difference between Endo and the other treatments with exclusion (MW tests: Exo1P: p=0.488; Exo3P: p=0.301). Interestingly, there is no significant difference either in the first period between the treatments with exclusion (MW tests: Exo1P *vs*. Endo: p=0.931; Exo1P *vs*. Exo3P: p=0.643; Endo *vs*. Exo 3P: p=0.703). This indicates that the effect of possible exclusion in terms of cooperation at the beginning of the game depends on whether a punishment mechanism is available or not, and not on the severity of this punishment (which differs from Neuhofer and Kittel (2015)).[22]

[19]Mann-Whitney tests - MW, hereafter: Exo1P: p=0.004; Endo: p=0.007; Exo3P: p <0.001)

[20]The values indicated in Table 1 are means of individual decisions but for the non-parametric tests we consider only one observation per group. For Mann-Whitney tests, the mean withdrawal over all periods for each group gives one independent observation, except when we consider only the first period, in which case the mean withdrawal of each group in this period gives one independent observation. All tests are two-tailed.

[21]In period 1 the mean withdrawal is 12.18 ECU in the Baseline, 7.73 ECU in Exo1P, 6.81 ECU in Exo3P and 7.56 ECU in Endo. MW tests: Exo1P: p=0.002; Endo: p=0.002; Exo3P: p <0.001.

[22]Individuals can probably better adjust their behavior to the severity of sanctions in the first period of a game when they are informed *ex ante* of the deterrence regime applied to a precise behavior (*e.g.*, De Angelo and Charness, 2012; Schildberg-Horisch and Strassmair, 2012). Note that Nikiforakis and Normann (2008) also observed no difference in the first period between treatments implementing different punishment ratios in a public goods game.

Table 1: Summary statistics

	Baseline	Exo1P	Exo3P	Endo
Average withdrawal (Groups with N>2)	16.94 (5.40)	10.13*** (8.13)	4.97 *** (6.44)	7.28*** (8.38)
Average withdrawal of non-excluded members (N>2)	16.94 (5.40)	7.38*** (7.58)	3.93*** (5.61)	5.76*** (7.85)
Average withdrawal of excluded members (N>2)	-	16.43 (5.39)	12.31*** (7.24)	15.86 (5.65)
Average number of votes in favor of exclusion (N>2) (with a max. of 12)	-	4.05 (3.19)	2.04* (2.42)	2.09* (2.52)
Average number of excluded subjects (with a max. of 4)	-	0.88 (1.30)	0.33*** (0.73)	0.43** (0.95)
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 (N>2)	-	1.64 (1.30)	1.29 (1.19)	1.35 (1.18)
Average payoff (in ECU, all players)	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: The Table reports mean values, with standard deviations in parentheses. It 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. The withdrawals per period are measured by the number of ECU taken from the group account. The average withdrawal of excluded members is at the period of exclusion. The average number of active members in favor of exclusion is the mean number of subjects who cast at least one vote to exclude another member when N>2. The average number of excluded members considers only the cases where a new subject has been excluded. The average length of exclusion is conditional on being excluded.

Regarding efficiency, as indicated in Table 1, including all periods and groups, an environment with possible exclusion does not change significantly the average payoffs compared to the Baseline,[23] but it significantly increases the payoffs of the non-excluded members.[24] Hence, exclusion does not systematically harm efficiency when it is followed by reintegration. This is consistent with Cinyabuguma *et al.* (2005), Sheremeta *et al.* (2011), and 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 section, cooperative players are less likely to be excluded.

Consider now exclusion. As visible in Table 1, individuals are willing to vote in favor of the exclusion of another member and they do so marginally more in the Exo1P treatment than in Exo3P (MW test: p=0.073) and in Endo (p=0.069). A regression analysis estimating random-effects logit models in which the dependent variable is the voting decision of subject i to exclude subject j in period t is reported in Table A.2 in Appendix 2. It shows that the willingness to exclude increases significantly (at the 1% level) when j withdraws more than the group average (taken as the group norm) with no difference across treatments, regardless of the severity of punishment (this is also visible in the Figure A.2 reported in Appendix 3). This table also shows that chronic free-riding increases significantly (at the 1% level) the likelihood of a vote for exclusion. Chronic free-riding is captured by the ratio between the current number of periods prior to t in which j has withdrawn more than the average of his group and the current number of periods prior to t in which j was active in the group (considering only the cases where N>2).[25]

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). To investigate the determinants of the choice of a given length of exclusion we estimated random-effects ordered logit regressions. The results are displayed in Table A.3 in Appendix 2. This Table shows that the vote for a specific length of exclusion is affected by the number of ECU taken by j relative to the average withdrawal in the group. There is also a significant effect of the reputation of the individual who is punished: a chronic positive deviation from the group average withdrawal leads to votes for a longer exclusion. When including this variable, the fact that j has already been excluded in the past does not affect the choice of a specific length of exclusion. 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. We analyze this behavior in sub-section 5.3. The group size has no significant effect.

To summarize, cooperation is higher in the presence of a risk of exclusion and increases

[23]MW tests: Exo1P *vs.* Baseline: p=0.329; Exo3P *vs.* Baseline: p=0.800; Endo *vs.* Baseline: p=0.591.

[24]MW tests: Exo1P *vs.* Baseline: p=0.028; Exo3P *vs.* Baseline: p <0.001; Endo *vs.* Baseline: p=0.088.

[25]Note also that time has no significant effect. This is in contrast with standard experiments on monetary punishment in public goods game in which sanctions occur from the very beginning the game to impose a norm in the group. Maier-Rigaud *et al.* (2012) also observe that ostracism occurs mainly at the beginning of the game. Moreover, the category of social preferences of the subject is never significant in any regressions. Most of the time, the other individuals characteristics are not significant either.

in the severity of the institution available. Both the probability to exclude and the severity of punishment increase not only in reaction to the importance of the deviation from the withdrawal norm of the group in the current period, but also in reaction to chronic free-riding over time. The choice of a longer exclusion depends also on the desire to take revenge after having suffered oneself from a more severe punishment.

5.2 Reintegration

The analysis of reintegration concerns mainly cases in which the group size is higher than 2. At the time of exclusion, the size of the group is, by design, higher than 2, and at the time of reintegration it is higher than 2 in 237 out of 244 observations. This means that in the large majority of cases, the structure of incentives is the same at the time of exclusion and at the time of reintegration and subjects face a social dilemma when taking their decisions. The analysis of reintegration patterns delivers two main findings.

Result 2: *Among punished individuals, a long exclusion has a higher disciplining effect on withdrawal levels in the reintegration period compared to a one-period exclusion, but this is observed only in the Exo treatments.* Hypothesis 2 is not supported in the Endo treatment.

Result 3: *Among punished individuals, the disciplining effect of a three-period exclusion (one-period exclusion, respectively) on the withdrawal level after reintegration is smaller (is similar) when the length of exclusion results from a vote compared to when the length of exclusion cannot be modulated and is fixed exogenously.* Hypothesis 4 is not supported when exclusion is short.

Support for Results 2 and 3. Figure 2 displays the difference between the number of ECU withdrawn 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 when N>2 (*i.e.*, when the game involves a social dilemma; see Figure A.3 in Appendix 3 for all groups regardless of their size).[26] We can see that in every treatment, the average difference with the group average is lower during the reintegration period than during the exclusion period.[27] The fact that subjects who have been excluded for a longer time in Endo withdraw more after reintegration relative to subjects who were excluded for a shorter duration may result from three motives. First, there is a selection effect: these subjects are more likely to be strong free-riders. Second, they may be willing to signal their anger to their group members after such a harsh punishment. Third, subjects excluded for a longer duration may attempt at making up after reintegration for lost earnings during exclusion.

Using regressions allows us to compare the effect of exclusion on the behavior of the excluded subjects after reintegration, controlling for the deviation of the subject's withdrawal from his group average at the time of exclusion, for the average withdrawal in the group in the period prior to reintegration and for the time of the decision. Table 2

[26] Although not illustrated in Fig. 2 to keep the comparison symmetric, subjects excluded for two periods in Endo take on average 0.33 ECU more than their group members in the reintegration period.

[27] 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.

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 include a variable that captures the deviation between the excluded subject's level of withdrawal and the average group withdrawal at the time of exclusion. We control for the average withdrawal of the other group members in t-1. 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. Finally, we include a dummy variable for the group size in t (equal to 1 when N=4 and 0 when N=3), a time trend and its squared value, and the same individual characteristics as in the previous tables. The data only include the observations of reintegrated subjects from periods 3 to 20 in groups that contain at least three members, but we do not include the observations corresponding to subjects excluded after period 17 (the length of exclusion is censored in Endo and this could bias the comparison).

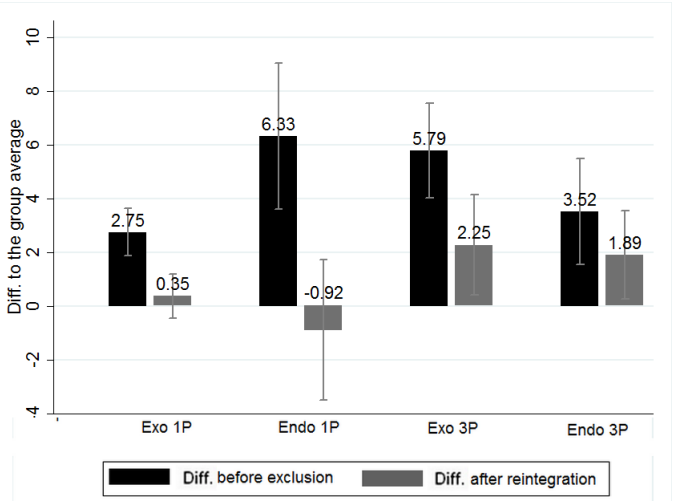


Figure 2: Difference between the number of ECU withdrawn by excluded subjects and by their group members in the exclusion period (left bars) and in the reintegration period (right bars), by treatment when N>2.

Model (1) in Table 2 shows the difference between the effect of any length of exclusion on withdrawal after reintegration compared to one-period exclusion in Exo1P. This analysis reveals a significant difference, at the 5% level, between Exo3P and Exo1P: reintegrated subjects withdraw less after three periods of exclusion than after one, controlling for their behavior at the time of exclusion and for group average in the period preceding reintegration. In contrast, choosing one-period exclusion in 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 is associated with a significantly (at the 5% level) higher level of withdrawal during the reintegration period. This supports Result 2. There is no significant difference in the effect of being excluded for two or for three periods on withdrawal after reintegration (chi^2 test: p=0.789). Model (1) also shows that the disciplining effect

of a one-period exclusion is the same in the Exo and the Endo treatments, supporting Result 3 against Hypothesis 4. Note that all the results hold if we include all group sizes in the regressions (with 244 observations instead of 237) (see Table A.4 in Appendix 2).

Table 2: Effect of the length of exclusion on the withdrawal (in ECU) of excluded subjects in the reintegration period

Dependent variable: Number of ECU withdrawn	All treatments with exclusion (1)	(2)
1P exclusion in Exo1P	*Ref.*	2.270
		(2.365)
1P exclusion in Endo	-2.270	*Ref.*
	(2.365)	
2P exclusion in Endo	0.720	2.990**
	(1.696)	(1.441)
3P exclusion in Endo	1.082	3.352**
	(1.865)	(1.601)
3P exclusion in Exo3P	-3.980**	-1.710
	(2.016)	(2.507)
$\sum_{-i=1,i\neq -i}^{n-1} w_{-i,t} - w_{i,t}$	-0.073	-0.073
at the time of exclusion	(0.119)	(0.119)
Group average withdrawal	0.304***	0.304***
excluding i in t-1	(0.111)	(0.111)
Group size in t	-0.081	-0.081
	(0.851)	(0.851)
Period	-0.399	-0.399
	(0.485)	(0.485)
Squared period	0.024	0.024
	(0.020)	(0.020)
Individual controls	Yes	Yes
Constant	2.649	0.379
	(4.527)	(4.720)
Observations	237	237
Clusters	35	35
R-squared	0.32	0.32

Notes: The model estimates random-effect GLS models with robust standard errors clustered at the group level in parentheses. The data include the observations from subjects reintegrated in periods 3 to 20 when the group contains at least three members, excluding subjects who were excluded after period 17. Individual controls include gender, age, student status, and category of social preferences. *** p <0.01, ** p <0.05, * p <0.10.

To control for the subjects' withdrawal behavior before exclusion and to compare them with non-excluded members, Table 3 reports random-effects GLS regressions with robust standard errors clustered at the group level in which 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. In model (1), we only consider the group members who are excluded in t-2 and reintegrated in t and the group

members who are active in both t-2 and t. Similarly in model (2), we only consider the group members who are excluded in t-4 and reintegrated in t and the group members who were active in both t-4 and t. These regressions exclude subjects that cannot be observed in both t and t-2 or t-4. If a subject has been excluded for one period (for three periods, respectively) and reintegrated in t, this indicates the difference in withdrawal levels between t, the period of reintegration, and t-2 (t-4, respectively), the period of exclusion.

Table 3: Determinants of the evolution of the level of withdrawal over time

Dependent variable:	$(wi_t - wi_{t-2})$	$(wi_t - wi_{t-4})$
Evolution of number	All 1P	All 3P
of ECU withdrawn	(1)	(2)
Reintegrated after exclusion	-5.165***	-5.966***
	(1.408)	(0.973)
Reintegrated after exclusion	0.962	2.175*
*Endo	(1.498)	(1.195)
Group average in t-1	0.045	-0.051
	(0.035)	(0.072)
$GA_{t-1} - GA_{t-2}$	0.102	-
	(0.078)	
$GA_{t-1} - GA_{t-4}$	-	0.322**
		(0.133)
Group size in t	-0.753	-0.412
	(0.0.731)	(0.710)
Period	-0.937***	-1.561***
	(0.175)	(0.311)
Squared period	0.047***	0.076***
	(0.008)	(0.013)
Individual controls	Yes	Yes
Constant	5.942*	7.304*
	(3.139)	(4.006)
Observations	981	934
Clusters	24	25
R-squared	0.14	0.24

Notes: This Table 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 in groups with more than two members, excluding subjects who were excluded after period 17. Individual controls include gender, age, student status, and category of social preferences. *** p <0.01, ** p <0.05, * p <0.10.

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 the subject was 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 could be modulated by the group members. 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. We add a dummy variable for the group size in t (equal to 1 when N=4 and to 0 when N=3), a time trend and its squared value, and controls for individual characteristics. The data include the data from subjects reintegrated in periods 3 to 20 in groups with N>2, excluding subjects excluded after period 17.

The two models show that controlling for the level and the evolution of the group average withdrawals, being reintegrated after being excluded for one or for three periods decreases the extraction level between the period of exclusion and the period of reintegration, compared to individuals who have not been excluded. In model (2) there is a marginally significant and positive difference between the Endo and the Exo treatments. There is a non-linear effect of time, indicating an improvement of cooperation over time but at a declining rate.

To summarize, when the length of exclusion is exogenous, a longer exclusion has a stronger disciplining effect on the withdrawal level of free-riders but the disciplining effect of a long exclusion is smaller when the severity of the actual sanction results from the vote of group members.

5.3 Retaliation

In this section we focus on retaliation behavior by examining the voting behavior of the excluded members after their reintegration, conditional on how long they have been excluded and by which procedure. Our last results are the following.

Result 4: *Subjects who have been excluded for a shorter duration are less likely to retaliate when they reintegrate their group than subjects excluded for a longer duration. This is found only when the length of exclusion has been chosen by group members.* This supports Hypothesis 3 for the Endo treatment but not for the Exo treatments.

Result 5: *The difference in the impact of long vs. short exclusion on the willingness to retaliate is larger when the length of exclusion is chosen by group members than when it is exogenous.* This supports Hypothesis 5.

Support for Results 4 and 5. An extreme form of retaliation consists for a previously excluded subject of voting to exclude all the other active group members after reintegration ("Total revenge"). In Exo1P 43.39% of the reintegrated subjects vote for excluding all the other group members in the reintegration period. This percentage is only 8.33% for reintegrated subjects excluded for one period in Endo. It is 26.56% in Exo 3P but it rises to 50% for those excluded for two periods in Endo and 45.71% for those excluded for three periods in Endo. The percentage in Exo1P is significantly higher than in Exo3P. The difference between Endo and the other treatments with exclusion is not significant

(35.80%, regardless of the length of exclusion).[28] One-period exclusion limits counter-punishment in Endo, but not when the length of exclusion is not chosen. Does the size of the group influence the willingness to vote for total revenge? When N=3 and the group size remains constant after reintegration, a subject who has been excluded knows for sure that the two other group members voted for his exclusion and he can target them directly. This is less precise when N=4 since people may have been excluded by two or by three group members. In fact, when N=4 total revenge occurs in 40.28% of the cases, whereas when N=3 it occurs in 27.45% of the cases. Thus, subjects vote for total revenge not only when they are sure of who in the group has made the decision to exclude.

We investigate the determinants of the excluded subjects' willingness to cast 0, 1, 2 or 3 exclusion votes when they reintegrate their group, using a random-effects Poisson regression with standard errors clustered at the group level (model (1) in Table 4). The dependent variable is a categorical variable equal to the number of exclusion votes by the reintegrated subject. The independent variables include dummy variables for each possible length of exclusion of i interacted with each treatment, setting Exo1P as the reference category. We control for the amount withdrawn by i in t, the average amount withdrawn by the other group members in t, the group size at the time of reintegration,[29] a time trend and its squared value, and the usual individual characteristics. The sample pools the data from all the treatments with exclusion but includes only the excluded subjects who reintegrate their group in period t. We only consider periods 3 to 17 and groups where exclusion is possible (N > 2).

Model (1) shows that members reintegrated after one-period exclusion in Endo cast significantly less exclusion votes than those excluded in Exo1P. Members reintegrated after three periods in Endo cast more exclusion votes than those excluded in Exo1P. But there is no significant difference between Exo1P and Exo3P. A longer exclusion increases significantly the number of exclusion votes cast by reintegrated members only in Endo. A similar regression conducted only on the Endo data indicates that the marginal effect of a two-period exclusion is 0.789*** (0.236) and the effect of a three-period exclusion is 1.100*** (0.192), compared to a one-period exclusion. The impact of the length of exclusion on retaliation is thus larger in Endo than in the Exo treatments.

To explore the determinants of voting for total revenge, model (2) reports the estimates of 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 reintegrated member votes for the exclusion of all the other active group members, and 0 otherwise. The independent variables are the same as in model (1). Controlling for individual and group average withdrawals, members excluded for one period in Endo are significantly less likely to vote for the exclusion of all the other members when reintegrating their group compared to similar subjects in Exo1P. There is no significant difference in the likelihood of total revenge between members excluded for one or for three periods when the length of exclusion is fixed exogenously. A similar regression conducted only on the data of the

[28] Two-tailed two-sample tests of proportions: Exo1P *vs.* Exo3P:p=0.017; Endo *vs.* Exo1P: p=0.246; Endo *vs.* Exo3P: p=0.235. 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.

[29] Alternatively we replaced the group size in t with the group size at the time of exclusion in these regressions and in the following ones. This variable was never significant.

Endo treatment indicates that the marginal effect of a two-period exclusion is 0.845*** (0.121) and the effect of a three-period exclusion is 0.821*** (0.117), compared to a one-period exclusion. The impact of the length of exclusion on total revenge is thus also larger in Endo than in the Exo treatments.

Table 4: Effect of the length of exclusion on the number of exclusion votes and on the likelihood of total revenge by reintegrated subjects

Dependent variables:	Number of exclusion votes	Total revenge
	(1)	(2)
1P exclusion in Exo1P	*Ref.*	*Ref.*
1P exclusion in Endo	-0.630***	-0.292***
	(0.180)	(0.073)
2P exclusion in Endo	0.026	0.181
	(0.196)	(0.201)
3P exclusion in Endo	0.275**	0.232
	(0.113)	(0.176)
3P exclusion in Exo3P	0.131	0.043
	(0.208)	(0.203)
Amount withdrawn	0.009	0.016**
by i in t	(0.007)	(0.007)
Group average withdrawal	0.030***	0.018*
in t excluding i	(0.012)	(0.009)
Number of exclusion	0.039	0.081
prior to t	(0.057)	(0.064)
Group size in t	0.456***	0.002
	(0.116)	(0.113)
Period	-0.052	-0.035
	(0.053)	(0.052)
Squared period	0.001	0.001
	(0.002)	(0.002)
Individual controls	Yes	Yes
Constant	-1.400	
	(0.582)	
Observations	323	323
Clusters	35	35
Log pseudo-likelihood	-479.24	-167.33

Notes: Model (1) reports the marginal effects from a random-effects Poisson model where the dependent varable is the number of exclusion votes cast by excluded members when they reintegrate the group. Model (2) reports the marginal effects from a random-effects logit regression where the dependent varable is the probability of the reintegrated members to vote for total revenge (*i.e.*, for the exclusion of all the 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. Standard errors (in parentheses) are clustered at the group level. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

To summarize, subjects excluded for a longer time are more likely to seek revenge than those 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 in itself that matters so much in triggering negative reciprocity, but the perceived intentions of the group.

6 Conclusion

The exclusion of non-compliers is a strong form of punishment. One question is to know whether the threat of exclusion is sufficient to discipline members to follow the norm of the community after rehabilitation. Another question is to know the extent to which people who reintegrate their group after being excluded are willing to cooperate with those who previously excluded them. The answer to this under-studied question may depend on the length of exclusion and on how this length has been determined. To address this question we have conducted a controlled experiment based on a taking game that has the properties of a social dilemma game when the group size is not too small. In this game, group members can vote for excluding their peers for different durations and the severity of punishment is either imposed exogenously or chosen by the group members.

Our findings show that individuals are willing to ostracize others, in particular those who withdraw more: larger deviations from the group average withdrawal behavior determine the vote for exclusion and to some extent the choice of the length of exclusion. People also punish more, and more severely, chronic defections. In return, a longer exclusion has a higher disciplining effect on cooperation after reintegration compared to a shorter exclusion, but only when the length of exclusion is not chosen by group members. The disciplining effect of a longer exclusion on the withdrawal level after reintegration is smaller when the length of exclusion results from a vote compared to when the length of exclusion cannot be modulated and is fixed exogenously. Another important dimension to be considered in the evaluation of the effects of exclusion is whether previously excluded individuals are willing to retaliate. We found that when group members choose the length of exclusion, individuals excluded for a shorter duration are less likely to retaliate when they reintegrate their group compared to people excluded for a longer duration. The difference in the impact of long *vs.* short exclusion on the willingness to retaliate is larger when the length of exclusion is chosen by group members than when it is exogenous.

Our results highlight the importance of a quick reintegration of the individuals who have been punished to sustain cooperation and limit anti-social behavior after reintegration when the community members choose the length of exclusion. Short exclusion has a positive effect when it can be interpreted as forgiving, which is less possible when there is only one possible length of exclusion. This is consistent with the 'reintegrative shaming' theory of Braithwaite (1989) and the results of Coricelli *et al.* (2014) showing in a tax evasion context that pardon reduces future violations whereas in the absence of pardon, offenders have no more moral motivation to comply once reputation is lost. Our results are also consistent with those of Mueller-Smith and Schnepel (2017) who show that court deferrals (which avoid formal convictions through probation) for first-time felony defendants reduce recidivism, whereas no effect is found for repeat offenders.

We acknowledge some limitations of our study. The differences in behavior between our Exo and Endo treatments may be driven by two differences in the design. First, in the Endo treatment the excluded individuals were able to compare the sentence they received with another, non-selected, sentence. In this context, learning that one receives the less severe sanction may help accept the sentence compared to a situation in which only one sentence is available. 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 effects. Similarly, the counter-effects of a long exclusion may result from more than one dimension. A longer exclusion may trigger negative emotions against group members who could have signaled their disapproval with a milder sanction. It may also encourage the excluded people to free-ride more on others and to vote for their exclusion when reintegrated in order to recover the lost earnings during exclusion or to reduce payoff inequality in a dynamic way. But the possibility to choose the length of exclusion in the Endo treatment helps also screen the free-riders. Those who are repeatedly less compliant may also be those who are more likely to behave aggressively towards others after reintegration. The fact that in our regressions we never found any effect of the type of social preferences on behavior speaks in favor of the first effects. But to weight each of these effects precisely, a random assignment of the length of exclusion and measures of emotions might be useful. Also, assigning the excluded members to different groups after reintegration would allow us to disentangle the effect of negative reciprocity from that of the willingness to recover a loss or to equalize payoffs. We leave this to further investigation.

References

[1] Ahn, T.-K., Isaac, R., and Salmon, T. C. Endogenous group formation. *Journal of Public Economic Theory 10*, 2 (2008), 171–194.

[2] Ahn, T.-K., Isaac, R., and Salmon, T. C. Coming and going: experiments on endogenous group sizes for excludable public goods. *Journal of Public Economics 93* (2009), 336–351.

[3] Akpalu, W., and Martinsson, P. Ostracism and common pool resource management in a developing country: Young fishers in the laboratory. *Journal of African Economies 21*, 2 (2011), 266–306.

[4] Ambrus, A., and Greiner, B. Democratic punishment in public good games with perfect and imperfect observability. *Mimeo* (2015).

[5] Balafoutas, L., Garcia-Gallego, A., Georgantzis, N., Jaber-Lopez, T., and Mitrokostas, E. Rehabilitation and social behavior: Experiments in prison. *Mimeo* (2017).

[6] Becker, G. S. *Crime and punishment: An economic approach.* Springer, 1968.

[7] Bernard, M., Fanning, J., and Yuksel, S. Finding cooperators: Sorting through repeated interaction. *Mimeo* (2017).

[8] Berninghaus, S., Bleich, S., and Güth, W. The consequences of (de)regulation on employment duration and efficiency: An experimental study. *Economic Inquiry 51*, 1 (2013), 1050–1065.

[9] Berube, D., and Green, D. The effects of sentencing on recidivism: Results from a natural experiment. *Paper presented at the Second Annual Conference on Empirical Legal Studies, New York* (2007).

[10] Bhuller, M., Dahl, G. B., Loken, K. V., and Mogostad, M. Incarceration, recidivism and employment. *NBER Working Paper 22648* (2016).

[11] 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.

[12] Bock, O., Baetge, I., and Nicklisch, A. Hroot: Hamburg registration and organization online tool. *European Economic Review 71* (2014), 117–120.

[13] Bolling, D., Pitskel, N., and Deen, B. e. a. Dissociable brain mechanisms for processing social exclusion and rule violation. *Neuroimage 54* (2011), 2462–2471.

[14] Bowles, S., and Gintis, H. The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theoretical Population Biology 65*, 1 (2004), 17–28.

[15] Braithwaite, J. *Crime, shame and reintegration.* Cambridge University Press, 1989.

[16] Brown, M., Falk, A., and Fehr, E. Relational contracts and the nature of market interactions. *Econometrica 72*, 3 (2004), 747–780.

[17] Chakravarty, S., and Fonseca, M. A. Discrimination via exclusion: An experiment on group identity and club goods. *Journal of Public Economic Theory 19*, 1 (2017), 244–263.

[18] Charness, G., Cobo-Reyes, R., and Jiménez, N. Efficiency, team building, and identity in a public-goods game. *UCSB working paper* (2011).

[19] Charness, G., and Yang, C.-L. Starting small toward voluntary formation of efficient large groups in public goods provision. *Journal of Economic Behavior & Organization 102* (2014), 119–132.

[20] 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.

[21] 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.

[22] Croson, R., Fatas, E., Neugebauer, T., and Morales, A. J. Excludability: A laboratory study on forced ranking in team production. *Journal of Economic Behavior & Organization 114* (2015), 13–26.

[23] Davis, B. J., and Johnson, D. B. Water cooler ostracism: Social exclusion as a punishment mechanism. *Eastern Economic Journal 41*, 1 (2015), 126–151.

[24] DeAngelo, G., and Charness, G. Deterrence, expected cost, uncertainty and voting: Experimental evidence. *Journal of Risk and Uncertainty 44*, 1 (2012), 73–100.

[25] 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.

[26] Eisenberger, N., Lieberman, M., and Williams, K. D. Does rejection hurt? an fmri study of social exclusion. *Science 302* (2003), 290–292.

[27] Falk, A., Huffman, D., and Macleod, W. B. Institutions and contract enforcement. *Journal of Labor Economics 33*, 3 (2015), 571–590.

[28] Fatas, E., Morales, A. J., and Ubeda, P. Blind justice. an experimental analysis of random punishment in team production. *Journal of Economic Psychology 31*, 3 (2010), 358–373.

[29] Feinberg, M., Willer, R., and Schultz, M. Gossip and ostracism promote cooperation in groups. *Psychological Science 25*, 3 (2014), 656–664.

[30] Fischbacher, U. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*, 2 (2007), 171–178.

[31] Fischbacher, U., Gächter, S., and Fehr, E. Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters 71*, 3 (2001), 397–404.

[32] Fischer, S., and Güth, W. Effects of exclusion on acceptance in ultimatum games. *Journal of Economic Psychology 33* (2012), 1100–1114.

[33] 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.

[34] Green, D. P., and Winik, D. Using random judge assignments to estimate the effects of incarceration and probation on recidivism among drug offenders. *Criminology 48*, 2 (2010), 357–387.

[35] Güth, W., I., L. M.-V., Sutter, M., and Van der Heijden, E. Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics 91* (2007), 1023–1042.

[36] Hirshleifer, D., and Rasmusen, E. Cooperation in a repeated prisoners' dilemma with ostracism. *Journal of Economic Behavior & Organization 12*, 1 (1989), 87–106.

[37] 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.

[38] Kling, J. R. Incarceration length, employment, and earnings. *American Economic Review 96*, 3 (2006), 863–876.

[39] 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).

[40] Kopanyi-Peuker, A., Offerman, T., and Sloof, R. Team production benefits from a permanent fear of exclusion. *Mimeo* (2017).

[41] Kuziemko, I. Going off parole: How the elimination of discretionary prison release affects the social cost of crime. *Mimeo* (2007).

[42] Landerso, R. Does incarceration length affect labor market outcomes? *The Journal of Law and Economics 58*, 1 (2015), 205–234.

[43] Lott, J. R. Do we punish high income criminals too heavily? *Economic Inquiry 30*, 4 (1992), 583.

[44] 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.

[45] Martinson, R. What works?-questions and answers about prison reform. *The Public Interest*, 35 (1974), 22.

[46] Masclet, D. Ostracism in work teams: a public good experiment. *International Journal of Manpower 24*, 7 (2003), 867–887.

[47] Masten, C., Eisenberger, N., and Borofsky, 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.

[48] Maurin, E., and Ouss, A. Sentence reductions and recidivism: Lessons from the bastille day quasi experiment. *IZA Discussion Paper n3990* (2009).

[49] Mueller-Smith, M., and Schnepel, K. T. Diversion in the criminal justice system: Regression discontinuity evidence on court deferrals. *Mimeo, available at https://www.dropbox.com/s/9qrexa7fqer1s7u/Diversion in the Criminal Justice System.pdf?dl=0* (2017).

[50] 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.

[51] Neuhofer, 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).

[52] Nikiforakis, N. Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics 92* (2008), 91–112.

[53] Nikiforakis, N., and Engelman, D. Altruistic punishment and the threat of feuds. *Journal of Economic Behavior & Organization 78* (2011), 319–332.

[54] Nikiforakis, N., and Normann, H.-T. A comparative statics analysis of punishment in public good experiments. *Experimental Economics 11* (2008), 358–369.

[55] Pressman, S. D., and Cohen, S. Does positive affect influence health? *Psychological Bulletin 131*, 6 (2005), 925.

[56] Riedl, A., and Aljaz, U. Exclusion and cooperation in social network experiments. *Mimeo* (2014).

[57] Riedl, A., Rohde, I., and Strobel, M. Efficient coordination in weakest-link games. *Review of Economic Studies 83*, 2 (2016), 737–767.

[58] Roach, M., and Schanzenbach, M. The effect of prison sentence length on recidivism: Evidence from random judicial assignment. *Mimeo* (2016).

[59] Schildberg-Hörisch, H., and Strassmair, C. An experimental test of the deterrence hypothesis. *Journal of Law, Economics, and Organization 28*, 3 (2012), 447–459.

[60] Selten, R. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory 4*, 1 (1975), 25–55.

[61] Sheremeta, R. M., J., T. S., and Jingjing, Z. Creating self-sustained social norms through communication and ostracism. *Mimeo* (2009).

[62] 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.

[63] Swope, K. J. An experimental investigation of excludable public goods. *Experimental Economics 5*, 3 (2002), 209–222.

[64] Sääksvuori, L. Intergroup conflict, ostracism, and the evolution of cooperation under free migration. *Behavioral Ecology and Sociobiology 68*, 8 (2014), 1311–1319.

[65] 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.

[66] Wesselmann, E. D., and Williams, K. D. *Ostracism and Stages of Coping.* Oxford: Oxford University Press, 2013.

[67] 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.

[68] 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.

[69] 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.

[70] Williams, K. D. *Ostracism: The power of silence.* Guilford Press, 2002.

[71] 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.

[72] Zapryanova, M. The effects of time in prison and time on parole on recidivism. *Mimeo* (2014).

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:

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

Example:

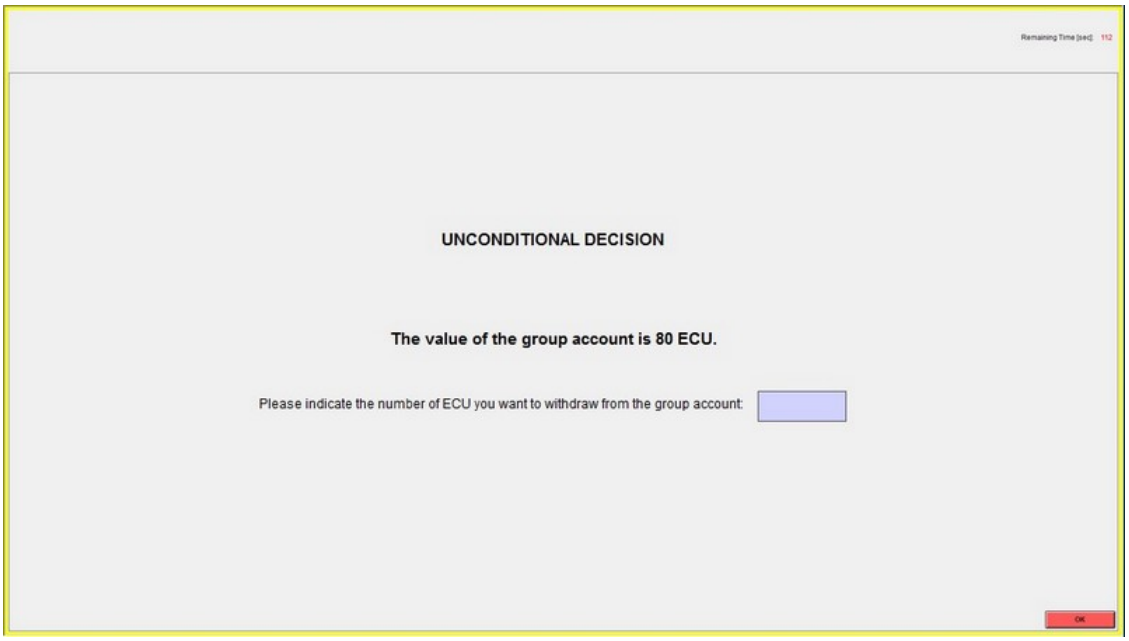
Suppose that you withdraw 10 ECU from the group account and each of the other 3 group members withdraws 1 ECU from the group account. The total amount left on the group account is 80 ECU minus the 10 ECU that you have withdrawn, minus 3 ECU withdrawn by the other group members, *i.e.* 67 ECU. In this 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 ».

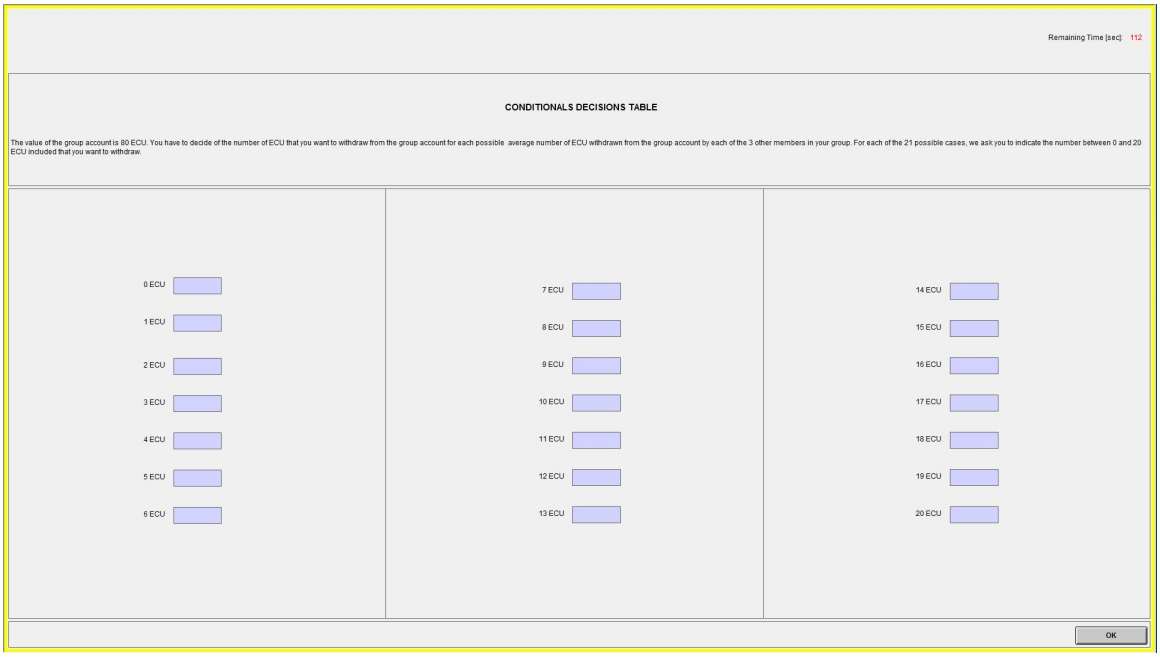
Unconditional decision: Here is an example of the screen you will get for an unconditional decision:



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:



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*(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*(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*(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*(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.

Instructions 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:

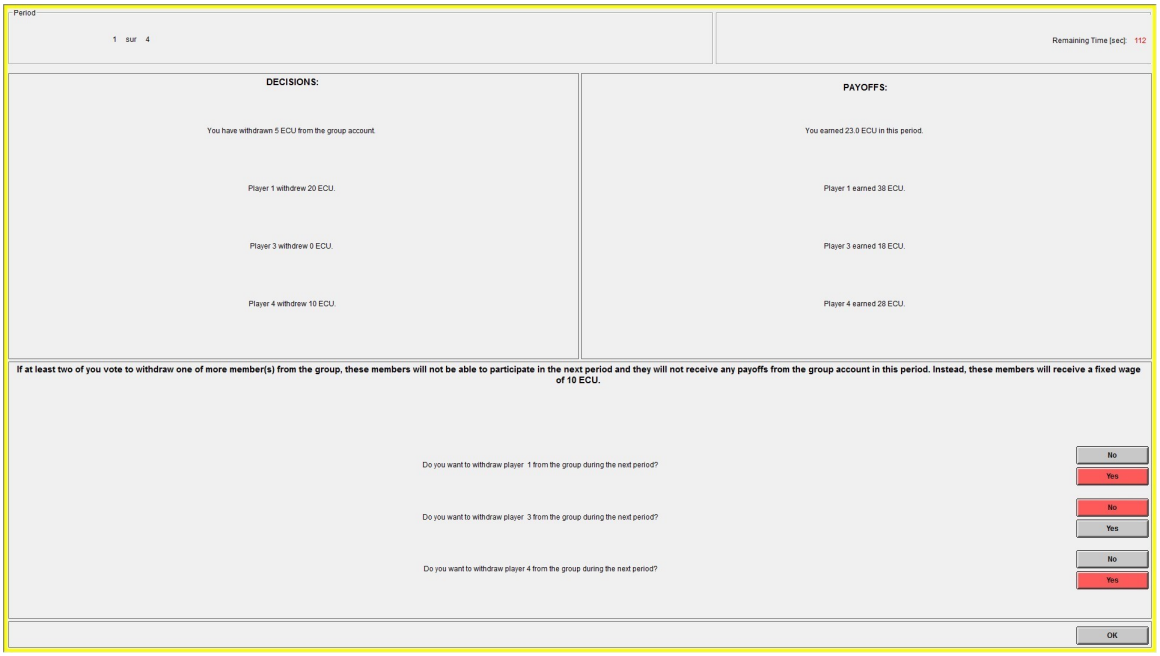
Decision rule

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:



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:

The amount you have withdrawn from the group account
+ 40% (value of the group account at the beginning of the period
− 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.

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.

Instructions 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:

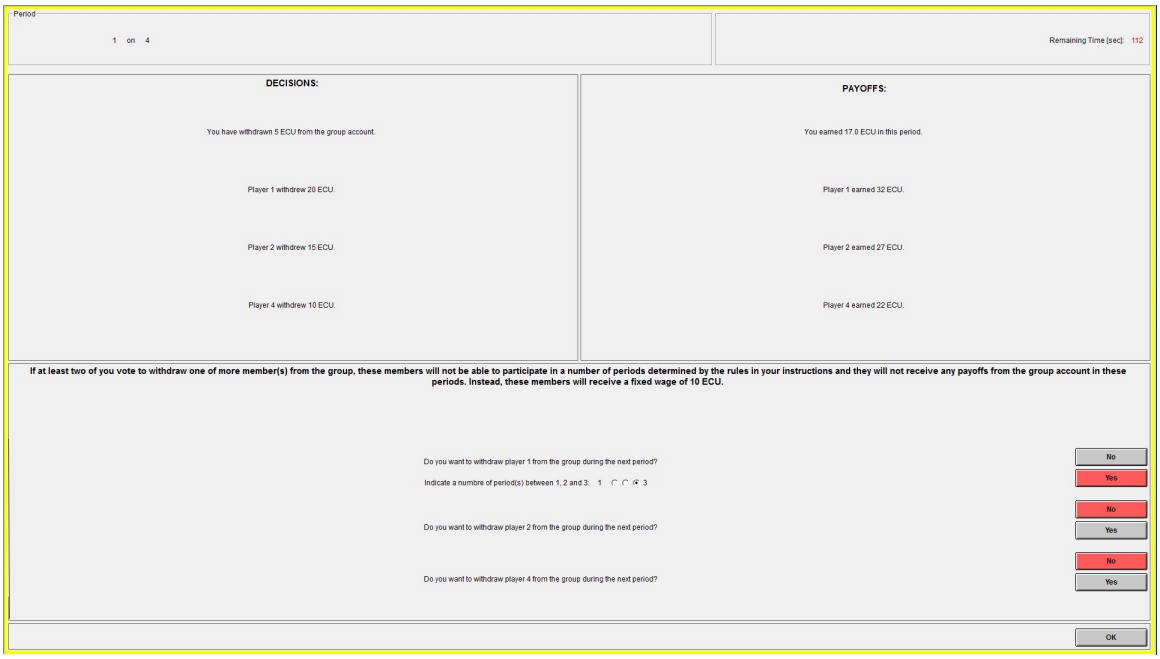
Decision rule
In this stage, the group members have the opportunity to remove one or several other group members for one, two or three following periods. Removing a member from the group means that he will not be able to participate in any of the two stages for the number of periods he has been removed: 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:



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:

The amount you have withdrawn from the group account
+ 40% (value of the group account at the beginning of the period
– 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.
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

Appendix 2.1 Subject pool

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 any 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 treatments is never significant (*vs.* Exo1P: p=1.000; *vs.* Exo3P: p=0.740).

Appendix 2.2 Determinants of exclusion

Table A.2 reports the marginal effects of the estimates of the vote for exclusion of a subject i, using random-effects logit regressions with standard 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 difference between j's withdrawal and the group average withdrawal (excluding j) in period t when j withdraws more than the group (max $(0; w_{j,t} - \sum_{-j=1,-j\neq j}^{n-1} w_{-j,t})$) and the difference between the group average withdrawal and j's withdrawal in t when j withdraws less than the group (max $(0; \sum_{-j=1,-j\neq j}^{n-1} w_{-j,t} - w_{j,t})$). The ratio between the current number of periods prior to t in which j has withdrawn more than the average group withdrawal ($w_{j,t} - \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t} > 0$) and the current number of periods prior to t in which j was active in the group (considering only the cases in which the group size was higher than two) captures the impact of chronic free-riding by j on i's willingness to vote for exclusion. The independent variables also include a control for whether j has already been excluded in the past and 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 xP exclusion (Y)" takes value 1 if i has been reintegrated in t after x period(s) of exclusion, in the Y treatment,

Table A.2: Determinants of an exclusion vote

Dep. variable: Decision to exclude j in t	All treatments (1)	with exclusion (2)
Exo1P	*Ref.*	*Ref.*
Exo 3P	-0.054	-0.047
	(0.059)	(0.058)
Endo	-0.062	-0.078
	(0.060)	(0.057)
Max $(0; w_{j,t} - \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t})$	0.044***	0.043***
	(0.010)	(0.010)
Max $(0; \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t} - w_{j,t})$	0.004	0.005*
	(0.003)	(0.003)
Ratio current nb of pos. dev. of j prior to t/	0.077***	0.079***
Nb of periods with j in the group prior to t	(0.022)	(0.021)
Exclusion of j prior to t	-0.027	-0.028
	(0.018)	(0.018)
i is reintegrated in t	0.114***	0.089**
	(0.041)	(0.044)
i is reintegrated in t	-	Ref.
after 1P exclusion (Exo1P)		
i is reintegrated in t	-	-0.087***
after 1P exclusion (Endo)		(0.031)
i is reintegrated in t	-	0.347
after 2P exclusion (Endo)		(0.242)
i is reintegrated in t	-	0.321
after 3P exclusion (Endo)		(0.245)
i is reintegrated in t	-	-0.037
after 3P exclusion (Exo3P)		(0.029)
Group size in t	-0.010	-0.010
	(0.018)	(0.018)
Period	0.002	0.003
	(0.003)	(0.003)
Squared period	0.000	0.000
	(0.000)	(0.000)
Individual controls	Yes	Yes
Observations	4197	4197
Clusters	32	32
Log pseudo-likelihood	-1321.18	-1300.54

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 he is excluded and i cannot vote to exclude j if j is already excluded). Pos. dev. for Positive deviation $(w_{j,t} - \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t}) > 0)$. Individual controls include gender, age, student status, and category of social preferences. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

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 include a time trend (period and period squared to control for a possible non linear effect of time). We also control for the group size and for individual characteristics (gender, age, student status and social preferences of i).

Table A.2 indicates that the more ECU j takes above his group average, the more likely subject i is to vote in favor of his exclusion. In this table, chronic free-riding is captured by the ratio between the current number of periods prior to t in which j has withdrawn more than the average group withdrawal and the current number of periods prior to t in which j was active in the group (considering only the cases in which the group size was higher than two). Chronic free-riding compared to the average withdrawal in the group increases the likelihood of i voting for j's exclusion. Controlling for j's withdrawal behavior, the probability to vote to exclude a group member is the same across treatments, regardless of the severity of punishment. Table A.2 also shows that i is more likely to vote to exclude another group member in t when he has been himself 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 he has been excluded for one period in Endo compared to when he has been excluded in Exo1P. Therefore, when it results from a vote, a short exclusion affects differently the willingness to exclude after reintegration.

Table A.3 reports the marginal effects from two random-effects ordered logit regressions aiming at investigating the determinants of the choice of the length of exclusion in the Endo treatment. Robust standard errors are clustered at the group level. The dependent variable is a categorical variable that equals 1 if individual i votes to exclude j for one period, 2 if i votes to exclude j for two periods, and 3 if i votes to exclude j for three periods. Table A.3 reports the marginal effects of the independent variables on the vote to exclude for one period in columns (1) and (2), two periods in columns (3) and (4), and three periods in columns (5) and (6). The independent variables include the difference between j's withdrawal and the group average withdrawal (excluding j) in period t when j withdraws more than the group (max $(0; w_{j,t} - \sum_{-j=1,-j\neq j}^{n-1} w_{-j,t})$) and the difference between the group average withdrawal and j's withdrawal in t when j withdraws less than his group (max $(0; \sum_{-j=1,-j\neq j}^{n-1} w_{-j,t} - w_{j,t})$). They include the ratio between the current number of periods prior to t in which j has withdrawn more than the average of his group and the current number of periods prior to t in which j was active in the group. They include a control for whether j has already been excluded in the past and a reintegration dummy that equals 1 if i is reintegrated in t and 0 otherwise. In model (2) reported in columns (2), (4) and (6), we add variables detailing the length of exclusion of i. The variable "i is reintegrated in t after xP exclusion" takes value 1 if i has been reintegrated in t after x period(s) of exclusion 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 include a time trend and its squared value to test whether exclusion votes are more likely at the beginning of the game when the expected benefits of punishment are potentially higher or to impose a norm as soon as possible. We control for group size and for individual characteristics (gender, age, student status and social preferences).

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

Dep. variable: Choice of i regarding the length of exclusion of j	Exclusion for 1 period (1)	1 period (2)	2 periods (3)	2 periods (4)	3 periods (5)	3 periods (6)
Max (0; $w_{j,t} - \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t}$)	-0.009*	-0.012*	-0.003*	-0.003*	0.013*	0.015**
	(0.005)	(0.006)	(0.002)	(0.002)	(0.007)	(0.007)
Max (0; $\sum_{-j=1,j\neq -j}^{n-1} w_{-j,t} - w_{j,t}$)	-0.006	-0.007	-0.002	-0.002	0.008	0.009
	(0.011)	(0.013)	(0.004)	(0.004)	(0.015)	(0.017)
Ratio current nb of pos. dev. of j prior to t/	-0.121*	-0.116*	-0.040**	-0.033*	0.162**	0.150*
Nb of periods with j in the group prior to t	(0.069)	(0.067)	(0.020)	(0.020)	(0.083)	(0.078)
Exclusion of j prior to t	-0.060	-0.046	-0.020	-0.013	0.080	0.060
(dummy)	(0.064)	(0.076)	(0.029)	(0.027)	(0.093)	(0.103)
i is reintegrated in t	-0.125*	0.148	-0.042***	0.042	0.167**	-0.192
(dummy)	(0.066)	(0.122)	(0.015)	(0.043)	(0.074)	(0.159)
i is reintegrated in t	-	-0.360**	-	-0.102*	-	0.466**
after 2P exclusion (Endo)		(0.179)		(0.061)		(0.208)
i is reintegrated in t	-	-0.362**	-	-0.103**	-	0.469**
after 3P exclusion (Endo)		(0.183)		(0.049)		(0.196)
Group size in t	-0.073	-0.070	-0.024	-0.020	0.098	0.090
	(0.072)	(0.074)	(0.025)	(0.022)	(0.095)	(0.094)
Period	-0.014	-0.017	-0.005	-0.005	0.019	0.022
	(0.012)	(0.013)	(0.004)	(0.004)	(0.016)	(0.016)
Squared Period	0.000	-0.000	-0.000	0.000	-0.000	-0.000
	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	337	337	337	337	337	337
Clusters	11	11	11	11	11	11
Log pseudo-likelihood	-263.15	-257.94	-263.15	-257.94	-263.15	-257.94

Notes: The Table reports the marginal effects from two random-effects ordered logit regressions with robust standard errors clustered at the group level in parentheses. The second model, corresponding to columns (2), (4), and (6), differs from the first model simply by controlling for the length of exclusion of player i. 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). Pos. dev. for Positive deviation ($w_{j,t} - \sum_{-j=1,j\neq -j}^{n-1} w_{-j,t}$) >0). Individual controls include gender, age, student status, and category of social preferences. *** p <0.01, ** p <0.05, * p <0.10.

Table A.3 shows that the vote for a specific length of exclusion is affected by the number of ECU taken by j relative to the average withdrawal in the group. There is a significant effect of the reputation of the subject who is punished: a chronic positive deviation from the group average withdrawal leads to votes for a longer exclusion. When including this variable, the fact that j has already been excluded in the past does not affect the choice of a specific length. 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.

Appendix 2.3 Reintegration

Table A.4 reports the marginal effects from random-effects GLS regressions with robust standard errors clustered at the group level.

Table A.4: Effect of the length of exclusion on the withdrawal (in ECU) of excluded subjects in the reintegration period (all group sizes)

Dependent variable: Number of ECU withdrawn	All treatments with exclusion (1)	(2)
1P exclusion in Exo1P	*Ref.*	2.990
		(2.348)
1P exclusion in Endo	-2.990	*Ref.*
	(2.348)	
2P exclusion in Endo	0.413	3.403**
	(1.638)	(1.447)
3P exclusion in Endo	0.862	3.352**
	(1.788)	(1.704)
3P exclusion in Exo3P	-3.988**	-0.998
	(1.988)	(2.507)
$\sum_{-i=1,i\neq-i}^{n-1} w_{-i,t} - w_{i,t}$	-0.062	-0.062
at the time of exclusion	(0.128)	(0.128)
Group average withdrawal	0.306***	0.306***
excluding i in t-1	(0.111)	(0.111)
Group size in t	-1.404	-1.404
	(0.980)	(0.980)
Period	-0.378	-0.378
	(0.440)	(0.440)
Squared period	0.023	0.023
	(0.019)	(0.019)
Individual controls	Yes	Yes
Constant	8.017	5.027
	(4.792)	(5.012)
Observations	244	244
Clusters	35	35
R-squared	0.31	0.31

Notes: The model estimates random-effect GLS models with robust standard errors clustered at the group level in parentheses. The data include the observations from subjects reintegrated in periods 3 to 20 when the group contains at least three members, excluding subjects who were excluded after period 17. Individual controls include gender, age, student status, and category of social preferences. *** p <0.01, ** p <0.05, * p <0.10.

The dependent variable is the number of ECU withdrawn by the excluded members when they reintegrate their group in the treatments with possible exclusion. This Table is similar to Table 2 except that all group sizes are included.

Appendix 3. Figures

Appendix 3.1 Evolution of average withdrawal over time, by treatment (including all the subjects and all group sizes)

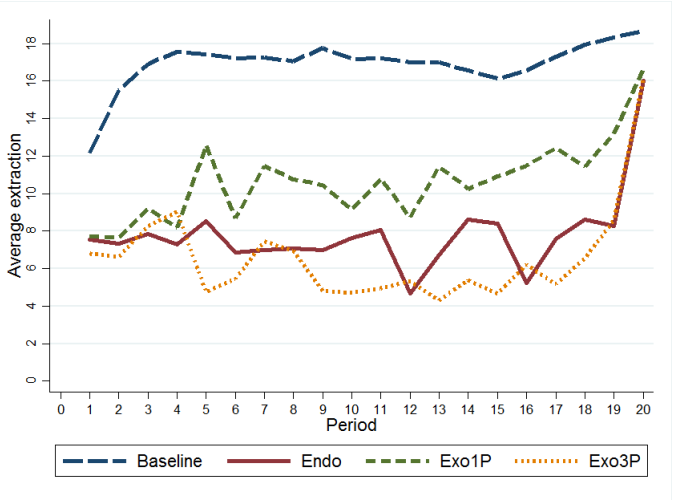
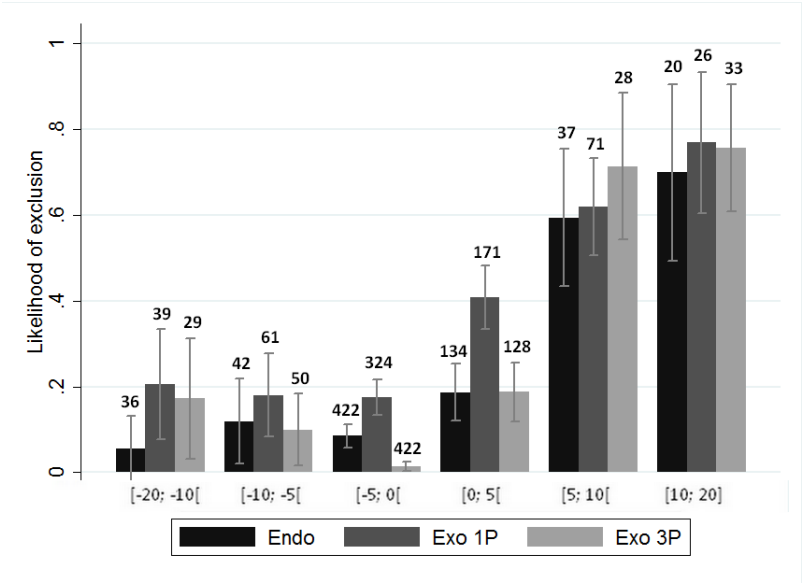


Figure A.1: Evolution of average withdrawal over time, by treatment (including all the subjects and all group sizes).

Appendix 3.2 Exclusion rate according to the difference between the individual withdrawal and the average withdrawal in the group (groups with N>2)

The following Figure displays the mean rate of exclusion of individuals depending on the difference between their withdrawal and the mean withdrawal in their group (excluding themselves), taken as the group norm. We consider the group average as the social norm in the group, based on the previous literature on sanctioning in social dilemma games. To support this point, in a post experimental questionnaire subjects were asked to report one or more reasons explaining their decision to exclude another participant at least once during the experiment. The proposed reasons were: (i) the participant was taking more than me, (ii) the participant was taking more than the average of the group, (iii) I wanted to signal that I would stop cooperating if the participant did not change her behavior, (iv) I was removed from the group myself and I wanted to take revenge, (v) other reasons. 33.33% of the subjects reported reason (i), 52.53% reported reason (ii), 51.52% reason (iii), and 25.25% reason (iv).) Each bar represents the proportion of members who, among those subjects, extracted x to y ECU more or less than their group average, have been excluded in each treatment. We can see that group members who extract on average 10 to 20 ECU more than the group average (bars on the far right are more likely to be excluded than those who extract on average less than 5 ECU more than the group average. This is observed in all treatments.



Notes: The numbers above the bars indicate the number of observations. A negative (positive) difference means that the subjet withdraw less (more) than the average of his group.

Figure A.2: Exclusion rate according to the difference between the individual withdrawal and the average withdrawal in the group, by treatment (N>2)

Appendix 3.3 Difference between the number of ECU withdrawn by excluded subjects and by their group members in the exclusion period and in the reintegration period (including all group sizes)

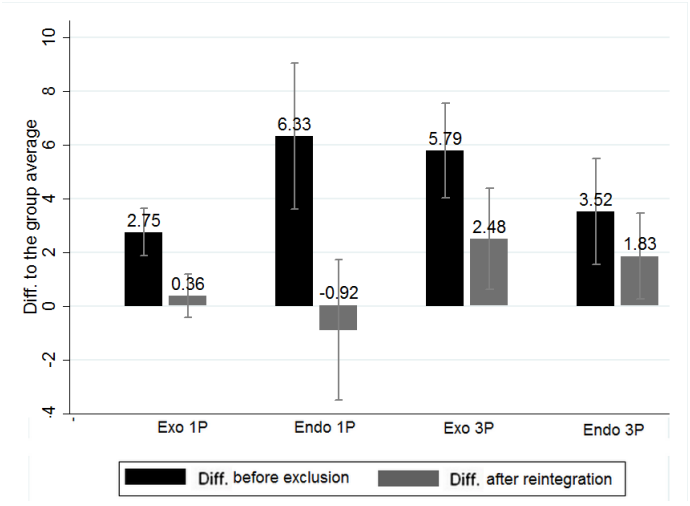


Figure A.3: Difference between the number of ECU withdrawn by excluded subjects and by their group members in the exclusion period (left bars) and in the reintegration period (right bars), by treatment (including all group sizes).