IZA

DISCUSSION PAPER SERIES

IZA DP No. 4643

Voting with Feet: Community Choice in Social Dilemmas

Özgür Gürerk
Bernd Irlenbusch
Bettina Rockenbach

December 2009

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Voting with Feet: Community Choice in Social Dilemmas

Özgür Gürerk
University of Erfurt

Bernd Irlenbusch
London School of Economics and IZA

Bettina Rockenbach
University of Erfurt

Discussion Paper No. 4643
December 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

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

IZA Discussion Paper No. 4643
December 2009

ABSTRACT

Voting with Feet: Community Choice in Social Dilemmas

Economic and social interactions often take place in open communities but the dynamics of the community choice process and its impact on cooperation of its members are yet not well understood. We experimentally investigate community choice in social dilemmas. Participants repeatedly choose between a community with and an alternative without punishment opportunities. Within each community a social dilemma game is played. While the community with punishment grows over time and fully cooperates, the alternative becomes depopulated. We analyze the success of this "voting with feet" mechanism and find that endogenous self-selection is key while slow growth is less decisive.

JEL Classification: C72, C92, H41

Keywords: cooperation, social dilemmas, community choice, punishment, voting with feet

Corresponding author:

Bernd Irlenbusch
London School of Economics
Department of Management
Houghton Street
London WC2A 2AE
United Kingdom
E-mail: b.irlenbusch@lse.ac.uk

The consumer-voter may be viewed as picking that community which best satisfies his preference pattern for public goods. Charles M. Tiebout (1956, p. 418)

Understanding the determinants and the extent of human cooperation is one of the most challenging questions in economics. Human cooperation in social dilemmas is particularly puzzling because the conflict between collective and individual interests creates the well-known free-rider problem (Garrett Hardin 1968, Robyn M. Dawes 1980, Elinor Ostrom 1999, Samuel Bowles 2004). In order to disentangle different motives for cooperation and defection, researchers have successfully studied behavior in controlled social dilemma experiments. In repeated interactions, cooperation turns out to be rarely stable and generally deteriorates to rather low levels over time (Douglas D. Davis and Charles A. Holt 1993, John Ledyard 1995, Rachel T. Croson 1998, Ostrom 1998, Urs Fischbacher and Simon Gächter, forthcoming). More recent studies identify possibilities to punish norm violators as forceful means to increase cooperation. Decentralized one-to-one punishment that reduces the income of the punished player is heavily used (Ostrom et al. 1992, Ernst Fehr and Gächter 2000, 2002) and is efficiency increasing in longer interactions (Gächter et al. 2008), even if punishing incurs costs for the punisher.

A common feature of the vast majority of previous experimental studies on social dilemmas is that the interaction framework is *exogenously* imposed. In reality, however, humans often vote *with their feet* between different institutional frameworks governing the interaction with others. For example, employees choose to work in a company with an appealing corporate culture that encourages employees' contributions, and to interact with others who are attracted by the same culture. People join different clubs, sports teams, and parties and commit to abide by their rules because they want to pursue certain activities and achieve certain goals with the help of others who join the same communities. Citizens move to different jurisdictions or even to different countries because these constitute a better fit for their preferences regarding public goods provision, ways of living, or the political system and because they prefer to live together with likeminded others who want to fulfill their citizen duties under the same system.[1] Along these lines, Tiebout (1956) suggests that individuals with similar preferences

[1] An illustrative example of a company with an impressive corporate culture that attracts highly qualified employees and convinces them to provide enormous contributions to its success is Google (David A. Vise 2006). For a prominent example in history when people voted with their feet for a political system, recall the large-scale migration in the 1950s, when thousands of East Germans flew to the West to benefit from and to contribute to the "industrial miracle". To quote from Time Magazine: "In the only kind of voting that remains to the East Germans—what one British diplomat calls voting with their feet—they have chosen to flee the country at a rate which for the past three months has averaged a startling 1,000 refugees a day." (21 Nov 1955).

for different bundles of local public goods sort themselves into communities governed by different institutions in expectation of interacting with others who have chosen the same institution. He argues that if communities are sufficiently heterogeneous and consumer-voters are fully mobile, voting with feet generates considerable efficiency gains in public goods provision. Although all this suggests that voting with feet is likely to play a crucial role in shaping cooperative behavior within a community, its actual role in social dilemma situations has not yet been satisfactorily investigated. What makes voting with feet so important? Voting with feet might enhance the quality of the match between preferences and institutions. Identification with an institution and the resulting commitment to the prevailing social norms of its community members can be assumed to be high *if* institutions are chosen. Presumably the self-selected founding members of an institution are able to create and establish an initial culture which most likely is decisive for how strong subsequent members will be committed to the institution.[2] In this paper, we experimentally study the impact and the dynamics of voting with feet on community choice.

To set our study on solid ground, we focus on the endogenous community choice between two "work-horse" institutions that separately have already been extensively discussed in the public goods literature. Experimental subjects repeatedly "vote with their feet" (henceforth VF) by choosing between joining a community with punishment possibilities (henceforth P) and joining a community with no punishment possibility (henceforth N). Within the chosen community members interact in a public goods game with all others who have chosen the same community. The behavior we observe provides illuminating insights. Initially, more than two-thirds of the subjects decide to avoid the community with punishment possibilities and vote with their feet for N. Over time, however, the proportion of subjects in the N community steadily decreases with an almost complete extinction towards the end of the experiment. The few subjects who initially choose P heavily punish free-riders and achieve almost full cooperation in P. Cooperation is surprisingly stable and even continues when the community grows large – in fact, in the end almost all subjects join P.[3] The high contribution levels render punishment unnecessary and almost full efficiency is achieved.

---

[2] Founding entrepreneurs, for example, have a very strong influence on the culture of the organization they create (Edgar H. Schein 2004).

[3] These observations are well in line with a previous study by Özgür Gürerk et al. (2006) in which the choice between a punishment-and-reward-institution and one with no sanctioning is investigated. We will discuss this study in more detail in Section 5.

How can we explain the evident success of the VF choice between P and N? We observe that in almost all sessions the P communities start off with only one third of the subjects, but grow large over time and ultimately reach almost 100 percent of the total population. This observation leaves room for two non-exclusive explanations. The VF choice allows for "self-selection" of the community members into the preferred institution and due to the different nature of the institutions – with and without punishment possibilities – the selection may be driven by different predispositions to cooperate. Particularly in the beginning, this selection process of like-minded people (Armin Falk et al. 2003, Gächter and Christian Thöni 2005) may initiate and foster a culture of high levels of cooperation in the punishment community. Later on, others with less cooperative attitudes might also be attracted simply by the success of the cooperative culture in P. A second explanation which is independent from the "self-selection" argument is that a group that starts small and grows slowly can coordinate better than a group that already starts at "full size". Evidence pointing into this direction is presented by Roberto A. Weber (2006), who finds that a slow growth path improves coordination in a coordination game.[4]

To disentangle these two possible explanations for the success of the VF mechanism, we report two additional control experiments. In the first control experiment, we simulate the same growth paths as they endogenously occurred in our main study VF, but we exogenously allocate subjects to the two institutions in each period. Hence institution allocations do not emerge from self-selection, but are exogenously imposed by the experimenter. We refer to this experiment as GX (growing groups with exogenous allocation). The comparison shows that contribution rates in GX are significantly lower than in VF indicating that self-selection of subjects plays a crucial role for the superior performance of the VF mechanism. A second control experiment eliminates the effects of an increasing growth path. Subjects are exogenously allocated into fixed-sized communities in which they remain for the entire experiment. We refer to this experiment as FX (fixed size groups with exogenous allocation). Contrasting FX with GX allows studying the impact of growing communities. The results of FX do not differ significantly from GX, neither in contributions nor in efficiency. This indicates that starting with a small group of subjects (and growing afterwards) does not per se foster high contributions. Instead, it seems that the group has to be composed of the "right"

---

[4] Note that if one assumes selfish preferences we do not have a coordination game since each participant has a dominant strategy to free-ride. However, if one considers the participants to have social preferences, coordination on one of several Pareto-ranked equilibria might be necessary (see our analysis in section 2).

subjects, who initially establish a cooperative environment through high contributions and rigid punishment. In our VF mechanism these subjects find their way into the P community quite early since they do not have to fear punishment. Hence, our findings suggest that the endogenous choice is key for the observed extremely high levels of cooperation. The self-selection of cooperative subjects (including some with a predisposition to punish defectors) into the punishment community and their success in establishing a cooperative culture despite initial personal payoff disadvantages seems to be the main driving force in VF.

In the next section, we introduce our model and section 2 provides predictions on community choice behavior. Section 3 deals with our experimental design and procedures. In section 4, we present the results and compare our findings to the two control experiments. Section 5, relates our findings to the literature, and section 6 concludes.

1. Community choice by voting with one's feet

In our model, twelve players choose between the two communities before interacting in a voluntary contributions situation with others who have chosen the same community. The non-punishment community (N) resembles the standard voluntary contribution mechanism. In the punishment community (P), players may additionally engage in costly punishment of other players after having observed their contributions. We consider a three-stage game consisting of a "voting with feet" stage (S0), a voluntary contribution stage (S1) and a punishment stage (S2).

1.1. "Voting with feet" stage (S0)

In S0 all twelve members of the population simultaneously choose one of the two communities N and P. Once all choices are completed, each player is informed about the number of players n_θ, $\theta \in \{N, P\}$, who have chosen the same community. Since the total number of players is common information, the size of the population of the other community is easily inferable. Notice that only the number of players and not their identities or histories of their play are revealed.

1.2. Contribution stage (S1)

In S1, a player i interacts only with those players who have chosen the same community. Each player is endowed with 20 monetary units (tokens) and may contribute an integer amount of g_i $(0 \le g_i \le 20)$ to a joint project. Players decide simultaneously on their own contribution. The amount not contributed remains in the player's private account.[5] The sum of all contributions $G_\theta = \sum_{j=1}^{n_\theta} g_j$ is multiplied by a_θ (with $1/n_\theta < a_\theta < 1$) and then consumed by each player, independent of the individual contribution g_i. The *marginal per capita return* (MPCR) a_θ – the return of each player from her own and others' contributions – depends on the community size n_θ. In order to give smaller communities the potential to be as productive as larger ones, we keep the *productivity* $R = n_\theta a_\theta$ constant for different group sizes by setting $R = n_\theta a_\theta = 1.6$. Thus, the return from the joint project for a player does not vary in n_θ if all members of a community symmetrically contribute a certain amount. As a consequence a_θ is a decreasing function in n_θ, as shown in Table 1.[6]

Table 1: Marginal per capita return a_θ

Community size n_θ	2	3	4	5	6	7	8	9	10	11	12
Marginal per capita return a_θ	0.80	0.53	0.40	0.32	0.27	0.23	0.20	0.18	0.16	0.15	0.13

After all players have taken their contribution decisions, they are informed about the individual contributions of each member in their own community, again without revealing identities.

1.3. Punishment stage (S2)

In S2 each player receives 20 additional monetary units independent of her community affiliation and her contribution in S1. Providing players in N with the same additional endowment eliminates incentives to choose P just for receiving the extra tokens from S2. For the members of N, in this period the game ends here. The total monetary payoff of player *i* in N is

(1) $$x_i^N = (20 - g_i + a_N G_N) + 20 .$$

[5] If only a single player joins a community, no joint project can be created and the total endowment of the player is automatically transferred to her private account. Therefore, this player has no decision in stages S1 and S2.

[6] Mark Isaac and James Walker (1988) examine the effects of different MPCRs in public goods experiments. They find significantly more free-riding behavior in their low MPCR treatment (0.30) than in the high MPCR condition (0.75). This suggests that in our setting cooperation is harder if a community grows large. Of course this observation emerges from a static comparison between treatments and does not reflect a dynamic change of the MPCR within a population.

In P all players simultaneously decide whether or not to punish other members of their community. All players are provided with the same punishment capacity, which is independent of their contributions g_i in S1. In total, each player may assign up to 20 tokens.[7] Player *i* can punish community member *j* by assigning punishment tokens t_{ij}. Each token assigned by player *i* to player *j* incurs a cost of 1 token for player *i* and reduces the payoff of player *j* by 3 tokens. Thus, the marginal cost of punishment is $c = 1/3$.[8] Let T^i denote the amount of tokens that player *i* assigns and T^{-i} denote the amount of tokens that player *i* receives from the other community members. The total monetary payoff of player *i* in P results in

(2) $$x_i^P = (20 - g_i + a_P G_P) + (20 - T^i - 3T^{-i}).$$

The expressions in parentheses represent the stage payoffs of S1 and S2, respectively. After S2 is completed, all players are informed about all other players' contributions, their punishment tokens assigned, their punishment tokens received and their resulting total payoffs. We model an open information flow between both communities, i.e., at the end of each period, members of P are informed about the contributions and payoffs in N and members of N receive the same information about P as members of P do. The game described so far is repeated 30 times (periods) involving the same twelve subjects.[9]

2. Theoretical Predictions

While in the N-community, full contributions by all players would be socially optimal, contributing zero is the dominant strategy for money-maximizing players (independent of the group size n_N). Thus, in the unique Nash equilibrium, each player free-rides and earns 40 tokens. Full contribution of each player (i.e., $g_i = 20$) maximizes the joint payoff of the community and each player's payoff equals 52 tokens, again independent of the group size

---

[7] It is worth noting that the threat of punishment is likely to be heavier in larger communities, as in total there are more punishment tokens available.

[8] This means it costs 1/3 token to reduce another player's income by 1 token. For punishment, the applied cost-to-effect ratio of 1:3 reflects that in general punishing someone is less costly than being punished (cf. Klaus Abbink et al. 2000, and Fehr and Gächter 2002). With a punishment leverage of 1:3, the punisher can reduce the absolute inequality in payoffs to his own disadvantage since punishment reduces the income of the punished player more than the own income is diminished.

[9] In our experiment, we consider a special case of a partner design in which not all members of the group necessarily interact in each period. For an investigation of the differences in behavior of strangers and partners in social dilemma situations, see, e.g., Croson (1996) or Claudia Keser and Frans van Winden (2000).

n_N. In P, a purely money-maximizing player will not punish in S2 since it is costly to do so. Rational players foresee this and refrain from contributing in S1. Hence, independent of n_P, players do not contribute and do not punish in the subgame-perfect equilibrium. In this case the total payoff is 40 tokens. As in N, the joint payoff is maximized and each player's payoff is 52 tokens when all community members fully cooperate. Thus, in a world of money-maximizing rational actors, a player is indifferent between N and P because in equilibrium identical payoffs of 40 will be achieved. Since our interaction is finite, backward induction suggests that in our repeated setting, players will also neither contribute nor punish.

Let us now assume that (at least some) players are not exclusively motivated by their own monetary payoffs, but have other-regarding preferences; as modeled, for example, in Fehr and Klaus M. Schmidt (1999). They suggest a utility function where player *i* weights inequality in payoffs to her disadvantage with a parameter α_i and inequality in payoffs to her advantage with a parameter β_i. They assume $0 \le \beta_i < 1$ and $\alpha_i \ge \beta_i$. If $x = (x_1, ..., x_{n_\theta})$ denotes the vector of the individual monetary payoffs of the n_θ community members, player *i*'s utility is described by

$$(3) \qquad U_i(x) = x_i - \alpha_i \frac{1}{n_\theta - 1} \sum_{j \neq i} \max\{x_j - x_i, 0\} - \beta_i \frac{1}{n_\theta - 1} \sum_{j \neq i} \max\{x_i - x_j, 0\}.$$

Fehr and Schmidt (1999) apply their model to a public goods setting with voluntary contributions and show that equilibria may exist in which *conditional cooperators* contribute strictly positive amounts to the public good.[10] In a first step, we consider the two communities N and P separately and examine the conditions for equilibria with positive levels of cooperation. In a second step, we extend this analysis by discussing the dynamics inherent in the endogenous choice process.

The analysis of the static consideration of the two communities is an extension of Fehr and Schmidt's examination, taking the varying community sizes into account. Adapting Fehr and Schmidt's original analysis to our model, we can show that players who are not sufficiently averse to advantageous inequality (i.e., players with parameters $\beta_i < 1 - R / n_N$) never

[10] A "conditional cooperator" is a player who reduces his disutility from advantageous inequality by contributing himself if other players also contribute. A necessary condition for tolerating some free-riders in a community without punishment is that the suffering from disadvantageous payoff inequality is not too high cf. Fehr and Schmidt (1999), Proposition 4 (c), p. 839.

contribute to the public good when punishment is absent, as in our N community. In addition, when the productivity parameter $R < 2$, as it is the case in our experimental parameterization as well as in the vast majority of similar experiments, we can show that in the N community, equilibria with positive contributions exist if and only if *all* players *i* are sufficiently averse to advantageous inequality, i.e., $\beta_i \geq 1 - R / n_N$ for $i = 1,..., n_N$. In all these equilibria, all players contribute the same amount. Thus, when $R < 2$, we can be sure that there is no equilibrium with positive cooperation levels in N, if there is at least one single player who is not sufficiently averse to advantageous inequality (i.e., $\beta_i < 1 - R / n_N$). The proof of these statements is given in Appendix B (Proposition 1). In our experimental parameterization the lower bound for β_i necessary to enable cooperation in equilibrium varies between 0.20 (for $n_N = 2$) and 0.87 (for $n_N = 12$) and thus becomes the more demanding, the larger the community is. Fehr and Schmidt propose an average $\beta = 0.315$ and according to empirical estimations of the inequality aversion parameters (e.g., Fehr and Schmidt 1999, p. 844; and Mariana Blanco et al. 2008), it is almost impossible to observe β-values that are sufficiently high to allow cooperation in larger communities. Thus, equilibria with positive contributions are highly unlikely in N.

In their analysis of institutions with punishment possibilities, Fehr and Schmidt (1999) assume, for the sake of simplicity, that all players are of one of two types: *enforcers* who conditionally cooperate and are ready to punish deviators or *payoff maximizers* with inequity parameters $\alpha = \beta = 0$. The adaptation of Fehr and Schmidt's model to the P community shows that equilibria with positive contributions exist if some players suffer sufficiently from disadvantageous inequality and credibly threaten to punish free-riders. In these equilibria, *all* players contribute an identical amount. We extend the analysis of Fehr and Schmidt and show that under reasonable assumptions equilibria with positive contributions can only exist if the number of payoff maximizers in P is low. To be precise, the number of payoff maximizers has to be strictly lower than the multiplicative inverse of the marginal punishment costs $1/c$. In our setting with $c = 1/3$ this implies that the P community can "afford" at most two payoff maximizers. If more than two payoff maximizers join P, no equilibrium with positive contribution levels exists and, interestingly, this is true independent of the community size n_P. This means that no matter how many members join P, more than two payoff maximizers destroy the possibility of an equilibrium with positive contributions. If there are no enforcers in P who threaten to punish the non-contributors, the situation is "equivalent" to the situation in N described above: equilibria with positive contributions only exist if all players are conditional cooperators with $\beta_i \geq 1 - R / n_P$, regardless how large the community size n_P is. A detailed analysis of this result is given in Appendix B, Proposition 2 and Corollary 1.

What does the Fehr-Schmidt model suggest for the dynamic case of voting-by-feet choice? Conditional cooperators who sufficiently dislike disadvantage inequality are likely to choose P because this gives them the possibility to punish free-riders. In P they might be able to sustain an equilibrium in which all players contribute. Hence the punishment possibility may well serve as a coordination device for conditional cooperators to gather in P. If conditional cooperators vote with their feet for P, cooperation payoffs would likely be higher in P than in N. This could attract other players to join P. However, with a growing community size, punishing a non-cooperator becomes less attractive, since although it would equalize payoffs towards the punished player, it would in fact increase the inequality towards the (many) players who contributed, but do not punish. For both communities, the chances of the existence of a cooperation-equilibrium tend to diminish the larger the community size is or becomes over time. Thus, even when one assumes that players are inequity averse it is highly unlikely that cooperation emerges in our population of 12 players.

3. Experimental Design and Procedure

Subjects were recruited for voluntary participation via the online recruitment system ORSEE (Ben Greiner 2004) and were randomly allocated to treatments. None of them had participated in a similar experiment before. On arrival, subjects were informed about the experimental procedure as well as the number of periods.[11] The experiment was programmed with *z-Tree* (Fischbacher 2007) and conducted in the computerized laboratory *eLab* at the University of Erfurt. Random reshuffling of the presentation order on the computer screens ensured that the identity of the players could not be traced over periods. In total 264 subjects participated, i.e., we collected 22 independent observations (8 in VF, 8 in GX, and 6 in FX) with 12 subjects each. An experimental session lasted for about 2 to 2.5 hours. Subjects' earnings were between 15 and 25 Euro.

[11] A translation of the instruction sheet is given in Appendix. Original instructions were written in German. They are available upon request from the authors.

4. Results

4.1. Results of the voting with feet experiment

As shown in Figure 1, over time an initially low acceptance rate of P steadily increases. While in the first period, on average roughly 4 out of 12 subjects choose P, in the last period this is the case for on average 11 subjects.[12] Figure 1 also shows that contributions in P of VF increase throughout the experiment. Compared with the first half, contributions in the second half increase significantly (p = 0.008). The most striking result, however, is that the P-communities establish and maintain almost perfect cooperation. Although the communities are quite large in the second half (about 9 to 12 players) and consequently the MPCR is relatively low, members of P contribute on average 19.5 tokens (while the average contribution in N is 3.4 tokens). This is especially noteworthy since previous studies (c.f. Isaac and Walker 1988) indicate that cooperation is much more difficult to establish if the community size is high and the MPCR is low.[13] Interestingly, in the last period 87 out of 88 subjects in P contribute the maximum amount of 20 tokens and the remaining one contributes 19 tokens. Hence, we do not observe an endgame effect in P of VF; in contrast the contributions even rise in the final period.

Community choice, contributions and punishment in the initial periods

The majority of subjects are initially reluctant to join P.[14] In the first period, significantly more subjects choose N (68.8%) than P (32.2%, binomial test, $p < 0.0001$). Yet, initial contributions in P (13.2 tokens) are significantly higher than the first period contributions in N (7.4 tokens) ($p < 0.0001$). Subjects with cooperative predispositions seem to self-select into P: In P 53.3% of the initial contributions are high contributions ($g \geq 15$) while only 20.0% of subjects contribute very low ($g \leq 5$). In N the distribution of contributions is quite different; while 45.5% contribute low only 15.2% contribute high.

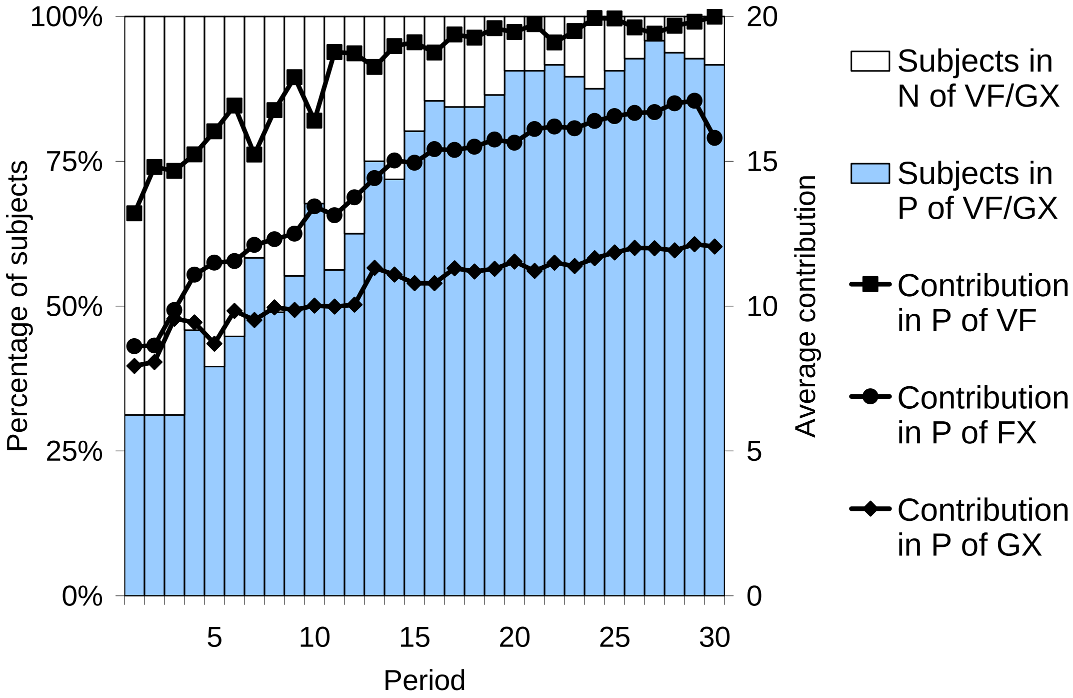
---

[12] In each of the 8 independent observations of VF, the group size of N, n_N, has a negative trend over time (determined by Spearman-rank-correlation coefficients). Hence a binomial test with these correlation coefficients clearly rejects that a negative trend is as likely as a positive trend (p = 0.016). All non-parametric statistical tests reported in this paper are two-tailed if not stated otherwise and take communities as units of independent observations. Comparisons within a treatment are tested with the Wilcoxon matched pairs tests. Comparisons across treatments are performed with the Whitney Mann U-tests.

[13] Jeffrey Carpenter (2007) observes that cooperation is not necessarily reduced with an increasing group size if mutual monitoring and punishment possibilities are available.

[14] This result is in line with Matthias Sutter et al. (2009), who let participants vote by unanimity rule whether they prefer a punishment institution, a reward institution, or an institution with no sanctions at all. For a detailed discussion, see section 5.

Figure 1: Community choice over periods and contributions in punishment communities



Subjects who initially choose P are not only highly cooperative but additionally ready to punish defectors. Already in the first period 87.5% of the high contributors (14 subjects) exert punishment tokens to discipline “less-contributors”, i.e., subjects who contributed strictly less than themselves. In the first five periods, subjects in P punish less-contributors in 80.6% of all possible cases. Subjects receive more punishment tokens the more they deviate from the community average.[15]

Behavior of incoming subjects to P

Subjects in the N community might be attracted by the high contributions in P. But at the same time they realize that the costs from punishing low-contributors may compromise the benefits from high cooperation. In fact, in the first half of the experiment the payoffs in P are slightly but not significantly lower than the payoffs in N (p = 0.148). The high initial contributions in P are “eaten up” by the high expenses for punishing those who contribute low.[16] Thus, contributing high and simultaneously free-riding on the punishment addressed to low contributors might seem appealing for incomers to P. It spares them from being punished

[15] All Spearman-rank-correlation-coefficients computed for each independent observation are positive. A binomial test with these correlation coefficients rejects that a negative correlation is as likely as a positive correlation (p = 0.016).

[16] This observation is in line with previous research (Martin Sefton et al. 2007, Fehr and Gächter 2002).

and "saves" the money for disciplining free-riders. Indeed, we observe that 76.3% of the subjects increase their contributions after switching from N to P.[17] 46.4% of the subjects switching from N to P are low contributors in N, but contribute $g \geq 15$ after the change to P. Moreover, 17.9% even "convert" from a complete free-rider to a full cooperator by increasing their contributions from 0 to 20 after the change to P. About two-thirds (65.6%) of incomers in P abstain from punishing in the very same period in which they enter P. Incoming subjects who contribute high but do not punish earn significantly more than they did in N before the change (average payoffs increase from 40.1 to 49.1, $p = 0.016$), while incoming subjects who contribute high and punish in P immediately after the change earn significantly less in the first period after the change in P than they did in N before the change (average payoffs decrease from 40.4 to 35.4, $p = 0.016$). There is no significant decrease in the percentage of punishers when the community size n_P increases.[18] Actually, for all community sizes n_P at least 50% of subjects are punishers.[19]

Payoffs

Punishment leads to increased contributions and higher "long-term" payoffs in P. In the second half of the periods the average payoff in P steadily increases towards the social optimum. The payoffs in the second half are significantly higher than in the first half ($p = 0.008$). The payoffs in N, however, constantly remain low and approach the Nash-equilibrium prediction in the end. In the second half of the experiment, payoffs in P (48.6 tokens) are significantly higher ($p = 0.016$) than payoffs in N (41.3 tokens).

Taking the perspective of a social planer one might be interested in the overall efficiency of the VF mechanism and ask for a joint evaluation of both communities. The overall efficiency (i.e., the sum of the players' gains from contributions reduced by their punishment expenses)

[17] We calculate the difference between the subject's contribution in period *t* and the contribution of that subject in period *t* -1 conditional upon the subject changed the community from period *t* -1 to period *t*.

[18] The burden of punishing is shared: 92 out of 95 subjects who inhabit P face at least one time a less-contributor in their community. In such a situation, 81.5% of them punish a less-contributor at least once, 69.6% punish more than once in such a situation. More than half of the subjects (55.4%) even punish 5 or more times.

[19] Incomers to N typically (66.8% of subjects) decrease their contributions after the change from P to N. 15.3% switch even from full cooperation to complete free-riding. This observation is well in line with the experimental findings by Falk et al. (2009). They show that the contribution behavior of subjects is influenced by social interactions with their "neighbors". Subjects who simultaneously belong to two different groups with disjoint group compositions exhibit conditionally cooperative behavior, i.e., the same subject contributes more if she is in a community with high contributors, but contributes less if she is in a community with low contributors. Note, however, that in our experiment subjects know that the members of one "neighborhood" can punish while those of the other cannot.

of the VF mechanism is low initially (74% of the possible maximum in periods 1-15), but steadily increases and reaches 94% in periods 16-30. These figures are remarkably high and can hardly be found in other public goods experiments with or without sanctioning possibilities.

What makes the VF mechanism so successful? In the light of our results, two explanations are at hand. The first is that self-selection into communities is key for success. It allows the subjects to join in groups of “like-minded” people and to establish a cooperative “culture” in P. The other non-exclusive explanation is the growth process of the P community, i.e., the mere fact that the community starts off with a small group of subjects and then grows to its maximum size. To investigate and to disentangle these two possible explanations we ran two control experiments, which will be described and discussed in the two following sections.

4.2. The effect of self-selection

We designed a control experiment with identical growth paths as in our main experiment VF, but without the possibility of self-selection. Instead, we exogenously allocated subjects to the communities. We refer to this experiment as the exogenous growth experiment (abbreviated GX). In this experiment, there was no community choice stage. Instead, all community choice vectors of subjects who participated in one session of the VF-experiment were randomly assigned to participants in one session of the GX-experiment. Technically, each session in GX exactly re-ran the community choices of the “mirror”-session in VF. We conducted 8 additional sessions with 12 new subjects each. Subjects were told that community affiliations may vary from period to period and will be announced privately at the beginning of a period.

In VF initial contributions are higher and punishment is more severe

Initial contributions in P of VF (13.2 tokens) are significantly higher ($p = 0.003$) than initial contributions in P of GX (8.7 tokens) (cf. Figure 1). In contrast to VF, the first period contributions in P and N are quite similar in GX. In both N and P of GX, only 16.7% contribute high while 43.3% contribute low in P and 59.1% in N. Compared to VF, the initial punishment behavior also differs in GX. In the first period, we observe only one subject who contributes high and punishes a less-contributor. This is significantly less than in VF where 14 subjects are high contributors and punish less contributors ($p = 0.004$). In the first five periods in P of GX subjects punish others who contributed less in 44.1% of the cases (P of VF: 80.6%, $p = 0.005$). Hence, the disciplining subjects in P of VF are not only more numerous than in P of GX, but they show also less mercy against defectors. These findings support the positive influence of the self-selection process on cooperation in the beginning phase of VF.

The impact of endogenous choice on incoming subjects to P in later periods

In GX 70.0% of subjects increase their contribution after switching from N to P (in VF: 76.3%). The incoming subjects in P of GX, however, increase their contributions less than subjects in the same situation in VF do (average increase in P of GX: 5.6 tokens, in P of VF: 10.3 tokens, p = 0.028). Only 6.7% of the subjects in P of VF who are punished in period t-1 but choose to remain in P in period t, decrease their contributions. These are significantly fewer (p = 0.062) than in GX (15.6%) where the decision to remain in P is predetermined. If the subjects, who have been punished in period t-1 and stay in P, increase their contributions in period t, then this increase is significantly lower (p = 0.065) in P of GX (on average 3.6 tokens) than the increase (5.3 tokens) in P of VF. On average, a punisher in P of GX contributes significantly less (p = 0.005) than a punisher in P of VF (11.1 versus 17.7 tokens). As a consequence, in the second half of the experiment, payoffs in P of GX are significantly lower than payoffs in P of VF (p = 0.099).

Figure 2: Average contributions in P of VF and P of GX

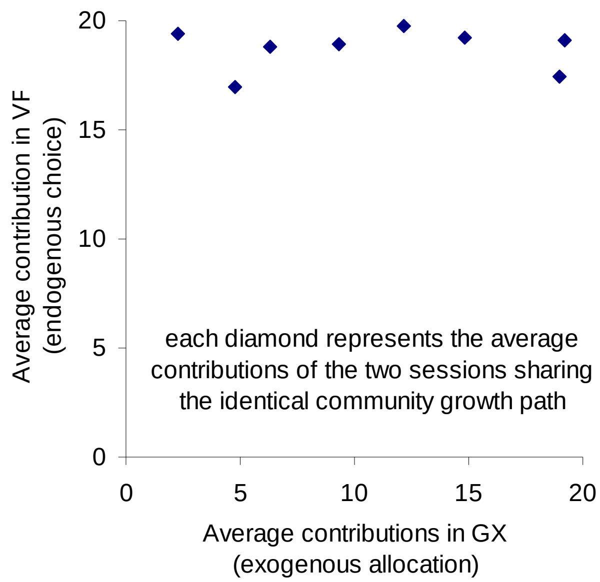


Figure 2 relates the average contribution of each P-community in VF to the average contribution of its “mirror” community in GX with identical (but exogenously imposed) growth-path. While the variability in contributions is low in P of VF (variance: 0.83) there is a huge difference in the variability of average contributions in P of GX (variance: 35.9). Overall, subjects in P of VF contribute significantly more than subjects in P of GX (18.7 tokens versus 11.0 tokens; $p = 0.016$). There is no significant difference between the overall contributions in N of VF (3.6 tokens) and in N of GX (3.9 tokens). Hence, the self selection of subjects into the communities seems to be not only initially relevant, but a crucial determinant of the success of the VF mechanism throughout the interaction.

4.3. The effect of community growth

Another characteristic of VF is the steady growth of the P communities, from a minority to the entire population. To study the pure effect of this community growth on cooperation we conducted a second control experiment with fixed sized communities (abbreviated FX). In this control experiment, we also involved 12 (new) subjects and again there was no community choice stage. Instead, in each of the 6 sessions we created two equally sized communities (P and N) consisting of 6 members each. The community compositions were fixed for the whole duration of the sessions. Comparing FX to GX provides us with insights about the impact of the community growth on cooperation.

First period contributions between P of GX and P of FX are not statistically different ($p = 0.564$). The same is true for N of GX and N of FX ($p = 0.169$). Also overall contributions between P and N-communities of GX and of FX are not significantly different ($p = 0.414$ and $p = 0.513$, respectively). Within each experiment, contributions in P are significantly higher than in N (VF: $p = 0.008$, GX: $p = 0.008$, and FX: $p = 0.031$, respectively).

With respect to punishing of less-contributors there is no significant difference between P of GX and P of FX ($p = 0.470$) in the first five periods. Overall, there is no difference in contribution behavior of punishers between GX and FX ($p = 0.491$). In the first half, there is no difference between payoffs in P of GX and payoffs in P of FX ($p = 1.000$). In the second half, however, payoffs in P of FX are higher than payoffs in P of GX ($p = 0.087$). The result that we do not observe higher cooperation in P of GX than in P of FX suggests that the effect of a P community that starts small and increases steadily is not substantial.[20] The main effect of the superior performance of the VF mechanism seems to be due to self selection of subjects into different communities rather than their growth.

[20] It might well be that a possible advantage of growing slowly in GX might be canceled out by a possible advantage of having a fixed group composition in FX (which arguably makes it easier to sustain cooperation). We do not have indications pointing into this direction.

4.4. Societies performances – from a social planner's point of view

Taking again the perspective of a social planner, we consider the efficiency of the entire mechanism with both communities N and P. Over 30 periods and both communities, the contributions in VF are significantly higher than in GX ($p = 0.038$) as well as in FX ($p = 0.029$). The contributions in GX and FX are not significantly different ($p = 0.662$). Hence, also from a social planner's perspective VF outperforms GX and FX. The payoffs in the second half of the periods are significantly higher in VF than the payoffs in GX ($p = 0.098$) and almost significantly higher than the payoffs in FX ($p = 0.108$).[21] Hence, in terms of efficiency, VF exhibits the best long-run performance. Although we do not have hard evidence the trend in payoffs suggests that the payoff differences between VF and the other two experiments would have been even more pronounced had the experiment been continued (cf. Gächter et al. 2008).

5. Related Literature

Recently a modest experimental literature on endogenous choice in social dilemma situations has emerged. One line of research investigates the endogenous choice of interaction partners in public goods settings[22], and another strand focuses on institution choice through voting.

To the best of our knowledge Karl-Martin Ehrhart and Keser (1999) are the first to use endogenous regrouping in a public goods experiment. They allow subjects to move freely from one group to another. In each group and period a simple public goods game (without punishment) is played. The MPCR is decreasing in the group size but a contribution to the public good in a larger group yields a higher group return than in a smaller group. They find that high contributors are chased by free riders. This results in an unstable sorting of high and low contributors and over time in a declining trend of contributions. Giorgio Coricelli et al. (2004) let subjects bid for the right to choose partners. As in Ehrhart and Keser (1999), in their unidirectional choice treatment free riders show a tendency to chase high contributors which leads to higher contributions than with random re-matching. Talbot Page et al. (2005) regroup subjects after each third period according to their expressed preferences. This

---

[21] The comparison between VF and FX might be considered not to be completely fair since in our setting the MPCR decreases with the community size which makes cooperation easier in smaller communities than in larger ones. While the majority of players join P in the last periods of VF, the community sizes in FX are fixed to 6.

[22] Partner selection and its effects on behavior are also explored in market interactions (Georg Kirchsteiger et al. 2005, and Martin Brown et al. 2004), and networks (Arno M. Riedl and Aljaz Ule 2003). Esther Hauk and Rosemarie Nagel (2001) study the pure effect of unilateral and mutual choice of partners in finitely repeated prisoners' dilemma games.

regrouping procedure increased contributions, both in a simple public goods setting and also in a public goods setting with punishment. In Matthias Cinyabuguma et al. (2005), subjects vote on irreversibly excluding others. Subjects who are not expelled contribute highly. Over time more and more subjects are banished which leads to a decrease in social efficiency. In a study by Gary Charness and Chun Lei-Yang (2008), subjects can decide whether they want to exit a group, to exclude other players by majority vote, or to merge with other groups (if 60% of the merging groups' members accept) before they play a public goods game. The greater the group size, the higher is the marginal social value, i.e., greater groups are more productive. Opportunities of regrouping together with the increase in the marginal social value raise contributions substantially. The increase in contributions, however, is smaller when the marginal social value is capped at a certain group size. Toh-Kyeong Ahn et al. (2008) also investigate the group formation under different rules: free entry and exit into a group, restricted (free) entry into and free (restricted) exit from a group. In the restricted entry and exit treatments, a majority vote decides on the entry and exit wish. Ahn et al. (2008) do not find significant differences between overall contributions. The different treatments, however, lead to groups of different sizes. Restricted entry leads to mid-sized and more effective groups while restricted exit and free entry/exit treatments lead to rather large and uncooperative groups. In all these studies subjects include or exclude interaction partners in static institutional settings. In contrast, in our voting with feet mechanism subjects do not actively choose their interaction partners, but opt for different institutions, that allow or do not allow for punishment. This at best allows for an indirect choice of interaction partners since choosing an institution means to interact with "likeminded" individuals in the sense that all have chosen the same institution. Additionally, in contrast to our study, in the mentioned studies productivities of groups increase with their size which is likely to make cooperation easier in larger groups.

Few studies focus on situations in which institution choice is the result of a voting process. Thorsten Decker et al. (2003) let subjects vote for different punishment mechanisms. After observing others' contributions, subjects individually submit their intended punishment level for each of the other three group members. Whether the highest, the medium, or the lowest punishment level should be applied is auctioned off in a first-price auction. Subjects prefer the rule implementing the lowest of the proposed punishment levels which can be interpreted as unanimity vote on punishment, i.e., all agree that the respective player should be punished as least as much as the lowest suggested level. Arhan Ertan et al. (2009) allow subjects to interact in non-punishment and punishment institutions before they decide by majority vote whether punishment shall be available in a future interaction and if yes to whom it can be directed. Many groups opt to restrict punishment against high contributors but allow punishment of low contributors. These groups achieve higher levels of cooperation than groups with unrestricted or no punishment. Stephan Kroll et al. (2007) let subjects vote on a non-binding minimum contribution. This mechanism alone, however, turns out not to be effective. Only if contributions below the non-binding minimum contribution can be punished or the minimum contribution becomes binding, contributions go up. In all three studies, players are not allowed repeatedly to choose to play the public goods game under different institutions so that self-selection into different institutions cannot occur.

Three studies on the endogenous formation of institutions are most related to ours. Michael Kosfeld et al. (2009) let players decide whether they are willing to participate in a sanctioning organization. All participants who declare their willingness to join the organization have to vote by unanimity rule whether the organization is actually implemented or not. If the organization is implemented, members are sanctioned for not contributing their full endowment. Outsiders of the organization are not sanctioned but nevertheless benefit from the contributions of all players. The data shows that a large majority of groups implement an organization by the final periods, despite the fact that it is costly. About 75% of these organizations even involve *all* players. A comparison with control treatments reveals that the opportunity to form organizations enhances group welfare by higher and stabilized contributions. Our approach is different in several aspects. First, we implement a setting in which punishment is decentralized. Therefore, our results do not rely on the creation of a global organization. Second, our public goods are local in the sense that players benefit from each others' contributions only within the same institution. Players joining different institutions play different public goods games.

Sutter et al. (2009) systematically investigate the effects of institution choice by voting between standard public goods, public goods with rewards, or public goods with punishment. Voting is costly, but not mandatory. Those who choose to vote repeat voting until they reach an unanimous decision. The vote determines the institution, under which *all* individuals subsequently have to interact for 10 periods, i.e., play includes also those players who decided not to vote. In different treatments the leverages, i.e., effectiveness, for punishment and rewards are varied. For comparison, "exogenous treatments" are also conducted, in which one of the institutions is exogenously imposed by the experimenter. Sutter et al. (2009) find that the reward institution is chosen almost exclusively, particularly when rewards and punishment have a high leverage. The punishment institution is rarely chosen and only when the leverage is low. When it is selected, however, it is the most successful institution in eliciting high contributions. This is true in comparison with all other endogenously selected institutions, but also in comparison with all exogenously imposed institutions. The study by Sutter et al. (2009) analyzes choices between a larger variety of institutions, while we are more interested in the dynamic aspects of a voting with feet mechanism. Thus, we focus on two institutions, a standard public goods setting and one with a high leverage punishment. Sutter et al. (2009) and Kosfeld et al. (2009) also differ in other respects from our approach. First, in their studies players cannot escape from each other, i.e., they are exogenously forced together to play the public goods game. We allow players to choose institutions every period which enables players to enter and exit. This leads to a dynamic evolution of the group composition. Thereby, we model an environment in which players have the freedom to escape, for example, if they are not satisfied with the institution or the treatment they receive from their respective partners. Secondly, we allow for a larger population which is three times as large as in the two other studies. Thus, in this respect we address a situation in which cooperation is particularly hard to achieve. Additionally, communities under each institution can grow endogenously and have a varying size and composition which is likely to make coordination on cooperative behavior even more difficult.

Gürerk et al. (2006) investigate endogenous choice of institutions in a similar experimental setting, but with a different pair of communities than in this study. In Gürerk et al. (2006), participants had the choice between a community with punishment *and* reward possibilities, and a community with no sanction possibilities at all (as the N community in this study). The results show some qualitative similarities to this study, with respect to the dynamics of endogenous choice and the final cooperation level reached. The present study differs in various aspects. First, in Gürerk et al. (2006) it is not clear whether the observed cooperation is due to the interplay of punishment and reward or whether punishment alone also leads to high cooperation and efficiency levels. Second, the current study is designed to investigate reasons of the superior performance of endogenous choice mechanisms. The presented two control experiments allow us to investigate separately the effect of the endogenous choice from the effect of slow community growth.

6. Summary and conclusion

In this study, we investigate community choice in social dilemma situations. In our voting with feet setting (VF) – where subjects are given the choice between a non-punishment community (N) and a community with punishment possibilities (P) – most subjects initially choose N, but over time P becomes increasingly populated. In advanced periods, almost all subjects join P. Despite severe punishment in the beginning, P leads to high efficiency levels with full contributions of all participants and no punishment in later periods. In a control experiment (GX), we show that communities with exogenous subject allocation but identical growth paths as in VF perform significantly less successful. With the help of a second control experiment (FX) we find that growing communities per se are not more successful than fixed-size ones, if subject allocation is exogenous. The synopsis of all three experiments suggests that the endogenous self-selection of subjects is an important key for the establishment and efficient maintenance of cooperation. In the beginning the punishment community P attracts the "right" subjects, willing to contribute high and ready to punish defectors. Although entry and exit is not restricted, they manage to establish high cooperation levels which sustain, even when the group grows large and the entire subject population ultimately joins P.

Particularly the observed behavior in advanced periods cannot easily be expected from our theoretical analysis. With standard selfish and myopic preferences there should be no punishment and thus it is dominant to free-ride in both communities N and P. Also if one assumes that players are inequity-averse, cooperation is very unlikely. As we have shown, cooperation in N is only possible if all community members are conditional cooperators. To sustain cooperation in P, only very few payoff-maximizers can be coped with, even if all the others are ready to enforce cooperation by exerting costly punishment. In the experiment, the initial periods of the P community come relatively close to this situation. Only a small number of players self-select into P and very few of them contribute low. Those are heavily punished by high contributors. This seems to create and stabilize a "cooperative culture" in P that is also adopted by subjects switching to P later on, even if they have contributed almost nothing in N before. Although this comes close to what our adaption of the Fehr and Schmidt (1999) model suggests for P, there is a remarkable difference in advanced periods: We observe far too many enforcers than could be expected from existing estimations of the inequity parameters (e.g. Blanco et al. 2008, Fehr and Schmidt 1999). These subjects, however, do not punish consistently all the time. It seems that they occasionally rely on others taking the burden of punishment. Nevertheless, free-riders have lower payoffs than both cooperators and enforcers.

Our findings emphasize and extend previous findings in several dimensions. They confirm the superior performance of punishment mechanisms by demonstrating that – also when subjects are free to enter and exit – efficiency of the punishment community in the advanced periods is significantly higher than without punishment. Efficiency in P under the voting with feet mechanism also turns out to be higher than when subjects are exogenously allocated to a P institution. It seems that if people vote with their feet to join an institution, they develop a certain degree of identification and commitment to the chosen community which in our setting induces newcomers in the P community – who have previously contributed nothing in the N community – not only to immediately contribute high but also to incur costs for punishing low contributors. This adoption of the established social norms of the chosen community helps to foster cooperation.[23] Thus, our findings highlight a so far undervalued feature of the voting with feet mechanism: In addition to the efficiency improvement from implementing the right match between consumers and public goods by the "consumer-voter […] picking that community which best satisfies his preference pattern for public goods" as suggested by Tiebout (1956), the adoption and thus the coordination on the existing norms in a new community improves efficiency even further if the community is joined voluntarily. This observation adds one essential piece to the puzzle why people often show such a surprisingly high willingness to cooperate in communities they have voluntarily joined, like companies, parties or jurisdictions.

In our setting the P institution seems to serve at least two purposes. First it constitutes a coordination device to attract the more cooperative subjects. Since in the beginning the cooperative subjects predominantly self-select into P they are able to initiate a cooperative culture. Second, it provides the tools necessary to sustain the cooperative culture by disciplining low contributors. It is an open question whether both features are essential for the success of the voting-with feet mechanism. As shown by Ehrhart and Keser (1999) having the choice between different communities appears not to be sufficient, if the tools to discipline free-riders are not available. They allow subjects to choose between communities with identical rule systems (N). Thus, in the first place it is difficult for cooperators to coordinate

[23] Sutter et al. (2009) report a similar observation. Subjects who choose to participate in the unanimity vote on the institutions (and are willing to exert some costs to do so) contribute more than those who abstain from voting. For a systematic overview of how participation in an institution may shape preferences, for example, by "conformist transmission", see Bowles (1998).

on joining a certain community since the communities are not distinguishable. But our guess would be that even if one facilitates coordination, for example, by giving different names to different communities voting with feet would not be successful. The reason is that in addition to gathering in one community cooperators need a tool to discipline free-riders to successfully establish a cooperative culture.

Our findings also raise questions for future research that are beyond the scope of this paper. Why are subjects initially reluctant to join P although it becomes so successful over time? Maybe some are hoping to find sufficiently many cooperative fellows in N so that cooperation would be possible without having to rely on punishment mechanisms? Or, do they simply fear the punishment? Switching to a different community is often not that easy. Thus, it would be informative to know how sensitive our findings are regarding the absence/existence of switching costs for joining a different community? What happens if individuals do not choose between exogenously given institutions, but instead are given the possibility to create their own alternative institutions? For feasibility reasons we concentrate on just two institutions. Introducing a larger variety of institutions would certainly come closer to Tiebout's vision. How would a larger variety of institutions affect cooperation success of the voting with feet mechanism? Providing answers to these questions would be promising to help to design voting with feet institutions that further reduce the downsides inherent to social dilemmas.

References

Abbink, Klaus, Bernd Irlenbusch, and Elke Renner. 2000. “The Moonlighting Game – An Experimental Study of Reciprocity and Retribution.” *Journal of Economic Behavior and Organization*, 42: 265-277.

Ahn, Toh-Kyeong, Mark Isaac, and Timothy C. Salmon. 2008. “Endogenous Group Formation.” *Journal of Public Economic Theory* 10(2): 171-194.

Blanco, Mariana, Dirk Engelmann, and Hans-Theo Normann. 2008. “A Within-Subject Analysis of Other-Regarding Preferences. Working Paper. Royal Holloway, University of London.

Bowles, Samuel. 1998. “Endogenous Preferences: The Cultural Consequences of Markets and other Economic Institutions.” *Journal of Economic Literature*, 26: 75-111.

Bowles, Samuel. 2004. *Microeconomics: Behavior, Institutions, and Evolution*. Princeton University Press.

Brown, Martin, Armin Falk, and Ernst Fehr. 2004. “Relational Contracts and the Nature of Market Interactions.” *Econometrica*, 72: 747-780.

Carpenter, Jeffrey. 2007. “Punishing Free-Riders: How Group Size Affects Mutual Monitoring and the Provision of Public Goods.” *Games and Economic Behavior*, 60(1): 31-52.

Charness, Gary, and Chun-Lei Yang. 2008. “Endogenous Group Formation and Public Goods Provision: Exclusion, Exit, Mergers, and Redemption”, Working Paper UCSB.

Cinyabuguma, Matthias, Talbot Page, and Louis Putterman. 2005. “Cooperation under the Threat of Expulsion in a Public Goods Experiment.” *Journal of Public Economics*, 89: 1421-1435.

Coricelli, Giorgio, Dietmar Fehr, and Gerlinde Fellner. 2004. “Partner Selection in Public Goods Experiments.” *Journal of Conflict Resolution*, 48(3): 356-378.

Croson, Rachel T.A. 1996. “Partners and strangers revisited.” *Economics Letters*, 53: 25-32.

Croson, Rachel T.A. 1998. “Theories of Altruism and Reciprocity: Evidence from Linear Public Goods Games.” Working Paper *98-11-04,* The Wharton School of the University of Pennsylvania.

Davis, Douglas D., and Charles A. Holt. 1993. *Experimental Economics*. Princeton

University Press.

Dawes, Robyn M. 1980. “Social Dilemmas.” *Annual Review of Psychology*, 5: 163-193.

Decker, Thorsten, Andreas Stiehler, and Martin Strobel. 2003. “A Comparison of Punishment Rules in Repeated Public Good Games.” *Journal of Conflict Resolution*, 47(6): 751-772.

Ehrhart, Karl-Martin, and Claudia Keser. 1999. “Mobility and Cooperation: On the Run.” Working paper 99s-24.CIRANO, Montreal.

Ertan, Arhan, Talbot Page, and Louis Putterman. 2009. “Who to punish? Individual Decisions and Majority Rule in Mitigating the Free Rider Problem.” *European Economic Review*, 53: 495-511.

Falk, Armin, Simon Gächter, and Urs Fischbacher. 2009. “Living in Two Neighborhoods - Social Interaction Effects in the Lab.” CEDEX Discussion Paper, 2009-01.

Fehr, Ernst, and Klaus Schmidt. 1999. “A theory of fairness, competition, and cooperation.” *Quarterly Journal of Economics*, 114(3): 817-868.

Fehr, Ernst, and Simon Gächter. 2000. “Cooperation and Punishment in Public Goods Experiments.” *American Economic Review*, 90(4): 980-994.

Fehr, Ernst, and Simon Gächter. 2002 “Altruistic Punishment in Humans.” *Nature*, 415, 137-140.

Fischbacher, Urs, and Simon Gächter. “Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Good Experiments.” Forthcoming in: *American Economic Review*.

Fischbacher, Urs. 2007. “z-Tree: Zurich Toolbox for Ready-made Economic experiments.” *Experimental Economics*, 10(2): 171-178.

Gächter, Simon, and Christian Thöni. 2005. “Social Learning and Voluntary Cooperation among Like-Minded People”, *Journal of the European Economic Association*, 3: 303-314.

Gächter, Simon, Elke Renner, and Martin Sefton. 2008. “The Long-Run Benefits of Punishment” *Science*, 322: 1510.

Greiner, Ben. 2004. “An Online Recruiting System for Economic Experiments.” In: Kurt Kremer, Volker Macho (eds.). Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63, Göttingen: Ges. Für Wiss. Datenverarbeitung: 79-93.

Gürerk, Özgür, Bernd Irlenbusch, and Bettina Rockenbach. 2006. “The Competitive

Advantage of Sanctioning Institutions." *Science*, 312(5770): 108-111.

Hardin, Garrett. 1968. "The Tragedy of the Commons." *Science*, 162: 1243-1248.

Hauk, Esther, and Rosemarie Nagel. 2001. "Choice of Partners in Multiple Two-Person Prisoner's Dilemma Games." *Journal of Conflict Resolution*, 45(6): 770-793.

Isaac, Mark, and James Walker. 1988. "Group Size Hypotheses of Public Goods Provision: An Experimental Examination." *Quarterly Journal of Economics*, 103: 179-199.

Keser, Claudia, and Frans van Winden. 2000. "Conditional Cooperation and Voluntary Contributions to Public Goods." *Scandinavian Journal of Economics*, 102(1): 23-39.

Kirchsteiger, Georg, Muriel Niederle, and Jan Potters. 2005. "Endogenizing Market Institutions: An Experimental Approach." *European Economic Review*, 49(7): 1827-1852.

Kosfeld, Michael, Akira Okada, and Arno M. Riedl. 2009. "Institution Formation in Public Goods Games." *American Economic Review*, 99(4): 1335-1355.

Kroll, Stephan, Todd L. Cherry, and Jason F. Shogren. 2007. "Voting, Punishment, and Public Goods." *Economic Inquiry*, 45(3): 557-570.

Ledyard, John. 1995. "Public Goods: A Survey of Experimental Research." J. Kagel and A. Roth, *Handbook of Experimental Economics*, Princeton University Press, 111-194.

Ostrom, Elinor. 1998. "A Behavioral Approach to the Rational Choice Theory of Collective Action." *American Political Science Review*, 92(1): 1-23.

Ostrom, Elinor. 1999. "Coping With the Tragedy of the Commons." *Annual Review of Political Science*, 2: 493-535.

Ostrom, Elinor, James Walker, and Roy Gardner. 1992. "Covenants with and without a Sword: Self-Governance is Possible." *American Political Science Review*, 86: 404-417.

Page, Talbot, Louis Putterman, and Bulent Unel. 2005. "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry, and Efficiency." *Economic Journal*, 115: 1032–1053.

Riedl, Arno M., and Aljaz Ule. 2003. "Exclusion and Cooperation in Social Network Experiments." Working Paper, University of Amsterdam.

Schein, Edgar H. 2004. *Organizational Culture and Leadership*. Jossey-Bass, San Francisco.

Sefton, Martin, Robert Shupp, and James Walker. 2007. "The Effect of Rewards and

Sanctions in the Provision of Public Goods.” *Economic Inquiry*, 45: 671-690.

Sutter, Matthias, Stefan D. Haigner, and Martin G. Kocher. 2009. “Choosing the Carrot or the Stick? – Endogenous Institutional Choice in Social Dilemma Situations.” University of Innsbruck, Working Paper.

Tiebout, Charles M. 1956. “A Pure Theory of Local Expenditures.” *Journal of Political Economy*, 64(5): 416-424.

Vise, David A. 2006. *The Google Story – Inside the Hottest Business, Media and Technology Success of Our Times*. Pan Macmillan, London.

Weber, Roberto A. 2006. “Managing growth to achieve efficient coordination in large groups.” *American Economic Review*, 96(1): 114-126.

Appendix A: Instructions for the experiment (Treatment VF)

General Information: At the beginning of the experiment, you will be randomly assigned to one of **2 subpopulations each consisting of 12 participants**. During the whole experiment, you will interact only with the members of your subpopulation. At the beginning of the experiment, **1000 experimental tokens** will be assigned to the experimental account of each participant.

Course of Action: The experiment consists of **30 rounds**. Each round consists of 2 stages. In Stage 1, the group choice and the decision regarding the contribution to the project take place. In Stage 2, participants may influence the earnings of the other group members.

Stage 1

(i) The Group Choice: In Stage 1, each participant decides which group she wants to join. There are two different groups that can be joined:

	Influence on the earnings of other group members
Group	A: No
	B: Yes, by assigning negative points

(ii) Contributing to the Project: In stage 1 of each round, each group member is endowed with 20 tokens. You have to decide how many of the 20 tokens you are going to contribute to the project. The remaining tokens will be kept by you.

Calculation of your payoff in stage 1: Your payoff in stage 1 consists of two components:

- **tokens you have kept** = endowment – your contribution to the project
- **earnings from the project** = 1.6 x sum of the contributions of all group members / number of group members

Thus, **your payoff in Stage 1** amounts to:

20 – your contribution to the project

+ 1.6x sum of the contributions of all group members / number of group members

The earnings from the project are calculated according to this formula for each group member. **Please note:** Each group member receives the same earnings from the project, i.e., each group member benefits from **all** contributions to the project.

Stage 2

Assignment of Tokens: In stage 2 it will be displayed how much each group member contributed to the project. (**Please note: Before each round a display order will randomly be determined.** Thus, it is not possible to identify any group member by her position on the displayed list throughout different rounds.) By the assignment of tokens you can reduce the payoff of a group member or keep it unchanged.

In each round each participant receives additional 20 tokens in stage 2. You have to decide how many from the 20 tokens you are going to assign to other group members. The remaining tokens are kept by yourself. You can check the costs of your token assignment by pressing the button *Calculation of Tokens*.

- Each **negative token** you assign to a group member **reduces her payoff by 3 tokens**.
- If you assign **0 tokens** to a group member her **payoff won't change**.

Calculation of your payoff in stage 2: Your payoff in stage 2 consists of two components:

- **tokens you kept** = 20 – sum of the tokens that you have assigned to the other group members
- **less the threefold number of negative tokens** you have received from other group members

Thus, **your payoff in Stage 2** amounts to: 20 – sum of the tokens that you assigned to other group members – 3x (the number of tokens you received from other group members)

Calculation of your round payoff: Your round payoff is composed of

	Your payoff from Stage 1	20 – your contribution to the project + 1.6 x sum of the contributions of all group members / number of group members
+	Your payoff from Stage 2	20 – sum of the tokens that you have assigned to other group members – 3 x (the number of tokens you have received from other group members
=	Your round payoff	

Special case: a single group member: If it happens that you are the only member in your group you receive 20 tokens in Stage 1 and 20 tokens in Stage 2, i.e., your round payoff amounts to 40. You neither have to take any action on Stage 1 nor on Stage 2.

Information at the end of the round: At the end of the round you receive a detailed overview of the results obtained in all groups. For every group member you are informed about her: Contribution to the project, payoff from the Stage 1, assigned tokens (if possible), received tokens (if possible), payoff from Stage 2, round payoff.

History: Starting from the 2nd round, in the beginning of a new round you receive an overview of the average results (as above) of all previous rounds.

Total Payoff: The total payoff from the experiment is composed of the starting capital of 1000 tokens plus the sum of round payoffs from all 30 rounds. At the end of the experiment, your total payoff will be converted into Euro with an exchange rate of 1 € per 100 tokens.

Please notice: Communication is not allowed during the whole experiment. If you have a question please raise your hand out of the cabin. All decisions are made anonymously, i.e., no other participant is informed about the identity of someone who made a certain decision. The payment is anonymous too, i.e., no participant learns what the payoff of another participant is.

We wish you success!

Appendix B: Theoretical Predictions (Intended for online publication only)

Fehr and Schmidt (1999) apply their model of inequity aversion to a public goods setting with voluntary contributions. They study the case in which no punishment is possible as well as the case with a punishment possibility. We adapt the results of Fehr and Schmidt (1999) to our N and P communities where the MPCR depends on the community size and derive special results for our setting.

Adaption of the Fehr and Schmidt (1999) results to the N community

Fehr and Schmidt (1999) show that in a public goods setting with voluntary contributions and no punishment possibility contributing zero is a dominant strategy for players with $a_N + \beta_i < 1$ (cf. the proof to the Proposition 4, part (a) in the Appendix of Fehr and Schmidt (1999), p. 860). Since in our model the MPCR varies with the community size while the productivity of the public good is constant, i.e., $R = n_N a_N$, we replace a_N by R / n_N. Thus, for players with $R / n_N + \beta_i < 1$ it is a dominant strategy to contribute nothing. In the following, we show that there is an equilibrium in N in which all players contribute zero and that equilibria with positive contributions exist only if *all* players are "conditional cooperators", i.e., their preferences satisfy $R / n_N + \beta_i \geq 1$.

Proposition 1. Assume that $R < 2$.

I. If for at least one player $R / n_N + \beta_i < 1$ is satisfied, the unique equilibrium in N prescribes free-riding of all players.

II. Equilibria with strictly positive contributions exist if and only if *all* players i satisfy $R / n_N + \beta_i \geq 1$. In all these equilibria, all players contribute the same amount.

Proof.

I. In analogy to Fehr and Schmidt (1999) assume that there are n_N players in N with contributions $g_1 \leq g_2 \leq ... \leq g_{n_N}$ and k>0 of these n_N players have $R / n_N + \beta_i < 1$. As mentioned above, for these k players it is a dominant strategy to contribute zero: $g_1 = \ldots = g_k = 0$. Suppose there exists a player $l > k$ who contributes the smallest positive amount $0 = g_{l-1} < g_l \leq g_{l+1} \leq ... \leq g_{n_N}$. Player l's utility is given by

(A1) $$U_l(g_l) = y - g_l + \frac{R}{n_N} g_l + \frac{R}{n_N}\sum_{j\neq l+1}^{n_N} g_j - \frac{\beta_l}{n_N - 1}\sum_{j\neq l+1}^{n_N}(g_j - g_l) - \frac{\alpha_l}{n_N - 1}\sum_{j\neq 1}^{l-1} g_l\ .$$

Rearranging the terms we obtain:

(A2) $$U_l(g_l) = y + \frac{R}{n_N}\sum_{j\neq l+1}^{n_N} g_j - \frac{\beta_l}{n_N - 1}\sum_{j\neq l+1}^{n_N} g_j - (1 - \frac{R}{n_N}) g_l + \beta_l \frac{n_N - l}{n_N - 1} g_l - \alpha_l \frac{l-1}{n_N - 1} g_l$$

The first three terms on the right-hand side of equation (A2) are equivalent to player *l*'s utility if she would deviate and contribute zero while the last three terms summarize the utility loss through contributing $g_l > 0$. Thus (A2) can be rewritten as:

(A3) $$U_l(g_l) = U_l(g_l = 0) - (1 - \frac{R}{n_N}) g_l + \beta_l \frac{n_N - l}{n_N - 1} g_l - \alpha_l \frac{l-1}{n_N - 1} g_l$$

Since $\alpha_l \geq \beta_l$, $l \geq k+1$, and $\beta_l < 1$, an upper bound for *l*'s utility from contributing is:

(A4) $$U_l(g_l) \leq U_l(g_l = 0) - (1 - \frac{R}{n_N}) g_l + \beta_l \frac{n_N - l}{n_N - 1} g_l - \beta_l \frac{l-1}{n_N - 1} g_l$$

$$\leq U_l(g_l = 0) - (1 - \frac{R}{n_N}) g_l + \beta_l \frac{n_N - 2(k+1) + 1}{n_N - 1} g_l$$

$$< U_l(g_l = 0) - (1 - \frac{R}{n_N}) g_l + \frac{n_N - 2k - 1}{n_N - 1} g_l$$

$$= U_l(g_l = 0) - \frac{(1 - R/n_N)(n_N - 1) - (n_N - 2k - 1)}{n_N - 1} g_l$$

Thus if

(A5) $$\frac{(1 - R/n_N)(n_N - 1) - (n_N - 2k - 1)}{n_N - 1} \geq 0$$

a deviation of player *l* to a contribution of zero is profitable. Equivalent transformation of (A5) yields that *l* has an incentive to deviate to a contribution of zero if and only if

(A6) $k \geq \frac{n_N - 1}{2n_N} R$.

However, since $0 \leq \frac{n_N - 1}{2n_N} < \frac{1}{2}$, the value $R/2$ is an upper bound for $\frac{n_N - 1}{2n_N} R$. Hence, if $k \geq R/2$ we can be sure that no equilibrium with strictly positive contributions exists. In other words, if the number of players *k* with $R/n_N + \beta_i < 1$, is at least $R/2$, all players contribute zero, independent of the community size n_N. If $R < 2$ (as in our experimental setting with *R* = 1.6) the presence of already one single player with $R/n_N + \beta_i < 1$ prevents cooperation in N, independent of community size n_N.

II. It remains to be shown that in case all players have $R/n_N + \beta_i \geq 1$ equilibria with strictly positive contributions exist and in all these equilibria, all players contribute the same amount.

First we show that under the assumption that all players in N satisfy $R/n_N + \beta_i \geq 1$ there exists a multiplicity of (pure strategy) equilibria in which all players contribute an identical amount of $g_N \in \{0...y\}$ to the joint project. If all players contribute g_N, then player *i* has no incentive to deviate to a lower contribution $g_i < g_N$ since for her (by assumption) the monetary benefit of withholding 1 unit is lower (or equal) than the total loss from deviation $R/n_N + \beta_i$. Player *i* has also no incentive to deviate to a higher contribution $g_i > g_N$ since in this case a contribution increase by 1 unit would cause a strictly positive utility loss of $1 - R/n_N + \alpha_i > 0$. Thus, if all players contribute the same amount, no player has an incentive to deviate, neither to a higher nor to a lower contribution.

Are there additional equilibria with non-identical contributions? Assume that there are two different contribution levels: *l* players contribute $g^L > 0$ while $(n_N - l)$ players contribute $g^H > g^L$. A player *j* with $R/n_N + \beta_j \geq 1$ is ready to increase her contribution if *all* other players contribute more than player *j* since for her the monetary benefit of withholding one monetary unit is lower (or equal) than her total loss in utility from deviation $R/n_N + \beta_j$. Thus, no player has an incentive to contribute less than g^L. An analogous argument shows that no player has an incentive to contribute more than g^H.

The situation with l players contributing g^L and $(n_N - l)$ players contributing g^H can only be an equilibrium if the players contributing g^L have no incentive to increase their contributions towards g^H while the players contributing g^H should not have an incentive to decrease their contributions towards g^L. In the following, we deduce the condition which has to be satisfied such that players with contributions g^H have no incentive to deviate to a lower contribution (Part A). Moreover, we deduce the condition which has to be satisfied such that players with contributions g^L have no incentive to deviate to a higher contribution (Part B).

Part A. If a high contributor i deviates from g^H by reducing her contribution by Δ, with $g^H - \Delta > g^L$, then her utility increases by $\Delta - \frac{R}{n_N}\Delta + \frac{\alpha_i}{n_N - 1} l\Delta - \frac{\beta_i}{n_N - 1}(n_N - l - 1)\Delta$. If this term is negative then a deviation to a lower contribution is not profitable. The term is negative if and only if

(A7) $$l \leq \frac{(R/n_N + \beta_i - 1)(n_N - 1)}{\alpha_i + \beta_i}.$$

This means that a high contributor does not decrease her contribution if the number of low contributors is not too high (cf. Fehr and Schmidt 1999, proof of Proposition 4 part c), p. 862).

Part B. If a low contributor i deviates from g^L by increasing her contribution by Δ, with $g^H - \Delta > g^L$, then she gains $-\Delta + \frac{R}{n_N}\Delta - \frac{\alpha_i}{n_N - 1}(l - 1)\Delta + \frac{\beta_i}{n_N - 1}(n_N - l)\Delta$.

If this term is negative then a deviation is not profitable. The term is negative if and only if

(A8) $$l \geq \frac{n_N(R/n_N + \beta_i - 1) + 1 - R/n_N + \alpha_i}{\alpha_i + \beta_i}.$$

Thus, a low contributor does not increase her contribution if the number of low contributors is not too low.

Hence, strategy combinations with two groups of conditional cooperators contributing different amounts can only be part of an equilibrium if l satisfies both conditions (A7) and

(A8). In the following we show that there is no such *l*.

$$\frac{n_N(R/n_N+\beta_i-1)+1-R/n_N+\alpha_i}{\alpha_i+\beta_i} \le l \le \frac{(R/n_N+\beta_i-1)(n_N-1)}{\alpha_i+\beta_i}$$

(A9) $$\Leftrightarrow \quad \frac{n_N(R/n_N+\beta_i-1)}{\alpha_i+\beta_i}+\frac{1-R/n_N+\alpha_i}{\alpha_i+\beta_i} \le l \le \frac{(n_N-1)(R/n_N+\beta_i-1)}{\alpha_i+\beta_i}$$

Because all players are assumed to have $R/n_N+\beta_i-1\ge 0$ and $1-R/n_N>0$ is satisfied in each public goods game, the lower bound of *l* is strictly greater than the upper bound of *l*. Hence, there is no *l* that satisfies (A9). This means that the situation with two groups of conditional cooperators who contribute different amounts cannot be part of an equilibrium. **Q.E.D.**

Adaption of the Fehr and Schmidt (1999) results to the P community (cf. Fehr and Schmidt 1999, Proposition 5, p. 841):

Proposition 2. Assume there are two types of players, n_P' "*conditional cooperative enforcers*" (short: "*enforcers*") with preferences that obey $R/n_P+\beta_i\ge 1$ and $c<\frac{\alpha_i}{(n_P-1)(1+\alpha_i)-(n_P'-1)(\alpha_i+\beta_i)}$ for $i\in\{1,...,n_P'\}$ and (n_P-n_P') players who are only interested in their own monetary payoff (short: "*payoff maximizers*"), i.e., $\alpha_j=\beta_j=0$ for all $j\in\{n_P'+1,...,n_P\}$. Then the following actions are part of a subgame perfect equilibrium in P:

I. On the equilibrium path all players contribute the same amount $g\in[0,20]$ and no punishment occurs in the punishment stage.

II. If, off the equilibrium path, one of the payoff maximizers chooses $g_j<g$ then each enforcer punishes the deviator with the punishment level $t_{ij}=(g-g_j)/(n_P'-c)$ while all other players do not punish.

Proof. By following backward induction reasoning, we first consider the punishment stage. Suppose that one of the payoff maximizers $j\in\{n_P'+1,...,n_P\}$ deviates and chooses $g_j<g$. We show that this deviator is punished by all enforcers with the punishment level t_{ij} as stated above and that this makes the deviation not profitable.

If in the punishment stage all enforcers choose a punishment level $t_{ij} = (g - g_j)/(n_P' - c)$, then deviator *j* obtains the same monetary payoff as each enforcer $i \in \{1,..., n_P'\}$. To see this consider the monetary payoffs of *j* and *i*:

(A10) $$x_j = y - g_j + \frac{R}{n_P}\left[(n_P - 1)g + g_j\right] - n'\frac{g - g_j}{n' - c}$$

(A11) $$x_i = y - g + \frac{R}{n_P}\left[(n_P - 1)g + g_j\right] - c\frac{g - g_j}{n_P' - c} - \frac{n_P' - c}{n_P' - c}(g_j - g_j)$$

The right hand side of (A11) can be rewritten as

(A12) $$y - g_j + \frac{R}{n_P}\left[(n_P - 1)g + g_j\right] - (n_P' + c - c)\frac{g - g_j}{n_P' - c} = x_j$$

which demonstrates that the deviating payoff maximizer achieves the same payoff as each enforcer. This payoff, however, is strictly lower than the payoff of a payoff-maximizer who did not deviate and contributed *g*. Thus, given t_{ij} a deviation of a payoff maximizer to a lower contribution is not profitable. Obviously, a deviation to a higher contribution level is also not profitable.

Now, we have to assure that enforcers' punishment strategies are credible, i.e., that an enforcer does not have an incentive to unilaterally reduce her punishment t_{ij}. If an enforcer reduces t_{ij} by δ she saves δc and experiences less disadvantageous inequality relative to the $(n_P - n_P' - 1)$ non-enforcers. This creates a utility gain of $[\alpha_i (n_P - n_P' - 1)\delta c]/(n_P - 1)$. On the other hand, the enforcer also experiences disutility from the disadvantageous inequality with respect to the defector *j* and advantageous inequality with respect to the other $(n_P - 1)$ enforcers who stick to the punishment t_{ij}. The disadvantageous inequality causes a utility loss of $\alpha_i (1 - c)\delta /(n_P - 1)$ whereas the advantageous inequality reduces the utility by $\beta_i (n_P' - 1)\delta\varepsilon /(n_P - 1)$. Thus the total utility loss from a reduction in t_{ij} is greater than the gain if

(A13) $$\frac{1}{n_P - 1}[\alpha_i (1 - c)\delta + \beta_i (n_P' - 1)\delta c] > \delta c + \alpha_i (n_P - n_P' - 1)\frac{\delta c}{n_P - 1}$$

holds. One can easily show that (A13) is equivalent to

$$\text{(A14)} \qquad c < \frac{\alpha_i}{(n_P - 1)(1 + \alpha_i) - (n_P' - 1)(\alpha_i + \beta_i)} \ ,$$

i.e., the condition we assumed in the Proposition. Obviously, a deviation to a higher punishment level is also not profitable. It would cause a monetary loss, disadvantageous inequality with respect to the other enforcers, would increase the disadvantageous inequality with respect to the non-punishing contributors and would cause advantageous inequality with respect to the punished player. Hence, the punishment level $t_{ij} = (g - g_j)/(n_P' - c)$ provides no incentives for deviation and is thus credible.

Do enforcers have an incentive to deviate in the contribution stage? Suppose the deviating enforcer reduces her contribution by $\delta > 0$. The deviator i gains $(1 - R/n_P)\delta$ in monetary terms but she experiences a disutility of $\beta_i\delta$ from the advantageous inequality with respect to all other players. Since, by assumption, $1 - R/n_P \leq \beta_i$ and since the player may additionally experience punishment in stage 2, this deviation does not pay. Hence, no enforcer deviates in the contribution stage either. On the other hand, choosing $g_i > g$ is not profitable for any player either, since it reduces the monetary payoff and increases inequality. **Q.E.D.**

Corollary 1. For the class of equilibria described in Proposition 2, enforcers can only exist if $(n_P - n_P') < 1/c$, i.e., the number of payoff maximizers is strictly lower than the reciprocal value of the cost of punishing.

Proof. Assume the existence of n_P' enforcers in P who satisfy (A14) which can be rephrased as

$$\text{(A15)} \qquad c(n_P - n_P') < 1 - \frac{c(n_P - 1) - c\beta_i(n_P' - 1)}{\alpha_i} .$$

By contradiction we show that $(n_P - n_P') < 1/c$ has to be satisfied. Assume $(n_P - n_P') \geq 1/c$. Then (A15) would imply that $\frac{c(n_P - 1) - c\beta_i(n_P' - 1)}{\alpha_i} < 0$. This, however, can never be the case because $\frac{c(n_P - 1) - c\beta_i(n_P' - 1)}{\alpha_i}$ is always strictly positive for each of the n' enforcers.

Hence $(n_P - n_P') < 1/c$ has to be satisfied. **Q.E.D.**

An intuition for this potentially unexpected implication is the following: By investing c in punishment, an enforcer reduces the monetary payoff of the deviator by exactly one unit, hence the inequality between the payoffs of the enforcer and the deviator decreases by $(1-c)$, i.e., the enforcer's disutility from being worse off than the deviator decreases exactly by $[\alpha_i/(n_P-1)](1-c)$. At the same time, the enforcer creates a payoff inequality of c units with respect to each non-enforcer who contributes but does not punish. This means that the enforcer suffers from a disutility $[\alpha_i/(n_P-1)]c$ with respect to each non-enforcer; in sum $[\alpha_i(n_P-n_P'-1)/(n_P-1)]c$. For punishment to be profitable for the enforcer, the utility gain from punishing must outweigh the disutility with respect to the non-enforcers, i.e., $[\alpha_i/(n_P-1)](1-c) > [\alpha_i(n_P-n_P'-1)/(n_P-1)]c$. This condition is equivalent to what Corollary 1 proposes: $(n_P - n_P') < 1/c$.

Corollary 1 implies for $c = 1/3$ that in an equilibrium of the class above the P-community can "afford" at most two payoff-maximizers independent of the community size n_P. If there are no enforcers in P who threaten to punish the non-contributors, the situation is "equivalent" to the situation in N described above: equilibria with positive contributions only exist if all players are conditional cooperators with $R/n_P + \beta_i \geq 1$, independent of n_P.