NBER WORKING PAPER SERIES

GUN FOR HIRE: DOES DELEGATED ENFORCEMENT CROWD OUT PEER PUNISHMENT IN GIVING TO PUBLIC GOODS?

James Andreoni
Laura K. Gee

Working Paper 17033
http://www.nber.org/papers/w17033

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2011

We would like to thank our colleagues at UCSD, participants in the Social Dilemmas Conference, participants at the ESA 2010 Conference, participants at the SWET 2011, Simon Gachter, Nikos Nikiforakis, and Anya Savikhina for helpful comments during the writing of this paper. Andreoni also gratefully acknowledges the National Science Foundation, grant 1024683, for financial support. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Gun For Hire: Does Delegated Enforcement Crowd out Peer Punishment in Giving to Public Goods?

James Andreoni and Laura K. Gee

NBER Working Paper No. 17033
May 2011
JEL No. C72,C91,C92,D7,H41

ABSTRACT

This paper compares two methods to encourage socially optimal provision of a public good. We compare the efficacy of vigilante justice, as represented by peer-to-peer punishment, to delegated policing, as represented by the "hired gun" mechanism, to deter free riding and improve group welfare. The "hired gun" mechanism (Andreoni and Gee, 2011) is an example of a low cost device that promotes complete compliances and minimal enforcement as the unique Nash equilibrium. We find that subjects are willing to pay to hire a delegated policing mechanism over 70% of the time, and that this mechanism increases welfare between 15% to 40%. Moreover, the lion's share of the welfare gain comes because the hired gun crowds out vigilante peer-to-peer punishments.

James Andreoni
Department of Economics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0508
and NBER
andreoni@ucsd.edu

Laura K. Gee
Department of Economics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0508
l1gee@ucsd.edu

An online appendix is available at:
http://www.nber.org/data-appendix/w17033

1 Introduction

The title character of the 1950s television western, Paladin, is described as a "gentleman" and "accomplished warrior" who "insists the *rule of law* be enforced, rejecting man-to-man frontier justice" (Hirschman, 2000). His calling card read simply, "Have Gun. Will Travel." The lawlessness of the "wild west" can be described in modern terms as "peer-to-peer punishment" in which scores are settled between parties, often with inefficient punishment. Paladin encouraged cooperative behavior by providing order through reason and, failing that, force. He was a gun for hire and was portrayed as a costly but superior alternative to shootouts, feuds, and endless retribution.

This paper provides a theoretical model and experimental analysis of Paladin. We show that (as seen on TV) a simple mechanism (a gun for hire) is an efficient and desirable substitute for lawless peer-to-peer punishment, and our experimental subjects show a clear preference for hiring such an enforcer. In particular, subjects endogenously hire the gun and discontinue the use of peer-to-peer punishments. This mirrors the movement from widespread dueling in the "wild west" toward the crowding out of such vigilante justice by delegating punishment authority.

This paper describes an experiment that compares the relative performance of peer-to-peer punishments and a delegated punishment mechanism. Previous work has shown that subjects in public goods games are willing to use costly punishments to reduce free riding behavior. Although a strong demand for peer-to-peer punishments exists in the laboratory setting, we observe little of this type of justice in the real world. Instead we often observe the development of delegated or appointed parties that sanction bad behavior. Consider the following examples: the homeowners' association, the building superintendent, the soccer coach, the department head, committee chair, the parent teacher association, and synagogue or church elders. These authorities are created and often funded by a subset of the people and institutions that they monitor. One suspects that these mechanisms arise because they are more efficient than the alternative of vigilante justice. We see people and institutions choosing a hired gun instead of punishing each other for infractions in the real world, but a thorough investigation of the two different mechanisms has not been conducted.

In an effort to demonstrate the potential for research in this area, we use a series of linear public goods games to examine whether subjects will choose to be governed by a delegated mechanism. In all of our games, subjects are randomly assigned to a group of four people, in which they are asked to allocate an endowment between a public good and a private good. We experiment with three enforcement regimes that players can use to discourage free riding. Subjects have three types of punishment conditions: peer-to-peer punishments only, a "gun for hire" punishment mechanism, or both peer-to-peer punishment with a "gun for hire".[1] The "hired gun" punishment mechanism is meant to be just one of any number of examples of small scale self-policing devices; it is a stylized version of the homeowner's association, or building superintendent. By looking at the peer-to-peer (P2P) and gun for hire (G4H) separately and jointly, we can identify their relative welfare effects.

[1]Consult (Andreoni and Gee, 2011) for a more general discussion of the "Hired Gun" mechanism. We will concentrate on a punishment regime's ability to discourage free riders. However the peer-to-peer regime actually allows for punishment of any subject, not only those who free ride. Previous papers have observed that sometimes players punish those who are contributing more than the average contribution in their group. We do not observe a large amount of this behavior, so we will treat peer punishment as a device to deter free riding behavior.

We find there is significant demand for the hired gun both when it is the only punishment option and when it is offered alongside peer-to-peer punishment. Welfare, as measured by group earnings net of costs of punishment, is significantly improved when groups can choose to hire a gun compared to when they can only peer-to-peer punish. Welfare is also improved when subjects can choose to hire a gun in addition to peer-to-peer punish compared to when they can only peer-to-peer punish. Furthermore, when both types of punishment are available and the gun is hired, the costs of peer-to-peer punishment decline precipitously. The hired gun mechanism has variable marginal costs while the peer-to-peer mechanism has constant marginal cost, so it is not clear ex-ante which regime will be less costly to the group. Groups choose to use the hired gun in spite of this cost uncertainty and in spite of the fact that hired gun has higher average costs than the peer-to-peer mechanism in our sessions.

In sum, when peer punishment is the only option, individuals use it, often with negative welfare consequences. When given the option of a centralized punishing mechanism, players prefer this to taking justice into their own hands; they cease to engage in peer punishment, and welfare improves dramatically. While our model and results are highly stylized, we will argue that the experimental observation is suggestive of a common real-world phenomenon, that is, inefficient social institutions (such as peer-to-peer punishment) can be easily supplanted by lower cost, more efficient mechanisms that delegate enforcement.

2 Background

In previous experiments on costly peer-to-peer punishment, subjects can pay a fee to reduce the payoff to another subject in their group *only* once. While this type of peer-to-peer punishment leads to higher contributions to the public good, the effects on group welfare (group earnings net punishment costs) have been ambiguous. Egas and Riedl (2008), Gachter, Renner and Sefton (2008), Herrmann, Thoni and Gachter (2008), Botelho, Harrison, Pinto and Rutstrom (2007), Fehr and Gachter (2002), Fehr and Gachter (2000), and Ostrom, Walker and Gardner (1992) all found decreases in net earnings in the short run, while Masclet, Noussair, Tucker and Villeval (2003) found that adding a single round of punishment increased net earnings.[2] If the peer-to-peer punishment is repeated over many periods (50 periods of play, rather than 10) with the *same* groups intact, then there is a welfare improvement (Gachter et al., 2008). In this case, it is possible that repeated interaction created reputation or reciprocity concerns that may have partially driven this result.

Notice that a single round of costly punishment does not take into account the possibility for revenge. When an opportunity for counter-punishment is added then net earnings are dramatically reduced, as found by Denant-Boemont, Masclet and Noussair (2007) and Nikiforakis (2008).[3] Hence, multiple rounds of costly punishment can create disastrous revenge cycles (Nikiforakis and Engelmann, 2011). One way to lower the costs is to allow non-monetary punishments, such as disapproval messages or expos-

[2]If punishments are only carried out when at least two members of the group request them, then over time there is a welfare gain (Casari and Luini, 2009). While, if players receive a noisy signal of other group member's behavior then the addition of punishment is not only detrimental to welfare, but also decreases contributions to the public good. See Grechenig, Nicklisch and Thoni (2010).

[3]Cinyabuguma, Page and Putterman (2006) found that if subjects are not given information about who specifically punished them, then net earnings increase. This restriction on the information basically makes revenge motivated second round punishments impossible.

ing only low contributors (Noussair and Tucker, 2005; Masclet et al., 2003; Savikhina and Sheremeta, 2010). The fact that people enjoy expressing their disapproval is convincingly shown by Fudenberg and Pathak (2010), who demonstrate that subjects still engaged in costly punishment even though it was not observed until the end of 10 rounds of play. In such a case, punishment logically could have no effect during the game. If people enjoy punishing, and if costly punishment is the only tool available, then the negative welfare effects of costly punishment are likely to be exacerbated by revenge cycles.

This literature suggests that to improve welfare, we need to curb the enjoyment of punishment and prevent peer-to-peer revenge cycles. When the streets are full of vendettas, and desparados are roaming the frontier looking for a fight for fun, what do the town folks do? They call Paladin. That is to say that a natural method for solving both these problems is "hiring" or "appointing" someone to discipline the group. Note that by delegated we don't necessarily mean someone outside the group, but simply mean a commonly recognized conduit for complaints, who monitors and metes out punishments. The punishments need not be more severe than those available by peer-to-peer punishment. The key is that discipline is centralized. Some previous work has already shown that central coordination of punishment can be welfare improving both theoretically (Boyd, Gintis and Bowles, 2010; Sigmund, Silva, Traulsen and Hauert, 2010; Steiner, 2007) and in experiments (Dickinson and Villeval, 2008; Yamagishi, 1986).

Yamagishi's experiment is most closely related to our "Gun For Hire". Yamagishi allowed subjects to first play the public goods game and then contribute to a punishment fund which punished the lowest contributor to the public good. Unlike our mechanism, Yamagishi's punishment size was not related to the size of deviation from compliance. Yamagishi finds public contributions were higher under punishment, but welfare was only improved under certain cost schemes. Although these results lend credence to the idea that these is a welfare gain from a delegated sanctioning mechanism, we believe that choosing the amount of punishment after choice of public contribution is fundamentally different than choosing to hire a delegated mechanism before the public goods game has taken place. We also see our study as improving on Yamagishi's insights by making punishments sensitive to the severity of the infraction.

If delegated punishment is the solution, will people voluntarily submit to a "hired gun"? Clearly many positive examples exist in the real world on both a large and small scale, such as the police regulating public safety, the EPA assessing fines for emissions, the PTA socially penalizing those who don't sell raffle tickets, the building superintendent speaking to the noisy neighbors, or the department chair cracking down on bad teaching. There have been some experiments in which subjects have been able to choose if they would like to be punished either by each other (Sutter, Haigner and Kocher, 2010; Ertan, Page and Putterman, 2009; Gurerk, Irlenbusch and Rockenbach, 2006; Botelho et al., 2007; Decker, Stiehler and Strobel, 2003) or by a third party (O'Gorman, Henrich and Van Vugt, 2009; Kosfeld, Okada and Riedl, 2008; Guillen, Schwieren and Staffiero, 2007; Tyran and Feld, 2006). These authors have found that, in some cases, subjects are willing to choose to allow punishing. Many of these experiments have made the implementation of a punishing mechanism monetarily costless. Monitoring, however, typically requires some resources or opportunity cost. By contrast, we make our punishment mechanism costly, but the cost is less than the gain realized through cooperation.

3 The Games

The experiment contains four different public goods games. We use the the linear public goods game with four players as the baseline and add three variations. The first variation assumes the world has a default state of no punishment of any kind, while the second and third variations assume peer punishment as the natural default state. The first variation adds a pre-play stage to the standard linear public goods game in which subjects can pay to implement a mechanism, a gun for hire, that punishes free riding behavior in the subsequent game. If subjects pay enough as a group, the mechanism will be implemented. If they do not, the game is the same as the baseline. The second variation adds a post-play stage in which subjects can pay a cost to punish other players. The third variation combines both the adaptations in a game with three stages. In the first stage, subjects choose whether to implement a mechanism that punishes free riding. In the second stage, they play the baseline public goods game. In the third stage, they can choose to peer-to-peer punish other players in their group.

Moreover, the gun for hire game is meant to be a single example of a low-cost but effective enforcement device. Previous work has used peer-to-peer punishment to represent phenomena like striking workers ostracizing strike breakers (Fehr and Gachter, 2000). We offer our gun for hire mechanism as a highly stylized representation of the naturally occurring delegated enforcement mechanisms, like the fines that a building superintendent might use to punish a noisy tenant. The gun for hire mechanism does appear to be naturally occurring (Rockenbach and Wolff, 2009).

Our research strategy here is to first look at a situation where the enforcement is delegated to a third party. If this does not result in crowding out of peer-to-peer punishment, then it is unlikely that weaker forms of delegation would do so. If we find significant crowding out of peer-to-peer punishment, however, then it suggests future research should explore, for instance, delegating the authority to someone within the group, or relaxing the "mechanism" to be a voluntary "rule" or "advice" given to the delegated authority. Hence, we see our study as a natural first step toward understanding how delegated enforcement can arise to supplant peer-to-peer punishment.

In the next section we describe the baseline and three variations as implemented in this experiment. For a general discussion of the "hired gun" mechanism see Andreoni and Gee (2011).

3.1 The Linear Public Goods (LPG) game

Subjects are given an "automatic payment" of \$1 (to control for within experiment income effects, as will be seen later), and an endowment of 5 tokens that they allocate between a public good and a private good. Each token invested in the public good pays a return of \$2 to all group members for an aggregate social return of \$8. Each token invested in the private good pays a return of \$3 to only the individual who made the investment. Let g_i be player i's contribution to the public good. The earnings for a subject for a period are:

$$\pi_i^{LPG} = 1 + 3(5 - g_i) + 2\sum_{j=1}^{4} g_j$$

A selfish profit-maximizing player would choose to set $g_i = 0$ and if all players are selfish they will each earn \$16. The group welfare maximizing level of contribution is

$g_i = 5$. If all players choose this amount, their earnings would be \$41 each. After all subjects have chosen g_i they are given anonymous information about the contribution to the public good, private good, and initial LPG earnings for each of their group members.

3.2 The Gun For Hire (G4H) game

In stage 1, each subject is given an endowment of 4 tokens worth \$0.25 each, making it equivalent to the value of the "automatic payment" in the basic LPG game. Subjects choose e_i, $0 \leq e_i \leq 4$, to contribute to the "hiring fund." If the sum of the 4 person group's contributions reach a threshold of 8 tokens, a delegated punishment mechanism will be implemented in stage 2. Subjects' stage 1 earnings equal the number of tokens they kept multiplied by \$0.25. Over-payments for hiring the gun are not refunded to the subjects. If the threshold for hiring is not met, subjects are refunded their e_i and earn \$1 in stage 1. Thus, if the gun is not hired, then we employ the "hired gun" mechanism (Andreoni and Gee, 2011). Andreoni and Gee (2011) provide a formal model and test of the "hired gun" mechanism which we summarize next.

3.2.1 What The Hired Gun Shoots

In the "hired gun" mechanism the administrator simply takes a deduction from the lowest contributor to the public good. The size of the bullet fired by the hired gun varies with the size of the infraction from the group behavior.[4]. The size of the deduction is set so as to make the lowest contributor to the public good just slightly worse off (in terms of net subgame payoff) than the second lowest contributor to the public good. The lowest contributor to the public good will earn the amount that the second lowest contributor earns minus a constant (the value of one unit of the private good, \$3).

Formally, let g_z denote the contribution of the lowest contributor to the public good, $g_z = min\{g_1, g_2, g_3, g_4\}$. If there is a tie for the lowest contributor, then all those who tied will be punished. Let g_y denote the second lowest contribution to the public good, $g_y = min\{g_1, g_2, g_3, g_4 \backslash g_z\}$. The size of the punishment will be the difference between the initial payoffs of player z and player y plus a constant, M. We set M equal to the cost from taking one token of the player's private good, so $M = \$3$. The punishment for player z is equal to:

$$P = \pi_z - \pi_y + 3 = 3(g_y - g_z) + 3$$

In the special case in which all the players choose the same level of contribution to the public good, but still give below full contribution ($g_i = g_j < 5 \ \forall i, j$), all the subjects are punished P_0. We set P_0 to \$3, the payoff from contributing a token to the private good. See Andreoni and Gee (2011) for the generalized model and proofs. Lastly if all 4 subjects contribute the full 5 tokens to the public good, then no one is punished. To summarize, when the gun is hired, the size of the shot fired is equal to:

$$P = \begin{cases} 3 & \text{if} \quad g_i = g_j < 5 \text{ for all } i, j \\ 0 & \text{if} \quad g_i = 5 \text{ for all } i \\ 3(g_y - g_z) + 3 & \text{if} \quad \text{for lowest contributor(s) in other cases} \end{cases}$$

In stage 1, subjects choose e_i and are told the sum of these contributions by their group, whether they have hired the punishment mechanism for stage 2, and their

[4]E.g. the punishment fits the crime (Andreoni, 1991)

stage 1 earnings. In stage 2 they are reminded whether they hired the punishment mechanism, and they make their choices of g_i. After all players have chosen g_i, they are given anonymous information about the contribution to the public good, private good, initial LPG earnings, size of punishment (if any), and final net payoffs for each of their group members.

3.2.2 G4H Subgame Equilibrium

Notice that any choice of $g_z < g_y$ will result in earning \$3 less than player y, and so it is strictly dominated by a choice of $g_i = g_y + \epsilon > g_y$, which will result in no punishment ($\epsilon > 0$). The best response of the lowest contributor is to change g_z to be just slightly higher than g_y. If all subjects are reasoning this way it is never a best response to set $g_i = 0$. Knowing that all subjects will not choose to set g_i to zero, a subject will choose g_i equal to the next discrete amount above zero, $g_i = 1$. But then knowing that everyone else is using similar reasoning, subjects will want to choose the next discrete amount above $g_i = 1$, and so they need to move to $g_i = 2$. In short, the best response for any player is to find what the lowest level of contribution is, and to set their contribution slightly above it. The only fixed point is full contribution to the public good $g_i = 5$.

When the gun is hired the game is like a p-beauty contest (Nagel, 1995) in reverse.[5] Each player is trying to guess the lowest amount given by the others in her group and then wants to give the closest contribution above that amount possible. This thought process eventually pushes all the players to contribute all of their endowment to the public good. Each player should choose $g_i = 5$ and will earn \$40 in the subgame.

3.2.3 How Should Players Behave In Stage 1?

When the gun is hired in stage 2, we expect subjects to use iterated dominance reasoning to contribute fully to the public good, and earn \$40. If the gun is not hired, we expect own-profit maximizing subjects to contribute 0 to the public good, and earn \$15. A subject should be willing to pay any amount less than or equal to the gain from hiring the gun (\$25), to hire the gun. We have set the total group cost of hiring the gun to only \$2 per group. Any combination of contributions summing to exactly \$2 will be an equilibrium of the stage 1 game (Bagnoli and Lipman, 1989; Bagnoli and McKee, 1991; Marks and Croson, 1998).

Any two players could pay for the punishment mechanism, so one could interpret the implementation of our mechanism as requiring 50% of the group to agree on implementation. The average cost of the gun per person should be \$0.50, and with the gun hired 2nd stage earnings should be \$40, resulting in average earnings of \$40.50 per subject in the G4H game.

3.3 The Peer-to-Peer (P2P) game

Our peer-to-peer punishment game is similar to that of previous experiments (see Fehr and Gachter, 2002; Cinyabuguma et al., 2006; Herrmann et al., 2008; Gachter et al.,

[5]The p-beauty contest is a game in which a group of subjects are all asked to choose a number between 0 and 100. The average of these numbers is computed and multiplied by a number p, typically $p < 1$. The person who guesses the number closest to p multiplied by the group average wins a prize. The only fixed point in this game is for all players to choose the number 0. In the case where $p > 1$ the only fixed point is the number 100.

2008). Subjects first play the LPG game (with an automatic payment of \$1 (again to control income effects), then are given anonymous information about the contribution to the public good, private good, and about initial LPG earnings for each of their group members. At this point, each player i can pay \$1 to assign a punishment point to another player j, which we write as p_{ij}. Each point assigned reduces player j's payoff by \$3.[6] Final payoff are given by the following expression:

$$\pi_i = 1 + 3(5 - g_i) + 2\sum_{j=1}^{4} g_j - \sum_{j \neq i} p_{ij} - 3\sum_{k \neq i} p_{ki}$$

Given that groups are randomly and anonymously rematched each period, own-profit maximizing subjects should choose to assign zero punishment points to all players and the game should be the same as the LPG game. The predicted outcome under own-profit maximizing behavior is $g_i = 0$ for all subjects, and final earnings per subject of \$16.

It is important to note that the own-profit maximizing equilibria predictions of the P2P and LPG games are the same, but that many previous works have found that subjects behave very differently in these two games. The fact that players engage in punishment at all is surprising, not only because it is not the equilibrium action, but more so because we do not observe much peer-to-peer punishment in many real world situations. One reason we observe such high amounts of peer punishment in the lab may be that players were never offered another alternative, such as hiring a delegated punishing mechanism in addition to peer punishments. Our final game allows the use of both a delegated punishment mechanism and peer-to-peer punishments.

3.4 The Gun For Hire and Peer-to-Peer (G4H/P2P) game

The last game combines the G4H and P2P games. In stage 1, subjects are given 4 tokens and they make contributions toward a hiring fund. If the sum of those contributions is greater than 8 tokens, then a gun is hired and subjects get \$0.25 for each token they kept. If the gun is not hired stage 1 earnings are \$1. Subjects are informed of their stage 1 earnings, group contributions to the hiring fund, and whether the gun has been hired. In stage 2, subjects get 5 tokens to contribute to either a public or private good. If the gun was hired, then the lowest contributor(s) to the public good will be punished by the delegated punishment mechanism. In stage 3, subjects are given anonymous details of group members' contributions to the public good, the private good, and their initial earnings before any punishments from the delegated mechanism. They also learn the size of punishment from the mechanism (if any), and the net earnings for each subject in their group. At this point, subjects can choose to assign peer-to-peer punishments to their group members. Again, subject i chooses an amount of punishment points to

[6]The punishment to cost ratio of 3:1 has been employed by many of the previous experiments (e.g. Fehr and Gachter, 2002; Gachter et al., 2008; Herrmann et al., 2008), while some others have employed a 4:1 ratio (Cinyabuguma et al., 2006). For a discussion of the constant ratio versus other punishment regimes see Casari (2005). Previous work has found that a cost to punishment ratio of no lower than 1:3 is necessary to raise public contributions and welfare (Nikiforakis and Normann, 2008; Egas and Riedl, 2008). There is the possibility of earning a negative payoff in the P2P game. Subjects were warned about the possibility of negative payoffs in the instructions and were told that they would never owe money at the end of the experiment; and that at minimum they would be paid \$7. In only 3 of cases did a subject earn a negative amount in a period.

assign to player j. Each point player i assigns costs player i \$1, and reduces the payoff of player j by \$3.

Again, own-profit maximizing subjects would assign zero punishment points, leading to predictions identical to the G4H game: subjects hire the gun in the first stage, and fully contribute in the second stage. Average per person earnings would be \$40.50 per person.

4 Procedures

There are two equally valid views of what is the "natural" baseline. The first is that the the LPG game is the baseline and the P2P is an intervention. The second takes vigilante justice as an ever present option, and so the baseline should be a game with peer-to-peer (P2P) punishments available. We conduct two sets of experiments using both the LPG, and the P2P games as baselines.

Each session involved 12 subjects and 20 periods: 10 periods of a baseline game (either LPG or P2P) followed by 10 periods of a game with punishment (either P2P, G4H, or P2P/G4H). Each treatment is a set of two games, and there are a total of 4 treatments: (1) LPG-P2P, (2) LPG-G4H, (3) P2P-P2P, and (4) P2P-G4H/P2P. Each treatment was conducted 3 times for a total of 36 subjects per treatment. We have a total of 144 subjects. Each session was conducted using z-tree software (Fischbacher, 2007), lasted under 90 minutes and subjects earned \$28 on average.

To minimize repeated game effects, participants were randomly and anonymously re-matched into a new group of 4 participants at the beginning of each period (see Andreoni, 1988). Subjects were given the instructions for the first 10 periods of play, a quiz, and then played that game for 10 periods. This is done again for the last 10 periods. To remove experimenter effects, all sessions were run by the same person. Subjects could earn up to \$46 in each period, so we paid subjects for only a single period of play. Each of subjects was informed that she would be paid for a randomly selected single period from the 20 periods in the session.[7]

The instructions were written in neutral language by referring to the public good as the "BLUE investment", the private good as the "RED investment", the delegated punishment mechanism as "the computer simulated administrator", and referring to all punishments as "deductions." Full instructions and screen shots are available from the authors.[8]

5 Results

The "natural" baseline for our experiments is either a world without any punishment options, the LPG game, or a world with only vigilante justice, the P2P game. Although we are most interested in the value added by different punishment regimes, we will take a moment to discuss behavior in periods 1-10.

[7] To choose the random period after the end of the 20th period, a subject was given a 20 sided die. The subject was asked to verify the die had 20 sides, and then to roll and announce the outcome on the die out loud.

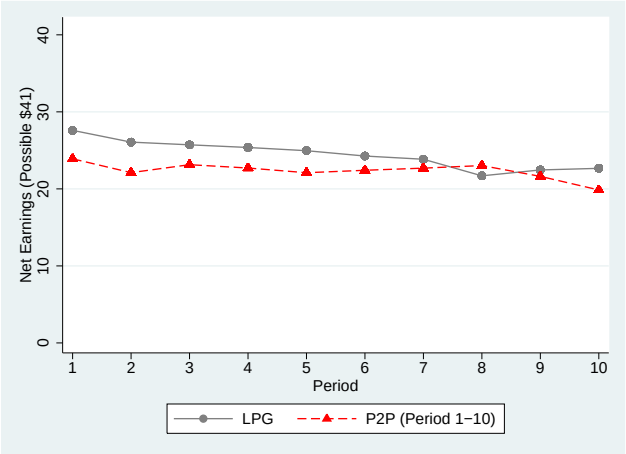
[8] `econ.ucsd.edu/~jandreon` or `econ.ucsd.edu/~l1gee`

5.1 Two Baseline Games: LPG and P2P

Average earnings were \$24.47 in periods 1-10 of the LPG game. The earnings decrease from an average of \$27 in the first period to only \$22 in the 10th period. We do not observe an end game effect in our sample, which is likely due to the random rematching of groups each period. The other baseline for comparison is a world with peer-to-peer punishment. When the P2P game was played in periods 1-10 the average earnings were \$22.36. The earnings decrease from \$23.90 to \$19.80 from Period 1 to Period 10. Again we do not observe an end game effect.[9]

Our two baseline worlds are not equal. Average earnings are statistically significantly higher in the LPG game than in the first 10 periods of the P2P game using a Kolmogorov-Smirnov test ($p = 0.05$).[10] Figure 1 shows that average earnings are higher in the LPG game in all but a single period. This result is consistent with previous work comparing the LPG setting to the P2P setting (see Gachter et al., 2008; Botelho et al., 2007; Fehr and Gachter, 2002, 2000; Ostrom et al., 1992). Also in line with previous results, the actual public contributions are higher in P2P (41% of endowment) as compared to LPG (33% of endowment). We do not observe the upward trend in public good contributions that others have seen, but the P2P session has a slower convergence toward free riding. In both games public contributions start at 46% of the endowment, but in LPG they fall to 26% of endowment while in P2P they fall to 36% of endowment by Period 10.

Figure 1: Average Per Subject Earnings during Period 1-10



[9]We expected that the behavior in the first 10 periods of LPG-P2P and our LPG-G4H sessions should be the same regardless of what is played in periods 11-20, because subjects are not given directions for the last part of the session until *after* they have completed periods 1-10. Averaging across all sessions we find that this is indeed true. Average earnings in the first 10 periods of the LPG-P2P were \$25.15 and in the LPG-G4H were \$23.78, but this difference is not statistically significant using a Two-sample Kolmogorov-Smirnov test ($p = 0.518$). We treat the average earnings in each of the 3 sessions of LPG-P2P and 3 sessions of LPG-G4H as an observation, so we are comparing 3 observation of LPG-P2P to 3 observations of LPG-G4H. Likewise we expected behavior in the beginning periods of the P2P-P2P, and P2P-P2P/G4H treatments to be similar. The average earnings in the first 10 periods of P2P-P2P (\$22.10) were very slightly lower than the average earnings in the first 10 periods of the P2P-P2P/G4H sessions (\$22.61), but the difference is statistically insignificant using a Two-sample Kolmogorov-Smirnov test ($p = 0.996$).

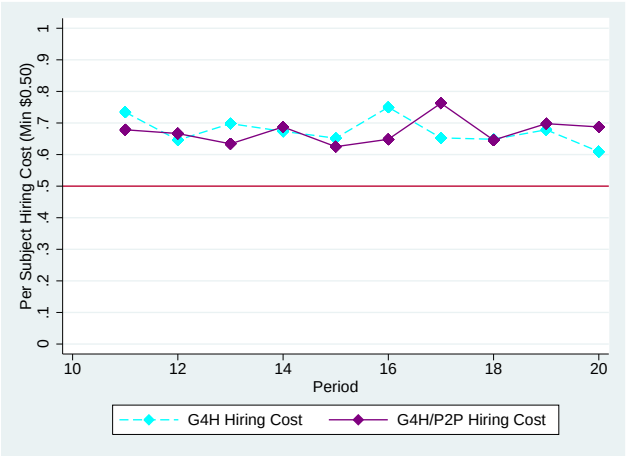
[10]We treat each session of each treatment as an observation, so we have 6 observations for the LPG (from 3 sessions of LPG-P2P and 3 sessions of LPG-G4H) and 6 observations for the P2P (from 3 sessions of P2P-P2P and 3 sessions of P2P-G4H/P2P).

5.2 Implementation: Will people hire a gun?

We see from periods 1-10 that, regardless of baseline, subjects could be earning more if they contributed more to the public good. We ask here, will subjects pay a small fee to hire the gun to subsequently raise contributions to the public good. Although the equilibrium of the G4H and G4H/P2P game is to hire the gun, subjects may not always immediately realize this fact. For subjects to hire the gun, they must believe that the cost of implementing the delegated punishment mechanism will be outweighed by the gains from reduced free-riding. Subjects appear to believe this - they hire the mechanism 85% of the time in last 10 periods of LPG-G4H, and 72% of the time in P2P-G4H/P2P.

There are multiple equilibria for the hiring stage, such that any combination of contribution to the hiring fund that total to exactly \$2 (8 tokens) is a Nash equilibrium. Yet, we only observe the groups paying exactly \$2 a mere 12% of the time. The average cost paid is \$2.68, which is about 34% higher than the equilibrium cost. Because the bulk of the analysis concerns average earnings per subject, we will look for average per subject hiring costs of \$0.50 (instead of \$2 per 4 person group). Figure 2 shows the average costs of hiring the gun in the last 10 periods, when the gun has been successfully paid for (so when at least \$2 has been raised by the group). Groups do not appear to converge toward the Nash predicted price of an average \$0.50 per person (\$2 per group) over time.

Figure 2: Average Individual Hiring Costs

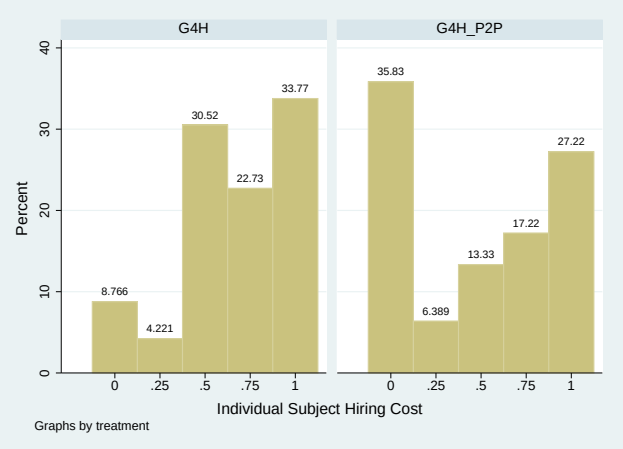


Note: Average per subject hiring costs in dollars by period in the G4H and G4H/P2P games. Red line is at \$0.50 which is equilibrium prediction.

Subjects may not converge toward paying exactly \$2 because some subjects always over-pay (pay more than \$0.50) for the hired gun. Figure 3 shows that in the last 10 periods of both the LPG-G4H and the P2P-G4H/P2P treatments a large percentage of players choose to pay their whole first stage endowment of \$1 for the hired gun. In both the G4H and the G4H/P2P treatment, payment over the "fair" contribution of \$0.50 occurs at least 40% of the time. Result 1 summarizes.

Result 1*: Subjects are willing to pay a cost to submit to a delegated punishment mechanism. In both the G4H and G4H/P2P treatments the delegated punishment mechanism is implemented over 70% of the time, and groups over-pay for this implementation in most cases.*

Figure 3: Distribution of Individual Hiring Costs Across Treatments



Note: Percentage of subjects paying \$0.00, \$0.25, \$0.50, \$0.75 and \$1.00 over all periods in the G4H and the G4H/P2P games.

5.3 Hiring Guns: Welfare in Periods 11-20 following LPG in Periods 1-10

Let us begin by looking at our sessions that had no punishments during the first 10 periods. Looking at Table 1 we see that average per person earnings in Periods 11 to 20 are higher in the LPG-G4H treatment (\$35.44 overall: \$38.12 when gun is hired and \$19.55 when not hired), than they are in the LPG-P2P treatment (\$30.69). Table 2 provides an overview of how average earnings are shaped in each treatment of this experiment. The variable G4H takes the value 1 when subjects are playing the G4H game, and zero when they are playing the P2P game after periods 1-10 of LPG. Playing the G4H game instead of the P2P game raises earnings by \$4.76 per period on average including when the gun was not hired.[11]

There are two possible reasons for the increased average earnings: increase average giving and decreased average punishment costs. Table 1 shows that average giving was nearly identical in P2P (4.11) and G4H (4.16). However, as Figure 4 this average masks a deal of heterogeneity across treatments. When the delegated punishment mechanism is hired, average giving is higher in G4H. As a result, punishment in G4H are small (\$1.35 per subject), especially in comparison to P2P (\$5.85 per subject). Thus, lower punishments are primarily responsible for the increased efficiency.

The gain in earnings between the two treatments is illustrated in Figure 5. In the left panel are earnings in the last 10 periods of the LPG-G4H treatment both with and without the hired gun, which is almost always higher than earnings with LPG-P2P punishments. In the right panel we see that when the gun is hired, average per subject earnings are always higher than those under P2P punishments.

[11]The same patterns of significance can be shown in Kolomogrov Smirnov test at the session level. Some may worry that selection into hiring the gun is driving this result, as Sutter et al. (2010) found that public contributions were higher when subjects endogenously chose the sanctioning mechanism versus having it exogenously imposed. We ran trials where subjects were randomly assigned into have the hired gun mechanism already implemented Andreoni and Gee (2011), a treatment we call LPG-GH. The level of earnings across the last 10 periods of the exogenously chosen mechanism (LPG-GH) were not statistically significantly different from those in the last 10 periods with the endogenously chosen mechanism (LPG-G4H) (Kolomogrov Smirnov test $p = 0.290$).

Table 1: **Average Earnings per Subject after LPG in Periods 1-10**

Game (Periods)	**Net Earnings (Dollars)**	**Public Good Contribution (5 tokens)**	**P2P Costs (All)**	**G4H Costs (All)**
LPG (1-10)	24.47	1.69	.	.
P2P (11-20)	30.69	4.11	5.86	.
G4H (11-20): All	35.44	4.16	.	1.35
Hired (85%)	38.12	4.74	0	1.57
Not Hired (15%)	19.55	0.71	0	0

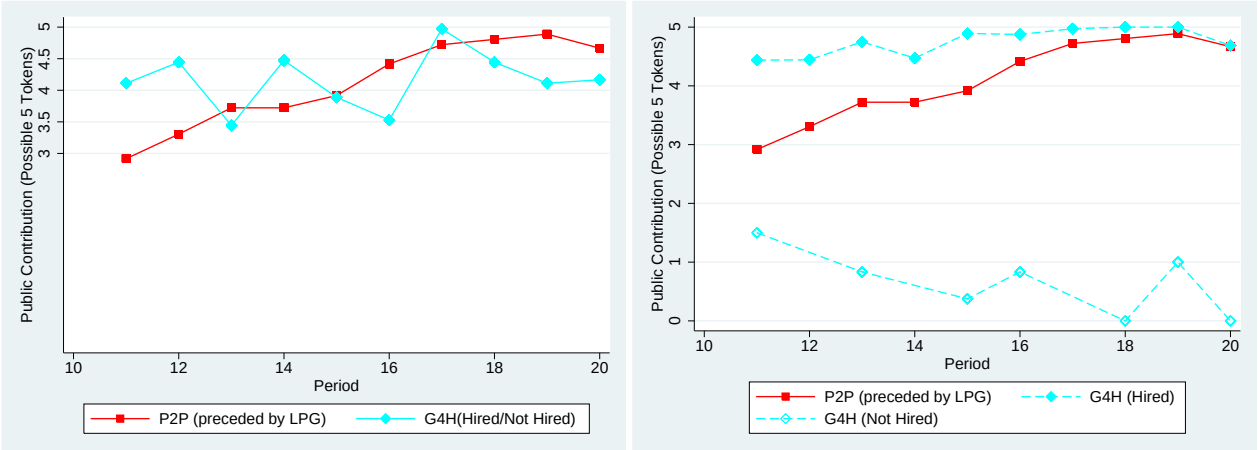
Note: 10 Periods of each game per Session, 3 Sessions, 3 Groups, 4 Subjects per Group

Table 2: **Determinants of Earnings**

	After LPG	**After P2P**
G4H	4.76***	
	(1.14)	
G4HP2P		8.98**
		(3.07)
Period	0.90***	0.49***
	(0.10)	(0.09)
Constant	16.77***	15.25***
	(1.76)	(2.59)
N	720	720
Wald Chi-Squared	97.14***	38.03***

Notes: Linear random effects models. Clustered standard errors in parentheses. Standard errors clustered by session. *** $p < .01$, ** $p < .05$, * $p < .10$ significance

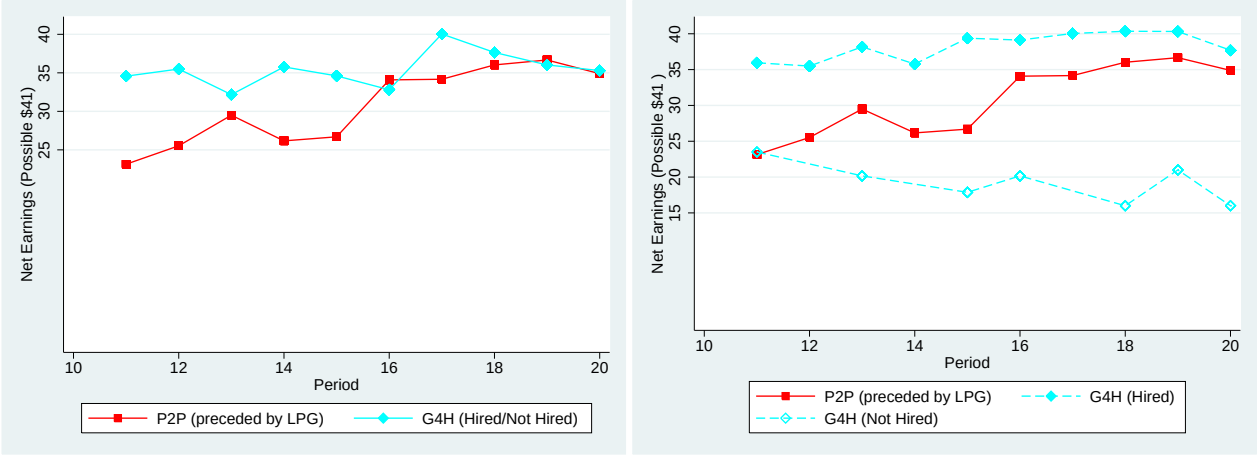
Figure 4: Contributions to the Public Good after LPG in Periods 1-10



***Result 2**: Welfare, as measured by average individual net earnings, is higher in the G4H treatment than the P2P treatment. The use of a delegated punishing mechanism both improves public contributions, and lowers costs as compared to allowing peer-to-peer punishments.*

One can see in both Figures 4 and 5 that the advantage of G4H over P2P diminishes with time, as shown by Fehr and Gachter (2000). As we see next, however, if P2P is

Figure 5: Average Per Subject Earnings after LPG in Periods 1-10



considered "natural", then the advantage of G4H grows rather than diminishes.

5.4 Starting From Vigilante Justice: Welfare in Period 11-20 following P2P in Periods 1-10

It has been argued that peer-to-peer punishment "plays an important role in real life" (Fehr and Gachter, 2000). If such peer-to-peer punishment is indeed natural and often occurring then we should use the P2P game as our baseline rather than the setting without any punishment opportunities. This brings us to a discussion of the sessions where subjects began with the ability to peer punish for the first 10 periods. Table 3 shows that average per person earnings in Periods 11 to 20 are higher in the G4H/P2P treatment ($31.87 overall: $36.14 when gun is hired and $20.77 when not hired), than they are in the P2P treatment ($22.89). Average net earnings are 40% higher in periods 11-20 in the P2P-G4H/P2P treatment than they are in the P2P-P2P treatment.

Additionally, in the regression reported in the second column of Table 2 one can see that the coefficient on the treatment dummy variable G4H/P2P is positive and significant. Playing the G4H/P2P game instead of the P2P game raises earnings by $8.98 per period on average (this is an average over when the gun was hired and when the gun was not hired).[12] Figure 6 shows that contributions to the public good in the last 10 periods of the P2P-G4H/P2P treatment seem to rise over time, while the public contributions stay relatively flat in the P2P-P2P treatment. In the left panel of Figure 7 we see that averaging over when the gun is hired and not hired subjects have higher per person net earnings in the G4H/P2P treatment than in the P2P treatment in every period. Additionally Figure 7 shows that earnings trend upwards for the G4H/P2P treatment, while they stay relatively flat in the P2P treatment. Result 3 summarizes.

Result 3*: Welfare, as measured by average individual net earnings, is higher in the last 10 periods of the P2P-G4H/P2P treatment than the P2P-P2P treatment.*

[12]The variable G4H/P2P takes the value 1 when subjects are playing the G4H/P2P game, and zero when they are playing the P2P game after periods 1-10 of P2P. The same patterns of significance can be shown in Kolomogrov Smirnov test at the session level.

Table 3: **Average Earnings per Subject after P2P in Periods 1-10**

Game (Periods)	**Net Earnings (Dollars)**	**Public Good Contribution (5 tokens)**	**P2P Costs (All)**	**G4H Costs (All)**	**Total Costs**
P2P (1-10)	22.36	2.09	4.07	.	4.07
P2P (11-20)	22.89	2.33	4.74	.	4.74
G4H/P2P (11-20): All	31.87	3.67	1.03	1.43	2.45
Hired	36.14	4.55	0.63	1.97	2.60
Not Hired	20.77	1.37	2.08	0	2.08

Note: 10 Periods of each game per Session, 3 Sessions, 3 Groups, 4 Subjects per Group

Figure 6: Contributions to the Public Good after P2P in Periods 1-10

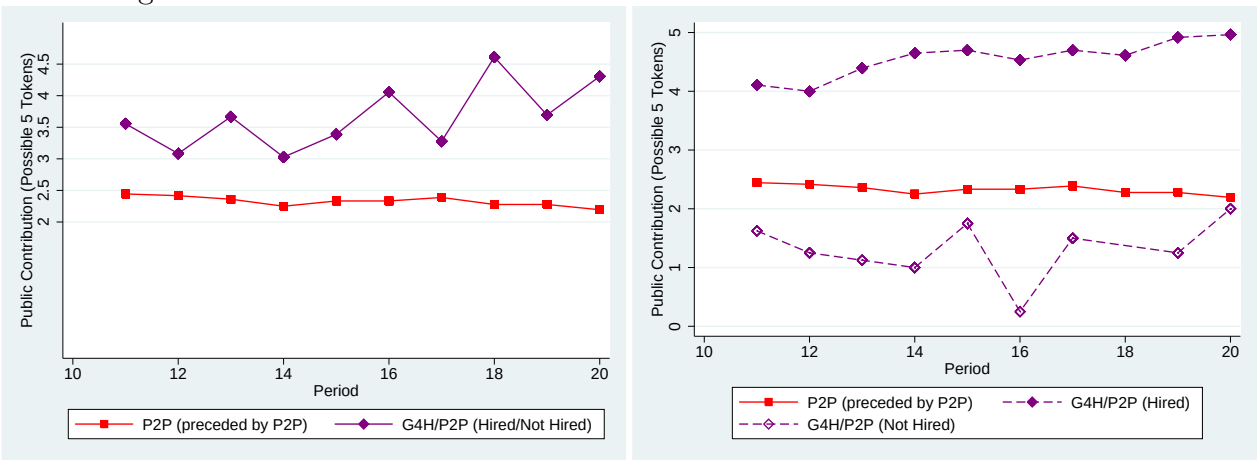
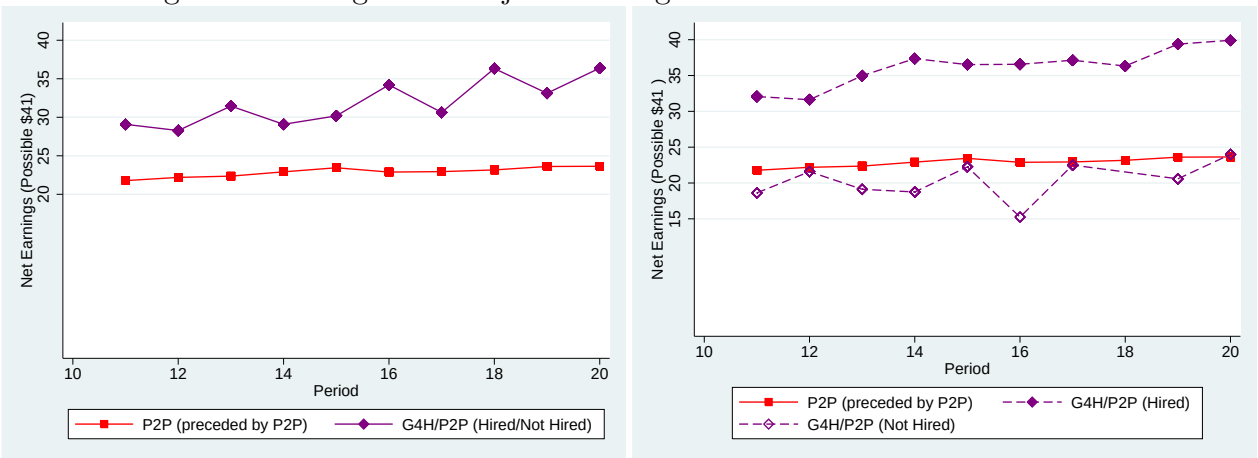


Figure 7: Average Per Subject Earnings after P2P in Periods 1-10



5.5 Does Delegated Enforcement Crowd Out Peer Punishment?

We have shown that our G4H mechanism is both implementable and welfare improving when compared to a P2P punishment regime. Next we show that hiring a gun crowds out the use of peer punishments. If delegated punishment crowds out peer-to-peer punishment, this may be welfare improving. Also if delegated punishment crowds out peer-to-peer punishment, this will in turn lower any possible motives for for peer-to-

peer revenge punishments.[13]

Looking back at the Table 3, we can compare the behavior of subjects who could only peer punish in periods 11-20 (P2P) to those who were also allowed to hire a gun and peer punish (G4H/P2P). When we make this comparison we see that peer punishment costs fall from an average \$4.74 in the P2P game to \$1.03 in the G4H/P2P game (\$0.63 when the gun is hired, and \$2.08 when the gun is not hired). The average use of peer punishment when it is the only option is over four times higher than when peer punishment is available alongside the option for a hired gun (\$1.03 versus \$4.74).

Some may wonder if the "hired gun" mechanism is simply a less expensive way to punish than the P2P. In the P2P it always costs \$0.33 to punish another player \$1. In the G4H and G4H/P2P games the ratio varies since the punishment depends on size of deviation, and can range as high as \$18. Although it may appear that paying only \$0.50 per person (\$2 in total) to punish \$18 is simply a great value, in our experiments the punishment was usually well below this \$18 size. On average it cost \$0.74 to punish \$1 in the G4H game, and \$0.51 to punish \$1 in the G4H/P2P game, and so punishment was actually cheaper under the P2P mechanism.

Comparing the use of peer punishment by hiring decision in the G4H/P2P treatment alone, we see the average costs of peer punishment fall about 70% when the gun is hired. In Figure 8 the solid line with circles shows the total costs (on average per subject) of peer punishments in the P2P game, while the two other lines with diamonds show these costs by hiring decision in the G4H/P2P game (dashed diamonds when the gun is hired, and solids diamonds when the gun is not hired). In both the P2P and G4H/P2P games the use of peer punishment is trending downward over time. In the P2P game the costs of punishment are always higher than in the G4H/P2P game, whether a gun has been hired or not. One would expect the use of peer punishments to fall when a gun has been hired, but it is especially surprising that the use of peer punishments is lower even when the gun is not successfully hired. It appears that merely giving the option for the Gun For Hire, even when that option is not exercised, decreases peer-to-peer punishments. Furthermore, it is noteworthy that in 4 of the 10 periods (after 10 periods of P2P) when the gun is hired peer punishment costs are equal to zero.

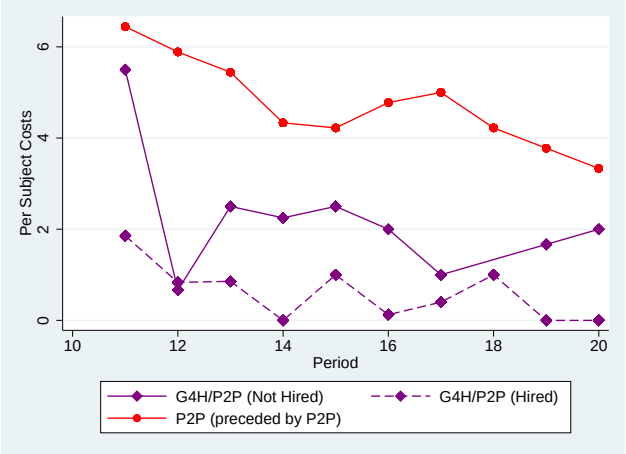
Result 4*: Use of peer punishments is over four times higher when the option for a delegated mechanism is not available. When the delegated mechanism is implemented, peer punishment converged to zero by period 19. Delegated punishment crowds out peer-to-peer punishment, resulting in an overall welfare gain.*

6 Concluding Remarks

Much of the previous work on punishment in public goods games has concentrated on asking whether groups can govern themselves through the use of peer-to-peer punish-

[13]We expect that the use of a delegated punishing mechanism would preclude revenge motives in subsequent rounds of punishment, although we have not allowed for multiple rounds of punishment in our current design. When the mechanism is levying fines, it is not possible for an individual to know who to take revenge on. Imagine if your neighbor was leaving garbage in the common areas of your building. You can either speak with your neighbor directly, or ask the superintendent to speak to your neighbor without mentioning your name. If you speak with your neighbor directly they may take offense, and they may "counter-punish" you by stealing your newspaper. On the other hand if your superintendent speaks with your neighbor, there is no way for your neighbor to know that you commissioned the punishment.

Figure 8: Average Per Subject Peer Punishment Costs after P2P in Periods 1-10



ments. This line of inquiry does not allow individuals to collectively agree to concentrate the punishment in a recognized authority. In this paper, we show that subjects willingly pay to delegate punishments in a linear public goods game. We offer a stylized version of delegated punishment in our "hired gun" mechanism (Andreoni and Gee, 2011). The mechanism has the properties that only the largest free rider is punished, the size of the punishment is related to the degree of defection from the group behavior, in equilibrium the mechanism is efficient in the sub-game, and the mechanism is relatively low cost.

When given the opportunity to hire a delegated punishment mechanism, we see the mechanism being implemented over 70% of the time in both the linear public good and peer-to-peer baseline worlds. The likely reason that subjects are so willing to submit to a costly outside authority is that they expect monetary gains from reduced free-riding. These expectations are well-founded, as can be seen by the 15% and 40% increase in welfare when comparing a peer-to-peer punishment regime to those with the option of a "gun for hire" regime. When subjects can only use peer punishments (P2P) the peer punishment costs are over four times those with a delegated mechanism (G4H/P2P). Lastly and most importantly, we find that when both punishment methods are available (G4H/P2P), subjects lower their use of peer punishments by 70%. The existence of a delegated punishing mechanism crowds out the use of peer punishments.

To our knowledge this paper is the first to allow subjects to choose between hiring a costly punishment mechanism and using peer-to-peer punishments. We have shown that players want to hire the delegated mechanism, and that the "hired gun" mechanism provides a low cost solution to the problem of free-riding. Interestingly, although the delegated mechanism is itself a public good it does not appear to suffer from the same level of free-riding as observed in the subsequent LPG game. The reason may be that the cost of hiring is fairly low as compared to the potential gains in payoffs. This is analogous to the way we pay taxes or fees to fund delegated punishing mechanisms in general. Often these fees and penalties are small as in our mechanism. Further research exploring how players react to changes in the cost of implementing the mechanism could be illuminating. Although formally our Hired Gun was an external third party, it is clearly an important and desirable next step for research to investigate a more general set of ways individuals can delegate authority. For instance, the recognized authority can be internal to the group, and the enforcers's conformity with enforcement rules

a choice variable. This would be most interesting, of course, in the default domain of peer-to-peer punishment. The ultimate research goal suggested by our study is to understand how easily small self-governed groups can innovate ways to avoid the inefficiencies of peer-to-peer punishment.

In sum, this paper illustrates that under reasonable conditions individuals prefer to pay to be governed by a delegated punishment mechanism rather than use peer-to-peer punishments. The gun for hire mechanism is just one example of a low-cost device that can deter free-riding behavior in a public goods game, improve welfare, and crowd out the use of deleterious peer-to-peer punishments.

In short, when Paladin comes to town, vigilante justice is driven out. Have gun. Will travel.

References

Andreoni, James, "Why free ride? : Strategies and learning in public goods experiments," *Journal of Public Economics*, 1988, *37* (3), 291 – 304.

_ , "Reasonable Doubt and the Optimal Magnitude of Fines: Should the Penalty Fit the Crime?," *The RAND Journal of Economics*, 1991, *22* (3), pp. 385–395.

_ **and Laura Gee**, "The Hired Gun Mechanism," 2011.

Bagnoli, M. and M. McKee, "Volutary Contribution Games: Efficient Private Provision of Public Goods," *Economic Inquiry*, 1991, *29.*

Bagnoli, Mark and Barton L. Lipman, "Provision of Public Goods: Fully Implementing the Core through Private Contributions," *The Review of Economic Studies*, 1989, *56* (4), pp. 583–601.

Botelho, A., G.W. Harrison, L.M.C. Pinto, and E.E. Rutstrom, "Social Norms and Social Choice," *Working Paper*, 2007.

Boyd, Robert, Herbet Gintis, and Samuel Bowles, "Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare," *Science*, 2010.

Casari, Marco, "On the Design of Peer Punishment Experiments," *Experimental Economics*, 2005, *8* (2), 107–115.

_ **and Luigi Luini**, "Cooperation under alternative punishment institutions: An experiment," *Journal of Economic Behavior and Organization*, 2009, *71* (2), 273 – 282.

Cinyabuguma, Matthias, Talbot Page, and Louis Putterman, "Can Second-Order Punishment Deter Perverse Punishment," *Experimental Economics*, 2006, *9* (3).

Decker, Torsten, Andreas Stiehler, and Martin Strobel, "A Comparison of Punishment Rules in Repeated Public Good Games: An Experimental Study," *Journal of Conflict Resolution*, 2003, *47* (6), 751–772.

Denant-Boemont, Laurent, David Masclet, and Charles N. Noussair, "Punishment, Counterpunishment, and Sanction Enforcement in a Social Dilemma Experiment," *Economic Theory*, 2007, *33* (1).

Dickinson, David and Marie-Claire Villeval, "Does monitoring decrease work effort?: The complementarity between agency and crowding-out theories," *Games and Economic Behavior*, 2008, *63* (1), 56 – 76.

Egas, Martijn and Arno Riedl, "The economics of altruistic punishment and the maintenance of cooperation," *Proceedings of the Royal Society B: Biological Sciences*, 2008, *275* (1637), 871–878.

Ertan, Arhan, Talbot Page, and Louis Putterman, "Who to punish? Individual decisions and majority rule in mitigating the free rider problem," *European Economic Review*, 2009, *53* (5), 495 – 511.

Fehr, Ernst and Simon Gachter, "Cooperation and Punishment in Public Goods Experiments," *The American Economic Review*, 2000, *90* (4), 980–994.

_ **and** _ , "Altruistic punishment in humans," *Nature*, 2002.

Fischbacher, Urs, "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 2007, *10* (2).

Fudenberg, Drew and Parag A. Pathak, "Unobserved punishment supports cooperation," *Journal of Public Economics*, 2010, *94* (1-2), 78 – 86.

Gachter, Simon, Elke Renner, and Martin Sefton, "The Long-Run Benefits of Punishment," *Science*, 2008, *322* (5907), 1510–.

Grechenig, Kristoffel, Andreas Nicklisch, and Christian Thoni, "Punishment Despite Reasonable Doubt: A Public Goods Experiment with Sanctions Under Uncertainty," *Journal of Empirical Legal Studies*, 2010, *7*.

Guillen, Pablo, Christiane Schwieren, and Gianandrea Staffiero, "Why feed the Leviathan?," *Public Choice*, 2007, *130*, 115–128. 10.1007/s11127-006-9075-3.

Gurerk, Ozgur, Bernd Irlenbusch, and Bettina Rockenbach, "The Competitive Advantage of Sanctioning Institutions," *Science*, 2006, *312* (5770), 108–111.

Herrmann, Benedikt, Christian Thoni, and Simon Gachter, "Antisocial Punishment Across Societies," *Science*, 2008, *319* (5868), 1362–1367.

Hirschman, Elizabeth, *Heroes, monsters and messiahs: movies and television shows as the mythology of American culture*, Andrews McMeel Publishing, 2000.

Kosfeld, Michael, Akira Okada, and Arno Riedl, "Institution Formation in Public Goods Games," *forthcoming, The American Economic Review*, 2008.

Marks, Melanie and Rachel Croson, "Alternative rebate rules in the provision of a threshold public good: An experimental investigation," *Journal of Public Economics*, 1998, *67* (2), 195 – 220.

Masclet, David, Charles Noussair, Steven Tucker, and Marie-Claire Villeval, "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism," *The American Economic Review*, 2003, *93* (1), 366–380.

Nagel, Rosemarie, "Unraveling in Guessing Games: An Experimental Study," *The American Economic Review*, 1995, *85* (5), 1313–1326.

Nikiforakis, Nikos, "Punishment and counter-punishment in public good games: Can we really govern ourselves?," *Journal of Public Economics*, 2008, *92* (1-2), 91 – 112.

_ **and Dirk Engelmann**, "Altruistic punishment and the threat of feuds," *Journal of Economic Behavior and Organization*, 2011, *In Press, Corrected Proof*, –.

_ **and Hans-Theo Normann**, "A comparative statics analysis of punishment inpublic-good experiments," *Experimental Economics*, 2008, *11*, 358–369. 10.1007/s10683-007-9171-3.

Noussair, Charles and Steven Tucker, "Combing Monetary and Social Sanctions to Promote Cooperation," *Economic Inquiry*, 2005, *43* (3), 649 – 660.

O'Gorman, Rick, Joseph Henrich, and Mark Van Vugt, "Constraining free riding in public goods games: designated solitary punishers can sustain human cooperation," *Proceedings of the Royal Society B: Biological Sciences*, 2009, *276* (1655), 323–329.

Ostrom, Elinor, James Walker, and Roy Gardner, "Covenants With and Without a Sword: Self-Governance is Possible," *The American Political Science Review*, 1992, *86* (2), 404–417.

Rockenbach, Bettina and Irenaeus Wolff, "Institution Design in social dilemmas: How to Design if you must?," *Working Paper*, 2009.

Savikhina, Anya and Roman M. Sheremeta, "Visibility of Contributions and Cost of Information: An Experiment on Public Goods," *Working Paper*, 2010.

Sigmund, Karl, Hannelore De Silva, Arne Traulsen, and Christoph Hauert, "Social learning promotes insistitutions for governing commons," *Nature*, 2010.

Steiner, Jakub, "A Trace of Anger is Enough: on the Enforcement of Social Norms," *Economics Bulletin*, 2007, *8*, 1–4.

Sutter, M., S. Haigner, and M. Kocher, "Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations," *Review of Economic Studies*, 2010, *77*.

Tyran, J. and L. Feld, "Achieving Compliance when Legal Sanctions are Non-deterrent," *Scandinavian Journal of Economics*, 2006, *108*, 1467–9442.

Yamagishi, Toshio, "The Provision of a Sanctioning System as a Public Good," *Journal of Personality and Social Psychology*, 1986, *51* (1), 110 – 116.