Appendix: Proofs and Supplementary Analysis

Gun For Hire: Delegated Enforcement and Peer Punishment in Public Goods

By James Andreoni and Laura Gee

Supplemental Material Not Intended for Publication

1 Generalized Model

In the LPG, each individual i is endowed with w_i. Each allocates g_i, $0 \leq g_i \leq w_i$, to the public good and consumes the rest, $w_i - g_i$. Each person receives one util for each unit of w_i kept, and α for each unit allocated to the public good by themselves and all others. So i's payoff is $\pi_i = w_i - g_i + \alpha \sum_{j=1}^{n} g_j$. Assuming, as is usual, $0 \leq \alpha < 1$ and $n\alpha > 1$ then the unique Nash equilibrium of this game is total free riding, $g_i = 0$ by all players, while the Pareto efficient allocation is complete contribution, $g_i = w_i$. Thus, full compliance with the socially proscribed optimum is for $g_i = w_i$ for all.[1]

We assume for our mechanism that members of the community commonly understand the socially desirable action of each individual, that is, they all agree that $w_i = g_i$ is socially desirable, and that deviation from this action is undesirable. Thus, define

$$d_i = w_i - g_i$$

as i's deviation from his assigned contribution. Define $d = (d_1, d_2, ..., d_n)$ as the vector of deviations from full compliance. Then payoffs to i can be written in terms of d_i rather than g_i :

$$\begin{aligned} \pi_i(d) &= d_i + \alpha \sum_{j=1}^{n} (w_j - d_j) \\ &= d_i - \alpha \sum_{j=1}^{n} d_j + \alpha W \end{aligned}$$

where $W = \sum_{i=1}^{n} w_i$. In this formulation, $d_i = 0$ for all i is clearly the socially desired outcome.

Define d_z as the *largest* element of the vector d, that is, $d_z \geq d_i$ for all i. Let S be the set of potential contributors and $L(d) \subseteq S$ be the set of contributors with the largest deviations. Finally, define d_y as the *second* largest deviation. In particular if $L(d) \subset S$, then d_y is the highest d_i for all i in $S \backslash L(d)$. However, if $L(d) = S$, then $d_i = d_j$ for all i and j and there is no second highest deviator.

[1] More generally, let $g = g^*$ be any among the set of possible Pareto Effecient allocations, where g_i^* is perhaps chosen by some focal ideal, such as equal division ($g_i^* = g_j^*$, all i, j) or equal sacrifice ($g_i^*/w_i = g_j^*/w_j$, all i, j).

Let $P(d_i, d)$ be the punishment required by the mechanism for an individual who chooses d_i when the vector of choices is d. Then the mechanism requires punishments such that $\pi_z(d_z, d) - P(d_z, d) < \pi_y(d_y, d)$, that is, the biggest deviator would rather have been the second biggest deviator. Assume for convenience, and realism, that d_i can only be chosen in discrete amounts, and for simplicity assume it takes integer values. Then, to assure punishments are as small as possible, set the punishment on the biggest deviator to be $P(d_z, d) = \pi_z(d_z, d) - \pi_y(d_y, d) + 1 = d_z - d_y + 1$. Note, if there are ties for the biggest deviator, then all must be punished. In this case, punishment need only be high enough to break indifference, that is the smallest integer unit of the private good, thus $P = 1$. The exception is if all tie by being fully compliant. In this case, of course, there are no punishments.

In general, therefore, define punishment as follows:

$$P(d_i, d) = \begin{cases} 1 & \text{if } \; L(d) = S, \; \text{ and } \; d \neq 0 \\ 0 & \text{if } \; L(d) = S, \; \text{ and } \; d = 0 \\ d_z - d_y + 1 & \text{if } \; L(d) \subset S, \; \text{ and } \; i \in L(d) \\ 0 & \text{if } \; L(d) \subset S, \; \text{ and } \; i \notin L(d) \end{cases}$$

It is easy to see that $d = 0$ is the unique equilibrium. Proof follows from elimination of dominated strategies. Let m' be the highest standard of compliance set for anyone (this allows for m' to be the same for all i). Then any $d_i = m'$ is sure to be the lowest compliance level and i is sure to receive a punishment of at least 1. As such d_i will never equal m'. In this case, anyone with compliance level m' will act as if their true compliance level is $m' - 1$. But if m' is never chosen, then $m' - 1$ is sure to be punished and anyone choosing it would be better off choosing $m' - 2$. Thus $d_i = m' - 1$ will never be chosen by any i. We can repeat this logic, eliminating dominated strategies until the only choice that is not eliminated is that $d_i = 0$ for all i. This demonstrates that $d = 0$ is the unique equilibrium.

2 Estimating Equations

Here we provide the estimating equations for regressions in the "Determinants of Earnings" Table. The estimating equation predicts individual earnings (after all punishments have been deducted) for person i during session j, E_{ij}, as a function of a constant, c, the period (11-20), P, and a dummy variable which takes the value 1 when the treatment is G4H, and 0 if the treatment is the baseline. There is a session level fixed effect given by S_j, and there are a total of 6 sessions for each column. There is also an individual error term given by ϵ_{ij}. In the first column the equation is:

$$E_{ij} = c + \beta_1 P + \beta_2[1\{G4H = 1\}] + S_j + \epsilon_{ij}$$

In the second column of Table 2 the estimating equation is:

$$E_{ij} = c + \beta_1 P + \beta_2[1\{G4H/P2P = 1\}] + S_j + \epsilon_{ij}$$

Here we report the results for clustering at the individual level.

Table 1: **Determinants of Earnings**

	After LPG	**After P2P**
G4H	4.76***	
	(0.77)	
G4HP2P		8.98***
		(1.07)
Period	0.90***	0.49***
	(0.10)	(0.08)
Constant	16.77***	15.25***
	(1.60)	(1.51)
N	720.00	720.00
Wald Chi-Squared	123.70***	104.83***

Notes: G4H= 1 if subject in the G4H condition in rounds 11-20, and G4H= 0 if subject in the P2P condition in rounds 11-20 after playing LPG in 1-10. Similarly, G4H/P2P= 1 if subject in the G4H/P2P condition in rounds 11-20, and G4H/P2P= 0 if subject in the P2P condition in rounds 11-20. Linear random effects models. Clustered standard errors in parentheses. Standard errors clustered by individual (72 clusters per regression). Significance *** $p < .01$, ** $p < .05$, * $p < .10$