+ Model
ARTICLE IN PRESS

Journal of Economic Behavior & Organization
Vol. xxx (2006) xxx–xxx

JOURNAL OF Economic Behavior & Organization

www.elsevier.com/locate/econbase

The demand for punishment

Jeffrey P. Carpenter*

Department of Economics, Munroe Hall, Middlebury College, Middlebury, VT 05753, USA

Received 21 January 2003; received in revised form 12 July 2004; accepted 9 May 2005

Abstract

While many experiments demonstrate that behavior differs from the predictions of traditional economic theory, they have not shown that economic reasoning is necessarily incorrect. Instead, these experiments illustrate that the preferences of homo economicus have been mis-specified. Modeled with social preferences, it may be rational for agents to forego material gains. Social dilemmas are examples in which punishment is not credible and yet people often pay to reprimand other participants. At the same time, we show that these people also react to changes in the price of punishing and income as if punishment was an ordinary and inferior good.

JEL classification: C72; C92; H41

Keywords: Public good; Social dilemma; Social preference; Experiment

1. Introduction

At this point in the evolution of experimental and behavioral economics, laboratory experiments have provided more new questions about economic behavior than answers. Instead of confirming the standard tenets of neoclassical economics, experiments have identified decision-making anomalies (Camerer, 1995), preference reversals (Tversky et al., 1990), and non-standard or "social" preferences (Camerer and Fehr, 2001; Carpenter, 2002; Charness and Rabin, 2002).[1] Expanding on the idea of social preferences, experiments have shown that instead of being selfish and myopic, average participants are much better described as trusting and trustworthy (Berg et al., 1995), fair (Güth et al., 1982; Fehr et al., 1993), and cooperative (Isaac et al., 1984), but, as

* Tel.: +1 802 443 3241; fax: +1 802 443 2084.
E-mail address: jpc@middlebury.edu (J.P. Carpenter).

[1] In fact, competitive markets is one of the few areas where experiments have come close to confirming existing theories (Davis and Holt, 1993).

0167-2681/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jebo.2005.05.004

motivation for what follows, it is important to stress that they can also be vindictive (Camerer and Thaler, 1995; Fehr and Gächter, 2000).

However, the fact that many economic models predict behavior that is at odds with what we observe in experiments may be because we have misspecified peoples' preferences, not because the methodology of economics is fundamentally flawed. Although people behave as if they have preferences for cooperation and retaliation, they may still react to incentives in ways predicted by standard economic logic. For example, if we hypothesize that peoples' observed preferences for cooperation operate like preferences for more standard consumption goods, then we might expect people to cooperate less when the implied price of cooperation increases, just as they tend to buy fewer ordinary goods when the price increases.

In addition to being predisposed to cooperate, recent experiments have demonstrated that people retaliate against perceived injustices, even when doing so is costly and the material benefits of doing so are small or nonexistent. This evidence (reviewed below) leads one to believe that many participants have a preference for punishing asocial behavior. In the experiment reported on herein, we test, in a controlled setting, whether such a nonstandard preference behaves according to standard economic reasoning.

This research is unique because it is among the first to examine explicitly whether standard economic tools can explain behavior motivated by the nonstandard preference to punish free riders. At the same time, this research is linked to other recent work in behavioral economics. One area of research examines the sacrifices that people are willing to endure to assure fair outcomes and, in this sense, examines the price responsiveness of fairness preferences. Examples of this literature include Eckel and Grossman (1996), Suleiman (1996), and Zwick and Chen (1999). In a second related project, Andreoni and co-workers (Andreoni and Vesterlund, 2001; Andreoni and Miller, 2002; Andreoni et al., 2003) empirically recover utility functions that are based on social preferences. Variants of these utility functions could, in principle, generate the sort of demand for punishment functions that we estimate below.

We proceed by briefly reviewing the literature on social dilemma experiments in which players were given the opportunity to punish each other. Hopefully, this review will convince the reader that cooperation and retaliation are robust behaviors. We then discuss the current experiment that was designed to examine whether peoples' preferences for punishment behave according to standard economic logic. Specifically, the experiment provides us with data which we use to estimate the demand for punishment. Our analysis indicates that punishment is both ordinary and inferior, but is also relatively inelastic with respect to both price and income.

2. Fairness, cooperation, and punishment

The first evidence of a preference for punishing asocial behavior came from one-shot ultimatum games in which a first-mover makes an offer to share a sum of money with a second-mover who accepts or rejects this offer (Güth et al., 1982; Camerer, 1995). Although any division of the pie can be supported as an equilibrium of this game, subgame perfection leads one to expect that the first-mover will receive all (or almost all) of the money because selfish second-movers will always accept small offers rather than reject them and get nothing. Despite this unambiguous prediction, nearly all small offers are rejected, and the most common explanation given by second-movers is that they are retaliating against greedy first-movers (Pillutla and Murnighan, 1996).

Punitive behavior has also been witnessed in social dilemma games in which individual and group incentives are at odds, and therefore, free-riding is expected from selfish players. One of the first of these experiments was conducted by Ostrom et al. (1992). In this common pool resource experiment players cooperate with each other by not extracting too much from an open-access and subtractable resource. Resource use is problematic because by extracting, one player imposes a negative externality on all the other players. Under these incentives, the authors showed that when costly punishment was allowed, cooperative players used it to regulate the behavior of over-extractors (i.e., free riders) and the gross efficiency of extraction increased, especially when communication was allowed too.

Considering positive rather than negative externalities, Fehr and Gächter (2000) tested whether costly punishment could curtail free riding in a public goods experiment. In the voluntary contribution mechanism players emit a positive externality every time they contribute to a group project, the benefits of which are shared by the entire group. Given this structure, selfish players should contribute nothing and free ride on the contributions of others. Fehr and Gächter's results mirror those of Ostrom et al. in that they find that many contributors are willing to pay to punish those who contribute less than the average. Further, the (theoretically incredible) threat to punish reduces free riding dramatically. These results suggest that when subjects punish free riders they are expressing a social preference for retaliation because they punish despite having to pay to do so and despite the negligible material benefits that are expected to follow punishment.

There are many recent extensions of the Fehr and Gächter results. For example, Bochet et al. (in press) confirm that punishment increases contributions, but they also find that it does not increase them as much or as efficiently (when one accounts for the payoff consequences to both punisher and target) as communication. Bowles et al. (2001) also find increases in contributions due to punishment and extend the literature by looking at the implications of group size and the return on the public good. Their data suggest that the amount of punishment received by free riders is increasing in both the size of the group and the return on the public good. The end result is that large groups that generate large externalities contribute at very high levels when punishment is allowed.

In an extension of Bowles et al. (2001), Carpenter (in press) examines the interaction of group size, public good productivity, and monitoring technology and finds that punishment (as a contribution elicitation mechanism) is sensitive to the structure of groups. When each group member can monitor and punish all the other members of her group, contributions are high regardless of group size and the return on the public good. However, the most efficient use of punishment occurs when group members are allowed to monitor and punish only half of the other group members (provided the monitoring subgroups overlap). In these situations, contributions are as high as when everyone monitors everyone else, but the amount of punishment received by free riders is just enough to get them to contribute at high levels. That is, there is less wasted punishment compared to the first technology. Lastly, when group members see and can punish just one other group member, punishment is not an adequate deterrent, and contributions collapse at a rate that is similar to the no-punishment control treatment.

There are also a few experiments that show that punishment affects free riders even when it imposes no material cost on them. Masclet et al. (2003) show that even "cheap talk" punishment is effective. Participants in the experiment assign punishment to free riders even though the punishment does not reduce the target's payoff. Amazingly, free riders respond by contributing more in the future. However, there is one small problem with the Masclet et al. design: it is costless to mete out punishment, and therefore there are equilibria in which strictly egoistic players punish along side those with preferences to punish. Carpenter et al. (2004a) adjust their protocol so that punishment is costly to the punisher but still imposes no material harm on the target and bring the resulting game to the field. Their participants are slum dwellers in southeast Asia who face the sort of social dilemmas on a daily basis that the game is meant to model (e.g., clean water and solid waste). Interestingly, while participants in Ho Chi Minh City and Bangkok both use punishment, it only has a significant effect on contributions in Ho Chi Minh City. The authors provide some evidence that these behavioral differences may be due to culture.

In sum, there is a lot of evidence that participants in social dilemma experiments will punish free riders. It also appears that free riders respond to punishment under a variety of conditions including games in which punishment imposes no material harm. Social disapproval appears to be enough to motivate some free riders. However, punishment is no panacea; there are some instances in which it works poorly (e.g., in small groups with weak externalities from contributions), and in other cases, it is not clear that punishment always improves the social efficiency of interactions. Simple communication (perhaps confounded by the social disapproval in many group members' voices) appears to elicit contributions more efficiently than force, and when force is available, it appears important to limit the amount of punishment that free riders receive. The following experiment extends this research by asking whether punishers adhere to the law of demand.

3. Experimental design

While the following experiment is based on the voluntary contribution mechanism (Isaac et al., 1984) to test whether we can explain punishment in terms of standard economic logic, we made a few changes. Our changes were designed to provide us with the data to estimate the demand for punishment. First, we allowed players to monitor and punish each other. Second, punishment was costly to impose, and the price of punishment changed during the course of the experiment. This feature allows us to estimate the price elasticity of the demand for punishment. Third, the level of provision of the public good during each round determines an income for each player from which players paid to punish each other. This feature allows us to estimate the income elasticity of demand. Also note that because players' earnings and the price of punishment varied over the course of the experiment we are able to analyze the demand for punishment using a (more powerful) within-subject design.[2] The specifics of our experiment are as follows.

Define the *price of punishment*, r, as the amount a punisher must pay in experimental monetary units (EMUs) to remove one EMU from the target. Our experiment was 15 periods long, and each session was split into five blocks, each block lasting three periods. The price of punishment varied from block to block such that $r \in \{0.25, 0.5, 1, 2, 4\}$. We ran two treatments to balance the effect of changing prices. In the decreasing price treatment r equaled 4 for the first three periods, meaning the punisher spent 4 EMUs to remove 1 EMU from the target, r equaled 2 in periods four through six, and so on until in periods thirteen through fifteen the price was 0.25. In the increasing price treatment r started at 0.25 and cycled upward to 4. Our players were randomly assigned to a treatment, and we ran a total of six sessions (three for each treatment). This design resulted in a total of 18 four-person groups.

We used the familiar *strangers* protocol (Andreoni, 1988) under which players are randomly reshuffled from group to group at the beginning of each period because we wanted to control, to some extent, for strategic reasons to punish. For example, players who remain in the same group may perceive that their payoffs will increase if they punish free riders early on. However, if the target of one's punishment is likely to be in a different group next period, participants should understand that the expected benefit of punishing will be negligible. This is especially true in the current experiment where each participant monitors and can punish only one other member of

[2] These features differ to a significant degree from the related work of Anderson and Putterman (2005).

her current group. Controlling for strategic punishment is important because doing so allows us to focus on punishment as the expression of a social preference.

The payoff function for the voluntary contribution mechanism was augmented to account for punishment.[3] Imagine groups of n players, each of whom can contribute any fraction of their w EMU endowment to a public good and keep the rest. Say player i free rides at rate $0 < \sigma_i < 1$ and contributes $w(1 - \sigma_i)$ to the public good, the benefits of which are shared equally among the members of the group.

Each player's contribution was revealed to one other player in the group who could punish this person at a price of r EMUs per sanction. Let rs_{ij} be the expenditure on sanctions assigned by player i to player j, and let s_{ki} be the sanctions player i receives from player k (the instructions explicitly mentioned that $j \neq k$); then the payoff to player i is

$$\pi_i = w[\sigma_i + nm(1 - \bar{\sigma})] - rs_{ij} - s_{ki}$$

where $\bar{\sigma} \equiv (\sum \sigma_i)/n$ is the average free riding rate in the group. The variable m is the marginal per capita return on contributions to the public good (see Ledyard, 1995). In all sessions n equaled 4, m was set to 0.5, and w was 25 EMUs.[4]

Because $1/n < m < 1$ the game without punishment is a social dilemma: group incentives are at odds with individual incentives. Each contributed EMU returns only 0.5 to the contributor, which means free riding is a dominant strategy, but if $\bar{\sigma} = 1$ then everyone is free riding fully and each player's payoff is lower than it would be if everyone contributed fully. The game is finitely repeated, which implies that subgame perfection predicts free riding on every round.

Notice that adding the possibility of punishment does not change the subgame perfect prediction. Because sanctions are costly to impose and any potential benefits from getting a free rider to contribute cannot be fully internalized by the punisher, punishment is incredible and therefore cannot be a component of any subgame perfect equilibrium. Without credible punishment, free riding is still subgame perfect.

As noted above, each player monitored and was able to sanction only one other member of the group. This design feature was added to control for other possible strategic or coordination reasons that might affect players' punishing propensities. For example, if each player monitors and can punish all the other members of the group, there are at least two problematic scenarios that may arise. First, from a strategic perspective, a player may be less likely to punish a free rider because she thinks she can free ride on the punishment of others. Second, a player may be less likely to punish because she cannot explicitly coordinate her punishment efforts with the rest of her group.[5] For example, she may feel that the free rider should be punished, but also that there is

[3] The instructions (Appendix A available on the JEBO website) referred to "reductions" with no interpretation supplied.

[4] While the current protocol is quite standard in the public goods literature in terms of group size, matching rule (partners versus strangers), and the return on contributions to the public good, the punishment and public good literature is too small yet to have any consensus on a protocol. The nexus of this literature, Fehr and Gächter used four person groups, a return of 0.4 per contribution, and examined both partners and strangers groupings. Despite this role model, more recent experiments have varied the group size between 3 and 10 members (compare Anderson and Putterman, 2003 and Carpenter et al., 2004a), or used returns on the public good between 0.3 and 0.75 (see Bowles et al., 2002; Bochet et al., 2003, and Sefton et al., 2001), and both the partners and strangers protocols are used in addition to the *complete* strangers protocol in which players are guaranteed never to be in exactly the same group twice. The most important innovation, however, has been to use the less complicated punishment mechanism described here and used in much of the more recent punishment literature. In Fehr and Gächter players paid to reduce a target's earnings by a certain percentage, implying that the "price" of punishment is not constant.

[5] This rationale is consistent with the "unresponsive bystander" hypothesis advanced in Latane and Darley (1970).

an appropriate level of punishment that fits the infraction. If she does not know or cannot estimate how much others will punish, she may withhold sanctions to be sure that the punishment does not exceed the offense. If all players see only one other player and knows that the person they are monitoring is not monitoring them, we control for any strategizing and coordination problems. People should only pay to punish if they wish to express their preferences.

4. Overview of the data

We recruited 72 participants (36% were female) from the undergraduate population at Middlebury College in our six experimental sessions. Participants were from a variety of majors, and most were between 18 and 22 years old. Because the experiment occurred near the end of the semester when students' opportunity cost of time is particularly high, we calibrated the game so that payments would be generous and so that we would not have any problems recruiting participants. On average, the students earned $26.26, including a $5 show-up fee. The typical session lasted a little less than an hour. We begin our analysis by giving the reader a broad sense of the data, and then we focus on our estimates of the demand for punishment and whether punishment is effective.

Reviewing previous punishment experiments (e.g., Fehr and Gächter, 2000) we see that the typical time path of contributions, averaging across treatments, starts near half the endowment and then increases at a decreasing rate. However, as seen in Carpenter (in press), punishment has less of an effect on contributions when players monitor only a small subset of their groupmates. Pooling across periods the two price change treatments exhibit nearly the same levels of contributions ($\bar{c}_{\text{increase}} = 7.55$, $\bar{c}_{\text{decrease}} = 7.59$), but such pooling does not account for the dynamics of contributions.[6] Fig. 1 illustrates the time paths from the current experiment. In one sense, the current contributions data look similar to the same monitoring treatment of Carpenter (in press) in that punishment seems to maintain initial contribution levels, at best.[7] However, in another sense, the current data is markedly different because contributions seem to be affected by whether the price of punishment is increasing or decreasing.

The treatment effect in our contributions data seems reasonable from an economic point of view. When the price of punishment starts at a relatively low level and then increases over the course of the experiment, contributions fall steadily and more dramatically than when the price is constant (as in Carpenter, in press). One explanation, which we will confirm below, is that our players based their punishment decisions on price as well as on how egregiously the target free rode. On the one hand, when the price increased over time players bought less punishment causing the threat of punishment to abate. This led to more free riding. On the other hand, when the price fell players responded by buying more punishment per offense. In this case, the effectiveness of punishment increased over the course of the experiment, and although we see an initial drop in contributions, they recover as the price of punishment continues to fall.

Without accounting for the time paths or the amount of free riding, the average amount of punishment purchased in the decreasing price treatment ($\bar{s}_{\text{decrease}} = 2.13$) is marginally significantly greater ($t = 1.72$, $p = 0.08$) than the average amount purchased in the increasing price treatment ($\bar{s}_{\text{increase}} = 1.55$). Fig. 2 presents the time paths of the average expenditure on punishment. Even though this graph does not control for other factors that might have affected our

[6] In fact the pooled averages are not significantly different ($t = 0.09$, $p = 0.93$).

[7] We cautiously make this comparison because, although the monitoring network was identical, the Carpenter et al. (2004a) experiment used the Fehr and Gä chter payoff function.

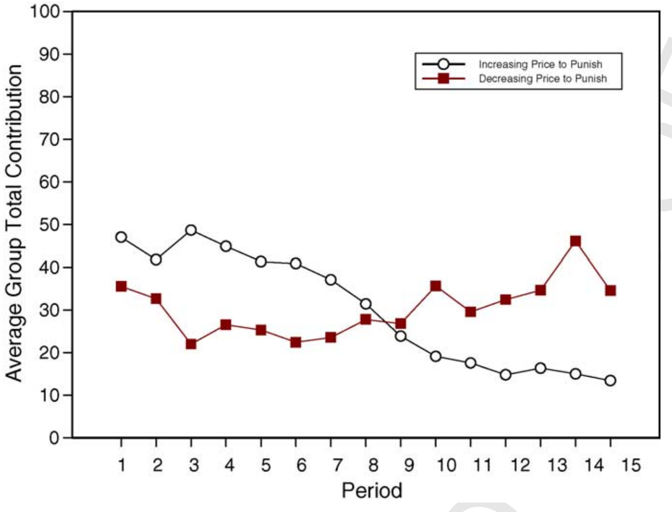

Fig. 1. The evolution of average group contributions over time (note: increasing price indicates that the price per sanction increased from 0.25 to 4 while decreasing price means the opposite).

players' punishment decisions (e.g., income or average level of free riding), it provides evidence consistent with the hypothesis that players reacted to the price of punishment and that this affected the credibility of punishment and the level of contributions. As the price increased, our players spent less on punishment. In fact, by the last three rounds of the increasing price treatment when it cost 4 EMUs to remove 1 EMU from the target, the players stopped punishing completely. In the other treatment, as the price fell, players spent more on and bought more punishment.

Because the instructions explicitly mentioned the order in which the price of punishment would change (see Appendix A on the JEBO website), one might worry that players anticipated

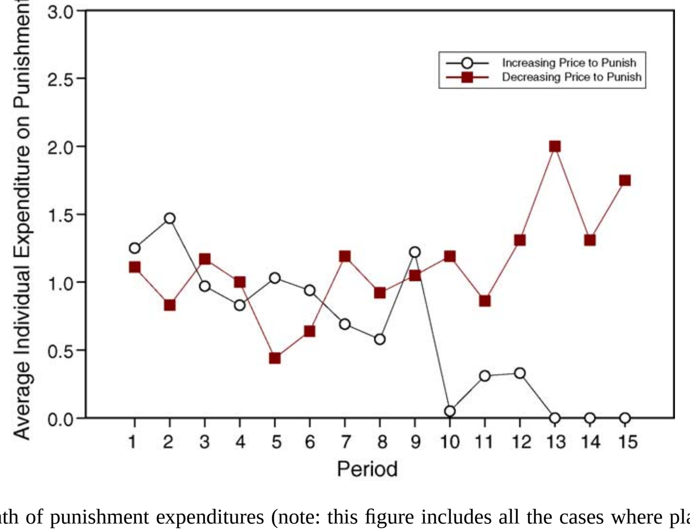

Fig. 2. The time path of punishment expenditures (note: this figure includes all the cases where players choose not to punish and does not control for how much free riding occurred).

and reacted in advance to the direction of the price change. For example, in the increasing price treatment, it might be reasonable to think that players spent more on early punishment than they would have had they not known that the price was going to increase or that players in the decreasing price treatment might have delayed punishment to later rounds when they knew it would be cheaper. If this is true then the slopes of the graphs in Fig. 2 are steeper than they would have otherwise been. In the next section we control for these differences when estimating the demand for punishment and see that they do not matter.

5. The demand for punishment

We now proceed by econometrically estimating the demand for punishment. One valuable benefit of using an experiment to elicit the data for our estimation is that we control for most of the problems that typically plague demand estimates. Specifically, simultaneity and identification are not problems for us because price is, by design, completely exogenous. However, we do face other issues. Because our experiment is 15 periods long, we generate a panel of data. To control for individual heterogeneity, all our regressions include random effects. Because there are a lot of observations where our players showed no preference for punishment, our dependent variable, the quantity of punishment purchased, is truncated from below at zero. For this reason, we use the Tobit procedure.[8] Finally, there is one criticism of the strangers matching protocol that has been, to this point, ignored in the literature. It could be the case that our point estimates will be biased by the fact that observations at the individual level are not independent because contaminants are generated when participants are randomly reshuffled into new groups at the beginning of each round. Of course, as is true of most procedural criticisms, this is an empirical question. In Appendix B (on the JEBO website) we offer a methodology for testing whether such a bias affects the coefficients that we calculate below. Saving the details for the readers of the appendix, the upshot is that we are confident that any strangers bias has a minimal effect on our estimates.

Before we present the fully controlled estimate of the demand for punishment, we present the reader with a graphical presentation of the main result. Fig. 3 illustrates the uncontrolled demand for punishment function based on a quadratic specification. As the reader can see, while the best fit is nonlinear, the slope of the demand function is relatively shallow, indicating that the reaction of participants who punish to price changes is rather small. However, there is some response, and it is important to note that Fig. 3 does not represent the behavior of the entire population of participants because there are a significant portion of our participants (32%) who did not punish at all. We return to the issue of punishing types in Section 6.

We can demonstrate the robustness of Fig. 3 by considering the regression results presented in Table 1. We build our econometric model in two stages. In stage one, we estimate the uncontrolled price and income elasticities. In stage two we add controls for how much the target free rode (the null hypothesis being one is punished more the more one free rides) and for how much the punisher free rode (the null, in this case, being that people who free ride less, punish more). Our definition of free riding is based on the results of Fehr and Gächter who show that people direct punishment at targets who contribute less than the group average. Hence, for our purposes *free riders* contribute less than the current group average, and *contributors* contribute at or above the average. In the

[8] Technically, punishment could also be limited by a participant's earnings. However, we can report that the "budget constraint" never bound our participants' desire for punishment. In all cases the difference between a participant's period earnings and her period expenditures on punishment was positive.

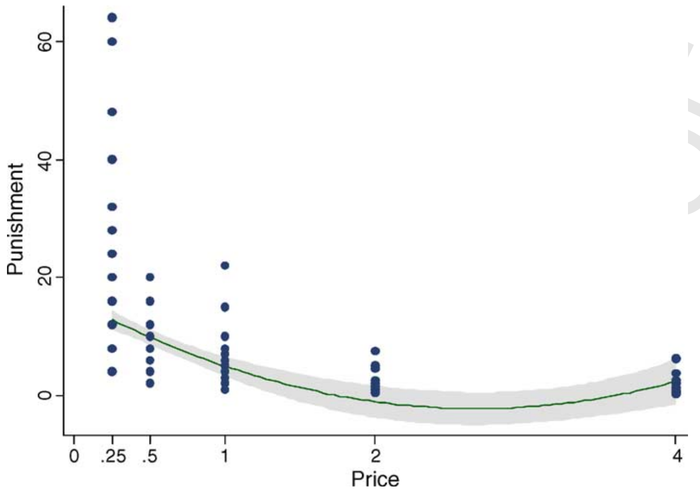

Fig. 3. The demand for punishment (note: each dot may represent several observations. The fitted values are based on a quadratic estimation of the effect of price on quantity. Shading represents the 95% confidence interval).

second stage we also control for the sex of the punisher, whether the punisher participated in the increasing price treatment or not, the differential effect of price changes in the increasing price treatment, and any time trend in our data. To control for the time trend we model a simple dynamic that says that the amount of punishment purchased by participant i in period t is a function of the contributions of the other group members, $C(\text{group total})_{-i,t-1}$, in period $t-1$.[9]

In the upper panel of Table 1 we report the marginal effects and standard errors of our regressions, and in the middle panel we report elasticities calculated at the regressor means.[10] Eq. (1) introduces our main result. Both the price and income elasticities are negative, which indicates that, given the average participant prefers to punish free riders, people react to economic incentives in what economists would consider reasonable fashion. In addition, demand appears to be slightly elastic with respect to price and inelastic with respect to income. Specifically, in the uncontrolled regression a 1% increase in price reduces the quantity of punishment demanded by 1.22% and a 1% increase in income decreases the amount of punishment demanded by 0.27%. At first blush, punishment appears to be ordinary and inferior.

Eq. (2) indicates that our initial elasticity estimates are not entirely robust to the inclusion of other punishment determinants. Part of the variation in punishment previously attributed to changes in price is actually caused by changes in how egregiously the target and the punisher free ride, but the coefficient on price (and its square) remain highly significant. At the same time, there

[9] Notice that we could have included period fixed effects as an alternative to modeling a dynamic, but this approach would be incorrect because doing so assumes there are time idiosyncrasies while what we need to control for appears to be (based on Figs. 1 and 2) a dynamic process.

[10] Calculating elasticities from Tobit coefficients is not straightforward because when one calculates the marginal effect, one has to account for the probability that a change in the regressor will push one past the “kink,” and the impact of a change in the regressor on the dependent variable, given it is uncensored. However, we can use the McDonald and Moffitt (1980) decomposition to calculate elasticities. With latent variable, p_{it} we have the following marginal effect:

$$\frac{\partial E(p_i|x_i)}{\partial x_{it}} = E(p_i^*|x_{i,} p_i^* > 0)\frac{\partial \Pr(p_i^* > 0)}{\partial x_{it}} + \Pr(p_i^* > 0)\frac{\partial E(p_i^*|x_i, p_i^* > 0)}{\partial x_{it}}.$$

Table 1
Dependent variable is the punishment inflicted on target$_{i,t}$

	(1) All types	(2) All types	(3) Free riders who punish cooperators	(4) Free riders who punish free riders	(5) Cooperators who punish free riders
Price$_{i,t}$	−3.00***(0.40)	−1.93***(0.49)	−0.10(0.63)	−3.30***(0.96)	−1.13(0.90)
Price$^2_{i,t}$	0.49***(0.09)	0.32***(0.09)	0.02(0.13)	0.52***(0.18)	0.17(0.15)
Income$_{i,t}$	−0.002***(0.0008)	−0.0007(0.001)	0.0001(0.0007)	−0.002(0.004)	−0.0004(0.002)
$C(\text{target})_{i,t} - C(\text{group average})_{i,t}$		−0.13***(0.02)	−0.0008(0.006)	−0.19***(0.05)	−0.13***(0.05)
$C(\text{punisher})_{i,t} - C(\text{group average})_{i,t}$		0.07***(0.02)	0.002(0.01)	0.04(0.04)	0.12***(0.05)
$C(\text{group total})_{-i,t-1}$		0.004(0.005)	0.001(0.005)	0.03**(0.01)	−0.007(0.006)
Female$_i$		0.10(0.40)	0.001(0.03)	−0.07(0.63)	−0.37(0.28)
Increasing price$_i$		−0.50(0.56)	0.02(0.12)	−1.27(1.39)	0.60(1.04)
Price$_{i,t}$ × Increasing price$_i$		−0.32(0.33)	−0.11(0.48)	0.25(0.79)	−0.99(0.65)
Price elasticity	−1.22	−0.79	−0.03	−1.03	−0.34
Income elasticity	−0.27	−0.09	0.02	−0.19	−0.04
Number of censored observations	845	795	44	234	211
N	1080	1008	56	322	308
Log likelihood	−1242.65	−998.21	−51.28	−414.67	−394.99
Wald χ^2	117	237	13	88	167
p-value	< 0.01	< 0.01	0.19	< 0.01	< 0.01

Note: (1) Tobits with individual random effects. (2) Coefficients are marginal effects. (3) Elasticities calculated at the regressor means. (4) The N in (2) is lower due to differencing. (5) Columns (3)–(5) restrict attention to three of five behavioral types identified in the data. (6) * indicates 0.10, ** indicates 0.05, *** indicates 0.01.

is a dramatic change in our estimate of the income elasticity. There is a simple explanation for why the income regressor shrinks. The punisher's income is highly correlated with her deviation from the group average contribution ($\rho = 0.75$, $p < 0.01$). Without controlling for the punisher's level of free riding, the income regressor picks up the variation due to both how much the punisher free rides and how much income is generated by her group.

Controlling for the other determinants of punishment, we continue to find that punishment is inferior based on the point estimate; however this estimate is actually not significantly different from zero, indicating that experimental income changes have little effect on the decision to punish. With respect to price, we find that punishment remains ordinary, but it is now less responsive to price. A one percent increase in the price of punishment leads to a 0.79% reduction in the quantity demanded. We conclude that the demand for punishment is ordinary, inferior, but inelastic.[11]

Eq. (2) also reveals other interesting facts about punishment, only some of which have been documented elsewhere. As is now common, we find that punishment decreases as the target increases her contributions ($p < 0.01$) and punishment is meted out primarily by people who themselves contribute a lot ($p < 0.01$). As for results one does not typically see in this literature, we find that the lag of the contributions by one's group mates has little effect on punishment decisions, females purchase more punishment (but not significantly more), and the increasing price treatment does not have a direct effect on the quantity of punishment purchased nor does it have a significant differential effect on the impact of price.

6. Types of punishers

Experimenters are becoming as interested in the heterogeneity of behavior as they used to be in average behavior (e.g., Fischbacher et al., 2001). Given this interest, we notice that there are six basic "punisher types" that can be found in our game. On one dimension players can contribute or not, and on another dimension, players can either not punish, punish free riders, or punish contributors. To implement this classification system, we sort players by their average contribution over the 15 periods (identifying those who contribute less than the average as free riders), and then for each individual, we regress the amount of punishment a player bought on the contribution of their target to get an estimate of each player's punishment propensity.[12] We find support for five of the six possible types; there are no players who contribute themselves and punish other contributors. The distribution of the five supported types is as follows: 5.56% are free riders who punish cooperators, 16.67% are free riders who never punish, 31.94% are free riders who punish other free riders, 15.28% are contributors who do not punish, and the remaining 30.56% are contributors who punish free riders. We refer to the first group as *principled free riders*, the third group as *hypocritical free riders*, and the last group as *principled cooperators*.

[11] Averaging across periods, treatments, and individuals those participants who punished spent 14% of their per period income on sanctions. Given this is a significant fraction of their earnings, a change in the price of punishment has a dramatic effect on their *real* budget constraints in the experiment. For this reason, we also decomposed our data into an income effect that is not picked up in our income elasticity, and a pure substitution effect. The relative size of the substitution effect accounts for more than 99% of the observed change. In other words, the income effect is negligible and the substitution effect is large.

[12] There were actually plenty of degrees of freedom to conduct these regressions because each individual generates 15 observations.

Returning to Table 1, we can assess whether individual types react differently to changes in price and income (as well as to the other determinants of punishment). In an unreported regression we stacked the data and created interaction terms to test whether any coefficient differences that we see in Table 1 are significant. At a minimum, the test of whether all the coefficients on the interactions were jointly different from zero is highly significant ($\chi^2 = 111.82$, $p < 0.01$), indicating some differences are important.

Notice that in the analysis of punishment, only three of the five types are important because two types never punish. First, in Eq. (3) we see that free riders who punish cooperators do not condition their punishment on any of the determinants of punishment. The fact that none of the regressors predict the behavior of this type of punisher may simply be due to the fact that we have few observations in this cell of the design. Alternatively, these principled free riders may feel as if they have "figured out" the game and are using punishment to indignantly signal to high contributors that they are being foolish. Another large portion of the players are hypocritical. In Eq. (4) we see that these hypocrites (those who free ride and punish other free riders) are more sensitive to price changes than any of the other types. The stacked regression suggests that our hypocrites are more sensitive to price changes than vengeful free riders, but the effect is only at the margins of significance ($p = 0.17$ for the first order comparison and $p = 0.16$ for the second order comparison). However, hypocrites are significantly more sensitive to price ($p < 0.10$ for both coefficients) than cooperators who punish free riders (Eq. (5)). It makes sense that although there may be some trace of morality to the punishment choices of the hypocrites, they are much more likely to shed their morals when the cost of punishing increases. At the same time, principled free riders and principled cooperators blindly follow established cultural norms (i.e., their responses are not significantly conditioned on the cost of punishment), albeit the norm followed is pro-social in only one of the two cases.

The other significant coefficients in Eqs. (3)–(5) provide more evidence supporting this normative story. Like principled cooperators, hypocrites condition their punishment on how much the target free rides while principled free riders do not. In addition, in Eq. (5) we see that those cooperators who are further above the group average punish more, but more interestingly, in Eq. (4) we see that only hypocrites are sensitive to the dynamic in contributions. The coefficient on the lag of the contributions of the others is positive (and significantly greater than the cooperator coefficient, $p = 0.03$), indicating that the more the others contributed in the past, the more hypocrites punish in the future. In sum, we find three types of punishers in our data: hypocritical free riders who are very sensitive to the parameters of the game and the reactions of the other players, principled cooperators who mostly condition their punishment on how badly a target free rides and care little for the payoff implications of punishing, and principled free riders who punish (almost indiscriminately) other group members, maybe because they want to express indignation for contributors.

7. Does punishment deter free riding?

Although it seems like an aside, to understand better the elasticities we calculated in Section 5, it is important to ask whether punishment is effective at attenuating free riding. This question is especially important if it turns out that punishment is not effective. In the pooled data, we find that the demand for punishment is ordinary, but inelastic, which suggests that the dominant force in our aggregate data is the behavior of the principled cooperators. We suspect these players of subscribing to a norm that compels them to sanction other players who are not acting in ways that contribute to group welfare. In other words (i.e., those of Elster, 1989) these players are not punishing for instrumental reasons. Instead principled cooperators punish for normative reasons, and as a result, they are less sensitive to the payoff implications of punishment.[13] Specifically, they do not punish to increase group welfare per se, they punish without much regard to the cost of doing so, and they punish even if it does not cause free riders to contribute more in the future.[14]

To examine whether punishment is effective in the current experiment, in Table 2 we look at the regression of individual contributions by player i in period t, $C(\text{individual})_{i,t}$, on the amount of punishment that this person received in period $t-1$. We include other regressors such as how much the other group members contributed last period to separate the effect of inertia from the effect of punishment, the player's sex, and controls for the direction of price changes and any differential effect of the price treatment on received punishment. We pool the data across punisher types in Eq. (1) and find that punishment has no effect on future contributions.[15] In fact, the only regressor with any significant explanatory power is $C(\text{group total})_{-i,t-1}$, which implies that increases in contributions in this experiment can be attributed to inertia, conformity (Carpenter, 2004), or conditional cooperation (Fehr and Fischbacher, 2003). Given the tiny coefficient on the lagged punishment, it must have been obvious to the participants that punishment did not deter free riding, yet as we see in Fig. 2, participants continued to purchase sanctions. As mentioned above, this can not be reconciled with the strategic use of punishment, but it is consistent with the normative motivations that we ascribe to principled cooperators.

In Eqs. (2) through (6) of Table 2, we dig a little deeper and examine the effect of being punished on our five behavioral types individually. Inertia seems to be the only common determinant of cooperation, although different types respond differently to the contributions of others It is interesting, for instance, that the inertial coefficients on the two cooperative types are three times as large as the coefficients on the three free riding types. These differences, significant in the stacked regression at the 2% level or better, suggest that cooperators are three times as motivated by conditional cooperation as free riders are; however, none of the punisher types are significantly motivated to contribute when, instead of inflicting it, they receive punishment.

There are two other interesting facts in the type-level regressions that are worth mentioning. Women who are vengeful free riders are also significantly more stubborn in their free riding. While not significant, the coefficient on the lag punishment regressor is negative, suggesting punishment makes these women free ride even more. This effect occurs in addition to the sex indicator being highly significantly negative with a very large marginal effect. This is the only substantial sex difference we find in our data.

Lastly, we see that principled cooperators (Eq. (6)) have significantly lower contributions in the increasing price treatment. A possible explanation for this coefficient is that principled cooperators "give up" at some point in the increasing price treatment. Early on they use punishment to repair the cracks in the dam that holds back the swell of free riding, but as the price of punishment

[13] Also see Carpenter and Matthews (2005) for a one-shot experiment that reinforces the normative interpretation of punishment.

[14] Actually, this model of behavior is not just speculation. For other evidence supporting this hypothesis, see Carpenter et al. (2004b), and for an experiment designed explicitly to test this idea see Carpenter and Matthews (2002).

[15] Finding that punishment is ineffective in a network in which each group member monitors and can punish only one other group member replicates the results of a similar treatment in Carpenter et al. (2004a).

Table 2
Dependent variable is $C(\text{individual})_{i,t}$

	(1) All types	(2) Free riders who punish cooperators	(3) Free riders who do not punish	(4) Free riders who punish free riders	(5) Cooperators who do not punish	(6) Cooperators who punish free riders
$C(\text{group total})_{i,t-1}$	0.17***(0.01)	0.09*(0.05)	0.05***(0.02)	0.10***(0.02)	0.19***(0.04)	0.27***(0.02)
Punishment received$_{i,t-1}$	0.06(0.10)	−0.41(1.01)	−0.13(0.12)	0.14(0.17)	0.56(0.65)	0.05(0.17)
Female$_i$	−0.28(1.40)	−3.90**(1.58)	−0.95(0.76)	0.10(0.94)	−1.00(2.98)	−1.49(1.08)
Increasing price$_i$	−0.59(1.12)	1.16(1.72)	−2.91***(0.48)	0.96(0.97)	3.63(3.12)	−3.52***(1.13)
Punishment received$_{i,t-1}$ × Increasing price$_i$	−0.15(0.19)	0.20(1.14)	0.19(0.18)	−0.05(0.27)	−0.46(0.91)	−0.36(0.57)
Number of censored observations	376	26	102	124	44	80
N	1008	56	168	322	154	308
Log likelihood	−2553.54	−129.78	−265.37	−786.82	−445.06	−445.06
Wald χ^2	247	12	200	41	27	162
p-value	< 0.01	0.04	< 0.01	< 0.01	< 0.01	< 0.01

Note: (1) Tobits with individual random effects. (2) Coefficients are marginal effects. (3) Columns (2)–(5) restrict attention to one of five behavioral types identified in the data. (4) * indicates 0.10, ** indicates 0.05, *** indicates 0.01.

increases, it becomes increasingly costly to keep the dam together, and at some point they simply get out of the way and stop contributing themselves.

8. Concluding remarks

At the beginning of this paper we pointed out that while laboratory experiments in economics have provided more puzzles than answers, we should not be too quick to conclude that the standard methodology of economics is inherently flawed. The results of the current experiment give us hope that after documenting and understanding anomalies such as social preferences, economic tools will remain informative. With this in mind, our analysis has demonstrated three things. One, we have replicated and extended the experiments suggesting that the average economic decision-maker will, at some personal cost, punish free riders who reduce the social efficiency of group interactions. Adding the current evidence to that of a number of other experiments illustrates that positing a preference for punishing free riders appears to be a reasonable addition to standard selfish preferences. Two, given we accept that people prefer to punish free riders, we have shown that the most basic economic analysis, the estimation of demand, illustrates that people react to price and income changes when they consider punishing free riders just as they react to changes in these variables when they consume more standard commodities. Specifically, the demand for punishment slopes downward and is relatively inelastic with respect to price and income. If punishment preferences are linked to normative behavior, then it makes sense that punishing behavior is relatively inelastic with respect to price and income because people punish primarily for social rather than economic reasons. Third, despite the relative inelasticity of the demand for punishment, we have shown that punishers are sensitive to the price of punishment but not sensitive to income changes that should allow one to punish more severely.

These results also dovetail nicely with some of the other experimental studies of mutual monitoring. While there have been a number of recent studies that have looked at the effects of different punishment mechanisms (see Section 2 and Decker et al., 2003), there has been little comparative static analysis. For example, we know next to nothing about the robustness of any given punishment scheme. For example, does the original Fehr and Gächter mechanism continue to elicit contributions if each punishment point removes 5% of the target's income instead of 10%? Likewise, while Sefton et al. find that punishment is better at controlling free riding than rewards, how much does this result depend on the relative magnitudes of punishments and rewards? Until there has been a systematic study of the determinants of punishing behavior similar to what Ledyard (1995) has done for public goods experiments, the punishment literature will remain a series of unconnected islands. The current experiment, however, extends previous work in which we fixed the punishment mechanism and varied the amount of information that group members have about each other (Carpenter, in press), the return on the public good, and the size of the groups (Carpenter, in press and Bowles et al., 2001).

The fact that many punishers in our experiment react little to the monetary consequences of their actions overlaps with similar situations in the real world including the interesting example of mutual monitoring among lobster fishermen along the Maine coast documented in Acheson (1988). In this example, fishermen monitor and punish others who extract too much from the local fishery, and when they do, they risk heavy monetary fines and imprisonment because their vigilante methods are often extreme and illegal (ranging from cutting trap lines to blowing up boats). In this sense, our finding that punishment is inelastic with respect to price is economically significant even if the exact estimate of the elasticity has little external validity.

Acknowledgement

I thank Julia Assael and Marla Weinstein for research assistance. I also thank James Andreoni, Steve Burks, Guillaume Frechette, Simon Gächter, Herb Gintis, Peter Matthews, Corinna Nölke, and two referees whose comments substantially improved this paper. This research is supported by Middlebury College and the National Science Foundation (SES-CAREER 0092953).

Appendix A. Participant instructions

You have been asked to participate in an experiment. For participating today and being on time you have been paid \$5. You may earn an additional amount of money depending on your decisions in the experiment. This money will be paid to you, in cash, at the end of the experiment. When you click the BEGIN button you will be asked for some personal information. After everyone enters this information we will start the instructions for the experiment.

During the experiment we will speak in terms of Experimental Monetary Units (EMUs) instead of Dollars. Your payoffs will be calculated in terms of EMUs and then translated at the end of the experiment into dollars at the following rate: 25 EMUs = 1 Dollar.

In addition to the \$5.00 show-up fee, each participant receives a lump sum payment of 10 EMUs at the beginning of the experiment.

The experiment is divided into 15 different periods. In each period participants are divided into groups of 4. The composition of the groups will change randomly at the beginning of each period. This means that in each period your group will consist of different participants.

Each period of the experiment has two stages.

A.1. Stage one

At the beginning of every period participants receive a 25 EMU endowment. In stage one participants decide how much of their 25 EMUs to contribute to a group project and how much to keep for themselves. Participants' payoffs are determined by the total contribution of their specific group and how much they individually keep.

To record their decisions, participants will type EMU amounts in two text-input boxes, one for the group project labeled GROUP ALLOCATION and one for themselves labeled PRIVATE ALLOCATION. These boxes will be yellow. Once a participant makes a decision, he or she will record his or her decision by clicking on the green SUBMIT button.

After all the participants have made their decisions, you will each be informed of your gross earnings for the period.

Participant gross earnings will consist of two parts:

(1) Earnings from the private allocation. Individuals are the only beneficiary of EMUs they keep. Specifically, each EMU a participant keeps increases that person's earnings by one.
(2) Earnings from the group project. Each member of a group gets the same payoff from the group project regardless of how much he or she contributed. The payoff from the group project is calculated by multiplying 0.5 times the total EMUs contributed by the members of the group.

Participant gross earnings can be summarized as follows:

$$1 \times (\text{EMUs you keep}) + 0.5 \times (\text{Total EMUs contributed by your group})$$

Let's discuss three examples.

Example 1. Say each member of a group contributes 15 of the 25 EMUs. In this case, the group total contribution to the project is $4 \times 15 = 60$ EMUs. Each group member earns $0.5 \times 60 = 30$ EMUs from the project. The gross earnings of each member will then be the number of EMUs kept, 25–15 = 10, plus the earnings from the group project, 30 EMUs, for each member. In total, each member would earn $10 + 30 = 40$ EMUs.

Example 2. Now say everyone in the group contributes 5 EMUs. Here the group total contribution will be 20 and each member will earn $0.5 \times 20 = 10$ EMUs from the group project. This means that the total earnings of each member of the group will be 20 (the number of EMUs kept) plus 10 (earnings from the group project) which equals 30 EMUs.

Example 3. Finally, say three group members contribute all their EMUs and one contributes none. In this case, the group total contribution to the project is $3 \times 25 = 75$ EMUs. Each group member earns $0.5 \times 75 = 37.5$ EMUs from the project. The three members who contributed everything will earn $0 + 37.5 = 37.5$ EMUs and the one member who contributed nothing will earn $25 + 37.5 = 62.5$ EMUs.

A.2. Stage two

In stage two participants will be shown the allocation decision made by one other randomly selected member of their group. Everyone's choice will be seen by exactly one other group member and the person you see is different from the person seeing you. In addition to seeing another group member's choice, at this stage participants can reduce the earnings of the group member they see, if they want to.

Participants will be shown how much one member of their group kept and how much this person allocated to the group project. Participants will also see their own allocation decision and this decision will be labeled 'YOU'.

At this point participants will decide how much (if at all) they wish to reduce the earnings of the other group member they are seeing. Participants reduce someone's earnings by typing the number of EMUs they wish to spend to reduce that person's earnings into the input-text box that appears below the other group member's allocation decision.

Participants can spend as much of their accumulated earnings as they want to reduce the earnings of the other group member. For each EMU spent by a participant the earnings of the other group member will be reduced by R EMUs. The value of R will change during the experiment.

[Price decrease] The experiment is divided into 5 blocks of 3 periods and the value of R will change every 3 periods according to the following sequence {0.25, 0.5, 1, 2, 4}. For example, during the first 3 periods of the experiment R will be 0.25 so spending 1 EMU will reduce the other group member's earnings by 0.25 EMUs. During the third block of periods R will equal 1 and spending 1 EMU will reduce the other group member's earnings by 1 EMU. During the final block R will equal 4 and spending 1 EMU will reduce the other group member's earnings by 4 EMUs.

[Price increase] The experiment is divided into 5 blocks of 3 periods and the value of R will change every 3 periods according to the following sequence {4, 2, 1, 0.5, 0.25}. For example, during the first 3 periods of the experiment R will be 4 so spending 1 EMU will reduce the other group member's earnings by 4 EMUs. During the third block of periods R will equal 1 and spending 1 EMU will reduce the other group member's earnings by 1 EMU. During the final block R will equal 0.25 and spending 1 EMU will reduce the other group member's earnings by 0.25 EMUs.

Consider this example: suppose someone spends 2 EMUs to reduce the earnings of the other group member when R is 0.5. This expenditure reduces the other group member's earnings by 1 EMU ($2 \times 0.5 = 1$). When participants have finished stage two they will click the blue DONE button.

Participant Net Earnings in each period will be calculated as follows:

(Gross earnings from stage one) – (R times the number of EMUs spent on reductions directed towards the participant) – (the participant's expenditure on reductions directed at someone else).

If you have any questions please raise your hand. Otherwise, click the red FINISHED button when you are done reading.

Appendix B. Assessing any strangers bias

In the *strangers* matching protocol all the participants in an experimental session are randomly re-matched into groups at the beginning of every decision-making period. This protocol is particularly valuable as a tool to gather more observations with the same number of participants as well as to examine the effects of learning while controlling, as much as is practical, for repeated game effects. Even though most experiments have end points that are common knowledge, the typical participant finds the difference between finitely repeated play and infinitely repeated play much less compelling than theorists do. As a practical matter this means that participants who interact in the same group (a.k.a., the *partners* protocol), despite knowing the exact number of interactions, fail to undertake the necessary backward induction required to see through pseudo-folk theorem like reasoning. As a result, researchers rely on the strangers protocol to come as close to a series of one-shot encounters as possible. However, as many researchers (and referees) have pointed out, when it comes to econometric analysis, the strangers protocol may cause violations of the assumption that observations at the individual level are independent. Taken to the extreme, such a critique implies that one can only conduct analyses at the session level because this is the level at which independence is guaranteed.

At first blush, this argument seems both correct and paralyzing for experimental research because either one needs to conduct a huge number of sessions or create protocols that can run with as few participants as possible. It is obvious that increasing the cost of experimental research by running many more sessions is a problem, but one should not disregard the incentive to run experiments on smaller groups in smaller sessions because there are many experiments for which group size and anonymity matter. However, as is true about most questions concerning experimental methodology, this is really an empirical question. While the logic is sound, if such dependence among individual decisions leads to no or only minor bias in the important point estimates, then it seems imprudent to neglect analyses at the individual level. To this point, most researchers simply ignore this stranger's bias and hope that their referees will not call them on it. However, in what follows, we offer two methods for assessing the magnitude of this potential bias.

In Table B.1 we report the results of two tests of the stranger's bias in our demand for punishment data. In columns (1) through (3) we compare three different estimates using our punishment data. Notice that these regressions do not contain all the same variables as our preferred model in

Table B.1
Dependent variable is the punishment inflicted on $\text{target}_{i,t}$

	(1) Tobit	(2) Tobit	(3) Tobit	(4) GLLAMM	(5) GLS
$\text{Price}_{i,t}$				−4.44***(0.54)	−4.46***(0.54)
$\text{Price}^2_{i,t}$				0.78***(0.11)	0.78***(0.11)
$\text{Income}_{i,t}$	0.36(0.44)	0.39***(0.15)	0.61***(0.13)	0.001(0.003)	0.001(0.002)
$C(\text{target})_{i,t} - C(\text{group average})_{i,t}$	−1.32***(0.49)	−1.01***(0.12)	−1.02***(0.12)	−0.23***(0.03)	−0.22***(0.03)
$C(\text{punisher})_{i,t} - C(\text{group average})_{i,t}$	0.15(0.60)	0.97***(0.21)	1.26***(0.18)	0.13***(0.03)	0.13***(0.03)
$C(\text{group total})_{i,t} - C(\text{group total})_{i,t-1}$		0.10***(0.04)		0.02*(0.01)	0.02*(0.01)
Female_i	2.18(4.51)	−0.09(2.48)	−0.03(2.55)	−0.31(0.50)	−0.31(0.48)
Increasing price_i				−1.09(0.90)	−1.13(0.89)
$\text{Price}_{i,t} \times$ Increasing price_i				0.23(0.51)	0.25(0.51)
Number of censored observations	50	795	795		
N	72	1008	1008	1008	1008
Log likelihood	−109.71	−1070.77	−1074.25	−2999.78	
Wald χ^2	9	149	143		292
p-value	0.06	< 0.01	< 0.01		< 0.01

Note: (1) Eqs. (1)–(3) are Tobits censored at 0. (2) The GLLAMM procedure stands for generalized linear latent and mixed models and incorporates random effects at both the individual and session levels. (3) The generalized least squares (GLS) model incorporates individual random effects. (4) * indicates 0.10, ** indicates 0.05, *** indicates 0.01.

Table 1. We do not include any of the price variables because they do not vary over their entire range in period one. For example, the first order price regressor takes only two values, 0.25 or 4. In column (1) we estimate the relationship using just the data from period one that implies there can be no stranger's bias affecting the coefficients. In columns (2) and (3) we use the data from periods two through fifteen. If the random reshuffling of partners causes problems in our data, we should be able to identify whether or not the effect is large enough to bias our point estimates by comparing the coefficients based on the period one data to the coefficients based on the remaining data. This is a simple Hausman test. What complicates the matter, however, is that learning may occur over the course of fifteen periods. This is why we compare column (1) to columns (2) and (3). Columns (2) and (3) differ in that column (2) controls (in an admittedly simplistic, but agnostic way) for learning by adding the lagged contributions of the other group members. As one can see, the coefficients look similar, but not exactly the same so the Hausman results are important. Comparing columns (1) and (3), which does not include a control for learning, the p-value on the Hausman test is 0.21 and we can not reject the hypothesis that the coefficients are the same (i.e., there is minimal stranger's bias). This result does not change much when we add the learning control. In this case we find $p = 0.27$. Based on this method, we conclude that the stranger's bias is not important in our data.

A second way to identify an effect of random re-matching (in which we use all our regressors) is to compare the standard method of analysis used in this literature, generalized least squares with individual random effects, which ignores any potential stranger's bias, to the generalized linear latent and mixed models (GLLAMM) approach. In the GLLAMM procedure we include individual random effects and session level random effects to capture the stranger's bias. As one can see, the difference in the coefficients between the GLLAMM model (column (4)) and the GLS model (column (5)) are almost imperceptible, which provides further evidence that the stranger's bias does not affect the coefficients of our estimates of the demand for punishment. Unfortunately, so far a GLLAMM model has not been developed for the Tobit regressor, but given the results of our two methods for assessing the magnitude of the stranger's bias in our data, we have confidence in the random effect Tobit regressions reported above.

References

Acheson, J., 1988. The Lobster Gangs of Maine. University Press of New England, Hanover.

Anderson, C., Putterman, L., 2005. Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism. Games and Economic Behavior 54, 1–24.

Andreoni, J., 1988. Why free ride? Strategies and learning in public good experiments. Journal of Public Economics 37, 291–304.

Andreoni, J., Castillo, M., Petrie, R., 2003. What do bargainers' preferences look like? Experiments with a convex ultimatum game. American Economic Review 93, 672–685.

Andreoni, J., Miller, J., 2002. Giving according to GARP: An experimental test of the consistency of preferences for altruism. Econometrica 70, 737–757.

Andreoni, J., Vesterlund, L., 2001. Which is the fair sex. Quarterly Journal of Economics 116, 293–312.

Berg, J., Dickaut, J., McCabe, K., 1995. Trust, reciprocity and social history. Games and Economic Behavior 10, 122–142.

Bochet, O., Page, T., Putterman, L., in press. Communication and punishment in voluntary contribution experiments. Journal of Economic Behavior and Organization.

Bowles, S., Carpenter, J., Gintis, H., 2001. Mutual monitoring in teams: theory and evidence on the importance of residual claimancy and reciprocity. mimeo.

Camerer, C., 1995. Individual decision making. In: Kagel, J., Roth, A. (Eds.), The Handbook of Experimental Economics. Princeton University Press, Princeton 588–683.

Camerer, C., Fehr, E., 2001. Measuring social norms and preferences using experimental games: A guide for social scientists. In: Henrich, J., Boyd, R., Bowles, S., Gintis, H., Fehr, E., et al. (Eds.), Foundations of Human Sociality: Experimental and Ethnographic Evidence from 15 Small-Scale Societies. Oxford University Press, Oxford 55–95.

Camerer, C., Thaler, R., 1995. Anomalies: Ultimatums, dictators and manners. Journal of Economic Perspectives 9 (2), 209–219.

Carpenter, J., 2002. Measuring social capital: adding field experimental methods to the analytical toolbox. In: Isham, J., Kelly, T., Ramaswamy, S. (Eds.), Social Capital and Economic Development: Well-Being in Developing Countries. Edward Elgar, Northampton 119–137.

Carpenter, J., in press. Punishing free-riders: How group size affects mutual monitoring and the provision of public goods. Games and Economic Behavior.

Carpenter, J., 2004. When in Rome: Conformity and the provision of public goods. Journal of Socio-Economics 33, 395–408.

Carpenter, J., Daniere, A., Takahashi, L., 2004a. Social capital and trust in southeast Asian cities. Urban Studies 41, 853–874.

Carpenter, J., Matthews, P., 2002. Social reciprocity, Middlebury College Department of Economics Working Paper 02–29.

Carpenter, J., Matthews, P., 2005. Norm enforcement: Anger, indignation, or reciprocity, Department of Economics, Middlebury College, Working Paper 05–03.

Carpenter, J., Matthews, P., Ong'ong'a, O., 2004b. Why punish? Social reciprocity and the enforcement of prosocial norms. Journal of Evolutionary Economics 14, 407–429.

Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. Quarterly Journal of Economics 117, 817–870.

Davis, D., Holt, C., 1993. Experimental Economics. Princeton University Press, Princeton.

Decker, T., Stiehler, A., Strobel, M., 2003. A comparison of punishment rules in repeated public goods games: An experimental study. Journal of Conflict Resolution 47, 751–772.

Eckel, C., Grossman, P.J., 1996. The relative price of fairness: Gender differences in a punishment game. Journal of Economic Behavior and Organization 30, 143–158.

Elster, J., 1989. The Cement of Society: A Study of Social Order. Cambridge University Press, Cambridge.

Fehr, E., Fischbacher, U., 2003. The nature of human altruism. Nature 425, 785–791.

Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. American Economic Review 90, 980–994.

Fehr, E., Kirchsteiger, G., Riedl, A., 1993. Does fairness prevent market clearing? An experimental investigation. Quarterly Journal of Economics 108, 437–459.

Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Economic Letters 71, 397–404.

Güth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. Journal of Economic Behavior and Organization 3, 367–388.

Isaac, R.M., Walker, J., Thomas, S., 1984. Divergent evidence on free-riding: An experimental examination of possible explanations. Public Choice 43, 113–149.

Latane, B., Darley, J., 1970. The Unresponsive Bystander: Why Doesn't He Help?. Appleton-Century-Crofts, New York.

Ledyard, J., 1995. Public goods: a survey of experimental research. In: Kagel, J., Roth, A. (Eds.), The Handbook of Experimental Economics. Princeton University Press, Princeton 111–194.

Masclet, D., Noussair, C., Tucker, S., Villeval, M.-C., 2003. Monetary and nonmonetary punishment in the voluntary contributions mechanism. American Economic Review 93, 366–380.

McDonald, J., Moffitt, R., 1980. The uses of Tobit analysis. Review of Economics and Statistics 62, 318–321.

Ostrom, E., Walker, J., Gardner, R., 1992. Covenants with and without a sword: Self-governance is possible. American Political Science Review 86, 404–417.

Pillutla, M., Murnighan, K., 1996. Unfairness, anger, and spite: Emotional rejections of ultimatum offers. Organizational Behavior and Human Decision Processes 68, 208–224.

Sefton, M., Shupp, R., Walker, J., 2001. The effect of rewards and sanctions in provision of public goods, Department of Economics Indiana University Working Paper.

Suleiman, R., 1996. Expectations and fairness in a modified ultimatum game. Journal of Economic Psychology 17, 531–554.

Tversky, A., Slovic, P., Kahneman, D., 1990. The causes of preference reversals. American Economic Review 80, 204–217.

Zwick, R., Chen, X.-P., 1999. What price for fairness? A bargaining study. Management Science 45, 804–823.