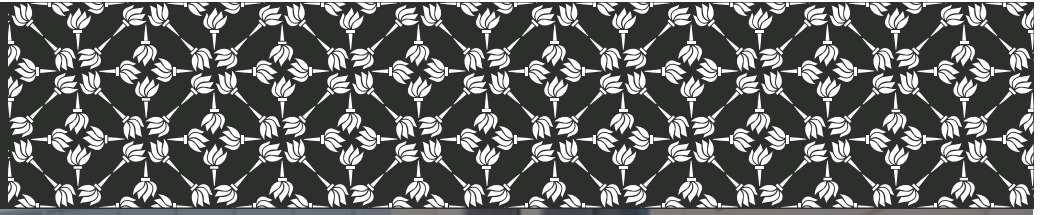


جامعة نيويورك أبوظبي

 NYU | ABU DHABI



Experimental Evidence Shows That Negative Motive Attribution Drives Counter- Punishment

**Manuel Muñoz-Herrera and Nikos
Nikiforakis**

**Working Paper # 0056
October 2020**

Division of Social Science Working Paper Series

New York University Abu Dhabi, Saadiyat Island P.O Box 129188, Abu Dhabi, UAE

<http://nyuad.nyu.edu/en/academics/academic-divisions/social-science.html>

Experimental evidence shows that negative motive attribution drives counter-punishment*

Manuel Muñoz-Herrera[†] Nikos Nikiforakis[‡]

Abstract

Evidence shows that the willingness of individuals to avenge punishment inflicted upon them for transgressions they committed constitutes a significant obstacle towards upholding social norms and cooperation. The drivers of the desire to counter-punish, however, are not well understood. We hypothesize that negative motive attribution – the tendency to assign negative motives to punishers for their actions – increases the likelihood of counter-punishment. We test this hypothesis in a lab experiment in which we exogenously manipulate the ability to attribute negative motives to punishers by having the punisher be either an unaffected third party or the victim of a transgression (second party). We show that individuals consider second-party punishment to be substantially more biased than an identical, payoff-equalizing punishment meted out by a third party. In line with our hypothesis, we find that second-party punishers are 66.3% more likely to be counter-punished than third-party punishers, and suffer a loss in earnings which is 64.6% higher, all else equal. Our findings have implications for designing mechanisms to uphold cooperation and reduce conflict.

Keywords: social norms, counter-punishment, altruistic punishment, cooperation

JEL Codes: C92; D70; H41

* The authors would like to thank Loukas Balafoutas and Aurelie Dariel for helpful comments, as well as Hamna Kahn and Marek Mihok for research assistance.

[†] Division of Social Science, New York University Abu Dhabi; email: manumunoz@nyu.edu

[‡] Division of Social Science, New York University Abu Dhabi; email: nikos.nikiforakis@nyu.edu

1. Introduction

Social norms often serve to enhance welfare by reducing externalities and conflict (e.g., Coleman, [1990](#); Elster, [1989](#); Young, [2015](#)). A key reason individuals often adhere to social norms in spite of opposing incentives is the fear of being punished by others (Sugden, [1986](#); Coleman, [1990](#)). Evidence from lab (Balafoutas et al. [2014a](#); Denant-Boemont et al., [2007](#); Hopfensitz and Reuben, [2009](#); Nikiforakis, [2008](#)) and field experiments (Balafoutas and Nikiforakis, [2012](#); Balafoutas et al. [2014b](#); Balafoutas et al. [2016](#)), however, shows that the willingness to avenge punishment reduces greatly people's desire to punish norm violators. In other words, the fear of counter-punishment emerges as a significant obstacle towards upholding social norms and cooperation.

Despite the importance of the issue, the determinants of counter-punishment are still not well understood. None of the models of social preferences used to explain the incidence of punishment can readily account for the willingness to counter-punish (Nikiforakis, [2008](#)).¹ What drives individuals to avenge punishment? A possible explanation is that counter-punishers act out of a desire to avenge a *perceived* wrong. But why would they perceive as wrong a punishment that is inflicted upon them for a transgression? One apparent reason is that they believe they did not commit a transgression and, hence, do not deserve to be punished.² Another explanation is that the punishment is deemed to be too harsh and unfitting of their transgression.³ Here, we consider a more subtle explanation – although no less important in our opinion – for the occurrence of counter-punishment: victims believe the *punisher* is biased, irrespective of his/her actions.

At the most fundamental level, for victims to regard punishers as biased, they must ascribe negative (ulterior) motives to them for their actions. In other words, victims must believe punishment was not driven purely by a concern for upholding a norm. For instance, a transgressor may believe the punisher acted the way he did because he dislikes members of her ethnic/racial/linguistic/gender group or because he carries a grudge against her for a past transgression. Herein, however, lies the challenge: just because a punisher *could have* other motives, does not mean he does (e.g., not all whites are racists, not all men are sexists, and not everyone holds

¹ For instance, in models of inequality aversion (Bolton and Ockenfels, [2000](#); Fehr and Schmidt, [1999](#)) and reciprocity (Rabin, [1993](#); Falk and Fischbacher, [2006](#)) individuals punish to restore equality or fairness, respectively. This leaves no room for counter-punishment, even off-the-equilibrium path. A common feature of these models is that individuals use their individual *outcomes* to decide whether to confront another person or not; the *motives* of others are not part of one's decision process.

² In line with this, Nikiforakis et al. ([2012](#)) show that the likelihood of counter-punishment is substantially higher when individuals hold conflicting normative views.

³ In line with this, evidence shows that the level and likelihood of counter-punishment increase with the severity of punishment (Denant-Boemont et al., [2007](#); Nikiforakis, [2008](#)).

grudges), or that ulterior motives affected the decision to punish. Punishment victims, therefore, are called to make a judgement about the (unobserved) motives of the punisher. Can they make the right judgment, or do they tend to over-ascribe negative motives to punishers? And if so, do negative motives increase the propensity to counter-punish?

We present results from a lab experiment designed to test the hypothesis that the desire to avenge punishment following a transgression is driven by a tendency to ascribe negative motives to punishers for their actions. Specifically, we hypothesize that individuals are more likely to perceive as biased a punishment that comes from an individual that *could be* acting out of ulterior motives – irrespective of whether the punisher actually acted in a biased way – and will be hence more inclined to retaliate such punishment. To test this hypothesis, we use a variant of a design developed by Fehr and Fischbacher (2004). In our experiment, a person can take away some of the money that was allocated by the experimenter to another participant (the “transgression”). An individual who, depending on the experimental condition can be either the victim of the transgression (second party) or an “unaffected” third party, has the opportunity to punish the transgressor. The transgressor is then given a chance to retaliate punishment.

As second-party punishment *could be* motivated by a desire to avenge a transgression against oneself, our hypothesis is that second-party punishment will be perceived as being more biased than third-party punishment *all else equal*, and hence more likely to trigger counter-punishment. To obtain clear test results for our hypothesis, we let participants play the game only once, thus ruling out any strategic motives for punishment and counter-punishment. Importantly, to prevent differences in punishment from affecting the willingness to counter-punish, we elicit the latter using the strategy method. This implies that any differences in the willingness to counter-punish our experiment, cannot be explained by differences in *outcomes* (i.e., the take rate or the extent of punishment). Differences in counter-punishment, therefore, must be attributed to differences in the attribution of motives.

The data lends strong support to our hypothesis. First, we show that the majority of participants (77%) perceives a *payoff-equalizing* punishment by a second-party following a transgression as being biased. The majority of participants (52%) also finds second-party punishment to be more biased than an *identical* third-party punishment (29% find it to be as biased). Second, in line with our hypothesis that negative motive attribution drives counter-punishment, we find that second-party punishers are 66.3% more likely to be counter-punished than third-party punishers for the same punishment, and suffer a loss in earnings which is 66% higher. These findings not only provide a novel explanation for the willingness of individuals to avenge punishment, but also suggest the important role that third parties have in upholding social norms and cooperation.

Previous evidence suggested that second parties are more likely to punish norm violations than third parties (e.g., Fehr and Fischbacher, 2004; Kurzban et al. 2007). However, this evidence is obtained in settings where counter-punishment is not possible. In our experiment, second and third parties punish transgressions equally in the presence of counter-punishment.

A related literature in psychology explores motive attribution and attribution bias. Attribution bias refers to the systematic errors people make when evaluating their own and others' behaviors (e.g., Tetlock and Levi, 1982). There are several different types of attribution bias. For instance, the “fundamental attribution error”, one of the most well-known of the attribution biases, refers to the tendency of individuals to overemphasize the role of personal (dispositional) factors in determining an outcome, while minimizing the influence of situational factors (Ross, 1977). Our study, however, is most closely related to studies exploring *motive* (as opposed to *outcome*) attribution (Heckhausen, 1991). Reeder et al. (2005) find that individuals tend to ascribe negative motives to people with whose opinions they disagree (e.g., about the motives behind the Iraq war). Mayz et al. (2014) show how adversaries in political and ethnoreligious intergroup conflict (Democrats/Republicans; Arab/Israelis) tend to attribute the aggression of their own group to in-group love more than out-group hate, but that of their opponents to out-group hate more than in-group love. Our study differs in several respects. In the controlled environment of a laboratory, we are able to show that the mere possibility of an individual acting on ulterior motives (i.e., when individuals are randomly assigned to roles), can affect how their actions are perceived. Importantly, we show that negative motive attribution can lead individuals to take costly actions, in particular, increase the extent to which they engage in costly counter-punishment.⁴ As we discuss in the concluding section, in certain circumstances, for instance when violators are outgroup members, the tendency to attribute negative motives to individuals can endanger cooperation and lead to costly escalation.

2. Research design

Our research design consists of a survey and an experiment. The survey was designed to evaluate whether individuals tend to attribute negative (ulterior) motives to punishers personally affected by a transgression, and evaluate whether they believe these second-party punishers are regarded as being more biased than third parties that were not personally affected by the transgression. The experiment was designed to test whether, in line with our hypothesis, second-

⁴ Note that we prefer not to use the term “bias” in our study as the data do not allow us to make claims about the existence of judgement errors.

party punishment is more likely to elicit counter-punishment than third-party punishment, all else equal. We first present the design of the experiment, and proceed to discuss the survey.

2.1 *The experiment*

The experiment consists of two conditions using a between-subject design. The conditions differ only in who gets to punish a transgression and, by extension, who is the target of counter-punishment. In the *Second-Party* condition (2P) subjects are randomly assigned the role of Player A or B, and then play the following three-stage game once. At the beginning of Stage 1, Players A and B are each endowed with an initial income of 30 Experimental Monetary Units (EMU).⁵ In Stage 1, Player A must decide whether to take 0, 5 or 10 EMU from Player B's endowment. In Stage 2, Player B must decide whether to reduce A's earnings by any positive percentage in 10% increments, contingent on Player A's take decision in Stage 1. We elicit the willingness to punish in Stage 2 using the strategy method, i.e., Player B has to indicate the level of punishment for each amount that A could have deducted from his income. Punishment is costly. Similar to Nikiforakis et al. (2012), Player B must pay 5 EMU to reduce player A's income, irrespective of the percentage chosen. For this, B is given an additional endowment of 5 EMU. Finally, in Stage 3, Player A must decide whether to counter-punish Player B. Income reduction in Stage 3 works exactly as in Stage 2. Player A receives 5 EMU and must state the (percentage) level of counter-punishment for all eleven possible levels of incurred punishment. The advantage of using this punishment technology for our purposes is that differences in behavior in stages 1 and 2 across conditions cannot affect individuals' ability to counter-punish.

The *Third-Party* condition (3P), is identical to condition 2P with the difference that a third party can punish the transgressor (i.e., someone who is not a victim of A's transgression in Stage 1). To facilitate comparisons between 2P and 3P, we follow the protocol used by Fehr and Fischbacher (2004). Specifically, Player B decides whether to reduce the income of a Player A from a *different* group, instead of the Player A that he was paired with in Stage 1. Thus, Player B is an unaffected third-party when deciding whether to reduce the income of that other Player A. Conversely, Player A's decision of reducing B's income (counter-punishment) is made towards the Player B who punished her (the independent observer), and not the player B she was paired with in Stage 1.

One might be concerned that, by using the matching procedure of Fehr and Fischbacher (2004), the punisher may be perceived as being an ingroup member of the victim of the transgression (e.g., the "Player B" group). If this is the case, then we would expect third-party punishers to be

⁵ We used the exchange rate: 1 EMU = 2 Emirati Dirhams (AED), which is approx. 0.55 USD.

perceived as more biased than they would have been perceived if we had recruited a third party whose sole purpose in the experiment was to decide whether to punish A in 3P. Hence, we would expect any differences in counter-punishment between 2P and 3P to be smaller in our setting. In this case, our design offers a harder test for our hypothesis.

2.2 The survey

The purpose of the survey is to provide support for the idea that individuals ascribe more negative (ulterior) motives to second-party punishers. In order to obtain clear evidence, we decided to administer a survey to a set of individuals drawn from the same pool of volunteers but who did not participate in the experiment.

Participants were presented with the following vignette:

“Imagine the following scenario. You are participating in a research study. You and another individual (Person B) are each given \$30 for your participation. When given the opportunity, you take \$10 away from Person B. Person B has the option to respond to your action and chooses to reduce your earnings by \$20. This does not have a monetary benefit for him/her, i.e., Person B does not get to keep the \$20.”

After reading the scenario, participants were asked to evaluate the action of Person B. Specifically, they were asked to state whether, in their opinion, the action of B was biased, somewhat biased, neither biased nor unbiased, somewhat unbiased, or unbiased. Importantly, the specific scenario was chosen such that the punishment equalized the earnings of Person A and Person B. In this regard, individuals have fewer reasons to consider the punishment biased.

Following that, participants were presented with another vignette:

“Now imagine a different scenario. You are participating in another research study. You and another individual (Person B) are each given \$30 for your participation. When given the opportunity, you take \$10 away from Person B. Person B does NOT have the option to respond to your action. Instead, a third participant (Person C), who observed what happened, can respond. Person C does not know Person B personally. Person C chooses to reduce your earnings by \$20. This does not have a monetary benefit for him/her, i.e., Person C does not get to keep the \$20.”

After reading the second scenario, participants were asked to compare it to the first one. In particular, they were asked whether the punishment was more, less or equally biased in the two scenarios. Importantly, the only difference to the first scenario is that punishment is meted out by Person C; the transgression and the size of the punishment are identical. We chose to use a within-

subject design for our survey as we believe it poses a more difficult test for our hypothesis. The reason is that respondents can see that the two instances of punishment are identical in all respects, except for who carries it out.

2.3 Procedures

A total of 207 individuals were recruited for the study. All individuals were drawn from the pool of volunteers of the Social Science Experimental Laboratory at New York University Abu Dhabi. The survey was completed by 111 individuals. A separate group of 96 students participated in the experiment (40 in 2P and 56 in 3P) that took place at the Social Science Experimental Laboratory at New York University Abu Dhabi.⁶ After all subjects made their decisions in each condition, all participants were asked to complete a standard Big Five Markers test (Goldberg, 1992). Each session lasted approximately 70 minutes and average earnings were 103 AED (approx. 28 USD). The experimental instructions can be found in Online Appendix C.

3. Results

3.1 Are second-party punishers perceived to be more biased than third-party punishers?

We start our analysis of the data by exploring the extent to which individuals find second- and third-party punishment to be biased. Panel A in Figure 1 shows that a clear majority of respondents (77%) found the decision of Person B to punish A for taking money from him to be biased.⁷ The data rejects the hypothesis that punishment is perceived as unbiased (Wilcoxon sign rank, $N=111$, two-tailed, p -value <0.001). Given that the punishment followed a transgression from Person A and was constructed to restore the equality in payoffs of A and B, the evidence strongly supports our hypothesis that the existence of ulterior motives for punishment (in this case, personal revenge for a transgression committed against B) leads individuals to attribute negative motives to the punisher, even if the punishment itself restores equality in payoffs between A and B.

⁶ In the absence of prior studies on the topic to inform our power calculations, we aimed to recruit 100 individuals for each treatment comparison. This would permit us to detect a 14 (28) percentage point difference in the willingness to counter-punish as significant at the 5-percent level 80-percent of the time using a one-tailed (two-tailed) χ^2 test. Unfortunately, the unexpected shutdown of laboratories due to the COVID-19 pandemic, prevented us from doing so. Our smaller sample would allow us to detect a difference of 18.5 (37) percentage points in the willingness to counter-punish as significant at the 5-percent level 80-percent of the time. Given that (i) it remains unknown when the forced shutdown will end, (ii) the difference in power is arguably small, especially in light of our directed hypothesis which means that one could reasonably rely on one-tailed tests, (iii) the fact that the difference in behavior we observe is much larger than anticipated, and (iv) that our sample is slightly larger than that in the analogous experiment by Fehr and Fischbacher (2004), we decided to write up our findings for publication.

⁷ For clarity in presentation, in Figure 1, we pool together the responses “biased” (51%) and “somewhat biased” (26%) in one category, and the responses “unbiased” (6%) and “somewhat unbiased” (7%) in another.

Panel B in Figure 1 compares perceptions of second- and third-party punishment. A majority of respondents (52%) found second-party punishment to be more biased, even though it was identical to third-party punishment from an outcome perspective. Only 19% of respondents viewed third-party punishment as being more biased. The data rejects the hypothesis that second- and third-party punishment are perceived to be equally biased/unbiased (Wilcoxon sign rank, $N=111$, two-tailed, p -value <0.001). The data, therefore, supports the alternative hypothesis that the mere possibility of ulterior motives taints individuals' perceptions of how biased a punishment incident is. Individuals are more prone to attribute negative motives to punishers that *could be* motivated by a personal grudge, even if the punishment is not different in magnitude and arguably fair. The answer to the question in the title of this subsection, therefore, is affirmative.

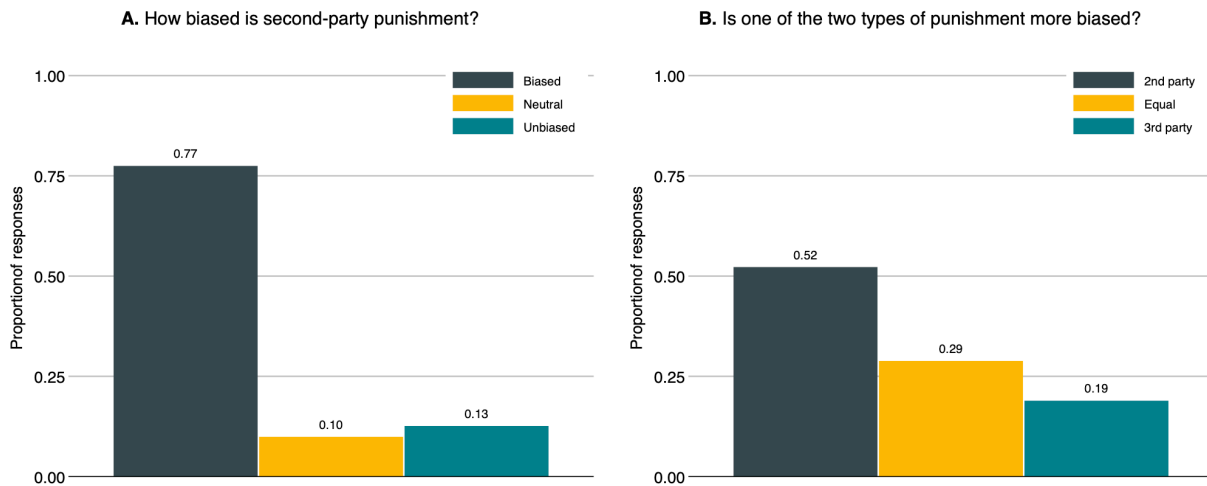


Figure 1. Survey responses on how biased second-party punishment is perceived to be, both independently (Panel A) and in comparison to third-party punishment (Panel B)

3.2 *Is second-party punishment more likely to trigger counter-punishment than third-party punishment?*

Having found support for the hypothesis that third-party punishment is perceived to be less biased than second-party punishment, we turn our attention to the question of how likely participants in our experiment are to retaliate each type of punishment. Figure 2 offers an overview of our main finding. The figure presents the mean amount of counter-punishment in the two experimental conditions as a function of the different levels of punishment. We observe a pronounced difference across conditions when punishment exceeds 50% of the target's first-stage earnings.⁸ Across the different levels of punishment, the mean amount of counter-punishment is

⁸ A discussion of individual counter-punishment strategies can be found in Online Appendix B.

64.6% higher in 2P (41.4% of the target’s earnings) than in 3P (25.2% of the target’s earnings). Using the mean amount of counter-punishment for each individual as an independent observation, this difference is highly significant (p -value <0.001 , two-tailed, Mann-Whitney test). This difference appears to be largely driven by an increased propensity to avenge punishment in 2P: punished individuals retaliate in 54.5% of the cases in 2P and only in 32.8% of the cases in 3P. That is, the likelihood of counter-punishment is, on average, 66.7% higher in 2P than in 3P. This difference is also highly significant (p -value <0.001 , two-tailed, χ^2 test).

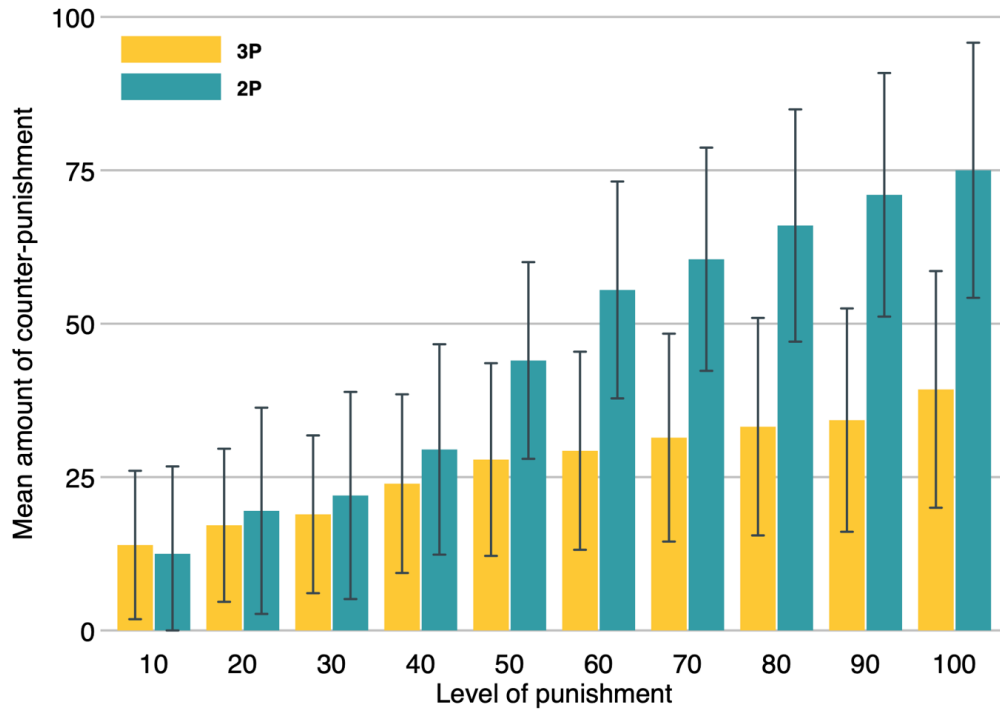


Figure 2. Mean amount of counter-punishment as a function of punishment

To explore in greater depth the determinants of counter-punishment in our experiment, we turn to a regression analysis. We consider three dependent variables: (i) the mean amount of counter-punishment, (ii) the likelihood of counter-punishment, i.e., a binary variable capturing whether counter-punishment was meted out or not, and (iii) the severity of counter-punishment, i.e., a discrete variable capturing the mean amount of counter-punishment conditional on counter-punishment occurring. Our independent variables are (i) *Punishment* – a discrete variable taking the values between 0 and 100 that captures the punishment inflicted on the individual, (ii) *Third-Party* – a dummy variable taking the value 1 if the individual is in the 3P condition and 0 if s/he is in the 2P condition, (iii) *Punishment * 3P* – the interaction term of the abovementioned independent variables, and (iv) *Amount Taken* – a variable taking the values 0, 5 or 10 depending on the amount

the individual took in Stage 1. We use linear regressions (when the dependent variable is binary, this amounts to using a linear probability model), with random effects clustered at the group level.

Table 1 presents the findings from the regression analysis. As in previous studies, we find that the amount of punishment has a strong positive effect on the mean amount of counter-punishment, the probability of counter-punishment, and its severity in condition 2P. As hinted in Figure 2, the variable *Third-Party* is insignificant as there are no appreciable differences when punishment is close to zero. On the other hand, the interaction term is highly significant in models (I)-(V), showing that both the mean amount of counter-punishment as well as the likelihood of counter-punishment increase at a slower rate in 3P. We find no significant association between the amount an individual took and his propensity to counter-punish across conditions. Our results are largely unchanged when we control for the Big-5 markers, which themselves are never statistically significant (see Table A1 in Online Appendix A).

Table 1. Determinant of counter-punishment

	<u>Counter-punishment</u>		<u>Choice</u>		<u>Severity</u>	
	I	II	III	IV	V	VI
<i>Punishment</i>	0.774*** (0.115)	0.774*** (0.115)	0.007*** (0.001)	0.007*** (0.001)	0.645*** (0.176)	0.639*** (0.176)
<i>Third-party (3P)</i>	7.961 (7.958)	7.030 (7.575)	0.052 (0.106)	0.044 (0.107)	9.705 (22.455)	9.094 (22.015)
<i>Punishment * 3P</i>	-0.484*** (0.143)	-0.484*** (0.143)	-0.005*** (0.001)	-0.005*** (0.001)	-0.068 (0.244)	-0.062 (0.240)
<i>Amount taken</i>		0.686 (1.143)		0.006 (0.013)		0.446 (1.014)
Constant	2.705 (5.818)	0.817 (7.381)	0.195** (0.077)	0.178** (0.088)	35.421** (16.018)	34.214* (17.406)
Observations	528	528	528	528	221	221
Groups	48	48	48	48	26	26

Columns I and II report results from linear regression models where the dependent variable is level of counter-punishment, as measured by the percentage reduction of Player A's income. Columns III and IV report results from linear probability models where the dependent variable is a dummy equal to 1 when counter-punishment occurs and 0 otherwise. Finally, columns V and VI report estimates from linear regression models where the dependent variable is the severity of counter-punishment, conditional on counter-punishment occurring. All regressions include random effects with errors clustered at the group level; * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

3.3 Other findings

The data presented in sections 3.1 and 3.2 address our research question concerning the determinants of counter-punishment. For completeness, in this section, we offer a brief discussion of behavior in Stage 2 (punishment), and Stage 1 (take rates).

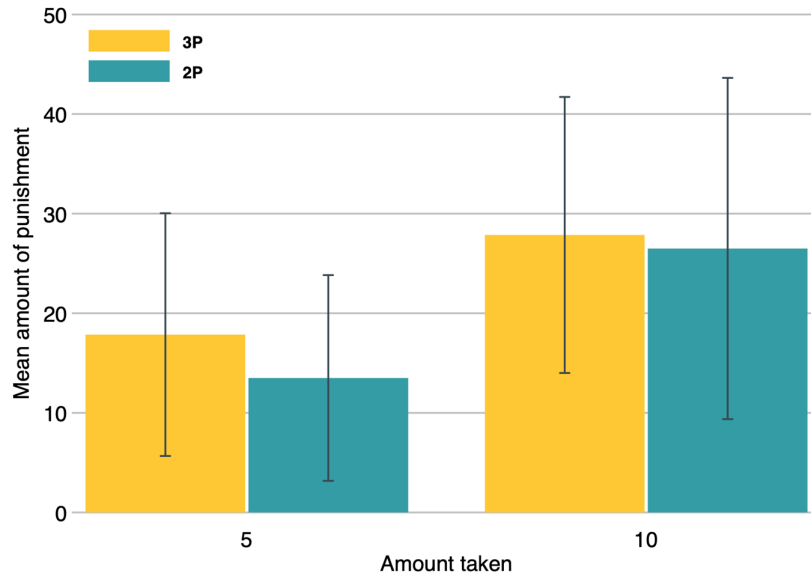


Figure 3. Mean amount of punishment as a function of the amount taken

Figure 3 presents the mean amount of punishment in each condition as a function of the amount taken by Player A in the first stage. As in previous studies, punishment increases in the amount taken by Player A in both 2P (p -value = 0.029, linear regression) and 3P (p -value = 0.008, linear regression). Perhaps most interestingly, unlike in the study by Fehr and Fischbacher (2004) where second parties were found to inflict substantially more punishment than third parties, we observe no significant differences in punishment levels across conditions (p -value = 0.747, two-tailed, Mann-Whitney test using the mean punishment of each individual as unit of observation). An interpretation of this finding is that second parties may be more inclined to punish to avenge a personal affront, *ceteris paribus*, but they hold back in anticipation of the greater level of counter-punishment meted out to second parties.⁹

Take rates are slightly higher in 3P (mean: 4.10, st. dev.: 4.72) than in 2P (mean: 2.75, st. dev.: 3.79), but the difference is far from being statistically significant (p -value = 0.38, two-tailed, Mann-Whitney test). The reason is that nearly half of Player As do not take away any of the endowment allocated to Player B in both conditions. The lack of a significant difference in take rates suggests that participants correctly anticipated the fact that the threat posed by punishment is very similar across conditions (see Figure 3).

⁹ Additional analysis of punishment behavior can be found in Online Appendix A.

4. Conclusions

This study provides novel experimental evidence suggesting that individuals have a tendency to attribute negative motives to punishers for their actions; a tendency that affects their willingness to counter-punish. Specifically, we found that individuals perceive the punishment of a norm violator to be more biased if the punishment is meted out by the victim of a transgression than a third-party. In line with this, we showed that individuals are substantially more likely to avenge second-party rather than a third-party punishment.

Apart from providing a novel explanation for the phenomenon of counter-punishment, our findings have implications for understanding and, by extension, avoiding conflict. Studies have found that individuals are more willing to punish norm violators that belong to outgroups, especially when the violation affects an ingroup member (Bernhard et al. [2006](#); Goette et al. [2006](#); Goette et al. [2012](#)). Although there is no evidence to date about how group membership affects the willingness to counter-punish, in light of our evidence, it seems plausible to assume that outgroup members will be more likely to perceive punishment for a norm violation as being biased and retaliate. To the extent that the willingness to counter-punish can be seen as a pathway to prolonged conflict (Engelmann and Nikiforakis [2015](#); Nikiforakis and Engelmann [2011](#); Nikiforakis et al. [2012](#)), our findings suggest a new channel through which group membership can affect conflict (Hewstone, [2002](#)).

Our findings may also help explain the prominent role third parties play in upholding social norms and cooperation in daily life. Third-party enforcers tend to be individuals in positions of relative power, which makes them relatively immune to the perils of counter-punishment. Our study, however, suggests that another reason for the central role played by third parties is that they are less likely to be attributed negative motives than individuals or parties directly affected by a transgression. All else equal, this makes it less costly for unaffected third parties to uphold social norms and cooperation. But in order to minimize the costs of norm enforcement in daily life, the abovementioned findings suggest that third parties must not be perceived as acting out of ulterior motives. This could happen if the third party is a member of the victim's ingroup.

References

- Balafoutas, L., Grechenig, K., Nikiforakis, N., 2014a. Third-party punishment and counter-punishment in one-shot interactions. *Econ. Letters* 122, 308-310.
- Balafoutas, L., Nikiforakis, N., Rockenbach, B. 2014b. Direct and indirect punishment among strangers in the field. *Proceedings of National Academy of Science, USA* 111, 15924-15927.
- Balafoutas, L., Nikiforakis, N., Rockenbach, B. 2016. Altruistic punishment does not increase with the severity of norm violations in the field. *Nature Communications* 7: 13327.
- Balafoutas, L., Nikiforakis, N., 2012. Norm enforcement in the city: A natural field experiment. *European Economic Review* 56 (8), 1773-1785.
- Bernhard, H., Fischbacher, U., Fehr, E. 2006. Parochial altruism in humans. *Nature* 442, 912-915.
- Bolton, G., Ockenfels, A. 2000. A Theory of Equity, Reciprocity, and Competition. *American Economic Review* 90, 166-193.
- Coleman, J. S. *Foundations of Social Theory*. Cambridge, Mass: Harvard University Press, 1990.
- Denant-Boemont, L., Masclet, D., Noussair, C., 2007. Punishment, counterpunishment and sanction enforcement in a social dilemma experiment. *Econom. Theory* 33 (1), 145–167.
- Elster, J. 1989. Social norms and economic theory. *Journal of Economic Perspectives* 3, 99-117.
- Engelmann, D., Nikiforakis, N. 2015. In the long-run we are all dead: on the benefits of peer punishment in rich environments. *Social Choice and Welfare* 45, 561-577.
- Fehr, E., Fischbacher, U., 2004. Third-party punishment and social norms. *Evol. Hum. Behav.* 25, 63-87.
- Falk, A., Fischbacher, U. 2006. A theory of reciprocity. *Games and Economic Behavior* 54, 293-315.
- Fehr, E., Schmidt, K. 1999. A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics* 114, 817-868.
- Goette, L., Huffman, D., Meier, S. 2006. The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups. *AEA Papers and Proceedings* 96, 212-216.
- Goette, L., Huffman, D., Meier, S., Sutter, M. 2012. *Competition Between Organizational Groups: Its Impact on Altruistic and Antisocial Motivations*. *Management Science* 58, 948-960.
- Goldberg, Lewis R. 1992. The Development of Markers for the Big-Five Factor Structure. *Psychological Assessment* 4, 26-42.
- Heckhausen H. 1991. *Attribution and Achievement Behavior*. In: Motivation and Action. Springer, Berlin, Heidelberg.
- Hewstone, M., Rubin, M., Willis, H. 2002. Intergroup bias. *Annual Review of Psychology* 53, 575-604.

- Hopfensitz, A., Reuben, E. 2009. The importance of emotions for the effectiveness of social punishment. *The Economic Journal* 119, 1534–1559.
- Kurzban, R., DeScioli, P., O'Brien, E. 2007. Audience effects on moralistic punishment. *Evolution and Human Behavior* 28, 75-84.
- Waytz, A., Young, L., Ginges, J. 2014. Motive attribution asymmetry for love vs. hate drives intractable conflict. *Proceedings of National Academy of Science, USA* 111, 15687-15692.
- Nikiforakis, N., 2008. Punishment and counter-punishment in public good games: Can we really govern ourselves? *J. Public Econ.* 92 (1–2), 91–112.
- Nikiforakis, N, Engelmann, D. 2011. Altruistic punishment and the threat of feuds. *Journal of Economic Behavior and Organization* 78, 319-332.
- Nikiforakis N., Noussair, C., Wilkening, T. 2012. Normative conflict and feuds: The limits of self-enforcement. *Journal of Public Economics* 96 (9-10), 797-807
- Rabin, M. 1993. Incorporating Fairness into Game Theory and Economics. *The American Economic Review* 83,1281-1302.
- Reeder, G., Pryor, J., Wohl, M., Griswell, M. 2005. On Attributing Negative Motives to Others Who Disagree With Our Opinions. *Personality and Social Psychology Bulletin* 31, 1498-1510.
- Ross, L. (1997). The Intuitive Psychologist And His Shortcomings: Distortions in the Attribution Process. *Advances in Experimental Social Psychology* 10, 173-220.
- Sugden, R. 1986. New Developments In The Theory Of Choice Under Uncertainty. *Bulletin of Economic Research* 38, 3307-3378.
- Tetlock, P., Levi, A. 1982. Attribution bias: On the inconclusiveness of the cognition-motivation debate. *Journal of Experimental Social Psychology* 18, 68-88.
- Young, P. 2015. The evolution of social norms. *Annual Review of Economics* 7, 359-387.

ONLINE APPENDIX

Experimental evidence shows that negative motive attribution drives counter-punishment

Manuel Muñoz-Herrera

Nikos Nikiforakis

Appendix A. Additional regressions

In the first three columns in Table A1, we report estimates from regression models similar to those in Table 1 in the paper, with the difference these models control also for personality traits (Big-5 markers). The dependent variable is the average amount of counter-punishment (see column I), the probability of counter-punishment (see column II), and the effective level of counter-punishment chosen by those who choose a positive level of counter-punishment (see column III). In the last three columns, we report models on the average amount of punishment (see model IV), controlling for the amount taken (see model V), as well as for personality traits (see model VI).

Table A1. Regressions on counter-punishment and punishment controlling for personality traits

	Counter-Punishment			Punishment		
	<u>Amount</u>	<u>Choice</u>	<u>Severity</u>	<u>Amount</u>		
	I	II	III	IV	V	VI
Punishment	0.774*** (0.115)	0.007*** (0.001)	0.662*** (0.162)			
Third-party (3P)	13.369 (10.426)	0.183 (0.134)	-2.332 (21.447)	5.424 (12.477)	33.278 (19.594)	19.043 (21.951)
Punishment X 3P	-0.484*** (0.144)	-0.005*** (0.001)	-0.067 (0.225)			
Amount taken	1.103 (1.440)	0.026 (0.019)	-0.883 (1.484)	0.920 (1.294)	3.850* (1.565)	3.036* (1.448)
Amount taken X 3P	-1.717 (2.151)	-0.042 (0.027)	2.948 (1.902)		-3.702 (2.207)	-2.704 (2.322)
Open	-0.915 (0.859)	-0.009 (0.010)	-0.387 (0.652)			-1.075 (1.065)
Conscientious	0.934 (0.607)	0.011 (0.007)	0.829 (0.768)			-1.108 (0.912)
Extrovert	0.004 (0.581)	0.001 (0.006)	-0.216 (0.585)			-0.028 (0.753)
Agreeable	-0.882 (0.820)	-0.011 (0.009)	-0.405 (0.827)			-1.741 (1.012)
Neurotic	0.685 (0.767)	0.002 (0.009)	0.757 (0.626)			1.547 (0.896)
Constant	15.142 (41.395)	0.444 (0.555)	20.806 (51.659)	37.290** (12.790)	14.500 (11.275)	129.752 (71.067)
Observations	528	528	221	50	50	50
Groups	48	48	26	25	25	25

Column I reports a linear regression model where the amount of counter-punishment, as measured by the percentage reduction of Player A's income, is the dependent variable. Column II reports a linear probability model where the dependent variable is a dummy equal to 1 when counter-punishment is positive and 0 otherwise. Column III reports a linear regression model where the dependent variable is the severity of counter-punishment, by taking into account only those decisions where counter-punishment was positive. That is why the number of observations in III are fewer than in first two regressions. Columns IV, V and VI report linear regression models where the effective level of punishment, measured by the percentage reduction of Player B's income when punishment was greater than zero, is the dependent variable. All regressions include random effects at the group level; * $p < 0.05$; ** $p < 0.01$.

Appendix B. Individual punishment and counter-punishment strategies

In this section we categorize the strategies participants followed in their punishment and counter-punishment choices. We run linear regression models for each individual participant and use the intercept and slope of the regressions to group strategies into one of three categories: No, Increasing and Maximal. A strategy is labeled *No*, when both intercept and slope are zero. This is the case when participant B chooses not to punish independently of the amount taken by A. Similarly, this is so when participant A chooses not to counter-punish, independently of the level of punishment chosen by B. A strategy is *Increasing*, when the slope is positive. This is the case when a participant punishes more the higher the amount taken or when the counter-punishment is higher the higher the level of punishment. Finally, a strategy is maximal when a participant B exert the highest level of punishment (100%) for any amount taken, or when a participant A exerts the highest level of counter-punishment (100%) for any level of punishment.¹ The frequencies of each strategy by treatment are summarized in Table B1.

Table B1. Frequency of punishment and counter-punishment strategies by treatment

	<u>Punishment</u>		<u>Counter-</u> <u>punishment</u>	
	2P	3P	2P	3P
<i>No</i>	0.50	0.46	0.25	0.60
<i>Increasing</i>	0.45	0.39	0.60	0.29
<i>Maximal</i>	0.00	0.11	0.15	0.11
<i>Other</i>	0.05	0.04	0.00	0.00

With respect to counter-punishment strategies, 60% of the participants choose strategy *No* in **3P** compared to 25% in **2P** (Wilcoxon Mann-Whitney, $N=48$, two-tailed, p -value=0.029). Unlike counter-punishment, there are no significant differences in the frequency of participants choosing not to punish: 50% in **2P** versus 46% in **3P** (Wilcoxon Mann-Whitney, $N=48$, two-tailed, p -value=0.809). That is why the average level of punishment, as well as the severity of the punishment effectively exerted, are indistinguishable between treatments.

¹ There is one participant in 2P who exerted a constant punishment of 30% and there is one participant in 3P who chose positive punishment levels but lower when the amount taken was 10 than 5. Neither of these two participants fit in any of the three main categories, and that is why they are labeled as “other”.

Appendix C. Experimental instructions

The experimental instructions were adapted from those of Fehr and Fischbacher.

Condition 2P

General Instructions

You are now taking part in an economic experiment. Please do not communicate with other participants during the experiment. Should you have any questions, please raise your hand and one of us will come to assist you. The amount of money you earn from this experiment depends on your decisions and those made by others in the experiment. It is therefore important that you take your time to understand the instructions. At the end of the experiment you will be asked to fill out a questionnaire.

Earnings

During this experiment **we do not deal with Emirati dirhams, but with Experimental Monetary Units (EMU)**. At the end of the experiment, the total amount of EMU you earned during the experiment will be converted into Emirati Dirhams at the rate of

$$1 \text{ EMU} = 2 \text{ AED}$$

The earned sum will consist of the show-up fee of **30 AED** plus any individual gains from the experiment. This total will be paid to you in cash in private, immediately after the completion of the experiment.

Should you have any questions, please raise your hand and an experimenter will come to assist you.

Instructions for Participant A

This experiment comprises 2 different types of Participants: Participants A and Participants B. **You are a Participant A**. At the start of the experiment, the computer will randomly place you into a group with an individual who has been assigned the role of Participant B.

The experiment consists of **three stages**. In the first stage, those who have been assigned the role of Participant A have to decide whether to take 0, 5 or 10 EMU from the individual with whom

they have been paired with. In the second stage, individuals who have been assigned the role of Participant B will have to decide whether they wish to **reduce the earnings** of Participant A by a certain amount. In the third and final stage of the experiment, those who have been assigned the role of Participant A will be asked whether they wish to reduce the earnings of Participant B.

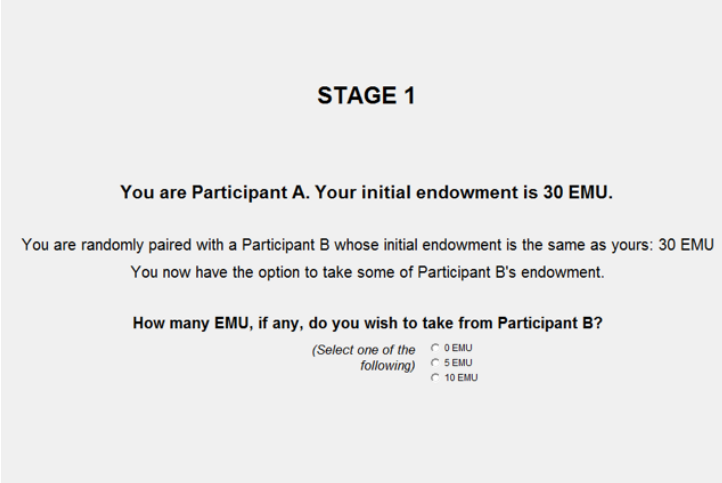
Neither during nor after the experiment will you be aware of the identities of any of the participants. Equally, other participants will never know with whom they were dealing. All earnings will be paid out anonymously at the experiment. No participant will ever learn how much you earned in the experiment.

Stage One

In stage one only individuals assigned the role of **Participant A** can make decisions. As mentioned, in stage one, you are grouped together with another participant. As a Participant A you get an endowment of **30 EMU** at the beginning of stage one. Participant B gets the same endowment of 30 EMU.

You must decide whether or not to take a portion of Participant's B income. **You can take 0, 5 or 10 EMU from Participant B.** If you decide to take 0 EMU, your earnings and the earnings of Participant B at the end of the first stage will be the same as the original endowment you received. If you take 5 EMU, your earnings at the end of the first stage will be 35 EMU, while the earnings of Participant B will be 25EMU. If you decide to take 10 EMU, your earnings at the end of the first stage will be 40 EMU, while the earnings of Participant B will be 20 EMU.

You will make your decision as displayed on the screenshot below:



The screenshot shows a grey background with the following text:

STAGE 1

You are Participant A. Your initial endowment is 30 EMU.

You are randomly paired with a Participant B whose initial endowment is the same as yours: 30 EMU
You now have the option to take some of Participant B's endowment.

How many EMU, if any, do you wish to take from Participant B?

(Select one of the following)

- 0 EMU
- 5 EMU
- 10 EMU

Stage Two

In stage two, only individuals assigned the role of **Participant B** can make decisions. In particular, **Participant B from your group** will be asked whether or not s/he wishes to reduce your income. Participant B will be informed of your decision in stage one. S/he will be awarded an additional **5 EMU**, which s/he can use to reduce your income by 10%, 20%, 30%... 90%, or 100%. The cost of reducing your income by a positive amount (i.e, more than 0%) is 5 EMU irrespective of the amount by which Participant B reduces your income (i.e., 10%, 20%, ... or 100%). If Participant B decides not to reduce your income, then the 5 EMU are added to his or her final earnings.

Stage Three

Your task in the third stage will be to **decide whether you want to reduce or leave unaffected the earnings of Participant B**; that is, the participant who was asked in stage two whether or not to reduce your earnings from stage one. Income reduction in this stage works exactly as in the second stage. You will be awarded **5 EMU**, which you can use to reduce Participant's B earnings by 10%, 20%, 30%... 90%, 100%. If you do not reduce Participant B's earnings, then the 5 EMU will be added to your final earnings.

Note that you will have to make your decision about how much you want to decrease Participant B's earnings **before you find out his or her decision in stage two**. You will have to state whether and by how much you wish to reduce B's earnings for each possible decision Participant B may have taken in the second stage. To make your decision you will use a screen as the one shown below.

STAGE 3

Your task in this stage is to decide whether you want to reduce the income of Participant B **in your group** , who may have reduced your income in Stage 2.

Please indicate for every possible action of Participant B in Stage 2 by how much, if at all, you would like to reduce his or her income.

<p>If Participant B in my group reduced my income in Stage 2 by:</p>	<p>I will reduce Participant B's income by the following percent. Do not use the "%" sign, and use increments of 10</p>
<p>0%</p>	<input style="width: 100%; height: 20px;" type="text"/>
<p>10%</p>	<input style="width: 100%; height: 20px;" type="text"/>
<p>20%</p>	<input style="width: 100%; height: 20px;" type="text"/>
<p>30%</p>	<input style="width: 100%; height: 20px;" type="text"/>
<p>40%</p>	<input style="width: 100%; height: 20px;" type="text"/>
<p>50%</p>	<input style="width: 100%; height: 20px;" type="text"/>
<p>60%</p>	<input style="width: 100%; height: 20px;" type="text"/>
<p>70%</p>	<input style="width: 100%; height: 20px;" type="text"/>
<p>80%</p>	<input style="width: 100%; height: 20px;" type="text"/>
<p>90%</p>	<input style="width: 100%; height: 20px;" type="text"/>
<p>100%</p>	<input style="width: 100%; height: 20px;" type="text"/>

On the screen you must indicate whether and by how much you want to reduce Participant B's earnings in the case he or she reduced your earnings by 0%, 10%, 20%... 100%.

Given that you will not know which one of the possible actions Participant B has decided on, any of your decisions might be used to determine your final earnings and the earnings of Participant B. Which of your decisions is actually realized depends on the actual decision of Participant B in stage two.

After you have made your decisions, you will be informed of your earnings from the experiment.

Instructions for Participant B

This experiment comprises 2 different types of Participants: Participants A and Participants B. **You are a Participant B.** At the start of the experiment, the computer will randomly place you into a group with an individual who has been assigned the role of Participant A.

The experiment consists of **three stages**. In the first stage, those who have been assigned the role of Participant A have to decide whether to take 0, 5 or 10 EMU from the individual with whom they have been paired with. In the second stage, individuals who have been assigned the role of Participant B will have to decide whether they wish to **reduce the earnings** of Participant A by a

certain amount. In the third and final stage of the experiment, those who have been assigned the role of Participant A will be asked whether they wish to reduce the earnings of Participant B.

Neither during nor after the experiment will you be aware of the identities of any of the participants. Equally, other participants will never know with whom they were dealing. All earnings will be paid out anonymously at the experiment. No participant will ever learn how much you earned in the experiment.

Stage One

As mentioned, in stage one, you are grouped together with another participant – Participant A. As a Participant B you get an endowment of **30 EMU** at the beginning of stage one. Participant A gets the same endowment of 30 EMU.

Participant A must decide whether or not to take a portion of your income. **Participant A can take 0, 5 or 10 EMU from you.** If s/he decides to take 0 EMU, your earnings and the earnings of Participant A at the end of the first stage will be the same as the original endowment you received. If s/he takes 5 EMU, your earnings at the end of the first stage will be 25 EMU, while the earnings of Participant A will be 35EMU. If s/he decides to take 10 EMU, your earnings at the end of the first stage will be 20 EMU, while the earnings of Participant A will be 40EMU.

Stage Two

Your task in the second stage will be to **decide whether you want to reduce or leave unaffected the earnings of Participant A.** You will be awarded an additional **5 EMU**, which you can use to reduce Participant's A earnings by 10%, 20%, 30%... 90%, 100%. The cost of reducing Participant A's income by a positive amount (i.e, more than 0%) is 5 EMU irrespective of the amount by which you reduce his/her income (i.e., 10%, 20%, ... or 100%). If you do not reduce Participant A's earnings, then the 5 EMU will be added to your final earnings.

Note that you will have to make your decision about how much you want to decrease Participant A's earnings **before you find out his or her decision in stage one.** You will have to state whether and by how much you wish to reduce A's earnings for each possible decision Participant A may have taken in the first stage. To make your decision you will use a screen as the one shown below.

Given that you will not know which one of the three actions Participant A has decided on, any of the three decisions might be used to determine final earnings of Participant A. Which of your decisions is actually realized depends on the actual decision of Participant A in stage one.

STAGE 2

You are Participant B. Your initial endowment is 30 EMU.

Your task in this stage is to decide whether you want to reduce the income of Participant A **in your group**.
Participant A's initial endowment was the same as yours: 30 EMU.
Participant A had the option to take some of your initial endowment in Stage 1.

Please indicate for every possible action of Participant A in Stage 1 by how much, if at all, you would like to reduce his or her income.

If Participant A took away from me in Stage 1:	I will reduce Participant A's income by the following percent. Do not use the "%" sign, and use increments of 10
0 EMU	<input type="text"/>
5 EMU	<input type="text"/>
10 EMU	<input type="text"/>

Stage Three

In stage three, **Participant A** will be asked whether or not s/he wishes to reduce your income. Income reduction in this stage works exactly as in the second stage. S/he will be awarded an additional **5 EMU**, which s/he can use to reduce your income by 10%, 20%, 30%... 90%, or 100%. If Participant A decides not to reduce your income, then the 5 EMU are added to his or her final earnings.

After you have made your decisions, you will be informed of your earnings from the experiment.

Instructions for Participant A

This experiment comprises different types of Participants: Participant A, Participant B, Participant C, Participant D etc. **You are a Participant A.** At the start of the experiment, the computer will randomly place you into a group with an individual who has been assigned the role of Participant B. Participant C will be placed into a group with Participant D, Participant E will be placed into a group with Participant F etc.

Neither during nor after the experiment will you be aware of the identities of any of the participants. Equally, other participants will never know with whom they were dealing. All earnings will be paid out anonymously at the experiment. No participant will ever learn how much you earned in the experiment. The experiment consists of **three stages**, whose procedure is described in detail below.

Stage One

In stage one, you as **Participant A** will be able to make decisions. As mentioned, in stage one, you are grouped together with another Participant B. As Participant A you get an endowment of **30 EMU** at the beginning of stage one. All the other participants get the same endowment of 30 EMU.

You must decide whether or not to take a portion of Participant B's income. **You can take 0, 5 or 10 EMU from Participant B.** If you decide to take 0 EMU, your earnings and the earnings of Participant B at the end of the first stage will be the same as the original endowment you received. If you take 5 EMU, your earnings at the end of the first stage will be 35 EMU, while the earnings of Participant B will be 25EMU. If you decide to take 10 EMU, your earnings at the end of the first stage will be 40 EMU, while the earnings of Participant B will be 20 EMU.

You will make your decision as displayed on the screenshot below:

STAGE 1

You are Participant A. Your initial endowment is 30 EMU.

You are randomly paired with a Participant B whose initial endowment is the same as yours: 30 EMU
You now have the option to take some of Participant B's endowment.

How many EMU, if any, do you wish to take from Participant B?

(Select one of the following) 0 EMU
 5 EMU
 10 EMU

Participant C from another pair also has the option to reduce or leave unaffected the endowment of Participant D by the same amounts – 0, 5 or 10 EMU. Similarly, Participant E has the option to reduce or leave unaffected the endowment of Participant F and so forth.

Stage Two

In stage two, **Participant D** from another group will be asked whether or not s/he wishes to reduce your income. Participant D will be informed of your decision in stage one. S/he will be awarded an additional **5 EMU**, which s/he can use to reduce your income by 10%, 20%, 30%... 90%, or 100%. The cost of reducing your income by a positive amount (i.e, more than 0%) is 5 EMU irrespective of the amount by which Participant D reduces your income (i.e., 10%, 20%, ... or 100%). If Participant D decides not to reduce your income, then the 5 EMU are added to his or her final earnings.

Participant B, from your group, will be asked whether s/he wishes to reduce the earnings of a Participant E in a third group. **Participant B cannot reduce your income**, but rather the income of Participant E in a third group.

Stage Three

Your task in the third stage will be to **decide whether you want to reduce or leave unaffected the earnings of Participant D**; that is, the participant from the **second pair** who was asked in stage two whether or not to reduce your earnings from stage one. Income reduction in this stage works exactly as in the second stage. You will be awarded **5 EMU**, which you can use to reduce Participant D's earnings by 10%, 20%, 30%... 90%, 100%. If you do not reduce Participant D's earnings, then the 5 EMU will be added to your final earnings.

Note that you will have to make your decision about how much you want to decrease Participant D's earnings **before you find out his or her decision in stage two**. You will have to state whether and by how much you wish to reduce D's earnings for each possible decision Participant D may have taken in the second stage. To make your decision you will use a screen as the one shown below.

STAGE 3

Your task in this stage is to decide whether you want to reduce the income of Participant D **in another group**.
Participant D from the other group may have reduced your income in Stage 2.
Please indicate for every possible action of Participant D in Stage 2 by how much, if at all, you would like to reduce his or her income.

If Participant D from the other group reduced my income in Stage 2 by:	I will reduce Participant D's income by the following percent. Do not use the "%" sign, and use increments of 10
0%	<input type="text"/>
10%	<input type="text"/>
20%	<input type="text"/>
30%	<input type="text"/>
40%	<input type="text"/>
50%	<input type="text"/>
60%	<input type="text"/>
70%	<input type="text"/>
80%	<input type="text"/>
90%	<input type="text"/>
100%	<input type="text"/>

On the screen you must indicate whether and by how much you want to reduce Participant D's earnings in the case he or she reduced your earnings by 0%, 10%, 20%... 100%.

Given that you will not know which one of the possible actions Participant D has decided on, any of your decisions might be used to determine your final earnings and the earnings of Participant D. Which of your decisions is actually realized depends on the actual decision of Participant D in stage two.

After you have made your decisions, you will be informed of your earnings from the experiment.

Instructions for Participant B

This experiment comprises different types of Participants: Participant A, Participant B, Participant C, Participant D etc. **You are a Participant B.** At the start of the experiment, the computer will randomly place you into a group with an individual who has been assigned the role of Participant A. Participant C will be placed into a group with Participant D, Participant E will be placed into a group with Participant F etc.

Neither during nor after the experiment will you be aware of the identities of any of the participants. Equally, other participants will never know with whom they were dealing. All earnings will be paid out anonymously at the experiment. No participant will ever learn how much you earned in the experiment. The experiment consists of **three stages**, whose procedure is described in detail below.

Stage One

In stage individuals assigned the role of **Participant A** can make decisions. As mentioned, in stage one, you are grouped together with Participant A. As Participant B you get an endowment of **30 EMU** at the beginning of stage one. All the other participants get the same endowment of 30 EMU.

Participant A must decide whether or not to take a portion of your income. **Participant A can take 0, 5 or 10 EMU from you.** If s/he decides to take 0 EMU, your earnings and the earnings of Participant A at the end of the first stage will be the same as the original endowment you received. If s/he takes 5 EMU, your earnings at the end of the first stage will be 25 EMU, while the earnings of Participant A will be 35 EMU. If s/he decides to take 10 EMU, your earnings at the end of the first stage will be 20 EMU, while the earnings of Participant A will be 40 EMU.

Participant C from another pair also has the option to reduce or leave unaffected the endowment of Participant D by the same amounts – 0, 5 or 10 EMU. Similarly, Participant E has the option to reduce or leave unaffected the endowment of Participant F and so forth.

Stage Two

Your task in the second stage will be to **decide whether you want to reduce or leave unaffected the earnings of Participant C**; that is, the participant from **another group** who was asked in stage one whether or not to reduce **Participant D's** income. You will be awarded an additional **5 EMU**,

which you can use to reduce Participant C's earnings by 10%, 20%, 30%... 90%, 100%. The cost of reducing Participant C's income by a positive amount (i.e., more than 0%) is 5 EMU irrespective of the amount by which you reduce his/her income (i.e., 10%, 20%, ... or 100%). If you do not reduce Participant C's earnings, then the 5 EMU will be added to your final earnings.

Note that you will have to make your decision about how much you want to decrease Participant C's earnings **before you find out his or her decision in stage one**. You will have to state whether and by how much you wish to reduce C's earnings for each possible decision Participant C may have taken in the first stage. To make your decision you will use a screen as the one shown below.

/

Given that you will not know which one of the three actions Participant C has decided on, any of the three decisions might be used to determine final earnings of Participant C. Which of your decisions is actually realized depends on the actual decision of Participant C in Stage One.

Stage Three

In stage three, **Participant C** will be asked whether or not s/he wishes to reduce your income. Income reduction in this stage works exactly as in the second stage. S/he will be awarded an additional **5 EMU**, which s/he can use to reduce your income by 10%, 20%, 30%... 90%, or 100%. If Participant C decides not to reduce your income, then the 5 EMU are added to his or her final earnings.

Also, Participant A from your pair can deduct income of a Participant in a third group – F. Participant A **cannot deduct your income**, but rather an income of another Participant in a third group.

After you have made your decisions, you will be informed of your earnings from the experiment.