CBESS Discussion Paper 13-07

On the Robustness of Emotions and Behavior in a Power-to-Take Game Experiment

By Fabio Galeotti

School of Economics and CBESS, University of East Anglia

Abstract
An important branch of economic research on emotions has used power-to-take game experiments to study the impact of negative emotions, such as anger, irritation and contempt, on the decision to punish. We investigate experimentally to what extent the findings of this literature are driven by the particular punishment technology adopted, and whether the experience and background of the participants affect behavior and emotions in this context. We found that (a) negative emotions do still play an important role when the potential relevant confound is removed from the punishment technology; (b) subjects display a similar behavior under a punishment technology with a constant and variable 'fine-to-fee' ratio; (c) previous experience mediates how contempt impacts on the decision to punish; and (d) non-UK students experience similar emotions to UK students, but generally appropriate more resources than UK students.

JEL classification codes
A12, C72, C91.
Keywords
Emotions, punishment, methodology, power-to-take, experiment.

Centre for Behavioural and Experimental Social Science
University of East Anglia
Norwich Research Park
Norwich NR4 7TJ
United Kingdom
www.uea.ac.uk/ssf/cbess

On the Robustness of Emotions and Behavior in a Power-to-Take Game Experiment

Fabio Galeotti*
University of East Anglia

*CBESS and School of Economics, University of East Anglia, Norwich NR4 7TJ, UK
E-mail: f.galeotti@uea.ac.uk; Tel.: +44(0)7531233121; fax: +44(0)1603456259

Version: June 2013

Abstract

An important branch of economic research on emotions has used power-to-take game experiments to study the impact of negative emotions, such as anger, irritation and contempt, on the decision to punish. We investigate experimentally to what extent the findings of this literature are driven by the particular punishment technology adopted, and whether the experience and background of the participants affect behavior and emotions in this context. We found that (a) negative emotions do still play an important role when the potential *relevant* confound is removed from the punishment technology; (b) subjects display a similar behavior under a punishment technology with a constant and variable 'fine-to-fee' ratio; (c) previous experience mediates how contempt impacts on the decision to punish; and (d) non-UK students experience similar emotions to UK students, but generally appropriate more resources than UK students.

Keywords: Emotions, punishment, methodology, power-to-take, experiment.

JEL classification numbers: A12, C72, C91.

1. Introduction

This study investigates the role played by punishment technology in driving the results of power-to-take game (PTTG) experiments, and tests to what extent these results can be attributed to negative emotions. In addition, it explores whether the experience gained from previous economic experiments and background of the participants affect how subjects behave and experience emotions in the context of the PTTG.

Over the last few decades, economists have started to pay greater attention to the complexity of emotions on economic scenarios and have been trying to capture the range of possible roles that emotions play in the decision-making process. For instance, emotions have been proposed as an explanation for important economic phenomena such as cooperation (e.g. Frank, 1988; Fehr and Gächter, 2002) and decision-making under risk (e.g. Loewenstein *et al*., 2001), and are seen to have important consequences for many other economic phenomena, such as inter-temporal choices (e.g. Rick and Lowenstein, 2008), competition (e.g. Kräkel, 2008), bidding behavior (e.g. Bosman and Riedl, 2004) and bargaining behavior (e.g. Pillutla and Murninghan, 1996). More recently, the advent of neuroeconomics has further pushed forward the interest of economists on the role played by emotions in the economic decision-making process (for a review on emotions and neuroeconomics, see Phelps, 2009).

An important branch of economic research on emotions has used experiments to study the impact of negative emotions, such as anger, irritation and contempt, on the decision to punish (Bosman and van Winden, 2002; Bosman *et al*., 2005; Ben-Shakhar *et al*., 2007; Hopfensitz and Reuben, 2009; Joffily *et al*., 2011). This stream of research started with the seminal work of Bosman and van Winden (2002) on the PTTG. In the PTTG, there are two players, the 'take authority' (with income Y_{take}), and the 'responder' (with income Y_{resp}). The game is divided into two stages. In the first stage, the take authority selects a take rate $t \in [0,1]$, which is the proportion of the responder's income that will be transferred to the take authority at the end of the game. In the second stage, the responder chooses a destroy rate $d \in [0,1]$, which is the proportion of Y_{resp} that will be destroyed. Therefore, the payoffs of the game are $(1-t)(1-d)Y_{resp}$ for the responder, and $Y_{take} + t(1-d)\, Y_{resp}$ for the take authority.

If the subjects are rational profit-maximizing agents, the responder should not destroy if the take rate is less than 1, and should be indifferent between all possible destroy rates if the take rate is 1. Hence, from backward induction, the take authority should select $t = 1 - \varepsilon$, where ε is an infinitesimal positive number. The PTTG can be interpreted as an ultimatum

game with continuous opportunities to punish[1] and can describe many economic situations where an agent can take away any part of the endowment of another agent (e.g. taxation, monopolistic pricing and principal-agent relationships) (Bosman and van Winden, 2002).

Most of the literature on the PTTG investigates the role played by emotions on responders' behavior through physiological (Ben-Shakhar *et al*., 2007) and self-report measures (Bosman and van Winden, 2002; Bosman *et al*., 2005). Both measures were found to be related to destruction decisions. In particular, participants who experienced intense negative emotions punished their counterparts more often and more severely. This result seems to identify negative emotions as the main driving force of the punishing behavior in this type of context.[2]

However, this finding may be driven by the technology of the punishment adopted. In particular, the punishing behavior may be simply explained by a non-constant 'fine-to-fee' ratio, defined as "the income reduction for the targeted subject relative to the cost for the subject who requested the punishment" (Casari, 2005:107). In the PTTG, the income reduction for the authority is tdY_{resp}, whereas the cost for the responder to punish is $d(1-t)Y_{resp}$. Therefore, the 'fine-to-fee' ratio is $\frac{tdY_{resp}}{d(1-t)Y_{resp}} = \frac{t}{1-t}$, where t – the proportion of the responder's income that will be transferred to the take authority – is an endogenous and non-constant parameter. This implies that the 'demand' for punishment is higher when t is higher (i.e. the 'fine-to-fee' ratio is higher). In other words, for high take rates the responder has a higher incentive[3] to punish her counterpart, whereas for low take rates the incentive is lower. Hence, the role of negative emotions might be overstated; when the offence is severe subjects experience strong negative emotions, but they punish because punishing is cheaper for increasing offences and not because they experience intense emotions. It thus becomes

[1] It is also worth noting that the endowments are allocated differently in the PTTG compared to the Ultimatum Game. From a traditional game theoretic point of view, this does not matter, but it may have important implications in terms of behavior.

[2] Other well-established findings from the experimental literature on the PTTG show that people appropriate almost 60% of responders' income, while only 20% of the responders destroy income and usually all of it (e.g. Bosman and van Winden, 2002; Reuben and van Winden, 2010). Small differences were observed between an effort treatment – where endowments were earned by doing a preliminary individual real effort task – and a no-effort treatment – where endowments were exogenously given by the experimenter (Bosman *et al*., 2005). A group version of the PTTG – where decisions were made by groups – presented the same results qualitatively as the no-group experiment (Bosman *et al*., 2006). Furthermore, in a three-player version of the PTTG with one take authority and two responders, Reuben and van Winden (2008) showed that responders who knew each other from outside the laboratory punish and coordinate more than strangers. The PTTG has also been used to study the influence of participation (Albert and Mertins, 2008), gender pairing (Sutter *et al*., 2009) and waiting time (Galeotti, 2013) on economic decision-making.

[3] Here the incentive is the cost reduction from punishing per unit of punishment.

important to test the robustness of the results of PTTG experiments against this possible source of confound. If a confound exists, we may need to reconsider the role played by negative emotions in this type of context.

In order to investigate to what extent the punishment behavior observed in previous studies on the PTTG is explained by the punishment technology rather than negative emotions, we conducted a laboratory experiment using students from our university. We varied the extent to which the punishment technology embedded a variable or a constant 'fine-to-fee' ratio. Emotions were assessed through self-report measures, as in previous studies.[4] Given the variety of cultural backgrounds of the students enrolled at our university, we ran separate sessions for UK students and non-UK students.[5] This allowed us to test, under certain homogeneous conditions (e.g. location of the laboratory, university training, etc.), the robustness of previous findings on the PTTG against possible differences in participants' backgrounds. We believe this is particularly crucial in our experiment since we are assessing the emotional responses of our subjects. Previous psychological and anthropological literature has shown that there are cultural differences in the elicitation and manifestation of emotions (for a review, see Mesquita and Frijda, 1992). In addition, we controlled for the experience of our subjects in prior economic experiments.[6] It is in fact reasonable to expect that subjects with more experience of the environment and the dynamics of laboratory experiments are more aware of what they should expect in an economic experiment and, therefore, they might experience less strong emotions and/or be better able to cope with their emotional urges than inexperienced subjects. If this is the case, we would be able to provide evidence, along with a possible explanation based on emotions, of inexperienced subjects being less profit-maximizing than experienced subjects in economic experiments.

[4] For a discussion on the reliability of self-reports in measuring emotions, see Bosman and van Winden (2002) or Hopfensitz and Reuben (2009).

[5] We recruited subjects using the software ORSEE (Greiner, 2004). In the non-UK sessions, we avoided recruiting students from Western societies. This was because we wanted to ensure the highest degree of cultural separation between UK and non-UK subjects. As a result, apart from one Australian participant, non-UK subjects consisted of Asian, South American, African and East European students.

[6] Note that there are generally two kinds of learning relevant to subjects who participate in economic experiments: 'experimental' learning and 'within game' learning (Friedman, 1969). 'Experimental' learning refers to the general experience that subjects acquire by participating in many different experiments not necessarily linked to each other. 'Within game' learning refers to a type of experience acquired in the particular experimental set-up or specific game-theoretical framework under examination. In our experiment, we controlled for and referred to the first type of learning, since none of our subjects had ever participated in a PTTG experiment before.

To give a brief overview of our results, we found that subjects who experienced intense negative emotions punished their counterparts more often and more severely, even controlling for the punishment technology. This finding supports the robustness of previous results on the PTTG. In addition, we did not find any significant difference in behavior if we changed the nature of the punishment technology. In contrast, we did observe behavioral differences between UK and non-UK subjects, particularly in the way they responded to changes in the punishment technology. Finally, we found that more experienced subjects punished less often and less severely when they experienced increasing contempt compared to less experienced subjects.

The paper is organized as follows. Section 2 presents the experimental design, Section 3 reports the results, then Section 4 discusses the results and concludes.

2. Experimental Design

We conducted the experiment between March and September 2012 at the University of East Anglia, with 282 students participating in the experiment over many sessions.[7] Each session lasted on average 50 minutes. No subject was allowed to participate in more than one session. Subjects received a show-up fee of £5 and earned on average £9.41 (around 15 US dollars). In order to ensure the greatest comparability of our experiment with previous literature, we tried to replicate, as close as possible, the experimental procedures adopted in previous PTTG experiments. In particular, we (a) conducted a paper-and-pencil experiment, (b) employed the same instructions and procedures as previous PTTG studies,[8] (c) avoided any particular or suggestive terminology during the sessions, such as 'take authority' or 'take rate', (d) adopted the same double blind procedure of Bosman and van Winden (2002) for the payments, and (e) assessed emotions on a 7-point Likert scale via self-reports after each subject learned about the decision of their counterpart. More details about the experimental procedure are provided below.

Upon arrival, each subject was randomly assigned the role of participant A (take authority) or participant B (responder) by drawing a letter from an urn, then randomly allocated to a computer workstation which was isolated from other workstations via partitions. Then the instructions were distributed and read aloud to help with common knowledge. Two

[7] Details of the socioeconomic characteristics of the subjects and experimental instructions are in the online appendix.

[8] Minor adjustments to the original instructions were made to fit them to our laboratory routines, monetary payments and comparability of our treatments.

individual computerized exercises followed in order to check the subjects' understanding of these instructions. Clarifications were individually provided to subjects with incorrect answers. After completion of these exercises, each participant A was randomly matched with a participant B by asking participant A to randomly choose a coded envelope which was linked to a certain participant B. Each participant A was then asked to fill in the take rate, that is the proportion of participant B's endowment that would be transferred to participant A at the end of the experiment, on the form that was placed inside the envelope. Afterwards, the envelopes containing the forms were collected and given to all participants B who were asked to complete the form with the destroy rate, that is the proportion of their endowment that will be destroyed. The envelopes with the forms inside were then given back to all participants A, who could take note of the decision of their corresponding matches. Subsequently, each subject was asked to fill in a questionnaire concerning emotions, expectations about the decision of their counterpart,[9] and personal information. Meanwhile, the envelopes were collected and handed to the cashier who was outside the laboratory and, hence, not present during the experiment. Subjects were then privately paid one-by-one outside the laboratory by the cashier.

As in Bosman and van Winden (2002), we assessed a list of eleven emotions, consisting of irritation, anger, contempt, envy, jealousy, sadness, joy, happiness, shame, fear and surprise. For each emotion, subjects were asked to state how much they felt the emotion on a 7-point Likert scale when they learned about the decision of their counterparts. The scale ranged from "no emotion at all" to "high intensity of the emotion" (Bosman and van Winden, 2002).

We employed a 2×2 factorial design crossing the nature of the 'fine-to-fee' ratio embedded in the punishment technology (constant or variable) with the cultural background of the subjects who participated in the sessions (UK or non-UK students). Note that subjects were not told about the nationality of their co-participants in the experiment. Hence, they did not know that they took part in a session with all UK or non-UK participants. The main features of the design and the number of independent observations[10] per treatment are summarized in Table 1. The two treatments under a variable 'fine-to-fee' ratio are exact

[9] Participant B was asked to indicate which percentage of his/her endowment he/she expected participant A would decide to transfer to himself/herself; participant A was asked to indicate which percentage of the transfer he/she expected participant B to destroy.

[10] An independent observation is a pair consisting of a responder and a take authority. Variations in the number of independent observations across treatments are due to different rates of attendance across sessions.

replications of previous PTTG experiments (i.e. Bosman and van Winden, 2002; Bosman *et al*., 2005), in one case with only UK subjects and in the other case with only non-UK subjects (consisting of Asian, South American, African and East European subjects).[11] In these treatments, the 'fine-to-fee' ratio was increasing in the take rate and ranged from 0 to infinite. The same separation of the subjects based on cultural background occurred in the treatments under a constant 'fine-to-fee' ratio. However, here the 'fine-to-fee' ratio was constantly equal to 2 and, therefore, independent of the take rate.[12] Most of the literature on punishment behavior in economics usually employs 'fine-to-fee' ratios ranging from 1 to 4 (e.g. Nikiforakis and Normann, 2008). In order to avoid extreme cases and, hence, push the results in one direction or the other, we opted for a value in the middle of the range, equal to 2.[13] Such a 'fine-to-fee' ratio has been extensively used in previous economic experiments (e.g. de Quervain *et al*., 2004; Cubitt *et al*., 2011) and, therefore, ensured a high degree of comparability with other studies on punishment behavior.

[Table 1 about here]

Figure 1 displays how the 'fine-to-fee' ratio evolved over different values of the take rate in the treatments under a variable 'fine-to-fee' ratio and under a constant 'fine-to-fee' ratio respectively. The graph clearly shows that under a variable 'fine-to-fee' ratio the effectiveness of the punishment increases exponentially as the take rate increases. As a consequence, subjects might punish simply because it is more 'convenient' to do so and not because they experience negative emotions (which is to be expected for increasing take rates). In other words, in this set-up, the idiosyncratic features of the punishment technology might

[11] The exception is one subject from Australia who played in the role of a take authority. It is also worth noting that there was a predominance of Chinese among the non-UK subjects (see the on-line appendix).

[12] In order to employ a constant 'fine-to-fee' ratio, we simply stated in the instructions that "for each 1% of his or her endowment that participant B decides to destroy, 10 pence of the transfer to participant A will be destroyed as well". In addition, we allowed subjects to deduct the cost of punishing from their show-up fee, if needed. For this reason, the show-up fee was set at the level of £5 to ensure that, at worst, subjects (in particular, participants B) could leave the experiment with £2.50 in their pockets. This ensured that participants B could punish participants A for any possible value of the take rate without incurring losses. A critical reader might argue that there might be a wealth effect at work due to the fact that under a variable 'fine-to-fee' ratio participants B could earn at worst £5, which is £2.50 more compared to what participants B under a constant 'fine-to-fee' ratio could, in the worst case, earn. In particular, participants B under a constant 'fine-to-fee' ratio might punish less even if they experience intense negative emotions, because they will earn less money if they punish too much. This potential confound goes systematically against our findings (i.e. punishment was frequent and driven by negative emotions under both conditions) and therefore, if anything, it makes the results stronger.

[13] A value of 2 means that the cost of punishing is half of how much the punishment reduces participant A's endowment.

induce an effect on behavior which can be confounded with that of negative emotions (for a discussion of confounds in experiments see Zizzo, 2011). This possibility is instead ruled out in the constant 'fine-to-fee' ratio treatments, where no confound can be attributed to the punishment technology.

[Figure 1 about here]

3. Results

Behavior of the take authorities. Table 2 displays the take rates of UK and non-UK subjects under both constant and variable 'fine-to-fee' ratios. The results are in line with previous PTTG experiments. However, there is evidence of behavioral differences between UK and non-UK subjects. In particular, non-UK subjects displayed statistically significantly higher take rates than UK subjects both in aggregate (Mann-Whitney $p = 0.023$) and under variable 'fine-to-fee' ratios (Mann-Whitney $p = 0.012$). If we compare the behavior under constant or variable 'fine-to-fee' ratios, subjects on average appropriated more resources under a variable 'fine-to-fee' ratio. This evidence, however, does not achieve statistical significance in aggregate (Mann-Whitney $p = 0.191$).[14] No significant difference occurred between UK and non-UK subjects under a constant 'fine-to-fee' ratio (Mann-Whitney $p = 0.351$). We can thus present our first result.

[Table 2 about here]

Result 1. Non-UK take authorities appropriated more resources than UK take authorities, in particular under a variable 'fine-to-fee' ratio.

A Tobit regression analysis confirmed this result.[15] Table 3 presents the outcomes of this regression. The dependent variable is the take rate. Explanatory variables include a dummy variable for the experimental sessions under a constant 'fine-to-fee' ratio, experience in previous experiments,[16] gender (Male = 1 for male subjects), nationality (non-UK = 1 for non-UK subjects), age, and an interaction term between the dummy for the constant 'fine-to-

[14] For UK subjects the difference was not statistically significant (Mann-Whitney $p = 0.732$). For non-UK subjects the difference was mildly statistically significant (Mann-Whitney $p = 0.098$).
[15] 10 observations were left-censored and 16 right-censored.
[16] The data for 'experience' was collected from the final questionnaire provided to the subjects. In particular, subjects were asked to indicate whether they had previously participated in "0", "1", "2" or "3 or more than 3" experiments.

fee' ratio and nationality.[17] The coefficient of the dummy variable *non-UK* is positive and statistically significant, whereas the coefficient of the interaction term is not statistically significant. This implies that non-UK take authorities took more money than UK take authorities both under a variable and constant 'fine-to-fee' ratio once covariates were controlled for. Finally, there is evidence that, *caeteris paribus*, older subjects took on average less than younger subjects.

[Table 3 about here]

Behavior of the responders. We define the *punishment rate* as the proportion of the amount taken by the take authority that was destroyed by the responder.[18] Table 4 displays the punishment rates for UK and non-UK subjects under both constant and variable 'fine-to-fee' ratios. Similarly to the behavior of the take authorities, non-UK responders displayed significantly higher punishment rates compared to UK responders (Mann-Whitney p = 0.045).[19] The difference is even larger if we consider only the punishment rates of those responders who destroyed (Mann-Whitney p = 0.021).[20] No significant difference occurred, on aggregate, between punishment rates of responders under a constant or variable 'fine-to-fee' ratio (Mann-Whitney p = 0.125).[21] However, the difference is mildly statistically significant if we consider only non-UK responders (Mann-Whitney p = 0.071).[22] Our second result follows.

[Table 4 about here]

Result 2. There is preliminary evidence that, contrary to UK responders, non-UK responders punished more harshly, particularly under a variable 'fine-to-fee' ratio.

[17] The results remain qualitatively the same if we exclude the demographic variables from the explanatory variables.
[18] The punishment rate coincides with the destroy rate under a variable 'fine-to-fee' ratio, and is equal to $2d/t$ under a constant 'fine-to-fee' ratio, that is the amount destroyed ($10\,pence \times d$) over the amount taken ($5\,pounds \times t$).
[19] The difference is larger under a variable 'fine-to-fee' ratio (p = 0.054), whereas it does not achieve statistical significance under a constant 'fine-to-fee' ratio (p = 0.346).
[20] If we consider only the punishment rates of responders who destroyed under a variable 'fine-to-fee' ratio, the difference is still significant (p = 0.037), but it is not significant under a constant 'fine-to-fee' ratio (p = 0.347).
[21] This is not the case if we consider only those responders who destroyed (Mann-Whitney p = 0.076).
[22] The difference is even larger if we consider only non-UK responders who destroyed (p = 0.022). For UK subjects, the difference is not statistically significant (Mann-Whitney p = 0.613), even if we consider only responders who destroyed (Mann-Whitney p = 0.975).

Because of several psychological reasons (e.g. inequality aversion, reciprocity), it is possible that the punishment rate depends on the amount taken by the take authority. We tested for this possibility by running a Tobit regression (see Table 5).[23] The dependent variable is the punishment rate. Once we include the take rate of the take authority among the explanatory variables, any treatment effect disappears. The only coefficient statistically significant is the variable take rate. In particular, the coefficient is positive, meaning that the take rate from the take authority negatively affected the punishment behavior of the responder and drove Result 2.

[Table 5 about here]

Result 3. Non-UK subjects punished more often and more severely under a variable 'fine-to-fee' ratio simply because they experienced higher take rates.

Role of emotions. We now turn to the main focus of this paper - that is, the role played by emotions in driving the punishment behavior of the subjects. First, it is worth pointing out that, as seen in the previous literature, different emotions captured similar underlying emotional states. In particular, anger is strongly positively correlated to irritation (Spearman $\rho = 0.81$ and 0.74, $p = 0.000$), envy to jealousy ($\rho = 0.84$ and 0.81, $p = 0.000$), and happiness to joy ($\rho = 0.86$ and 0.72, $p = 0.000$) for UK and non-UK subjects respectively.

In order to study what explains the different emotions experienced by the subjects, we ran some ordered logit regressions,[24] one for each emotion, where the dependent variable was the emotion of interest, whereas the independent variables were the take rate received from the take authority, a dummy variable which took value 1 when a constant 'fine-to-fee' ratio was employed, experience, gender (Male = 1 for male subjects), age, nationality of the subjects (non-UK = 1 for non-UK), and an interaction term between the dummy for the

[23] There were 87 left-censored observations and 19 right-censored. We also tried a logit regression where the dependent variable was a dichotomous variable taking value 1 if the responder destroyed and 0 otherwise. The results are similar to those presented in the paper. However, this approach omits much of the information about the punishment rate and, therefore, is less preferred than the approach based on the Tobit model. Note also that the qualitative results of our Tobit regression do not change if we do not include the demographic variables among the explanatory variables.

[24] Robust standard errors were employed to control for heteroscedasticity. Due to some subjects failing to report all the emotions, we have 1 missing observation for sadness, shame and envy (140 observations instead of 141), and 3 missing observations for contempt (138 observations instead of 141).

constant 'fine-to-fee' ratio and nationality. The results of these regressions are shown in Table 6.[25]

[Table 6 about here]

As in previous PTTG studies, negative emotions (in particular, anger, irritation, contempt and envy) are significantly positively related to the take rate. Similarly, positive emotions (in particular, happiness and joy) are significantly negatively related to the take rate. In contrast to the previous literature, we also found that surprise is negatively related to the take rate.[26] We can thus derive the following result.

Result 4. Higher take rates resulted in subjects experiencing more intense negative emotions and less intense positive emotions. This evidence is robust regardless of the background of the subjects (UK versus non-UK students) and the type of punishment technology employed.

We now consider the punishing behavior of the subjects and to what extent it can be explained by negative emotions. To do so, we ran a battery of Tobit regressions.[27] For each regression, the dependent variable was the punishment rate. Explanatory variables included the emotion of interest, a dummy for the 'fine-to-fee' ratio (fine-to-fee = 1 for a constant 'fine-to-fee' ratio), nationality (non-UK = 1 for non-UK subjects), experience in economic experiments, age, and gender (Male = 1 for male subjects) of the subjects, and interaction terms between the emotion of interest and the dummy for the 'fine-to-fee' ratio, nationality and experience respectively. The results of these regressions are displayed in Tables 7 and 8.[28]

[25] The qualitative results do not change if we do not include the demographic variables among the explanatory variables.
[26] The coefficients of the other explanatory variables are mostly not significant. We briefly mention here those which are significant. In particular, subjects with increasing experience in laboratory experiments experienced less fear ($p = 0.012$) and joy ($p = 0.028$). Older subjects experienced less fear ($p = 0.037$), anger ($p = 0.048$) or sadness ($p = 0.052$). Non-UK subjects under a constant 'fine-to-fee' ratio were on average sadder ($p = 0.029$). Finally, male subjects experienced more sadness ($p = 0.052$).
[27] Similar results are obtained if we run logit regressions and employ a dummy variable as the dependent variable, taking value 1 if the responder destroyed and 0 otherwise. However, this analysis would lack accuracy since we lose the information about the intensity of the punishment.
[28] If we exclude the demographic variables from the explanatory variables, the results remain qualitatively the same.

[Tables 7 and 8 about here]

As in previous PTTG experiments, an increase in the intensity of negative emotions induced responders to punish more. This result is statistically significant for irritation, contempt and shame, and is robust also under a constant 'fine-to-fee' ratio and against differences in the background of the subjects (none of the interactions between the aforesaid negative emotions and the variables *fine-to-fee* and *non-UK* respectively were statistically significant).[29] This brings us to the following result.

Result 5. Independent of the punishment technology employed and the background of the subjects, responders who experienced increasing negative emotions punished their counterparts more.

Finally, subjects with increasing experience in economic experiments were able to better cope with contempt, as they punished significantly less when they experienced such emotions compared to inexperienced or less experienced subjects. Note that experience in previous economic experiments did not eliminate the effect of contempt on punishing behavior, but only reduced it. Hence, our final result occurs.

Result 6. The impact of contempt on the decision to punish was lessened for subjects who had more experience in economic experiments.

4. Discussion and conclusion

The contribution of this paper to the experimental literature in economics is threefold. First, we investigated whether previous findings about emotions and behavior in the PTTG were confounded by the punishment technology. Second, we tested the robustness of these findings against possible differences due to the different backgrounds of the subjects. Third, we studied whether experience in previous economic experiments affects how emotions drive the punishing behavior of the subjects.

[29] It is worth noting that non-UK subjects punished significantly more when they experienced increasing surprise. We also ran the regressions without the interaction terms. The results for irritation, contempt and shame did not change. Interestingly, the coefficient of anger becomes statistically significant ($\beta = 18.62$, $p = 0.001$) as well as the ones for happiness ($\beta = -13.22$, $p = 0.079$) and joy ($\beta = -14.53$, $p = 0.064$).

With respect to the first contribution, our results provide clear-cut evidence of the importance of negative emotions in the PTTG. In particular, irritation, shame and contempt appear to be important driving forces for the punishing behavior of the responders, even once we controlled for the punishment technology and cultural background of the subjects. In addition, we did not find any significant difference in behavior between treatments under a variable or constant 'fine-to-fee' ratio once covariates were controlled for. To begin with, the take authorities did not seem to anticipate the fact that, under a constant 'fine-to-fee' ratio, responders had a lower incentive to punish when the take rates were higher. A possible explanation is that the take authorities felt guiltier to appropriate too much money when the 'fine-to-fee' ratio was constant because it was more costly for the responders to punish compared to a situation with a variable 'fine-to-fee' ratio. Alternatively, the take authorities might have adopted a general norm of fairness on how to split the resources that, conditional to no destruction from the responders, should apply equally to all of the treatments. Another alternative explanation is that the take authorities did not want to trigger a negative emotional response from the responders because it could in turn trigger a punishment. If the emotional response is totally independent of the punishment technology (which, as our data suggests, is indeed the case), then the take authority, independently of the punishment technology, should select the same take rate in order to avoid the punishment.

If we turn to the behavior of the responders, they displayed similar punishment rates under a constant or variable 'fine-to-fee' ratio, once we controlled for other covariates. This means that the punishing behavior of the responders was not affected by the punishment technology. This result is consistent with the hypothesis that negative emotions drive the punishing behavior of the subject. In particular, subjects who experienced strong negative emotions punished even if it was not in their best interest to do so and also when the incentives to punish were not positively correlated with the offence.

Our second contribution concerns the effect of the subjects' backgrounds on behavior. We found behavioral differences between UK students and non-UK students. Non-UK students in the role of take authorities appropriated more resources than UK students, particularly when the punishment technology embedded a variable 'fine-to-fee' ratio. As a consequence, responders punished more often and more severely in the sessions with non-UK students, most likely because of psychological motives such as reciprocity (Rabin, 1993) or inequality aversion (Fehr and Schmidt, 1999). It is more difficult to pin down a unique explanation of why non-UK students appropriated more resources than UK students. We mention here three. First of all, non-UK students might be more sensitive to social distance

since, as a minority, they are more likely to be matched with students from a different country than their own.[30] Several experimental studies have indeed shown that as social distance increases people become more anti-social (e.g. Buchan and Croson, 2004; Charness and Gneezy, 2008).

A second explanation might be related to potential cultural differences between Western societies and non-Western societies. Note that almost all the non-UK students who played in the role of take authority were from non-Western societies (in particular East Asia) and they were also those who displayed the highest take rates. This however is a rather speculative explanation, since it is based on an extreme cultural separation. In addition, we were not able to control for other unobserved variables which might instead better explain the difference in behavior. For instance, non-UK students are more likely to represent the wealthiest subset of the population of their countries of origin, and this might explain their different behavior.

A third explanation is that non-UK students might have found it more difficult to understand the instructions because of the language gap. If this was the case, then the non-UK students should have made more mistakes than the UK subjects in the comprehension questionnaire. However, the hypothesis that the number of mistakes was the same between the two populations was not rejected (Mann-Whitney $p = 0.627$). In addition, both UK and non-UK students with incorrect answers were equally provided with individual clarification in order to ensure that everyone understood the instructions. Furthermore, we should have observed more variance in the take rates of non-UK students than UK students if the former did not correctly understand the instructions. This is because the lack of understanding would have resulted in a more random behavior. However, we did not find any evidence of any difference in variance between the take rates of the two populations (Levene's robust $p = 0.219$).[31] Finally, students who are non-native English speakers were asked to provide formal evidence of their English language skills in order to be initially admitted to the university, and these standards are quite high. Therefore, it is unlikely that the non-UK students completely failed to understand the experimental instructions or did not understand them very well.

[30] Note that subjects were not aware of being in a session with only UK or non-UK students since (a) we did not inform them about our treatment manipulations; (b) subjects usually arrived at the experiment in dribs and drabs, and each of them was immediately seated at a workstation after registration; (c) workstations were isolated by partitions so no-one could figure out the composition of the participants in a given session.

[31] The null hypothesis of no difference between variances was also not rejected under the alternative formulations proposed by Brown and Forsythe of the Levene's test statistic (Brown and Forsythe, 1974).

Finally, we found that experience in previous economic experiments had very little impact on our experiment. In particular, the more experienced subjects did not appear to cope better with their emotional urges when they drove subjects to inefficient behaviors compared to less experienced subjects. Only the impact of contempt on punishing behavior appeared to be lower for the more experienced subjects.

To conclude, our findings significantly contribute to the current state of the experimental literature in economics, particularly from a methodological point of view. First, our results provide clear-cut evidence of the importance of emotions in the PTTG, and the marginal role that a potential *relevant* confound such as the punishment technology plays in driving the results of previous studies. Second, although we are not able to provide a definite explanation, we found that the non-UK students behaved systematically differently to the UK subjects in our experiment. This result should invite experimenters to place greater attention on the composition of their subjects' pool, particularly when the experiment is conducted with university students who possess different cultural and socio-economic backgrounds, in order to avoid spurious conclusions or wrong generalizations. Finally, we found only little evidence that *experimental learning* works as a potential moderator for the impact of negative emotions on punishing behavior and, therefore, it does not seem to be particularly crucial for experiments that study the emotional basis of economic behavior.

Acknowledgements

I wish to thank Abhijit Ramalingam, Ernesto Reuben, Daniel J. Zizzo and seminar participants at the University of East Anglia for useful feedback and advice, Melanie Parravano for experimental assistance in the laboratory, and the Centre for Behavioural and Experimental Social Science (CBESS) and the University of East Anglia for financial support. The usual disclaimer applies.

References

Albert, M., & Mertins, V. (2008). Participation and decision making: a three-person power-to-take experiment. *Justus-Liebig-University Giessen Discussion Paper Series in Economics 05-2008*.

Ben-Shakhar, G., Bornstein, G., Hopfensitz, A., & van Winden, F. (2007). Reciprocity and emotions in bargaining using physiological and self-report measures. *Journal of Economic Psychology*, *28*(3), 314-323.

Bosman, R., & Riedl, A. (2004). Emotions and economic shocks in a first-price auction: an experimental study. *University of Amsterdam Working Paper*.

Bosman, R., & van Winden, F. (2002). Emotional hazard in a power-to-take experiment. *Economic Journal, 112*(476), 147-169.

Bosman, R., Hennig-Schmidt, H., & van Winden, F. (2006). Exploring group decision making in a power-to-take experiment. *Experimental Economics, 9*(1), 35-51.

Bosman, R., Sutter, M., & van Winden, F. (2005). The impact of real effort and emotions in the power-to-take game. *Journal of Economic Psychology, 26*(3), 407-429.

Brown, M. B., & Forsythe, A. B. (1974). Robust Tests for Equality of Variances. *Journal of the American Statistical Association*, 69, 364–367.

Buchan, N., & Croson, R. (2004). The boundaries of trust: own and others' actions in the US and China. *Journal of Economic Behavior & Organization, 55*(4), 485-504.

Casari, M. (2005). On the design of peer punishment experiments. *Experimental Economics, 8*(2), 107-115.

Charness, G., & Gneezy, U. (2008). What's in a name? Anonymity and social distance in dictator and ultimatum games. *Journal of Economic Behavior & Organization, 68*(1), 29-35.

Cubitt, R. P., Drouvelis, M., & Gächter, S. (2011). Framing and free riding: emotional responses and punishment in social dilemma games. *Experimental Economics, 14*(2), 254-272.

de Quervain, D. J.-F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. (2004). The neural basis of altruistic punishment. *Science, 305*(5688), 1254-1258.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*(6868), 137-140.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics, 114*(3), 817-868.

Frank, R. H. (1988). *Passion within reason*. New York: W.W. Norton and Company.

Friedman, J. W. (1969). On experimental research in oligopoly. *The Review of Economic Studies, 36*(4), 399-415.

Galeotti, F. (2013). An experiment of waiting time and punishing behavior. *Economics Bulletin*, 33(2), 1383-1389.

Greiner, B. (2004). The online recruitment system ORSEE 2.0 - A guide for the organization of experiments in economics. *Working Paper Series in Economics 10, University of Cologne, Department of Economics*.

Hopfensitz, A., & Reuben, E. (2009). The importance of emotions for the effectiveness of social punishment. *Economic Journal, 119*(540), 1534-1559.

Joffily, M., Masclet, D., Noussair, C. N., & Villeval, M. (2011). Emotions, sactions and cooperation. *IZA Discussion Paper No. 5592*.

Kräkel, M. (2008). Emotions in tournaments. *Journal of Economic Behavior & Organization, 67*(1), 204-214.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin, 127*(2), 267-286.

Mesquita, B., & Frijda, N. H. (1992). Cultural variations in emotions - a review. *Psychological Bulletin, 112*(2), 179-204.

Nikiforakis, N., & Normann, H. T. (2008). A comparative statics analysis of punishment in public-good experiments. *Experimental Economics, 11*(4), 358-369.

Phelps, E. A. (2009). The study of emotion in neuroeconomics. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics* (pp. 233-250). London: Elsevier.

Pillutla, M. M., & Murnighan, J. K. (1996). Unfairness, anger, and spite: emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes, 68*(3), 208-224.

Rabin, M. (1993). Incorporating fairness into game-theory and economics. *American Economic Review, 83*(5), 1281-1302.

Reuben, E., & van Winden, F. (2008). Social ties and coordination on negative reciprocity: the role of affect. *Journal of Public Economics, 92*(1-2), 34-53.

Reuben, E., & van Winden, F. (2010). Fairness perceptions and prosocial emotions in the power to take. *Journal of Economic Psychology, 31*(6), 908-922.

Rick, S., & Loewenstein, G. F. (2008). The role of emotion in economic behavior. In M. Lewis, J. M. Haviland-Jones, & L. Feldman-Barrett (Eds.), *The handbook of emotion* (3rd ed.). New York: Guilford.

Sutter, M., Bosman, R., Kocher, M. G., & van Winden, F. (2009). Gender pairing and bargaining - beware the same sex! *Experimental Economics, 12*(3), 318-331.

Zizzo, D. J. (2011). Confounds and objectives in economic experiments. *SSRN Working Paper*.

Figures and Tables

Figure 1: Patterns of variable and constant 'fine-to-fee' ratios

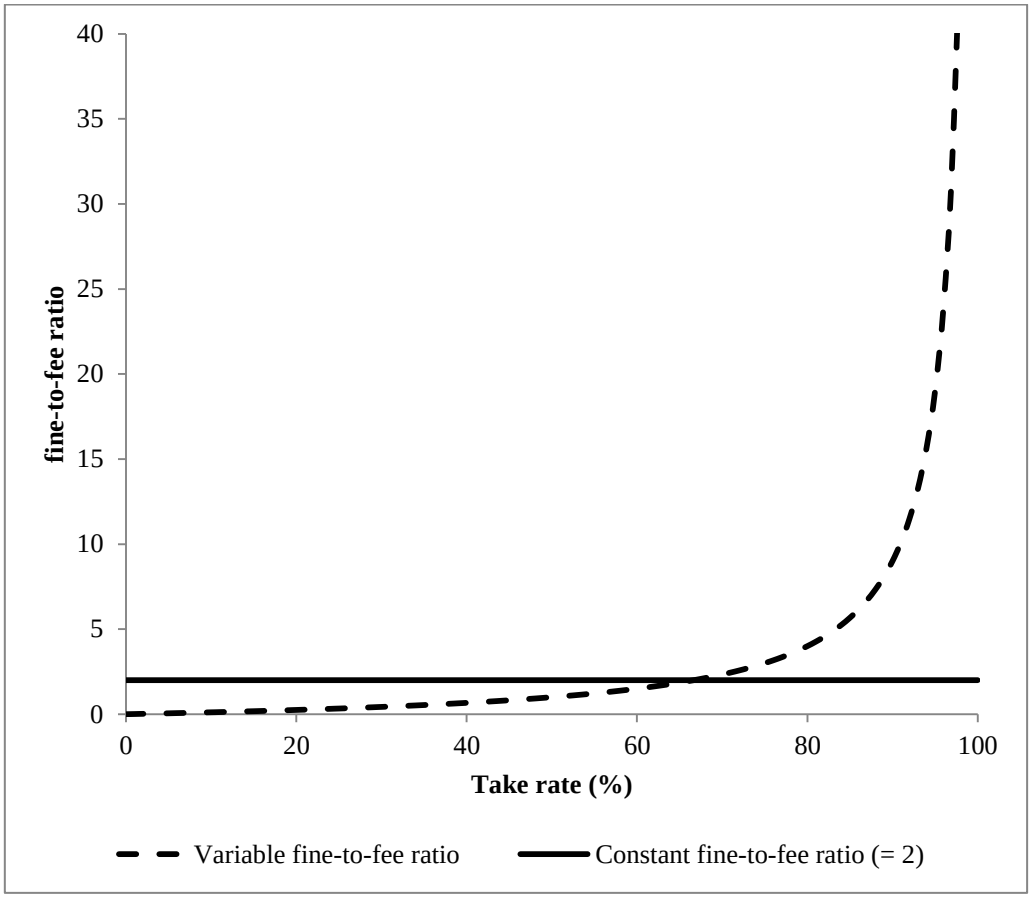

Table 1: Features and independent observations of the experimental treatments

	Variable ratio	**Constant ratio**
UK subjects	37	34
non-UK subjects	33	37

Table 2: Take rates

	Variable fine-to-fee ratio				**Constant fine-to-fee ratio**			
	n.	**Mean**	**Median**	**St. dev.**	**n.**	**Mean**	**Median**	**St. dev.**
UK	37	57	60	26.89	34	54.34	60	30.11
non-UK	33	72.27	80	20.71	37	62.70	60	25.35
Total	70	64.2	70	25.20	71	58.70	60	27.85

Table 3: Tobit regression on take rate

	Take rate	
	b	p
Constant ratio	-3.232	0.657
non-UK	22.450***	0.003
Experience	-0.115	0.966
Constant ratio × non-UK	-8.483	0.422
Male	4.673	0.378
Age	-1.719***	0.008
Constant	91.046****	0
Obs	141	
Pseudo R-Square	0.015	
Df	135	
Prob > F	0.008	

Notes: Tobit regression.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001.

Table 4: Punishment rates

	Variable fine-to-fee ratio				**Constant fine-to-fee ratio**			
	n.	**Mean**	**Median**	**St. dev.**	**n.**	**Mean**	**Median**	**St. dev.**
UK	37	17.86	0	34.58	34	13.17	0	25.82
non-UK	33	41.39	30	44.96	37	20.60	0	32.76
Total	70	28.96	0	41.25	71	17.04	0	29.67

Table 5: Tobit regression on punishment rate

	Punishment rate	
	b	p
Take rate	2.292****	0
Constant ratio	-27.561	0.358
non-UK	41.742	0.152
Experience	-5.243	0.581
Constant ratio × non-UK	-11.128	0.785
Male	-29.768	0.158
Age	-3.341	0.136
Constant	-72.883	0.282
Obs	141	
Pseudo R-Square	0.068	
Df	134	
Prob > F	0	

Notes: Tobit regression.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001.

Table 6: Ordered logit regressions on emotions

	Fear		Envy		Anger		Sadness		Happiness		Shame	
	b	p	b	p	b	p	b	p	b	p	b	p
Take rate	0.01	0.409	0.01*	0.09	0.03****	0	0.01	0.247	-0.04****	0	0.01	0.516
Constant ratio	0.22	0.688	-0.24	0.589	-0.02	0.966	-0.68	0.119	0.08	0.885	0.19	0.779
non-UK	1.08*	0.057	0.13	0.801	0.39	0.413	0.56	0.224	0.37	0.479	0.75	0.229
Experience	-0.37**	0.012	-0.12	0.394	-0.02	0.906	-0.19	0.196	-0.05	0.787	-0.1	0.607
Constant ratio × non-UK	0.06	0.932	-0.01	0.988	0.25	0.69	1.44**	0.029	-0.24	0.733	0.34	0.683
Male	-0.28	0.482	0.19	0.575	0.41	0.283	0.66*	0.052	0.29	0.477	0.42	0.341
Age	-0.08**	0.037	-0.07	0.113	-0.07**	0.048	-0.06*	0.052	0.01	0.799	0.04	0.329
Obs	141		140		141		140		141		140	
Pseudo R-Square	0.05		0.02		0.07		0.06		0.07		0.04	
Prob > F	0.01		0.26		0		0		0		0.24	

	Irritation		Contempt		Joy		Jealousy		Surprise	
	b	p	b	p	b	p	b	p	b	p
Take rate	0.03****	0	0.01**	0.037	-0.02***	0.005	0.01	0.166	-0.01**	0.049
Constant ratio	0.28	0.449	0.21	0.657	0.13	0.799	-0.19	0.629	0.16	0.695
non-UK	-0.49	0.288	0.47	0.293	0.39	0.46	0.37	0.458	0.25	0.604
Experience	-0.08	0.584	0.05	0.743	-0.31**	0.028	-0.07	0.581	-0.15	0.234
Constant ratio × non-UK	-0.05	0.943	0	1	-0.06	0.933	0.29	0.638	0.28	0.659
Male	0.52	0.142	0.19	0.584	0.14	0.706	0.63*	0.055	-0.22	0.501
Age	-0.07	0.116	0.01	0.848	0.02	0.602	-0.09*	0.051	0.02	0.687
Obs	141		138		141		141		141	
Pseudo R-Square	0.07		0.02		0.04		0.03		0.02	
Prob > F	0		0.3		0.04		0.15		0.08	

Notes: Ordered logit regressions with robust standard errors. * p<0.1, ** p<0.05, *** p<0.01, **** p<0.001.

Table 7: Tobit regressions on punishment rate

	Fear		Envy		Anger		Sadness		Happiness	
	b	p	b	p	b	p	b	p	b	p
fine-to-fee	-103.96**	0.013	-58.60	0.218	-38.52	0.366	-65.54	0.138	-74.47	0.069
non UK	96.36**	0.021	40.18	0.398	16.99	0.682	111.27**	0.019	80.11**	0.041
experience	-30.46	0.12	-23.92	0.274	-0.77	0.97	13.60	0.547	1.12	0.947
Age	-4.52*	0.089	-4.37	0.11	-2.29	0.344	-3.68	0.15	-4.11	0.118
Male	3.46	0.881	1.65	0.943	-16.25	0.458	-0.72	0.975	-6.00	0.79
fear	-20.74	0.36								
fine-to-fee×fear	27.39	0.105								
non UK×fear	-18.73	0.265								
experience×fear	9.12	0.195								
envy			-15.25	0.327						
fine-to-fee×envy			2.94	0.797						
non UK×envy			6.56	0.55						
experience×envy			3.36	0.515						
anger					21.85	0.15				
fine-to-fee×anger					-4.02	0.693				
UK×anger					6.92	0.479				
experience×anger					-2.01	0.661				
sadness							29.33	0.147		
fine-to-fee×sadness							5.73	0.655		
non UK×sadness							-21.37	0.12		
experience×sadness							-7.11	0.234		
happiness									0.67	0.969
fine-to-fee×happiness									13.54	0.37
non UK×happiness									-10.94	0.455
experience×happiness									-6.64	0.301
Constant	134.89	0.138	140.60	0.146	-20.51	0.813	-16.62	0.858	88.27	0.259

Notes: Tobit regressions. * p<0.1, ** p<0.05, *** p<0.01, **** p<0.001.

Table 8: Tobit regressions on punishment rate

	Shame		Irritation		Contempt		Joy		Jealousy		Surprise	
	b	p	b	p	b	p	b	p	b	p	b	p
fine-to-fee	-29.64	0.414	-37.13	0.412	-13.71	0.721	-68.72*	0.086	-75.34	0.109	-50.11	0.269
non UK	65.45*	0.08	57.46	0.204	32.33	0.402	77.76**	0.044	71.01	0.132	-42.42	0.344
experience	12.92	0.422	19.44	0.403	30.58	0.115	7.91	0.666	-24.33	0.283	25.14	0.296
age	-4.50*	0.078	-3.16	0.197	-3.86	0.114	-3.68	0.148	-4.15	0.128	-3.99	0.115
male	-10.39	0.631	-17.47	0.438	-5.58	0.795	-7.37	0.744	-3.32	0.887	-4.22	0.849
shame	62.126**	0.012										
fine-to-fee×shame	-14.01	0.418										
non UK×shame	-7.57	0.698										
experience×shame	-11.74	0.111										
irritation			36.67**	0.042								
fine-to-fee×irritation			-5.4	0.627								
non UK×irritation			1.896	0.861								
experience×irritation			-7.25	0.179								
contempt					53.04**	0.012						
fine-to-fee×contempt					-12.77	0.299						
non UK×contempt					7.45	0.542						
experience×contempt					-14.90**	0.018						
joy							7.5	0.694				
fine-to-fee×joy							12.11	0.429				
non UK×joy							-10.01	0.507				
experience×joy							-9.95	0.131				
jealousy									-10.45	0.504		
fine-to-fee×jealousy									6.64	0.538		
non UK×jealousy									-2.41	0.822		
experience×jealousy									3.34	0.529		
surprise											7.65	0.661
fine-to-fee×surprise											-1.88	0.87
non UK×surprise											29.20**	0.017
experience×surprise											-9.14	0.111
Constant	-9.22	0.907	-67.97	0.474	-58.17	0.476	61.84	0.441	123.4	0.209	41.38	0.647

Notes: Tobit regressions. * p<0.1, ** p<0.05, *** p<0.01, **** p<0.001.