

**DISCUSSION PAPERS**  
**Department of Economics**  
**University of Copenhagen**

**06-08**

**Social Ties and Coordination on Negative  
Reciprocity: The Role of Affect**

**Ernesto Reuben, Frans van Winden**

**Studivstræde 6, DK-1455 Copenhagen K., Denmark**  
**Tel. +45 35 32 30 82 - Fax +45 35 32 30 00**  
**<http://www.econ.ku.dk>**

# **SOCIAL TIES AND COORDINATION ON NEGATIVE RECIPROCITY: THE ROLE OF AFFECT\***

*ERNESTO REUBEN*

*UNIVERSITY OF COPENHAGEN*

[ERNESTO.REUBEN@ECON.KU.DK](mailto:ERNESTO.REUBEN@ECON.KU.DK)

*FRANS VAN WINDEN*

*UNIVERSITY OF AMSTERDAM*

[F.A.A.M.VANWINDEN@UVA.NL](mailto:F.A.A.M.VANWINDEN@UVA.NL)

ABSTRACT: This is an experimental study of a three-player power-to-take game where a proposer is matched with two responders. We compare a treatment in which subjects are anonymous to each other (strangers) with one in which responders know each other from outside the lab (friends). We focus on the responders' decisions, beliefs, and emotions. We find that friends punish the proposer more than strangers, and that they are more likely to coordinate their punishment (without communication). Both punishment and coordination are explained by the responders' emotional reactions. Furthermore, the responders' expectations are better predictors of emotions and destruction than their fairness perceptions.

---

\* A previous version of this paper circulated under the title "Reciprocity and emotions when reciprocators know each other". We are thankful for feedback from participants at the Public Choice Society meeting in Baltimore, the ESA meeting in Amsterdam, and the Tilburg Symposium on Psychology and Economics.

# 1. Introduction

Over the last decade, various papers have emphasized the importance of emotions in decision making (e.g. Elster, 1998; Loewenstein, 2000). Emotions may help us explain behavior that is hard to reconcile with existing economic models. For example, Fehr and Gächter (2002) mention emotions as a possible explanation of non-strategic punishment in public good settings. If this were the case, emotion research could then be helpful in improving our understanding of the circumstances facilitating cooperation in social dilemma settings. Other studies have focused on the effects of emotions on preferences (Hirshleifer, 1987; Loewenstein, 1996), the value of emotions as producers of reliable signals (Frank, 1987), the use of emotions as way of harnessing political support (Glaeser, 2005), and emotional hazard as a source of efficiency losses (van Winden, 2001). Others have proposed adaptations of utility functions reflecting emotions. For example, Erard and Feinstein (1994) assume that tax evasion causes disutility because of shame or guilt. Fehr and Schmidt (2000), in discussing the inequality aversion model, argue that individuals experience a disutility if their income is below the income of others because they may experience emotions such as envy.

However, the emotional aspects referred to in these models have not been tested empirically. There is very little research in economics that actually investigates which emotions are at stake in the situation of interest, and what precisely their impact is. If existing models would accurately predict observed behavior this would not be that important. Unfortunately, we have not yet reached this blissful state, as will be evidenced also by the results presented in this paper. Therefore, we think that a much greater effort in investigating the economic significance of emotions, taking seriously the findings of psychology and neuroscience, is warranted. A model based on an incorrect view may lead to incorrect conclusions. For instance, if the driving force behind an individual's decision to punish is anger but we incorrectly model it as envy, we will make wrong inferences regarding the action tendencies and other characteristics of the emotions at stake, and are consequently likely to make wrong predictions. In this case, for example, one might wrongly expect the individual to be in favor of some redistribution, instead of just opposing an intentional act (Lazarus, 1991), and further neglect that people may feel more anger in public than in private (Jakobs et al., 1996).

Recently, a number of experimental studies have begun to *explicitly* investigate the links between emotions and negative reciprocity (e.g. Bosman and van Winden, 2002; Sanfey et al., 2003; Quervain et al., 2004; Bosman et al., 2005). They find substantial support for the economic significance of emotions, using various methods of measurement, and explain the responder's decision to negatively reciprocate as a tradeoff between the emotional satisfaction of punishing an opportunistic act and the (more cognitive) reward of a monetary gain.

The main purpose of this paper is to extend this kind of analysis to cases where there is more than one responder and where the responders either do or do not know each other. In order to do this in a tractable setting we use a three-person version of the power-to-take game (Bosman and van Winden, 2002). In this game, a proposer (or the take authority) can make a claim on the resources of one responder. Subsequently, the responder can destroy any part (including nothing and everything) of her own resources. In this paper, we alter the game so that the take authority can make a claim on the resources of two instead of one responder. Having multiple responders makes this game more realistic for social environments characterized by appropriation of which the power-to-take game captures important aspects, such as taxation, common agency or monopolistic selling (see Bosman and van Winden, 2002). For illustration, one might think of a tax authority selecting an income tax rate, while tax payers can shirk the income tax base (at a cost to themselves). Furthermore, we investigate whether knowing each other as responders affects the emotional and behavioral response.

The step from one to two responders is not at all trivial. On one hand, in the relationship between the proposer and a responder, two new factors are introduced that may affect the outcome of the game. First, if the benefit from punishing the proposer is independent of who does the punishing, an externality is introduced which opens up the opportunity to free ride with less punishment. Previous work on emotions does not tell us whether angry individuals will satisfy their desire to harm the take authority if some one else does the punishing. Furthermore, the possibility of taking money from two instead of one responder makes the situation between the proposer and each of the responders considerably more unequal. If this triggers a more intense emotional response then more destruction could result.

On the other hand, the presence of a second responder introduces another set of considerations, namely, the relationship between the two responders. Responders

may care for their relative payoffs, and in addition, they may also care for how they will feel about each other's behavioral response. In both cases, behavior may be affected by the belief of what the other responder is going to do. If a responder cares for relative payoffs,<sup>1</sup> then her belief regarding the destruction decision of the other responder will determine whether her destruction increases or decreases the expected payoff difference between the two of them. This could lead to either more or less destruction. Alternatively, if a responder is concerned with how her action will be viewed by the other responder, then how much the other responder will destroy becomes, also from this perspective, a reference point for evaluating her decision. For example, if a responder destroys much less than the other, she might be seen as not standing up to unfair behavior, and conversely, if she destroys much more, she might be considered foolish for overreacting. *A priori* it is not clear what the behavioral consequences will be.

Beliefs and emotional responses are likely to be affected by the kind of relationship that exists between the responders. For example, people put different weights on the opinions of others depending on the type and strength of the relationship between them. We care more for what our close friends think of us than what a complete stranger might think. Furthermore, it is to be expected that individuals also differ in the importance assigned to payoff differences depending on who the other person is. It would be natural to expect, for instance, that a responder would mind more if a friend gets a lower payoff than if a stranger does. Although there are a few experiments that study the effects of social distance by having subjects interact across different countries (e.g. Charness et al., 2003), there is practically no work on the effect of real social ties on the behavior of individuals in economic experiments.<sup>2</sup> Hence, it is hard to predict the behavioral consequences of responders being friends instead of strangers. Since ties seem to play an important role in collective action (see Chong, 1991) and are potentially relevant in many economic situations (e.g. work environments), the issue whether and how they affect behavior is in fact of much wider interest. We have therefore decided to give it a prominent place in our experimental design.

---

<sup>1</sup> As in the models of Levine (1998), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), and Falk and Fischbacher (2005).

<sup>2</sup> An exception being Abbink et al. (2002).

The paper is organized as follows. In Section 2 we present the experimental design and link it to related studies. Section 3 describes the experimental procedures. Results are presented in Section 4. Section 5 discusses the main results in relation to the existing literature. Section 6 concludes.

## 2. Design and Related Literature

For our study we use a three-person version of the power-to-take game (Bosman and van Winden, 2002). In this one-shot game, one subject, who can be considered as the ‘take authority’ (with endowment  $E^{take}$ ), is matched with a pair of other subjects, the ‘responders’ (each with an endowment  $E_i^{resp}$  where  $i \in \{1,2\}$  indicates the responder). The game consists of two stages. In the first stage, the randomly chosen take authority decides on the ‘take rate’  $t_i \in [0,1]$ , which is the part of responder  $i$ ’s endowment after the second stage that will be transferred to the take authority. In the second stage, both responders decide simultaneously to destroy a part  $d_i \in [0,1]$  of their own endowment. For the take authority the payoff of the game equals her endowment plus the transfer from each of the responders, i.e.  $E^{take} + t_1(1 - d_1)E_1^{resp} + t_2(1 - d_2)E_2^{resp}$ , while responder  $i$ ’s payoff equals the part of her endowment that she does not destroy minus the amount transferred to the take authority, i.e.  $(1 - t_i)(1 - d_i)E_i^{resp}$ . In order not to introduce too many behavioral issues at a time, in our experiment take authorities can only select a uniform take rate (that is,  $t_1 = t_2 = t$ ) and all the endowments are equal ( $E_1^{resp} = E_2^{resp} = E^{take}$ ).<sup>3</sup>

To study the impact of responders knowing each other, the experiment consists of two treatments, one where responders are anonymous to each other (‘strangers’ treatment), and one where responders know each other (‘friends’ treatment). By comparing the results from the strangers treatment with earlier experiments involving only one responder, we can observe whether the presence of another responder appears to make a difference. By comparing the results from the strangers and friends treatments we can establish whether the existence of a tie between the responders makes a (further) difference. Furthermore, by using self-reports as research method for measuring emotions (Clare and Robinson, 2002) we

---

<sup>3</sup> The power-to-take game differs in three ways from the well-known ultimatum game. First, in the power-to-take game each participant has an endowment. Second, in this game only the endowment of the responder(s) is at stake. And third, the responders can destroy any amount of their endowment.

can determine which emotions are important in these settings and analyze their explanatory value for observed behavior.

The simplicity of our design facilitates the study of the influence of emotions on behavior. First, each responder makes only one decision. This is useful since emotions can impact various decisions and it might be hard to disentangle which emotion influenced which decision. Second, responders cannot influence each other's monetary payoffs. Therefore, we are able to observe how a responder feels about the decision of the other responder without interference of any effect the other responder might have had on the first responder's income.

Our work is related, on the one hand, to studies exploring the economic significance of emotions and, on the other hand, to studies investigating how the presence of others affects decision-making. Although still small in number, there are some studies explicitly dealing with emotions to explain responder behavior in the kind of game investigated in this paper. However, they are all restricted to the one-proposer-one-responder case. A relatively early paper exploring this issue is Pillutla and Murnighan (1996). Responders in an ultimatum game experiment were asked, after each of a series of offers they had to accept or reject, to answer the open-ended question "How do you feel?" All offers were predetermined and afterwards a lottery selected one to determine actual payoffs. Answers to the feeling question were rated for expressions of the emotion of anger, and the rejection of offers was found to be related to this measure of anger. Their analysis also suggests that emotional reactions provide the critical link that determines when fairness perceptions tend to affect behavior. Bosman and van Winden (2002) introduced the power-to-take game with the specific purpose of explicitly investigating the importance of emotions for negative reciprocity in a situation of appropriation. In several experiments they had responders self-report on their feelings, but now concerning a whole list of different emotions (positive as well as negative) and with 7-point scales for them to indicate the felt intensity of the respective emotion (see also van Winden, 2001; Bosman et al., 2005). In addition, they asked for responders' expectations (regarding the take rate). Their results show that the destruction of own resources by responders is related to the intensity of experienced negative emotions (particularly, contempt, irritation, and anger), which in turn is positively related to the actual take rate and negatively to the expected take rate.

Recently, for both games evidence has been found of a biological substrate for the negative reciprocity exhibited by responders. Sanfey et al. (2003), using fMRI of ultimatum game players, find that ‘unfair’ offers elicited activity in brain areas related to both emotion and cognition, and significantly heightened activity in an area related to emotions in case of rejection.<sup>4</sup> Regarding the power-to-take game, Ben-Shakhar et al. (2004), using skin conductance as physiological measure of emotional arousal as well as self-reports, find that both self-reported anger and physiological arousal are related to destruction, with frustrated expectations playing an important and consistent role. Moreover, the self-reported measures of emotions appeared to be correlated with the physiological measures, which is reassuring for the use of self-reports in the study of reciprocity.

In this study we take the important next step in this line of research by studying what happens when a third person is introduced. In this respect, our work is related to papers on three-person ultimatum games. For instance, various authors have conducted experiments using ultimatum games that involve an inactive dummy player (Güth and van Damme, 1998; Kagel and Wolfe, 2001; Bereby-Meyer and Niederle, 2005). They find that responders seem to concentrate on their own as well as the proposers’ payoffs and mostly ignore the welfare of the dummy players. Knez and Camerer (1995) use the strategy method to observe if a pair of responders playing with the same proposer condition their acceptance on the amount offered to the other responder. They find that about half of the responders will condition their response on the income the other responder would get. Riedl and Vyrastekova (2003) ran a three-person ultimatum game experiment in which they varied the effect the rejection of one responder has on the payoffs of another responder. They find that responders are more likely to reject proposals if this does not negatively affect their standing with respect to the other responder. However, all these experiments were not designed for an analysis of emotions and their explanatory value. Hence, important variables from that perspective, such as expectations, were not measured. Our experimental design is a first shot at exploring head-on the affective side of reciprocity in case of multiple potential reciprocators.

Psychological studies suggest that people may react quite differently emotionally to the same situation when others are present, and the more so if the other

---

<sup>4</sup> In a similar study, Quervain et al. (2004) show that the effective punishment of norm violators produces activity in areas of the brain associated with the processing of rewards.



person is a friend rather than a stranger (see e.g. Jakobs et al., 1996; Jakobs et al., 1999). We want to explore the economic relevance of this literature by investigating whether the presence of another responder in the power-to-take game and the nature of the relationship between the responders has an effect not only on their emotional responses but also on their behavior and, furthermore, whether the emotional response is linked to behavior.

### **3. Experimental Procedures**

The computerized experiment was run in November 2003 and May 2005 in the CREED laboratory of the University of Amsterdam. In total 189 subjects, almost all undergraduate students from the University of Amsterdam, participated in the experiment. About 42% of the subjects were students of economics. The other 58% were students from various fields such as biology, political science, law, and psychology. About 46% of the subjects were female. Subjects received a show-up fee of 5 euros, independent of their earnings in the experiment, and 10 euros as endowment. On average, subjects were paid out 13.52 euros. The whole experiment took about one hour.

The experiment consisted of two treatments: a ‘strangers’ treatment, where the two responders in the game did not know each other, and a ‘friends’ treatment, where the responders knew each other. Subjects were allowed to sign up only if they did so as a pair, that is, they had to provide the name of someone they knew and with whom they would take part in the experiment. If a subject signed up with someone else but nevertheless showed up alone to the experiment, he or she was not allowed to participate. In this way we hoped to recruit subjects with social ties. This approach, which is similar to the one used by Abbink et al. (2002), gives the opportunity to employ individuals with stronger bonds than one can establish in the laboratory. In an attempt to measure the strength of each pair’s social tie, we asked each individual to describe the type of relationship they had with their partner and how frequently they saw each other.

After arrival in the lab’s reception room, each pair of subjects drew a card to be randomly assigned to two seats in the laboratory. Once everyone was seated the instructions for the (one-shot) power-to-take game were read, followed by a few exercises to check the subjects’ understanding of the procedures (a translation of the instructions is provided in Appendix A). After these exercises the subjects were

informed, by opening an envelope on their desk, which role (that of ‘take authority’ or ‘responder’) they had been assigned in the game. The game was framed as neutral as possible, avoiding any suggestive terms (like ‘take authority’). Subsequently, the subjects were randomly assigned into groups of three. In the strangers treatment, subjects that came together to the experiment were assigned to *different* groups. Hence, complete anonymity was ensured since none of the members of the group knew who the other group members were. In the friends treatment, each group included a take authority and a pair of responders who signed up together for the experiment. Consequently, in this treatment anonymity was ensured between take authorities and responders but not between the responders themselves. In other words, take authorities knew that the responders in their group were people that came together to the experiment, but they did not know which responders they were. Similarly, responders knew that the other responder in their group was the individual with whom they came to the experiment whereas they did not know the identity of the take authority. The group assignment was clearly explained in the instructions.

Subjects then played the three-person power-to-take game via the computer.<sup>5</sup> During the game, subjects were asked to fill out a few forms indicating not only their decisions but also how they felt, which take rate they expected, and which take rate they considered to be fair. Since in this paper we will concentrate on the responders’ behavior, Figure 1 shows the precise order in which the responders’ decisions, emotions and expectations were measured. Note that we asked subjects to report what they expected others to do before they observed their actual behavior. As in Bosman and van Winden (2002) subjects’ emotions towards other players were measured through self-reports after the subject observed what the others did. We asked for the fair take rate, at the end, in the debriefing questionnaire.

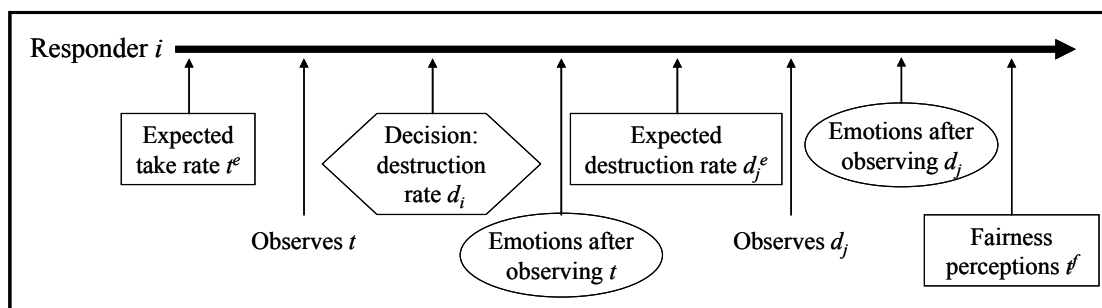


FIGURE 1 – SEQUENCE OF EVENTS FOR RESPONDERS

<sup>5</sup> The experiment was programmed and conducted with the software z-Tree (Fischbacher, 1999).

Expectations were measured by asking subjects to indicate the most likely value for  $t$  or  $d_j$ .<sup>6</sup> In addition to the point estimate, we asked subjects to indicate on a 7-point scale how confident they were of their expectation. Emotions were measured by providing subjects with a list of fourteen emotion names and asking them to report on a 7-point scale with what intensity they experienced each emotion. The scale ranged from “no intensity at all” (1) to “very intensely” (7). The list included the following emotions: admiration, anger, contempt, disappointment, envy, gratitude, guilt, irritation, joy, pride, regret, sadness, shame, and surprise. A variety of emotions were included to avoid pushing subjects in a particular direction.

## 4. Results

In this section we present and analyze the decisions that were taken and the emotions that subjects experienced. Furthermore, we investigate whether the reported emotions help explain the behavior of the responders in both treatments. A summary of the individual data is provided in Appendix B.

### 4.1. Observed behavior

On average, the take rate was 58.6% in the strangers treatment and 62.3% in the friends treatment. Interestingly, these take rates are very similar to the 60.0% mean take rate reported in the comparable one-responder power-to-take game (Bosman et al., 2005). The similarity between the take rates in both treatments leads to our first result.

**RESULT 1:** *Take rates do not differ between the friends treatment and strangers treatment.*

*Support:* Using Wilcoxon-Mann-Whitney (WMW) and Kolmogorov-Smirnov tests, the hypothesis that the take rates are drawn from the same distribution cannot be rejected ( $p = 0.78$  and  $p = 0.42$ ).<sup>7</sup>

Turning now to the responders, in both treatments, a considerable number of responders destroyed some or all of their endowment. In the strangers (friends)

---

<sup>6</sup> We decided to measure expectations in this way since subjects might have difficulty in reporting a probability distribution of a continuous variable (over the interval  $[0,1]$ ).

<sup>7</sup> Throughout the paper, unless it is otherwise noted, we always use a Wilcoxon-Mann-Whitney test. Furthermore, all tests in the paper are two-sided.

treatment 21.4% (40.0%) of the responders destroyed a positive amount. On aggregate, responders in the strangers (friends) treatment destroyed 13.2% (29.4%) of their endowment. For both measures, these results are respectively below (strangers treatment) and above (friends treatment) the ones obtained for the one-responder game. In Bosman et al. (2005) 37.5% of the responders destroyed a positive amount, while on aggregate, 24.70% of the endowment was destroyed. The difference in destruction between strangers and friends is even starker if we concentrate on responders who faced high take rates, that is, take rates that are above the average take rate. This leads us to our second result.

*RESULT 2: Friends destroy more and more frequently than strangers. This difference is due to high destruction rates among friends when faced with high take rates.*

*Support:* Using WMW tests, one can reject the hypothesis that friends and strangers destroy equal quantities ( $p = 0.02$ ) and equally often ( $p = 0.03$ ). Among the responders who faced an above-average take rate, strangers destroyed on aggregate less than friends, namely 29.5% vs. 67.3% ( $p = 0.01$ ). Moreover, only 30.0% of strangers destroyed some of their endowment whereas 78.6% of friends decided to do so ( $p = 0.01$ ). There are no significant differences between friends and strangers for responders who faced a below average take rate ( $p > 0.85$ ). This result is partly driven by the fact that friends are more likely to destroy *all* of their endowment than strangers: 24.3% of the friends destroyed everything while only 8.9% of the strangers did so ( $p = 0.03$ ).<sup>8</sup>

These two results suggest that, whereas at high take rates friends are likely to destroy more and more frequently than strangers, the behavior of take authorities does not depend on whether they are facing a pair of friends or a pair of strangers. This can be further substantiated by looking at the earnings of the take authorities. In the strangers treatment, take authorities choosing a high take rate earned on average 2.25 euros

---

<sup>8</sup> It would certainly be of interest to know if the strength of the social tie between responders has an effect on destruction. Unfortunately, we have very little variation in the two variables used to measure the strength of social ties, and hence, we cannot make a meaningful analysis. Roughly 60% of all pairs reported their type of relationship as a “friendship” while the other 40% was evenly distributed among four categories. Similarly, 60% of all pairs reported their frequency of contact as “very frequent” while the other 40% was evenly distributed among three categories.

*more* than take authorities who chose a low take rate. They also faced more risk, however, in the sense of a higher variance in earnings. In contrast, in the friends treatment, take authorities who chose a high take rate earned on average 4.04 euros *less* than those choosing a low take rate, even though they too faced a larger variation in earnings. Hence, while in the strangers treatment it might make sense to choose a high take rate and risk some variation in income, in the friends case this is clearly an inferior choice. Nevertheless, it turns out that the proportion of take authorities choosing a high take rate is roughly the same in both treatments.

#### *4.2. Determinants of behavior*

To investigate what is motivating a responder to destroy, we estimated a multivariate tobit model for the probability of destruction, using as explanatory variables: demographic data (gender and area of study), the take rate, the expected take rate, the perceived fair take rate, and treatment dummies. In addition, we checked for interaction between the explanatory variables and used any significant interaction term (see Table C1 in Appendix C). The following result is obtained.

**RESULT 3:** *Destruction is positively related to the take rate and to the difference between the actual and the expected take rate.*

*Support:* Judged by the signs of the (weakly) significant coefficients ( $p < 0.10$ ), it appears that responders who are likely to destroy some or all of their endowment are responders who: faced high take rates and/or experienced large positive differences between the take rate and their expected take rate. Furthermore, friends have a bigger coefficient than strangers for the influence of the take rate on the amount of destruction (Wald test,  $p = 0.05$ ). This mirrors our previous result where we found that friends destroy more frequently than strangers when faced with high take rates. Contrary to what one would expect given the emphasis on fairness in the literature, but in line with the findings of Pillutla and Murnighan (1996), it is less clear whether the variable measuring the difference between the actual take rate and the fair take rate is a significant determinant of destruction. Finally, the regression also suggests that women may destroy less than men do.

#### *4.3. Experienced Emotions*

The intensity scores of emotions reveal that subjects experienced a variety of emotions. Concentrating on how emotions influenced the destruction decision gives us a clear result.

RESULT 4: *Destruction is positively (negatively) related to the intensity of experienced negative (positive) emotions.*

*Support:* We estimated a tobit model for each emotion separately, with destruction as the dependent variable and censored at  $d = 0$  and  $d = 100$ . The resulting coefficients and their level of significance are presented in Table 1.

The table shows that, in both treatments, similar negative (as well as positive) emotions are related to destruction behavior. WMW tests give further support for this finding: responders who destroyed reported significantly higher intensities of anger, contempt, disappointment, and irritation, and significantly lower intensities of joy and gratitude ( $p < 0.07$ ). Furthermore, for none of these emotions can we find a significant difference in the intensity scores between the two treatments (for both responders who destroyed and responders who did not,  $p > 0.21$ ).

**TABLE 1 – TOBIT REGRESSIONS FOR EMOTIONS WITH DESTRUCTION AS THE DEPENDENT VARIABLE**

Explanatory variable	Strangers		Friends	
	Coef.	Std. Err.	Coef.	Std. Err.
gratitude	-44.888*	21.338	-87.711**	28.685
anger	33.073*	16.724	54.561**	17.787
disappointment	32.495*	16.831	56.900**	19.427
irritation	29.179*	15.298	55.219**	16.181
joy	-44.862*	23.672	-68.441**	21.996
contempt	30.091*	15.990	55.870**	16.425
admiration	-35.294	21.812	-95.017**	37.292
sadness	14.682	18.829	49.216*	25.170
regret	50.786	33.930	11.786	28.716
envy	23.359	16.224	15.280	16.318
surprise	13.721	15.103	23.085	16.512
shame	-38.207	47.329	18.045	29.285
pride	-10.829	16.398	25.921	18.117
guilt	-10.412	33.129	-22.905	36.298

\*\* Significant at the 5 percent level. \* Significant at 10 percent level.

Having found that destruction is related to experienced emotions, the question arises what explains the different emotional responses. To answer this question we estimated a multivariate ordered probit model, with the intensity of different positive and negative emotions as the dependent variable. For the analysis we concentrated on the emotions that turned out to be good predictors of destruction in both treatments (i.e.

anger, contempt, disappointment, gratitude, irritation, and joy). Again, we used as explanatory variables: demographic data (gender and area of study), take rate, expected take rate, perceived fair take rate, and treatment dummies. In addition, any significant interaction term was included. The following result is obtained.

*RESULT 5: The intensity of negative (positive) emotions is positively (negatively) related to the take rate and to the difference between actual and expected take rates.*

*Support:* In all the regressions for negative (positive) emotions the coefficient of the take rate has a positive (negative) sign. Furthermore, it is weakly significant in all cases ( $p < 0.10$ ) except for disappointment. The same holds true for the coefficient of the difference between the take rate and the expected take rate ( $p < 0.05$ ). For illustration, in Table C2 we present the results of one such regression. In this regression we used as dependent variable the average of the three anger-like emotions: anger, irritation, and contempt. The regression also points at a gender effect. It appears that female subjects are more likely to report lower intensities of negative emotions.<sup>9</sup>

Summarizing, the same variables that are good predictors of destruction behavior (see Result 3) are also good predictors of the intensities of experienced emotions. In combination with the finding that emotions are good predictors of destruction, this suggests the following intuitive explanation for destruction: the higher the take rate and the larger the difference between the take rate and the expected take rate, the stronger the intensity of anger-like emotions experienced by a responder, which in turn makes it more likely that she will destroy in order to punish the take authority. These findings are consistent with, but strengthen and extend, those obtained for the one-responder power-to-take game.

Further evidence that is easily explained with an emotion-driven account of destruction (but is hard to explain otherwise) is the time responders take to make their decision. In our experiment, responders that destroyed a positive amount not only reported higher intensities of negative emotions, they also took more time to make a decision (t test,  $p = 0.07$ ). However, if we focus on responders who destroyed

---

<sup>9</sup> In the regressions of individual emotions we find that women are more likely to report lower intensities of contempt and higher intensities of gratitude and joy.

everything, we find that, even though they reported the highest intensities of negative emotions, they did not take more time to decide than responders who did not destroy (t test,  $p = 0.43$ ). In other words, the slowest responders turn out to be those who reported intermediate intensities of negative emotions and destroyed intermediate amounts. Standard economic theory gives us no reason to think why making the decision to destroy requires more time than making the decision not to destroy. However, research on emotions suggests the following. At low intensities of negative emotions a decision can take little time because there is no real conflict between the (cognitive) interest to earn as much money as possible and the (emotional) urge to punish the proposer. At higher intensities, this conflict does arise and hence one would expect subjects to take more time in order to sort it out. However, if the intensity of the negative emotions becomes very high it can push subjects over a threshold beyond which they are less prone to think and simply follow the emotion's action tendency, which entails less time to reach a decision (Frijda, 1986; Frijda, 1988; Goleman, 1995).

Nevertheless, there is an important aspect of the data that is not explained by the emotional reaction of responders towards proposers. As was pointed out in Result 2, at high take rates friends destroy more and more frequently than strangers. However, we do not find that at high take rates the emotional reaction of responders differs between treatments. This means that friends and strangers are equally angry and unhappy at high take rates ( $p > 0.33$ ). A more detailed look at the data reveals that the disparity between destruction and anger is caused by the fact that angry strangers appear to destroy less frequently and smaller amounts than angry friends.<sup>10</sup> This can be seen in Figure 2.

As can be seen in the figure, in both treatments the ratio of non-angry and angry responders is roughly the same (2.11 to 1 for strangers and 2.04 to 1 for friends). Moreover, the frequency of destruction among non-angry responders is quite similar (15.8% for strangers vs. 21.3% for friends,  $p = 0.52$ ). However, the frequency of destruction among angry responders is considerably different: whereas only 33.3% of the angry strangers destroyed something, 78.3% of angry friends decided to do so ( $p = 0.01$ ). If we look at the amounts destroyed we also find a difference. On

---

<sup>10</sup> We classified subjects as angry if they reported an intensity of 6 or more (on the 7-point scales) for any of the following emotions: anger, irritation or contempt. However, similar results are obtained with different definitions of being angry.



aggregate, angry strangers destroyed 30.0% of their endowment while angry friends destroyed 73.5% ( $p = 0.01$ ).

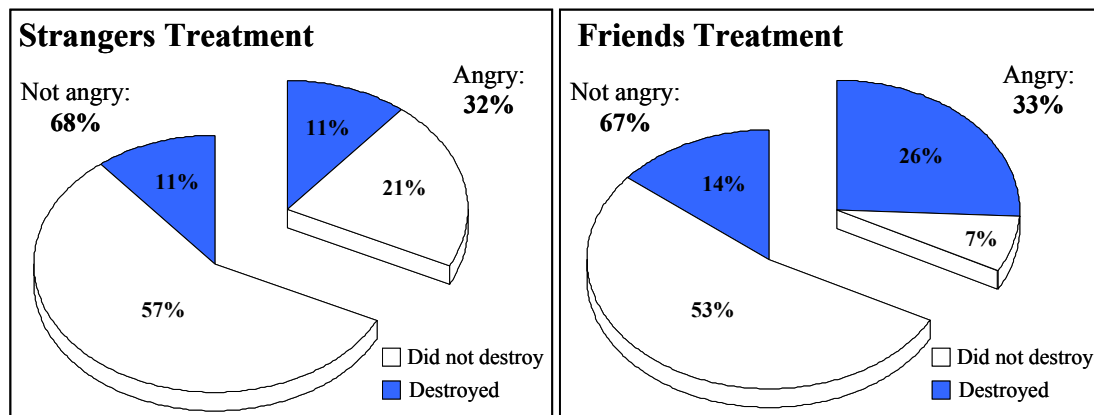


FIGURE 2 – DESTRUCTION BY ANGRY AND NON-ANGRY RESPONDERS

#### 4.4. Coordination

In order to explain this difference, we further analyze the behavior and emotional response of the responders. What we are interested in is to see whether pairs of friends behaved markedly different than pairs of strangers. The following results are obtained.

**RESULT 6:** *Friends are better at coordinating destruction than strangers.*

*Support:* In both treatments a comparable number of pairs of responders coordinate on similar destruction rates (i.e. within 10 percentage points of each other), specifically, 60.7% of the pairs of strangers, and 62.9% of the pairs of friends. However, if we concentrate on pairs in which at least one of the two responders destroyed, we find a large difference between treatments. Among these pairs, in the strangers treatment only 8.3% coordinate on similar destruction rates whereas 42.8% do so in the friends treatment ( $p = 0.04$ ). There is not a significant difference in the case of pairs where at least one responder did not destroy ( $p = 0.90$ ). This result can also be observed if we look at the correlation between the destruction rates within pairs of responders. The correlation coefficient in the friends treatment is significantly higher than in the strangers treatment, 0.560 vs.  $-0.175$  ( $z$  test,  $p = 0.01$ , this includes pairs of responders in which there was no destruction).

A possible explanation for the better coordination of friends compared to strangers is that friends tend to be more alike (which could also explain why they are friends in the first place). However, sharing a similar preference for destruction cannot explain

why we observe such a big difference between the destruction rate in the friends treatment (where pairs of friends played in the same group) and the strangers treatment (where pairs of friends played in separate groups). A more plausible explanation for the better coordination of friends is that, first, they are better at predicting each other's behavior and, second, that they have a preference for coordinating on the same action.

Overall, half of the responders accurately predict the destruction rate of the other responder in their group. Specifically, 53.4% of the strangers and 54.3% of the friends correctly predict the destruction rate of the responder they were paired with (within 10 percentage points). Although we do not see that in general friends predict better than strangers, we do find that friends are better at predicting positive destruction. In total, 37.5% of the strangers and 50.0% of the friends thought the other responder would destroy a positive amount. Among these responders, in the strangers treatment, only 4.7% of them correctly predict the other's destruction rate. Friends do much better with 34.3% of them making an accurate prediction ( $p = 0.01$ ). There are no significant differences between treatments if we look at responders who thought the other would not destroy ( $p = 0.39$ ). The better predicting ability of friends is also evident if we look at the correlation between the actual and the expected destruction rate. The correlation coefficient is significantly higher in the friends treatment (0.518 vs.  $-0.036$ ,  $z$  test,  $p = 0.01$ ).

Lastly, we also find evidence within the strangers treatment indicating that people are better at predicting the destruction rate of their friends. In this treatment subjects also came together as friends but they were assigned to different groups. Nevertheless, they were assigned to the same role.<sup>11</sup> Hence, after informing responders of the take rate faced by their friend, we were able to ask them to predict their friend's destruction rate. If we compare each responder's ability to predict the behavior of their friend vs. the behavior of the responder they were paired with, we find that they do considerably better when predicting their friend's destruction rate. Specifically, 53.6% of responders correctly predict the destruction rate of the other responder in their group, whereas 73.2% of them correctly predict the destruction rate of their friend. The better accuracy of responders when predicting the behavior of

---

<sup>11</sup> Subjects did not learn that they were assigned to the same role until they reached the debriefing questionnaire. The experiment's instructions simply gave no information concerning the role assigned to their friend.

friends is due to better predictions of positive destruction rates. Only 4.8% of responders correctly predict the positive destruction of the other responder in their group. In contrast, 30.8% of them accurately predict the positive destruction rate of their friend ( $p = 0.04$ ).

Even though friends predict destruction better than strangers, this should not lead to more coordination among friends unless responders within a group care about each other's destruction rate. Looking at the emotional response between responders demonstrates that both friends and strangers care about what the other responder does. However, there is one important difference. Whereas strangers wish to avoid being the one who destroys more, friends wish to coordinate on the same destruction. This leads to the following result.

*RESULT 7: The emotional response towards the other responder's behavior facilitates the coordination of destruction among friends but not among strangers.*

*Support:* To back up this result we compare differences in emotional intensity scores across two sets of responders.<sup>12</sup> The first set consists of responders who destroyed more than the responder they were paired with. For convenience, we will call them the 'punishers'. The second set consists of responders who destroyed less than the other responder, which will be labeled the 'acquitters', for short. In the strangers treatment, punishers reported higher intensities of disappointment, irritation, and sadness than acquitters ( $p < 0.07$ ). Given this negative emotional response, it stands to reason that, *ceteris paribus*, strangers would prefer to be among the acquitters rather than the punishers. In contrast, in the friends treatment, punishers reported similar emotional intensities as acquitters for all emotions ( $p > 0.29$ ). Hence, given the choice, friends unlike strangers might be indifferent between being in the punisher or acquitter position. Further support is obtained if we compare acquitters with paired responders who destroyed the same amount: the 'coordinators'. Compared to coordinators, acquitters in the friends treatment reported, on one hand, more anger, irritation, and disappointment, and on the other hand, less joy and gratitude ( $p < 0.08$ ). Therefore, given the choice, friends would presumably prefer to be coordinators

---

<sup>12</sup> Since we are interested in explaining the difference between friends and strangers, we concentrate on the emotions that, in this analysis, exhibited significantly different patterns across treatments, namely: admiration, anger, disappointment, gratitude, irritation, joy, and sadness.

rather than acquitters. In contrast, in the strangers treatment, acquitters reported similar emotional intensities as coordinators for all these emotions ( $p > 0.10$ ). Thus, from this point of view, strangers would be indifferent between being a coordinator and being an acquirer.

Summarizing, in view of the just discussed differences in emotional responses between responders, if subjects anticipate their emotional response then destroying seems to be more risky for strangers than for friends since it might leave them in the ‘punisher’ position inducing additional negative emotions. The very negative emotional reaction experienced by strangers in the punisher position is also evident across treatments. Strangers feel more anger, irritation, and disappointment than friends when they destroy more than the other responder ( $p < 0.04$ ). This may explain why, even when very angry, strangers often decide not to destroy. Friends, on the other hand, seem to have a strong preference for coordination. This is evident from their emotional response. Friends who coordinate experienced more admiration, gratitude, and joy than strangers who coordinate ( $p < 0.01$ ). Consequently, angry friends may be much more inclined to destroy, especially if they believe that the other responder will also destroy. Moreover, the fact that they also predict each other’s behavior better further facilitates them to obtain the positive emotional boost of coordination.

However, we should also discuss a potentially alternative explanation of destruction behavior. Since in our experiment friends had the possibility of interacting after the experiment, side payments were possible. So, one could perhaps argue or conjecture that the stronger coordination among friends who destroy is due to these side payments and not because of any differences in emotional responses. We have not succeeded ourselves in finding a convincing explanation, side-stepping emotions, in which side payments may lead to more coordination among responders. Nonetheless, in order to test if side payments played a role, we asked subjects in the debriefing questionnaire, first, if they intended to share their earnings after the experiment, and, second, if the possibility of sharing earnings after the experiment affected their decision. If side payments would indeed boost coordination one would expect more coordination among responders who answered positively to one or both of the abovementioned questions. However, we find no significant difference in coordination between those who answered positively or negatively to any of these two

questions. This is also true if we look only at angry responders, that is, the responders who acted noticeably different across the two treatments ( $p > 0.35$ ). Hence, we tentatively conclude that although side payments were possible they played no significant role in the game.

#### 4.5. Expected take rates and fairness perceptions

Expectations about the take rate turned out to have an important influence on the intensity of emotions and destruction behavior. This is further illustrated by Figure 3. Responders are divided into two groups: ‘optimists’, that is, people who expected a lower take rate than the one they faced (observations above the diagonal), and ‘pessimists’, who expected a higher take rate than the one they faced (observations below the diagonal). It is easy to see from the figure that destruction, and especially high destruction (at least 50%), is carried out almost exclusively by ‘optimists’.<sup>13</sup> A plausible explanation is that, first, subjects form an expectation of what the average take authority will do and, then, judge the degree of (un)kindness of the take authority by the deviation from this reference point. Negative deviations are judged as unkind and are punished by destruction.

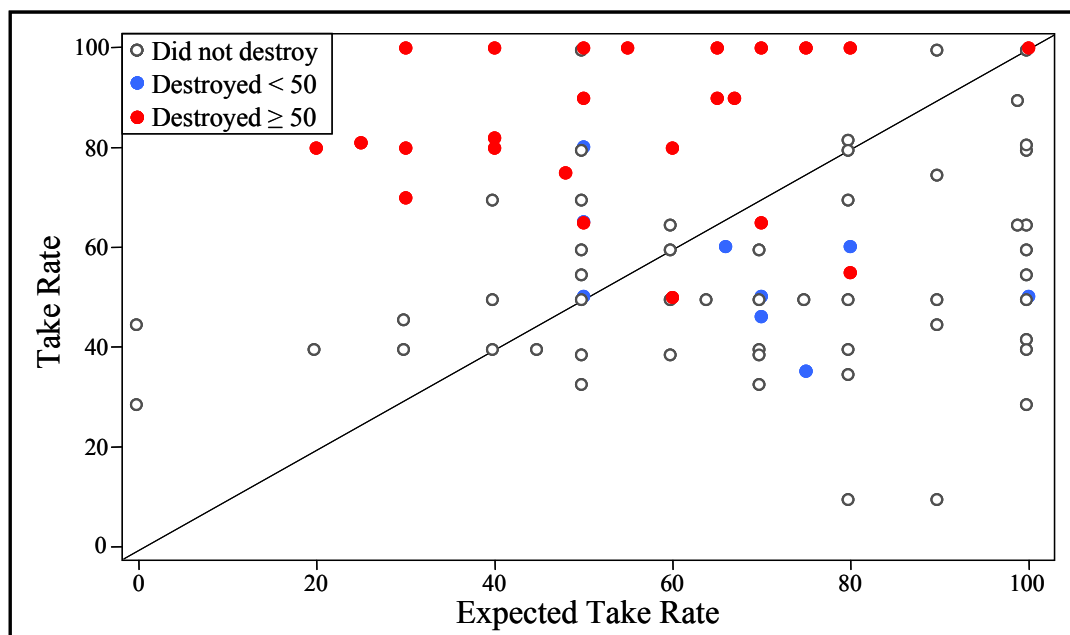


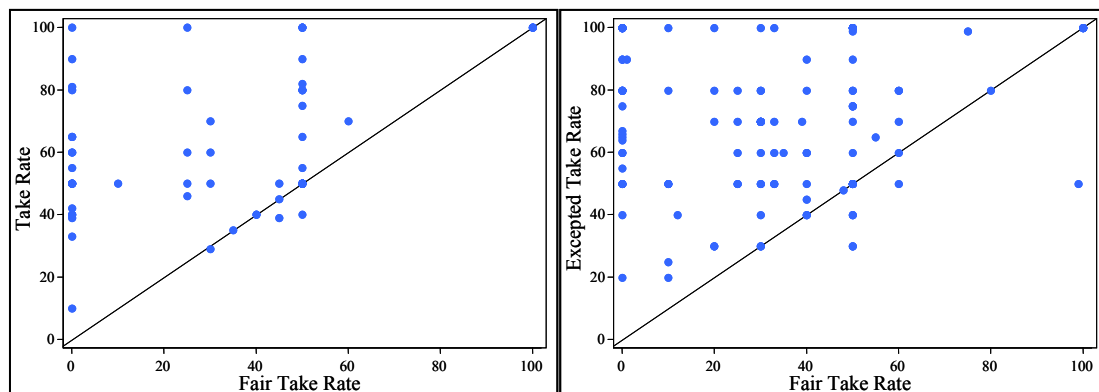
FIGURE 3 – SCATTER PLOT OF ACTUAL AND EXPECTED TAKE RATES

Instead of having as reference point the expectation of what the take authority will do, a responder may (also) be affected by what she thinks the take authority *should* do, that is, the reference point could be the take rate that she considers fair. However, as

<sup>13</sup> Of the responders who destroyed (at least 50%) 65.0% (82.8%) are optimists.

noted in the discussion of Result 3, the responder's fair take rate is not as clearly related to destruction as the expected take rate is. For example, once we control for the effect of the take rate by looking only at take rates in the third quartile, we find that, although responders who experienced an above average difference between the take rate and the fair take rate destroyed more frequently than the rest, this difference is not statistically significant ( $p = 0.38$ ). In contrast, if we do the same test for the difference between the take rate and the expected take rate, we find a higher frequency of destruction among responders who experienced an above average difference ( $p = 0.08$ ).

It appears that deviations from the expected take rate are more important than deviations from the fair take rate. This is the case even though it is argued that equity considerations are especially important in highly asymmetric situations with complete information (Smith, 1976; Fehr et al., 1993). This is not to say that fairness perceptions do not play a role in the responders' decision-making. It may be that fairness plays a more indirect role than usually envisaged. Suggestive in this respect is the following analysis. Focusing on the difference between the expected take rate and the fair take rate shows that in only 5.6% of the cases this difference was negative. In other words, the overwhelming majority of responders expected a higher take rate than the one they considered fair (reflected by the average fair take rate being lower than the average expected take rate). An almost identical pattern is seen if we look at the relationship between the take rate chosen by proposers and the take rate proposers considered fair (only 4.8% of proposers chose a take rate that was lower than their fair take rate). Indeed, a glance at the scatter plot of the fair and chosen take rates or the fair and expected take rates suggests that individuals seem to use their fair take rate as a lower bound for their choice or expectation (see Figure 4).



**FIGURE 4 – SCATTER PLOTS OF FAIR AND ACTUAL OR EXPECTED TAKE RATES**

To conclude, it seems that proposers are using the take rate they consider to be fair as a reference point for the determination of the optimal take rate. Similarly, responders may use their fair take rate as a reference point to form an expectation of what the real take rate will be. Once this expectation has been formed, it is a deviation from the expected take rate that triggers the high intensities of anger that motivate responders to destroy.

## **5. Discussion**

In this section we further discuss our results and relate them to recent modeling attempts to account for reciprocity in decision-making. First, we emphasize the fact that negative reciprocity seems to be particularly motivated by anger. Second, we draw attention to three important aspects of our results that are either missing or unsatisfactorily modeled in the existing approaches. Namely, the role of expectations, the importance of reference points, and the effects of social ties.

### *5.1. Anger induced punishment*

Although it might seem intuitively obvious, it is not until recently that economists have started to realize that punishment of opportunistic behavior may be motivated by anger. Knowing this is important because a model based on incorrect motivations may lead to incorrect predictions. For instance, if the driving force behind an individual's decision to punish is anger but we incorrectly model it as envy (as suggested in Kirchsteiger, 1994 and Fehr and Schmidt, 2000), we will make wrong inferences regarding the action tendencies and other characteristics of the emotions at stake, and are consequently likely to make wrong predictions. In this case, one might ignore that anger, as opposed to envy, is elicited by acts perceived as intentional (Haidt, 2003) and hence overlook the important role that intentions have on punishment behavior (Falk et al., 2000; Charness and Levine, 2003). Furthermore, one could neglect that anger's action tendency is to attack as opposed to the reduction of disparities (Lazarus, 1991), and hence, expect individuals that were treated unfairly to be in favor of compensation (e.g. from a third party) as opposed to harming the unfair person.

Recent research has demonstrated that individuals who are treated unfairly derive satisfaction from successfully punishing the offending party (Quervain et al., 2004). This raises the important question whether one could interpret punishment as simply another good that can be consumed to increase one's utility. If so, this would allow us to apply standard theoretical economic analysis to an otherwise puzzling

phenomenon (Anderson and Putterman, 2004; Carpenter, 2004; and for a critical view see van Winden, 2001). In this respect, it is important to point out that, even if anger was triggered by unfair behavior (e.g. deviations from equality or a maximin norm), the goal of angry individuals is to harm the other party, and not, through punishment, to correct unfair material outcomes.<sup>14</sup> This explains why we see individuals punishing, even when it is impossible to reduce income inequalities. For example, in the power-to-take game used in this paper, destruction leads to lower income inequality only when the take rate is above 50.0%. However, we still see 13.2% of subjects destroying a positive amount at take rates that are equal to or below 50.0%. Another less studied aspect of punishment that can be accounted for with an anger-motivated explanation is that in some cases, subjects punish more than the amount that is needed to equalize earnings. For instance, Hopfensitz and Reuben (2005) study a game in which some subjects have precisely this opportunity. They find that 31.3% of the subjects punished to the point where the offending subject had lower earnings than the punisher did. Similar results can be seen in public goods games with punishment such as in Fehr and Gächter (2002). In this sense, outcome based models of social preferences such as Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) miss an important characteristic of punishment behavior.

There are various known characteristics of anger that have not been theoretically modeled (e.g. people tend to feel more anger in public than in private, Jakobs et al., 1999). However, here we would like to concentrate on one important component of anger that is not yet modeled satisfactorily, namely, the fact that anger is affected by the expectations of what happened in the past.

## *5.2. Expectations about what happened*

Current models that attempt to capture fairness motivations ignore the role of unfulfilled expectations on negative reciprocity. Roughly speaking, models that incorporate fairness notions fit within two approaches, an outcome-based approach and an intention-based approach. In the former, fairness is related to differences in monetary outcomes (e.g. Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). In the intention-based approach, individuals are considered unfair, or more precisely

---

<sup>14</sup> In this respect, as is argued by Carpenter and Matthews (2005), there is an important difference between anger-induced punishment by the affected individual and indignation-induced punishment by an unaffected third party.



unkind, depending on how their choices affect the final outcome of a game, and the desirability of this outcome compared to other feasible ones.<sup>15</sup> Though both approaches have proven to be quite successful in explaining a range of different experimental results, they fail to explain the effect of the expected take rate on the responders' decision to destroy.

Our experimental results indicate that a responder's expected take rate is an important explanatory variable of whether she destroys or not. From a psychological point of view one can expect that someone would feel higher intensities of anger in the case of high expectations that were proved wrong as opposed to the case of low expectations that were confirmed (Ortony et al., 1988). Nonetheless, this simple and intuitive reaction is not modeled in either the outcome-based or the intention-based approach.

In the outcome-based approach expectations are not modeled at all. That is to say, in these models, once responders observe the real take rate, the take rate they expected is no longer relevant. Hence, in these models responders who destroy differ from responders who do not destroy only in how much they care about payoff differences. In other words, their preferences differ. Since we generally think of these preferences as fixed (at least in the short-run), we might tend to think that we are observing stable types of individuals (such as selfish or fair-minded) that will behave as such across games and time. Furthermore, individual behavior would be less susceptible to manipulation through framing or by observing others. If on the other hand, expectations explain a large part of the heterogeneity we observe in experiments, then we should be much more careful in predicting how subjects will behave across games and time.<sup>16</sup> It would also imply that framing or information that affects expectations can have a big impact on the behavior of individuals. Furthermore, since expectations, unlike preferences, may change substantially in the

---

<sup>15</sup> Papers in this approach include Rabin (1993), Dufwenberg and Kirchsteiger (2004), and Falk and Fischbacher (2005) who combine both the outcome-based and the intention-based approaches.

<sup>16</sup> We do not dismiss the possibility of consistent differences between people based on how strongly they feel about fairness norms. However, our results suggest that they can explain only part of responders' behavior.

short-run, behavior might adjust faster than a model based solely on preferences would predict.<sup>17</sup>

In the intention-based approach, which uses psychological game theory (Geanakoplos et al., 1989), expectations play an important role, but not in the way we have discussed here. As in the outcome-based approach, in these models once a responder observes the take rate, her expected take rate has no longer an effect on her decision. The only expectation that has an effect is the responder's expectation of what the take authority expects the destruction rate to be (which determines the kindness of the take authority's choice). Furthermore, in these models attention is focused on equilibria when individuals *correctly* anticipate the actions of others. However, this raises the question, to what extent these models capture experienced emotions (see Elster, 1998).

Some of our results are in line with the theoretical predictions of the models of Rabin (1993) and Dufwenberg and Kirchsteiger (2004). For instance, we find that even when expectations are fulfilled, people that faced above average take rates (more unkindness) experienced higher intensities of disappointment and lower intensities of gratitude, joy and admiration.<sup>18</sup> The reason is that the intensity of emotions does not only depend on the degree of unexpectedness of an event, but also on the extent to which the individual's interests are affected (see e.g. Ortony et al., 1988), which relates to the take rate as such. More research is needed, though, to see whether this effect survives in the longer run. With experience, people not only learn to anticipate what others will do, they may also habituate to situations and may gradually become less emotional about anticipated adverse events.

Although fulfilled expectations might be a plausible assumption for the long-run, in many situations there is simply not enough time to learn what others will do. In these circumstances, understanding the emotional reactions to deviations from expected actions might prove very useful for predicting how individuals will behave. Furthermore, in cases in which the long-run outcome is heavily influenced by the initial situation, emotions experienced when expectations are still unfulfilled can have

---

<sup>17</sup> For example, one would expect that if a responder consistently sees high take rates, she will eventually update her expectation upwards. In such a case, assuming that the observed relationship between expectations and emotions holds over time, one might observe less destruction.

<sup>18</sup> WMW tests ( $p < 0.07$ ), where "fulfilled expectations" means that the expected take rate equaled the actual take rate (within a range of plus or minus 5 percentage points).

a crucial effect on long-run behavior. Moreover, emotions may have an additional and lasting effect since they can also affect the way individuals learn from experiences (Bower, 1992) and therefore affect the path through which the long-run outcome is obtained.

### *5.3. Reference points*

Another neglected aspect in the modeling of fairness is the income that is taken as a reference point. If we compare the results of our experiment with the results of experiments that use the related ultimatum game, it is easy to see that any theory that wishes to explain both sets of results will have to take into account the possibility of changing reference points. As reported in Camerer (2003), in ultimatum games the proposers' share of the mean offer is usually between 60% and 70% and the share of the median offer is usually between 50% and 60%. Furthermore, there is barely any destruction (rejection) if the proposer offers to keep less than 60%, while on aggregate about 50% of the income is destroyed if the proposer offers to keep more than 80%. In our power-to-take game, the mean take rate is 60.7% and the median take rate is 50%. Moreover, only 4.1% of the responders' income is destroyed at take rates below 60%, while responders destroyed 63.2% of their income at take rates above 80%.<sup>19</sup> However, this similarity in results is misleading since in the power-to-take game only the responder's income is at stake. Hence, if one compares the shares of the total income that the proposer walks away with, it is much higher in the power-to-take game than in the ultimatum game.

These results seem incompatible with all the theories using an outcome-based approach.<sup>20</sup> These theories assume that people dislike unequal distributions of income. In the ultimatum game, the more a proposer cares about income differences the more likely it is that she will make a generous offer to the responder. Hence, if we increase the endowment of the proposer relative to the endowment of the responder (as we do in the power-to-take game) one should observe on average lower take rates. Since in ultimatum games about half of the subjects offer a fifty-fifty split, the

---

<sup>19</sup> The results of the one-responder power-to-take game investigated by Bosman et al. (2005) show a mean (median) take rate of 59.9% (60%), while 7.5% (58.3%) of the responders' income is destroyed at take rates below 60% (above 80%).

<sup>20</sup> This is also true for Falk and Fischbacher (2005) who combine both kinds of approaches.

percentage of ‘fair minded’ subjects is usually thought to be around that number.<sup>21</sup> For this reason, it is striking that in our experiment, even though proposers could offer a fifty-fifty split by choosing a take rate of 0%, none actually choose to do so.<sup>22</sup> This large difference between the two games is hard to explain by any theory assuming that subjects are focusing on the experiment’s total income distribution. As suggested by Bolton and Ockenfels (2005), in these and similar cases, distribution-based fairness models work rather well if one simply adjusts them to take into account only the ‘relevant’ income distribution. However, if we are to correctly predict behavior across a wide range of situations, we must incorporate into these theories an explicit way in which individuals adjust their reference incomes (Ariely et al., 2003).

In this respect, the intention-based fairness models perform better than the outcome-based models. In these models, individuals use a fairness benchmark that depends solely on feasible outcomes. More specifically, it depends on how the take rate affects the responder’s payoff in comparison to the highest and lowest payoff the responder could get. Since the responder’s possible payoffs are not affected by any additional income the proposer might have, these models do not predict that proposers will choose lower take rates in the power-to-take game.

#### *5.4. Social Ties*

As shown in the previous section, social ties can have a considerable impact on behavior. Not only did friends react differently to higher take rates, their emotional reaction towards one another was also very different. However, none of the models discussed so far incorporate an effect of social ties. Doing so is important since not all meaningful economic interaction occurs between strangers. In cases such as interaction at the work floor, informal credit institutions, scientific research, and political participation, more interaction might actually occur between friends than between strangers. In fact, as argued by Rosenblat and Mobius (2004), technological advances that reduce communication costs, such as the internet, can make interaction among groups of friends even more important.

In our experiment, angry friends managed to coordinate destruction much more frequently than angry strangers. If we consider the emotional reactions between

---

<sup>21</sup> For example, Fehr and Schmidt (1999) estimate that about of 60% subjects would prefer an equal split over any other income distribution.

<sup>22</sup> In other power-to-take experiments the number of proposers choosing a take rate of 0% is smaller than 10% (Bosman and van Winden, 2002; Bosman et al., 2005).

responders, as discussed in Result 7, this is not surprising. Note that angry strangers who intend to destroy their income face a situation akin to a collective action problem. Our findings suggest that they would like to see the take authority punished but want to avoid being the player in the game that is left with the lowest payoff. We do not know the precise amount of satisfaction that subjects derive from each of the possible outcomes. Nevertheless, judging by the responders' emotional reactions, it would not be farfetched to model the angry strangers' situation as a prisoner's dilemma or a stag hunt game. In either case, destruction is unlikely.<sup>23</sup> In contrast, in the angry friends case, the observed desire of responders to coordinate on the same action combined with an impulse to destroy makes their situation noticeably different. Angry friends can be modeled as playing a coordination game in which destruction not only gives them the highest payoff but is also the risk-dominant choice, and hence, the most attractive option.

We will discuss two natural ways of allowing for social ties in the two modeling approaches. The first way is to assume that friends are better than strangers at predicting what the other responder will do. The second way is to assume that friends, as opposed to strangers, care for each other's utility.

In the outcome-based approach, knowing what the others will do translates to friends knowing each other's preferences. This might indeed lead to a situation in which strangers destroy less than friends. To see this, note that if the income difference between two responders is important, responders will wish to destroy similar amounts. In this case, uncertainty about the other's preferences makes a responder's destruction decision much more difficult. If, as suggested by Fehr and Schmidt (1999), individuals prefer advantageous inequality to disadvantageous inequality, then destroying less becomes more attractive since it gives a higher payoff if responders fail to coordinate. Hence, since strangers face more uncertainty about the other's actions, on average, they will be likely to destroy less than friends. Nonetheless, this line of thought fails to describe one important aspect of the data, namely that the difference between friends and strangers occurs at high take rates. In the outcome-based approach, at high take rates the main concern of responders is to lower the income difference between themselves and the proposer. This is because, when facing a high take rate, unilateral destruction will only create small income

---

<sup>23</sup> This is certainly true for the prisoners' dilemma. In the stag hunt game, destroying is the risk-dominated action and hence less likely, especially if subjects are risk adverse.

differences between the responders. Hence, under these conditions both friends and strangers would destroy very similar amounts.

Assuming that friends care for each other's utility might be a more promising way for the outcome-based approach to explain why friends destroy more than strangers when facing high take rates. If responders care for the *utility* of the other responder, then in addition to receiving disutility because the proposer has a higher income than they do, they will also receive disutility because the proposer has higher income than their friend does. This leads to a stronger desire to destroy, even at high take rates. In this respect, investigating the precise effect of an interdependent utility function might prove a fruitful line of research.

In the intention-based approach, knowing better what the other responder will do does not affect behavior. The reason is that, in all the models within this approach, individuals care about each other's income only if they can affect it through their actions. Since in this experiment responders could not affect each other's income, these theories predict that knowing what the other responder will do has no effect on behavior. Therefore, as with the outcome-based approach, a more helpful way of incorporating social ties might be to assume that friends care for each other's wellbeing. In the current models, an individual evaluates the kindness of others by looking at how their actions lead to a higher or lower payoff for the individual. The friends case could be modeled by allowing individuals to include into their evaluation the way others treat their friends. Unfortunately, due to the complexity of these models an interdependent utility function might produce a model that is very difficult to analyze.

## **6. Conclusion**

An important goal of this paper was to extend the explicit study of emotions in economic decision-making to cases with more than two players. In addition, we looked at how social ties can affect emotional reactions and behavior. For this purpose we used a three-person power-to-take game. The experimental data demonstrate that high intensities of anger-like emotions and low intensities of joy-like emotions induce responders to destroy their income. We also find that friends and strangers as responders experience different emotional reactions towards one another, which leads to more destruction in the case of angry friends.

Our results indicate that the study of emotions helps explain observed behavior. Anger-like emotions appeared to be the main driving force behind the decision to destroy income. Furthermore, by observing the emotional reaction between subjects we could explain why friends are able to coordinate on destruction more frequently than strangers. Without investigating these emotional responses, the precise mechanism by which social ties affected the subjects' choices would have remained unclear.

In this paper we have also emphasized the role of expectations in determining the subjects' emotional responses. One interesting issue for future research concerns the interaction between expectations and social norms. If social norms are based on the actual behavior of the majority of individuals in a society, then expectations may be largely fulfilled in many well established situations. However, when faced with new circumstances in which a social norm is not clearly defined, the initial expectations of individuals might have an important effect on the behavior that later becomes a norm. Hence, norms might be very susceptible to the initial conditions in which they are formed.

In addition to expectations, we would like to emphasize the importance of studying social ties. Our experiment showed clear differences in emotional reactions depending on the presence of a social tie. In some situations this could lead to very different behavior that might be economically relevant. For example, the emotionally boost that friends receive from coordination might be one of the reasons people prefer to interact with friends rather than strangers. Results from the literature on social distance suggest that this type of preferences can lead to segregation, inefficient outcomes, and conflict between groups (Schelling, 1978; Borjas, 1995; Glaeser et al., 1996; Akerlof, 1997).

The behavioral differences induced by social ties could prove especially important if one is thinking on the effectiveness of various policies. It might be the case that a given policy would improve a situation only if individuals (do not) share social ties. Investigating the emotional responses may reveal the precise mechanism through which social ties affect people's choices and could therefore help us predict the effects of different policies. To give an example, the emotional responses of friends indicate a strong desire to coordinate their actions. In this case, a coordination mechanism such as the possibility to make decisions sequentially (as in Potters et al., 2005) might be more effective among friends than among strangers.

So far, the effects of social ties have received little attention in experimental investigations. To some extent, this neglect is due to the difficulty of creating strong social ties in controlled environments. The usual ingredients of complete anonymity, no face-to-face communication, and a short time period of interaction, produce an environment in which meaningful social bonding is difficult. Nevertheless, the design we have used suggests that it is possible to include social ties in experiments and to acquire insights into how to model them.



## Appendix A – Instructions

These are the instructions for the ‘friends’ treatment. The instructions for the ‘strangers’ treatment were very similar, and are available upon request.

### *Instructions (translation from Dutch)*

In order to sign up for this experiment, you had to sign up together with a second participant. For convenience, we will refer to this second participant as your partner. In the experiment each of you will be assigned to a 3-person group, that is, you plus two other participants. We will explain how groups are formed later on.

Throughout the experiment, the type of decision you make will depend on your position in your group. Some of you will be positioned to move first, and some of you will be positioned to move second. Participants moving first will be referred to as As while participants moving second will be referred to as Bs.

Before the experiment started each desk was assigned either an A or a B. Therefore, by randomly assigning the yellow cards (in the reception room), each participant was randomly assigned a position. Once you are informed which position has been assigned to you, the corresponding letter will appear on the top-right part of the screen.

The 3-person group that you belong to depends on your position as well as on the position of others in the following way:

- Your group (including yourself) consists of one A and two Bs.
- If you are a B:
  - Then, the other B in your group is your partner (the person with whom you signed up).
  - The other participant will be a randomly chosen A.
- If you are an A:
  - Then your partner (the person with whom you signed up) is also an A and thus he/she is not in your group.
  - The other two participants will be a randomly selected pair of Bs that signed up together for the experiment.

Note: Each group, and thus also your group, was formed randomly in the sense that the A in the group does not know who the B’s are, and similarly the B’s do not know who the A is.

### *The experiment*

At the beginning of the experiment each participant – this includes all A participants and all B participants – will receive 10 euros as his/her initial endowment. The experiment consists of two phases. In phase one, only the A participant must make a decision. Similarly, in phase two, only the B participants must make a decision. Hence, every participant makes only one decision. In addition to the decision, during the experiment you will be asked to answer a few questions.

#### *Phase one: A chooses a percentage*

In this phase, A must choose a percentage and type it into the corresponding field on the screen. This percentage determines how much of the money of each B in the group after phase two, will be transferred to A. The percentage chosen by A must be an integer between 0 and 100 (inclusive). If you wish to make any calculations, you can use the calculator located on your desk.

Once you are satisfied with your decision, you have to confirm it by clicking on the button “Ready”. Note that all decisions are final; once you have clicked on “Ready” there is no way of changing your choice. Once A has completed phase one, phase two begins.

#### *Phase two: each B chooses a percentage*

At the beginning of this phase, each B is informed of the percentage chosen by A. At this point, each B must also choose a percentage and type it into the corresponding field on the screen. This percentage determines how much of his/her initial endowment will B destroy. The percentage chosen by B must be an integer between 0 and 100 (inclusive). Hence, the transfer from each participant B to participant A will be based on the endowment of B that is left.

Once you are satisfied with your decision, you have to confirm it by clicking on the button “Ready”. Note that all decisions are final; once you have clicked on “Ready” there is no way of changing your choice. Once each person has made his/her decision, phase two ends.

#### *Payoffs*

After phase two, all participants will be informed of the amount of money they have earned during the experiment. You will also be informed of the amount of money earned by the other two participants in your group.

### *Example of how to calculate your payoffs*

We will now give an example for the purpose of illustration. Remember that all participants in your group have an initial endowment of 10 euros. Suppose that in phase one participant A decides that 30% of the endowment of each participant B will be transferred to him/her (participant A). In phase two, each B can destroy part or everything of his/her initial endowment. Suppose that both Bs decide to destroy 0% percent of their initial endowment. The transfer from each B to A is then equal to 3 euros (30% of 10 euros). The earnings of each B are equal to 7 euros (namely, the initial endowment of 10 euros minus the transfer of 3 euros). The final endowment of A is equal to 16 euros (namely, the initial endowment of 10 euros plus twice a transfer of 3 euros).

Now suppose that in this example, one of the B participants decides to destroy 50% of his/her initial endowment. In this case, his/her transfer to A is only 1.5 euros (namely, 30% of the endowment that was not destroyed, i.e. is 30% of 5 euros). The earnings of A are equal to 14.5 euros (namely, the initial endowment of 10 euros plus 3 euros transferred from the B who destroyed 0% plus 1.5 euros transferred from the B who destroyed 50%). The earnings of the B who destroyed 0% are again 7 euros, and, finally, the earnings of the B who destroyed 50% are 3.5 euros (namely, 50% of the initial endowment of 10 euros minus the transfer of 1.5 euros).

### *In summary*

In the experiment you will be divided into groups of 3, each consisting of one A and two Bs (who signed up together for the experiment). The roles of A and B were randomly and anonymously assigned by drawing your table number. Each participant receives 10 euros as an initial endowment. Then there are two phases. In phase one, A decides on a percentage that indicates how much of the endowments of each B after phase two will be transferred to A. In phase two, each B decides what percentage of his/her initial endowment will be destroyed.

If you have any questions now, please raise your hand. If you do not have any questions, please click on “Ready”. Note that once you click on “Ready” you will not be able to go back to the instructions. Next, we will ask you to answer a few questions in order to familiarize you with the calculation of your earnings.

## Appendix B – Summary of individual data

TABLE B1 – TAKE AND DESTRUCTION RATES BY GROUP AND TREATMENT

Group #	Friends			Strangers		
	$t$	$d_1$	$d_2$	$t$	$d_1$	$d_2$
1	29	0	0	10	0	0
2	33	0	0	39	0	0
3	35	20	0	39	0	0
4	40	0	0	50	0	0
5	40	0	0	50	0	0
6	40	0	0	50	0	0
7	40	0	0	50	0	0
8	40	0	0	50	0	0
9	42	0	0	50	0	0
10	45	0	0	50	0	0
11	46	5	0	50	0	0
12	50	0	0	50	0	0
13	50	0	0	50	0	1
14	50	0	0	50	0	70
15	50	0	0	50	10	0
16	50	5	0	50	0	10
17	55	100	0	50	10	0
18	60	0	0	50	50	0
19	60	0	0	55	0	0
20	60	0	25	70	0	0
21	60	20	0	75	100	0
22	65	0	50	80	0	0
23	65	25	0	80	0	100
24	65	50	0	80	100	0
25	70	100	0	81	90	0
26	80	20	100	82	0	100
27	80	100	0	100	0	0
28	90	0	100	100	100	0
29	90	90	100			
30	100	50	100			
31	100	100	100			
32	100	100	100			
33	100	100	100			
34	100	100	100			
35	100	100	100			

Where:  $t$  = take rate,  $d_i$  = destruction rate of responder  $i$ .

**TABLE B2 – MEAN EMOTIONAL INTENSITY OF RESPONDERS REGARDING THE TAKE AUTHORITY**

<b>Emotions of <i>i</i></b>	<b>Friends</b>		<b>Strangers</b>	
	<b><math>d_i = 0</math></b>	<b><math>d_i &gt; 0</math></b>	<b><math>d_i = 0</math></b>	<b><math>d_i &gt; 0</math></b>
admiration	2.7 (1.9)	1.3 (0.7)	2.6 (1.8)	1.9 (1.3)
anger	2.3 (1.6)	4.1 (2.2)	2.8 (2.0)	4.0 (1.9)
contempt	2.2 (1.7)	4.5 (2.2)	2.6 (2.1)	3.8 (2.2)
disappointment	2.7 (1.7)	4.3 (2.1)	3.0 (2.0)	4.3 (1.9)
envy	3.3 (1.7)	3.7 (2.3)	3.2 (1.9)	3.8 (2.3)
gratitude	3.5 (2.1)	1.5 (0.9)	3.1 (1.9)	1.9 (1.4)
guilt	1.4 (0.9)	1.4 (0.8)	1.6 (1.0)	1.6 (0.9)
irritation	2.9 (1.9)	5.2 (2.1)	3.4 (2.2)	4.6 (2.2)
joy	3.6 (2.0)	1.6 (1.3)	3.0 (1.8)	1.9 (1.2)
pride	2.3 (1.4)	2.8 (2.3)	2.8 (2.1)	2.7 (1.9)
regret	1.4 (1.0)	1.6 (1.2)	1.4 (0.9)	2.1 (1.2)
sadness	1.6 (1.1)	2.2 (1.5)	2.1 (1.7)	2.4 (1.7)
shame	1.4 (0.9)	1.6 (1.3)	1.4 (0.8)	1.3 (0.6)
surprise	4.0 (2.1)	4.5 (1.9)	3.7 (2.1)	4.1 (2.2)

Numbers between brackets are standard deviations.

**TABLE B3 – MEAN EMOTIONAL INTENSITY OF RESPONDERS REGARDING THE OTHER RESPONDER**

<b>Emotions of <i>i</i></b>	<b>Friends</b>			<b>Strangers</b>		
	<b><math>d_i &gt; d_j</math></b>	<b><math>d_i = d_j</math></b>	<b><math>d_i &lt; d_j</math></b>	<b><math>d_i &gt; d_j</math></b>	<b><math>d_i = d_j</math></b>	<b><math>d_i &lt; d_j</math></b>
admiration	3.1 (2.2)	3.8 (1.9)	3.1 (2.2)	2.2 (1.5)	2.5 (1.6)	2.5 (1.9)
anger	2.3 (2.3)	1.1 (0.3)	1.6 (1.0)	3.3 (2.1)	1.8 (1.5)	2.2 (1.6)
contempt	1.7 (1.3)	1.1 (0.3)	1.3 (0.7)	2.7 (2.1)	1.5 (1.1)	1.9 (1.0)
disappointment	2.6 (2.4)	1.1 (0.3)	1.6 (1.1)	4.1 (1.9)	1.9 (1.7)	2.2 (1.6)
envy	2.3 (2.1)	1.1 (0.3)	1.5 (1.1)	2.9 (1.7)	1.7 (1.3)	2.6 (1.9)
gratitude	2.7 (2.1)	3.2 (2.1)	2.0 (1.3)	1.6 (1.0)	1.8 (1.3)	1.8 (1.5)
guilt	1.5 (1.1)	1.2 (0.8)	1.4 (0.9)	2.4 (1.6)	1.3 (0.9)	2.0 (1.5)
irritation	2.5 (2.3)	1.1 (0.2)	1.6 (1.2)	4.3 (1.9)	2.0 (1.8)	2.8 (2.0)
joy	2.6 (2.0)	5.0 (1.3)	3.2 (2.1)	1.9 (1.5)	3.2 (2.0)	2.2 (1.7)
pride	3.1 (2.1)	4.8 (1.9)	2.5 (1.9)	2.2 (1.5)	3.6 (2.2)	2.0 (1.5)
regret	1.3 (0.7)	1.2 (0.7)	1.8 (1.5)	1.8 (1.2)	1.4 (0.9)	1.5 (1.0)
sadness	1.9 (1.9)	1.1 (0.5)	1.3 (0.6)	2.1 (1.4)	1.7 (1.5)	1.3 (0.8)
shame	1.9 (1.8)	1.1 (0.3)	1.9 (1.6)	1.8 (1.1)	1.3 (0.9)	1.7 (1.2)
surprise	4.7 (1.7)	2.1 (1.4)	4.1 (2.5)	4.8 (2.4)	2.0 (1.6)	3.7 (2.2)

Numbers between brackets are standard deviations.

## Appendix C – Regressions

Tobit model with the destruction rate as the dependent variable (censored at  $d = 0$  and  $d = 100$ ).<sup>24</sup>

TABLE C1 – TOBIT MODEL ESTIMATING THE AMOUNT OF DESTRUCTION

Variable	Coefficient	Std. Error	<i>p</i> value
Friends × Take Rate	2.922	1.129	0.011
Strangers × Take Rate	2.076	1.073	0.055
Take Rate – Expected Take Rate	1.506	0.612	0.015
Take Rate – Fair Take Rate	1.046	0.637	0.103
Economist	-15.905	29.575	0.592
Female	-47.494	29.431	0.109
Constant	-208.562	75.959	0.007
Number of obs. = 126		LR $\chi^2(6)$ = 75.71	
Log likelihood = -150.429		Prob > $\chi^2$ = 0.000	

Dummy variables: Friends: 1 if friends treatment, 0 otherwise; Strangers: 1 if strangers treatment, 0 otherwise; Economist: 1 if economics student, 0 otherwise; Female: 1 if female, 0 if male.

Ordered probit model with the average of the three anger-like emotions (anger, irritation, and contempt) as the dependent variable.

TABLE C2 – ORDERED PROBIT MODEL ESTIMATING THE INTENSITY OF ANGER-LIKE EMOTIONS

Variable	Coefficient	Std. Error	<i>p</i> value
Take Rate	0.0146	0.0071	0.039
Take Rate – Expected Take Rate	0.0143	0.0039	0.000
Take Rate – Fair Take Rate	0.0002	0.0037	0.957
Economist	0.2178	0.2007	0.278
Female	-0.3185	0.1930	0.099
Friends	-0.0018	0.1891	0.992
Number of obs. = 126		LR $\chi^2(6)$ = 47.49	
Log likelihood = -322.955		Prob > $\chi^2$ = 0.000	
Dummy variables as in Table C1.			

<sup>24</sup> Given the possibility errors are correlated among friends, we also ran the regression using robust standard errors and clustering within friend pairs. We get very similar results. The main difference is that the *p* value of the coefficient for ‘Take Rate – Fair Take Rate’ decreases to 0.045. Hence, although the fair take rate does reach a 5% level of significance, it still has a smaller effect than the expected take rate.

## References

- Abbink, K., B. Irlenbusch, and E. Renner (2002). Group size and social ties in microfinance institutions. CeDEx Discussion Paper, University of Nottingham.
- Akerlof, G.A. (1997). Social distance and social decisions. *Econometrica* 65: 1005-1027.
- Anderson, C.M. and L. Putterman (2004). Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism. *Games and Economic Behavior*, forthcoming.
- Ariely, D., G. Loewenstein, and D. Prelec (2003). Coherent arbitrariness: Stable demand curves without stable preferences. *The Quarterly Journal of Economics* 118: 73-106.
- Ben-Shakhar, G., G. Bornstein, A. Hopfensitz, and F. van Winden (2004). Reciprocity and emotions: Arousal, self-reports, and expectations. CREED Working paper. University of Amsterdam.
- Bereby-Meyer, Y. and M. Niederle (2005). Fairness in bargaining. *Journal of Economic Behavior and Organization* 56: 173-186.
- Bolton, G. and A. Ockenfels (2000). A theory of equity, reciprocity, and competition. *American Economic Review* 90: 166-193.
- Bolton, G. and A. Ockenfels (2005). A stress test of fairness measures in models of social utility. *Economic Theory* 25: 957-982.
- Borjas, G.J. (1995). Ethnicity, neighborhoods, and human capital externalities. *American Economic Review* 85: 365-389.
- Bosman, R., and F. van Winden (2002). Emotional hazard in a power-to-take game experiment. *The Economic Journal* 112: 147-169.
- Bosman, R., M. Sutter, and F. van Winden (2005). On the impact of real effort and emotions in power-to-take experiments. *Journal of Economic Psychology* 26: 407-429.
- Bower, G. (1992). How might emotions affect learning? In Christianson, S.A. (Ed.) *Handbook of emotion and memory: Research and theory*. Hillsdale: Erlbaum.
- Camerer, C. (2003). *Behavioral Game Theory*. New Jersey: Princeton University Press.

- Carpenter, J. P. (2004). The Demand for Punishment. Working paper. Middlebury College.
- Carpenter, J. P. and P. Matthews (2005). Norm Enforcement: Anger, Indignation or Reciprocity. Working paper. Middlebury College.
- Charness, G. and D. Levine (2003). The road to hell: An experimental study of intentions. Working paper 14-03. University of California, Santa Barbara.
- Charness, G. and M. Rabin (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics* 117: 817-869.
- Charness, G., E. Haruvy, and D. Sonsino (2003). Social distance and reciprocity: The Internet vs. the laboratory. Working paper. University of California, Santa Barbara.
- Chong, D. (1991). *Collective Action and the Civil Rights Movement*. Chicago: University of Chicago Press.
- Clore, G.L. and M.D. Robinson (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin* 128: 934-960.
- Dufwenberg, M. and U. Gneezy (2000). Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior* 30: 163-182.
- Dufwenberg, M. and G. Kirchsteiger (2004). A theory of sequential reciprocity. *Games and Economic Behavior* 47: 268-298.
- Elster, J. (1998). Emotions and economic theory. *Journal of Economic Literature* 36: 47-74.
- Erard, B. and J.S. Feinstein (1994). Honesty and evasion in the tax compliance game. *RAND Journal of Economics* 25: 1-19.
- Falk, A. and U. Fischbacher (2005). A theory of reciprocity. *Games and Economic Behavior* forthcoming.
- Falk, A., E. Fehr, and U. Fischbacher (2000). Testing theories of fairness: Intentions matter. Working paper No. 63. University of Zürich.
- Fehr, E. and S. Gächter (2002). Altruistic punishment in humans. *Nature* 415: 137-140.
- Fehr, E. and K. Schmidt (1999). A Theory of fairness, competition and cooperation. *The Quarterly Journal of Economics* 114: 817-868.
- Fehr, E. and K. Schmidt (2000). Theories of fairness and reciprocity: Evidence and economic applications. In Dewatripont, M., L. Hansen, and S.T. Turnovsky



- (Eds.) *Advances in Economics and Econometrics - 8th World Congress, Econometric Society Monographs*. Cambridge: Cambridge University Press.
- Fehr, E., G. Kirchsteiger, and A. Riedl (1993). Does fairness prevent market clearing? An experimental investigation. *The Quarterly Journal of Economics* 108: 437-459.
- Fischbacher U. (1999). Zurich toolbox for readymade economic experiments, experimenter's manual. Working Paper No. 21. Institute for Empirical Research in Economics, University of Zürich.
- Frank, R. H. (1987). If Homo Economicus could choose his own utility function, would he want one with a conscience. *American Economic Review* 77: 593-604.
- Frijda, N.H. (1986). *The emotions*. Cambridge: Cambridge University Press.
- Frijda, N.H. (1988). The laws of emotion. *American Psychologist* 43: 349-358.
- Geanakoplos, J., D. Pearce, and E. Stacchetti (1989). Psychological games and sequential rationality. *Games and Economic Behavior* 1: 60-79.
- Glaeser, E.L. (2005). The political economy of hatred. *The Quarterly Journal of Economics* 120: 45-86.
- Glaeser, E.L., B. Sacerdote, and J.A. Scheinkman (1996). Crime and social interactions. *The Quarterly Journal of Economics* 111: 507-548.
- Goleman, D. (1995). *Emotional intelligence: Why it can matter more than IQ*. New York: Bantam Books.
- Güth W. and E. van Damme (1998). Information, strategic behavior and fairness in ultimatum bargaining: An experimental study. *Journal of Mathematical Psychology* 42: 227-247.
- Haidt, J. (2003). The moral emotions. In Davidson, R. J., K. R. Scherer, and H. H. Goldsmith (Eds.) *Handbook of Affective Sciences*. Oxford: Oxford University Press.
- Hirshleifer, J. (1987). On emotions as guarantors of threats and promises. In Dupre J. (Ed.) *Latest on the Best: Essays on Evolutions and Optimality*. Cambridge: MIT Press.
- Hopfensitz, A. and E. Reuben (2005). The Importance of Emotions for the Effectiveness of Social Punishment. Discussion paper 05-075/1. Tinbergen Institute.

- Jakobs, E., A.S.R. Manstead, and A.H. Fischer (1996). Social context and the experience of emotion. *Journal of Nonverbal Behavior* 20: 123-142.
- Jakobs, E., A.S.R. Manstead, and A.H. Fischer (1999). Social motives, emotional feelings, and smiling. *Cognition and Emotion* 13: 321-345.
- Kagel, J.H. and K.W. Wolfe (2001). Tests of fairness models based on equity considerations in a three-person ultimatum game. *Experimental Economics* 4: 203-219.
- Kirchsteiger, G. (1994). The role of envy in ultimatum games. *Journal of Economic Behavior and Organization* 25: 373-389.
- Knez, M. and C.F. Camerer (1995). Outside options and social comparison in a three-player ultimatum game experiments. *Games and Economic Behavior* 10: 65-94.
- Lazarus, R.S. (1991). *Emotion and Adaptation*. New York: Oxford University Press.
- Lerner, J.S. and D. Keltner (2000). Beyond valence: Toward a model of emotion-specific influences on judgment and choice. *Cognition and Emotion* 14: 473-492.
- Levine, D. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics* 1: 593-622.
- Loewenstein, G.F. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes* 65: 272-292.
- Loewenstein, G. F. (2000). Emotions in economic theory and economic behavior. *American Economic Review* 90: 426-432.
- Ortony, A., G.L. Clore, and A. Collins (1988). *The cognitive structure of emotions*. Cambridge: Cambridge University Press.
- Pillutla, M.M. and J.K. Murnighan (1996). Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes* 68: 208-224.
- Potters, J., M. Sefton, and L. Vesterlund (2005). After you: Endogenous sequencing in voluntary contribution games. *Journal of Public Economics* 89: 1399-1419.
- Quervain, D.J.F., U. Fischbacher, V. Treyer, M. Schellhammer, U. Schnyder, A. Buck, and E. Fehr (2004). The neural basis of altruistic punishment. *Science* 305: 1254-1258.
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review* 83: 1281-1302.

- Riedl, A. and J. Vyrastekova (2003). Responder behavior in three-person ultimatum game experiments. Working paper. University of Amsterdam.
- Rosenblat, T.S. and M.M. Mobius (2004). Getting closer or drifting apart? *The Quarterly Journal of Economics* 119: 971-1009.
- Sanfey, A.G., J.K. Rilling, J.A. Aronson, L.E. Nystrom, and J.D. Cohen (2003). The Neural Basis of Economic Decision-Making in the Ultimatum Game. *Science* 300: 1755-1758.
- Schelling, T.C. (1978). *Micromotives and Macrobehavior*. New York: Norton.
- Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In Higgins E.T. and R. Sorrentino (Eds.) *Handbook of motivation and cognition: Foundations of social behavior*. New York: Guilford Press.
- Smith, V.L. (1976). Experimental economics: Induced value theory. *American Economic Review* 66: 274-279.
- van Winden, F. (2001). Emotional hazard exemplified by taxation-induced anger. *KYKLOS* 54: 491-506.