On Blame and Reciprocity: An Experimental Study*

Boğaçhan Çelen† Columbia University

Mariana Blanco‡ Universidad del Rosario

Andrew Schotter§ New York University

July 18, 2013

Abstract

The theory of reciprocity is predicated on the assumption that people are willing to reward kind acts and to punish unkind ones. This assumption raises the question as to how to define "kindness." In this paper we offer a novel definition of kindness based on a notion of blame. This notion states that in judging whether player i is kind or unkind to player j, player j has to put himself in the position of player i and ask if he would act in a manner that is worse than i does. If player j would act in a worse manner than player i acted, then we say that player j does not blame player i. If, however, player j would be nicer than player i was, then we say that player j blames player i. We consider this notion a natural, intuitive and empirically functional way to explain the motives of people engaged in reciprocal behavior. After developing the conceptual framework, we test this concept by using data from two laboratory experiments and find significant support for the theory.

JEL Classification Numbers: A13, C72, D63. Keywords: Altruism, blame, reciprocity, psychological games.

*We appreciate very useful comments of Pierpaolo Battigalli, Martin Dufwenberg, Colin Camerer, and Erkut Özbay. We are grateful to the participants of the CESS Experimental Economics Lunchtime Seminar, 2009 North-American ESA Conference, Amsterdam Workshop on Behavioral & Experimental Economics, Cornell University Behavioral Economics Workshop, Rutgers University, Brown University Microeconomics Seminar and SfED 2012 Winter Conference, University of Birmingham, University of Exeter, Melbourne Univerity, Oxford University, Vienna University, Universitat de Barcelona, Universitat Pompeu Fabra, Universitat Autònoma de Barcelona, Freie Universität Berlin and California Institute of Technology for comments. We also acknowledge the partial financial support of the Center for Experimental Social Science at NYU.

†C.E.S.S., New York University, and Graduate School of Business, Columbia University, 3022 Broadway, 602 Uris Hall, New York, NY 10027. E-mail:bc2132@columbia.edu, url: http://celen.gsb.columbia.edu/.

‡Faculdad de Economía, Universidad del Rosario, Calle 14 # 4-80, Oficina 207, Bogotá, Colombia. E-mail: mariana.blanco@urosario.edu.co, url: http://mbnet26.googlepages.com/home/.

§C.E.S.S. and Department of Economics, New York University, 19 W. 4th Street, New York, NY 10012. E-mail: andrew.schotter@nyu.edu, url: http://homepages.nyu.edu/∼as7/.

1 Introduction

Recent years have witnessed a growing literature on the theory of reciprocity. Founded on the seminal work of Rabin [17]—further extended by Falk and Fischbacher [11], Dufwenberg and Kirschteiger [7] and other scholars—the theory of reciprocity is predicated on the assumption that people are willing to reward kind acts and to punish unkind ones.[1] This approach raises the question of how to define "kindness." In this paper, we offer a novel definition of kindness based on a notion of blame.

Put most simply, the notion of blame states that in judging whether player i is kind or unkind to player j, player j has to put himself in the position of player i, and ask if he would act in a manner that is worse than player i does, under identical circumstances. If player j would act in a worse manner than player i does, then we say that player j does not blame player i for his behavior. If, however, player j would have been nicer than player i was, then we say that "player j blames player i" for his actions—i.e. player i's actions were blameworthy.

This way of viewing kindness is distinctly different from other theories in a number of ways. Following the criteria that were discussed in Schotter [18], our approach leads to an endogenous, context-dependent, and process-oriented theory. It is endogenous because players judge the actions of others by their own standards and not by some exogenous standard imposed by the analyst. The theory allows the standards, which people use to judge the actions of others, to differ from person to person depending on their personal norms. Indeed, actions that bother a person may not bother other people at all, or those actions that strike you as being fair may be very upsetting to others. This feature differentiates our theory from the theories that impose an exogenous norm in order to determine what is considered kind, nice, or fair. Blame is self-referential: It only matters what you would have done in your opponent's situation and not how the actions of others are compared to some exogenous norm.

Another important feature of our approach is that the theory is sensitive to the institutional setting. For instance, actions that are blame-free in a prison may certainly be blameworthy in civilian life. One cannot judge the behavior of people in isolation—we need to know the institutional setting they are in. This is fundamentally different than the theories that model players' preferences independent of the context. For example, in a leading paper Levine [16] takes this approach to analyze experimental evidence in ultimatum, centipede, and public good experiments. Gul and Pesendorfer [15] lay the foundations of interdependence between behavioral types, independent of the environment decision-makers interact. Finally, blame judges the actions of people that lead to outcomes and not

[1] For a comprehensive survey on reciprocity see Sobel [19].

merely the outcomes themselves. So, it is a process-oriented theory, and it differs from those theories that are outcome-based in that respect.

To put some flesh on this notion of blame and to differentiate it from other theories of reciprocity, let us consider a few examples of how our analysis differs from those of other approaches.

1.1 Inequity Aversion and Blame

Consider an Ultimatum game played between two players, p (Proposer) and r (Receiver), whose preferences exhibit inequity aversion à la Fehr and Schmidt [13] or Bolton and Ockenfels [3]. Let $u_i(x_p, x_r)$ represent the preferences of player i when the final allocation is x_p and x_r for the Proposer and the Receiver, respectively.

Given these preferences, in the subgame perfect equilibrium of the game the Receiver rejects an offer (x_p, x_r) if and only if $u_r(x_p, x_r) < u_r(0, 0)$. In our formulation of preferences that exhibit blame, the Receiver blames—hence perhaps rejects—an offer if that offer is less generous than the one he would have made, had he been in the Proposer position. Hence, according to our hypothesis, the Receiver compares the offer (x_p, x_r) to the offer he would have made if he were the Proposer, say (x_p^*, x_r^*). If the offer (x_p, x_r) is more generous than (x_p^*, x_r^*), then he accepts the offer, otherwise he blames. If blame causes a sufficiently high disutility for the Receiver, he will reject the offer.

1.2 Kindness and Blame

In Rabin [17]'s theory of fairness, given his beliefs about what player i thinks he is going to play, player j judges player i's action as being unkind if it leads to an expected payoff, which is less than a given proportion of the maximum total payoff available to him. In other words Rabin [17]'s definition is based on an exogenously imposed split-the-difference norm where player j judges player i's action as unkind if it determines a payoffs less than $1/2$ of the possible payoffs he could have been provided with. But what if player i's action led to a payoff for player j that was only $1/3$ of the total available but player j, if he was in player i's position, would have given his opponent even less, say $1/6$. Under what circumstance should player j be upset with player i or think that his action was unkind? While he may not like his payoff, he certainly understands player j's actions, and in fact, compared to what he would have done, he must even consider player i to be more kind.

In a similar vein, Charness and Rabin [5] define a *demerit parameter* which captures how a player feels towards his opponent. This parameter is determined by comparing the behavior of an opponent to what a *decent person* would do in his position. In this approach,

therefore, an opponent's action is considered in relation to an exogenously determined social norm, "a decent person", while in our theory, the standard used to judge others' behavior is endogenously defined by the player himself.

The point, therefore, is that feelings of justice, fairness and kindness are subjective and must emanate from the person doing the evaluation. They should not be imposed from the outside using some other standard.[2]

1.3 Interdependent Preferences and Blame

Another popular theory of reciprocity, which is pioneered by Levine [16], assumes that a player's preferences depend on the *types* of other players.[3] The types are private information and the game is modeled as a Bayesian game. The types reflect the *niceness* of a player. The utility that a player receives from an outcome is a function of the player's *direct* and *adjusted* utility. The direct utility (u_i) is simply player i's material payoff while the adjusted utility (v_i) takes into account how nice his opponent is. More precisely, Levine [16] posits a utility function, which is a generalized version of the following form:

$$v_i = u_i + \frac{a_i + a_j}{2} u_j$$

where, $-1 < a_i, a_j < 1$ are the niceness parameters (types) of players i and j respectively. Note that player i's utility is an increasing function of player j's type, meaning that the nicer player j is, the more he cares about player j's direct utility.

According our theory, however, this judgement is relative. Player i perceives player j as nice only if player j has taken an action that he does not blame, i.e., j's action was nicer than the action he would have taken if he were in j's position. This distinction is important. For example, let $a_i = 0.9$ and $a_j = 0.8$: Both players i and j are nice but player i is *nicer* than player j. In the context of our theory, player i would blame player j for not being as nice as he would be in his position. However, according to Levine [16], player j is considered nice regardless of player i's type. Put differently, in our model niceness is a relative concept, while in Levine [16] it is an absolute concept.

1.4 Overview and Summary

As our introductory discussion indicates, the essence of blame theory involves the examination of a counter factual, i.e. imagining what you would have done if you were in the

[2] In addition to the papers mentioned above, papers which imposes exogenous norms include Dufwenberg and Kirchsteiger [7], Falk and Fischbacher [11].

[3] Gul and Pesendorfer [15] provide a more general model, which captures Levine [16] as a special case.

position of the person whose actions you are judging. Although real world data does not always lend itself to such observations, in the lab it is possible to allow subjects to play all roles in a game anonymously and then test to see if their behavior is consistent with the blame theory.

The experiment reported in this paper does just that. We take a simple dictator game and implement it in two stages. Although the subjects knew that there were two stages, they did not know what would transpire in the second stage, until after they made their decision in the first stage. In the first stage, the subjects split 10 tokens between themselves and an anonymous other person in the room. In the second stage, subjects were randomly matched with another person in the lab. After they were matched, they were offered, as a Receiver, the amount their matched pair member sent as a Sender in the first stage. Subjects did not have the option of rejecting proposals, but could, if they wished, at no cost to themselves punish their pair member by reducing his payoff by 1 token. Note that this design places subjects both in the role of a Sender and a Receiver in the first and second stages, respectively. Hence, in the second stage, when a subject receives an offer he is able to compare the offer that he receives to the offer that he made as a Sender in the first stage.

According to the theory of blame we expect that subjects will only punish if the offer that they receive is less than the offer they made in the first stage. Such a prediction differs from the predictions of the theories that we discussed vide supra. Our data suggests that a non-negligible part of the population behave according to the prediction of blame theory. Furthermore, blame theory is consistent with subjects' behavior more often than other theories are.

This paper is organized as follows. Section 2 and Section 3 introduce the *blame* concept in a more rigorous manner and provides an appropriate equilibrium concept, respectively. In Section 3, we also compute the equilibria of a formal example to elaborate more the equilibrium concept. Section 3 explains our experimental design and Section 4 derives the theoretical predictions of various theories. We state the results in Section 5 and conclude.

2 Blame in Games

In order to introduce the notion of blame, we will follow the general framework of Battigalli and Dufwenberg [1]. Our discussion will be confined to two-player, finite-horizon, multi-stage games with observable actions under complete information without chance moves. As Battigalli and Dufwenberg [1] point out, this class includes simultaneous moves games, perfect information games, and repeated games as special cases. For a discussion of the generalization to imperfectly observable actions, chance moves, and asymmetric information we refer the reader to Section 6 in Battigalli and Dufwenberg [1].

Our purpose in this section is to provide a rigorous definition of how blame can be introduced into psychological games. Since we follow Battigalli and Dufwenberg [1] closely in our analysis, the reader can refer to that paper for a more general theoretical analysis of dynamic psychological games, of which our paper is a special case.

Players, actions, and histories. Consider a multi-stage game consisting of two players $i = 1, 2$. The set of histories $\mathcal{H}$ is comprised of the initial history h^0, as well as all finite sequences of players' actions. At each history h, $A_i(h)$ denotes the finite set of actions available to player i. If $A_i(h)$ is singleton we say that player i is not active. Also, a history is *terminal* if and only if $A_i(h)$ is empty for both $i = 1, 2$. We refer to a terminal history as an *outcome* and denote the set of all outcomes by H. In order to distinguish an outcome from a non-terminal history, we denote an outcome by z. Each outcome z is associated with a *material payoff* for each player. The function $\pi_i : H \mapsto \mathbb{R}$ determines player i's material payoff $\pi_i(z)$ at the outcome z.

Strategies. For a player i, a *behavioral strategy* is a function σ_i, which maps each history h from the set of all non-terminal histories $\mathcal{H} \setminus H$ to a probability distribution over the set of actions available to player i at h: $\Delta(A_i(h))$. As in Battigalli and Dufwenberg [1], although we allow randomized choices for a player i, we interpret them as player j's beliefs about about player i's actions. In other words, player i does not actually randomize his choices, but the behavioral strategy σ_i is interpreted as player j's first order belief about player i. We will denote a pure strategy of player i by s_i.

Note that each strategy profile $\sigma = (\sigma_1, \sigma_2)$ induces a unique probability distribution over H. For a given σ, $\mathsf{E}_\sigma(\pi_i)$ denotes player i's expected material payoff induced by σ.

Preferences and Blame. Our definition of blame revolves around the question of *what a player would do if he were in the other player's position*. We formulate this question by carefully taking a player's higher order beliefs into account.

The literature of *psychological games* pioneered by Geanakoplos et al. [14] provides the appropriate framework to study problems that make explicit reference to players' beliefs. Geanakoplos et al. [14] assume that in games, the utility of a player not only depends on the outcome of the game, but also on his beliefs about the other players' strategy, beliefs about other players' beliefs and so on.

Although Geanakoplos et al. [14]'s novel approach allows us to study problems that conventional game theory is incapable of, it has its own restrictions. For instance, Geanakoplos et al. [14] allow only initial beliefs to enter into a player's utility. This assumption disregards the possibility that changes in players' beliefs can affect players' preferences during the course of the game. Moreover, Geanakoplos et al. [14] allow a player's own beliefs to determine his preferences but not the beliefs of other players. Battigalli and Dufwenberg [1] extend the theory in all these directions and accommodate the possibility

that players update their beliefs at each decision node along a path in a sequential game. This is a particularly important extension since it allows one to address dynamic psychological effects that can be observed during the course of a game. In our formulation of blame we will adopt a slight modification of this approach.

Our definition of blame makes reference to three levels of *beliefs*:

1. *Player i's belief about player j's strategy.* The role of this belief is relatively straightforward: It represents i's beliefs about player j's action at each history.

2. *Player i's belief about player j's belief about his strategy.* What will his beliefs be when player i puts himself in player j's position? This belief determines the answer: It tells us player i's belief about what player j thinks he (player i) plays. Consequently, it determines his beliefs (about his own strategy) when he puts himself in player j's position.

3. *Player i's belief about his strategy in player j's position.* Finally, player i needs to understand what he would do in player j's position. This is exactly what this belief tells us.

In our discussion earlier, we mentioned that players update their beliefs at each history. In order to capture that, in the remainder of the paper, for each belief, we will specify the history as a superscript. So, conditional on the event that the history h' is reached, we denote the above beliefs by $\hat{\sigma}_{ij}^{h'}, \hat{\sigma}_{iji}^{h'}$ and $\hat{\sigma}_{ii}^{h'}$, in the order we discussed above. For instance, $\hat{\sigma}_{ij}^{h'}(h)$ is a probability distribution over $A_j(h)$ representing player i's beliefs about player j's actions when he is at the history h'. We will discuss how these beliefs are updated later in this section. Finally, we denote player i's beliefs at h' by $\mu_i^{h'} := \left(\hat{\sigma}_{ij}^{h'}, \hat{\sigma}_{iji}^{h'}, \hat{\sigma}_{ii}^{h'}\right)$, and profile of players' beliefs by $\mu^{h'} := (\mu_1^{h'}, \mu_2^{h'})$.

We are ready to define *blame* and introduce preferences that makes explicit reference to blame at a history h'. Suppose that player i's strategy is σ_i and his beliefs are $\mu_i^{h'}$ at history h'. That is, player i believes that player j plays $\hat{\sigma}_{ij}^{h'}$ and that if he were in player j's position he would play $\sigma_{ii}^{h'}$. Also, when he puts himself in player j's position, his belief about player i's (his) strategy would be $\hat{\sigma}_{iji}^{h'}$. In other words, when player i considers himself in player j's position, he holds the belief that player i's strategy is $\hat{\sigma}_{iji}^{h'}$ and that he would play $\hat{\sigma}_{ii}^{h'}$. As a result, if player i were in player j's position, he would create an expected material payoff of $\mathsf{E}_{(\hat{\sigma}_{iji}^{h'}, \hat{\sigma}_{ii}^{h'})}(\pi_i)$ for the player who plays in his position (i.e. for himself). On the other hand, if player i's belief about player j's strategy is $\hat{\sigma}_{ij}^{h'}$, he believes that player j gives him an expected material payoff of $\mathsf{E}_{(\hat{\sigma}_{iji}^{h'}, \hat{\sigma}_{ij}^{h'})}(\pi_i)$. The difference between player i's belief of what player j thinks player i's expected material payoff is, and his expected material payoff when he plays against himself in the position of player j, is the source of player i's *blame*.

DEFINITION 1. *At history h', given his beliefs μ_i, player i is said to blame player j if*

$$\delta_i\big(\mu_i^{h'}\big) := \mathsf{E}_{\left(\hat{\sigma}_{iji}^{h'}, \hat{\sigma}_{ii}^{h'}\right)}(\pi_i) - \mathsf{E}_{\left(\hat{\sigma}_{iji}^{h'}, \hat{\sigma}_{ij}^{h'}\right)}(\pi_i) > 0. \tag{1}$$

Let us clarify the intuition behind the definition with the following statement from player i: *I blame player j because the expected material payoff that I think he expects that I will get when I play against him is less than my expected material payoff if I played against myself in his position. In other words, if I was in his position I would be nicer to a player in my position than he is to me.*

We will assume that each player i is endowed with a moral (or behavioral) type $\mathcal{B}_i$. Players' types are common knowledge. Let us denote the psychological utility function of player i with type $\mathcal{B}_i$ by $\psi_i(z, \mu_i^{h'}; \mathcal{B}_i)$ at history h'. In order to explicitly address the role of blame in players' preferences, we represent them by a function $u_i(z, \delta_i(\mu_i^{h'}); \mathcal{B}_i)$ with the transformation given in Definition 1, i.e. $u_i(z, \delta_i(\mu_i^{h'}); \mathcal{B}_i) := \psi_i(z, \mu_i^{h'}; \mathcal{B}_i)$. We posit that blame is a cause of disutility for a player. That is, a player prefers an outcome when he does not blame the other player to the same outcome when he blames the other player. In short, we assume that u_i is decreasing in δ_i.

When player i considers himself in player j's position he makes reference to his preferences in that position. We capture player i's preferences in player j's position by a function u_{ij}, which we define as

$$u_{ij}(z; \mathcal{B}_i) =: \psi_{ij}(z, \mu_i^{h'}, \mathcal{B}_i) := u_j(z, 0, \mathcal{B}_i).$$

This definition asserts that when player i puts himself in player j's position, he adopts player j's preferences with his own type $\mathcal{B}_i$, without any blame. In other words, when a player puts himself in the position of the other player, he keeps his innate characteristics $\mathcal{B}_i$. Moreover, he does not question whether he would blame himself, so, when he considers himself in j's position playing against himself, he simply sets $\delta_i = 0$ in his utility.

Let us elaborate more with the following example:

$$u_i(z, \mu_i^{h'}; \mathcal{B}_i) := v_i\big(\pi_i(z)\big) + \big(b_i - f\big(\delta_i(\mu_i^{h'})\big)\pi_j(z),$$

where v_i is a non-decreasing function of i's material payoff, $b_i \geq 0$, and f is a non-negative and increasing function such that $f(0) = 0$ for all $\delta \leq 0$. This specification indicates that player i's utility is determined by the weighted sum of a non-decreasing function of his material payoff, and a proportion of player j's material payoff. The function v_i simply captures how player i evaluates his material payoff. The term $\big(b_i - f\big(\delta_i(\mu_i^{h'})\big)$ determines the weight attached to player j's material payoff. The first term b_i is the weight that player i

puts on player j's material payoff without any blame consideration. However, the second term captures the idea that if player i blames player j, the overall weight decreases. If player i does not blame his opponent, he assigns a constant non-negative weight b_i to his opponent's material payoff. As player i blames more, the weight he assigns to player j's material payoff decreases and in fact when it is high enough (i.e. $f(\delta) \geq b_i$) player i becomes antagonistic to player j. Note that in this specification player i's type is $\mathcal{B}_i = (v_i, b_i)$.

Suppose that player j's preferences are similar:

$$u_j(z, \mu_j^{h'}; \mathcal{B}_j) := v_j\big(\pi_j(z)\big) + \big(b_j - f\big(\delta_j(\mu_j^{h'})\big)\pi_i(z).$$

Then, player i's preferences in player j's position is

$$u_{ij}(z; \mathcal{B}_j) := v_i\big(\pi_j(z)\big) + b_i\pi_i(z).$$

3 Equilibrium

The equilibrium concept that we define in this section is based on Geanakoplos et al. [14], Dufwenberg and Kirchsteiger [7], and Battigalli and Dufwenberg [1]. The equilibrium concepts that are introduced in these studies are comprised of the two usual key elements: sequential rationality and belief consistency. The idea behind sequential rationality is standard: Given players' beliefs, at each conceivable history of the game, each player's strategy is a best response to the strategy of the other player.

On the other hand, belief consistency requires more attention since beliefs determine preferences. Geanakoplos et al. [14] require that the beliefs that players hold at the initial node of the game are consistent with—i.e. identical to—the equilibrium strategies of the game. This approach disregards the possibility that players can revise their beliefs in the course of the game. Clearly, a change in a player's beliefs can possibly affect behavior since different beliefs induce different preferences. Dufwenberg and Kirchsteiger [7], and Battigalli and Dufwenberg [1] tackle this problem. More precisely, they allow the players to update their beliefs about the strategy of the other player at each history along the path and make their decisions accordingly.

We follow Dufwenberg and Kirchsteiger [7], and Battigalli and Dufwenberg [1] and define the equilibrium concept by taking dynamic belief updating into account. More precisely, at any history h, players revise their beliefs in such way that they are consistent with this particular history h. Let us introduce more notation in order to formally explain updating.

For any $h' \in \mathcal{H} \setminus \{h^0\}$, let us define $\mathcal{H}_{h'} \subset \mathcal{H}$ such that $h \in \mathcal{H}_{h'}$ only if $h' = (h, \iota)$ for some sequence of actions ι.[4] Put differently, $\mathcal{H}_{h'}$ is the set of all histories on the path of h'. Since players have perfect information of what happened in the past, for any $h \in \mathcal{H}_{h'}$ there exists a unique action $a_h^{h'}$ such that $(h, a_h^{h'}) \in \mathcal{H}_{h'}$. That is, $a_h^{h'}$ is the action that follows history h on the path to h'.

Now we can formally define the updating rule. Let $\hat{\rho}$ be a belief of a player i, i.e., $\hat{\rho}$ is either $\hat{\sigma}_{ij}, \hat{\sigma}_{iji}$ or $\hat{\sigma}_{ii}$. At a history $h' \in \mathcal{H} \setminus \{h^0\}$, the updated belief is denoted by $\hat{\rho}^{h'}$ and defined as

$$\hat{\rho}^{h'}(h) := \begin{cases} \mathbf{1}_{a_h^{h'}} & \text{if } h \in \mathcal{H}_{h'}, \\ \hat{\rho}(h) & \text{otherwise,} \end{cases}$$

where $\mathbf{1}_a$ is the Dirac measure, indicating a single atom at the action a. In short, conditional on reaching history h', a player's belief about the action following history h is the unique action on the path from h to h' if h is on the path to h'. So, beliefs about what happened in the past are consistent with the fact that the history h' is reached. Otherwise, if h is not on the path to h', we assume that beliefs remain the same.[5]

Before we define the equilibrium concept, which we call the *Sequential Blame Equilibrium* (SBE), let us define the expected psychological utility of player i, given his beliefs, pure strategy s_i, and player j's strategy.

$$\mathsf{E}_{(s_i,\sigma_j^*)}\big(u_i(z, \delta_i(\mu_i^h); \mathcal{B}_i)\big) := \int u_i(z(s_i, s_j), \delta_i(\mu_i^h); \mathcal{B}_i) dP(s_j|\sigma_j^h),$$

where $P(s_j|\sigma_j^h)$ is the probability of player j's pure strategy induced by σ_j^h.[6] Similarly, the expected psychological utility of player i in player j's position, given his beliefs, his pure strategy in player j's position, s_i, is:

$$\mathsf{E}_{(\hat{\sigma}_{ii},s_j)}\big(u_{ij}(z; \mathcal{B}_i)\big) := \int u_{ij}(z(s_i, s_j); \mathcal{B}_i) dP(s_j|\hat{\sigma}_{ii}^h).$$

DEFINITION 2. *The profile (σ^*, μ^*) is a SBE if for $i, j \in \{1, 2\}$ and for each history $h \in \mathcal{H}$, the followings hold:*

(i) $P(s_i|\sigma_i^{h*}) > 0 \Rightarrow s_i \in \arg\max_{s_i} \mathsf{E}_{(s_i,\sigma_j^*)}\big(u_i(z, \delta_i(\mu_i^{h*}); \mathcal{B}_i)\big)$,

(ii) $P(s_j|\hat{\sigma}_{ii}^{h*}) > 0 \Rightarrow s_j \in \arg\max_{s_j} \mathsf{E}_{(\hat{\sigma}_{ii},s_j)}\big(u_{ij}(z, ; \mathcal{B}_i)\big)$,

[4] We follow the notational convention and write (h, ι) to refer to a sequence of actions which starts with h and then followed by the sequence ι.

[5] Although, at a given history h we leave players' beliefs unchanged at the histories that are not on the path to h, it is possible to consider alternative updating rules. In the current formulation, we implicitly assume that the fact that a player follows a strategy which is consistent with the history h does not offer any reason to change the beliefs about his strategy at the histories that are not on the path to h.

[6] Given the assumption that a behavioral strategy is independent across histories $P(s_j|\sigma_j^{h'}) := \prod_{h \notin \mathcal{H}_{h'}} \sigma_j(h)(s_j(h))$.

(iii) $\hat{\sigma}^*_{ij} = \sigma^*_j$, *and* $\hat{\sigma}^*_{iji} = \sigma^*_i$.

The SBE has three requirements. The first is sequential rationality: At each history, the player who is supposed to take an action, given his beliefs and the other player's strategy, achieve the maximum utility by following his strategy. The second requirement is about the sequential rationality of the strategy $\hat{\sigma}_{ii}$: At each history, the strategy that a player believes he would have played in other player's position, maximizes his utility in that position. The final requirement of SBE is consistency of belief: Players' beliefs are derived from equilibrium strategies.

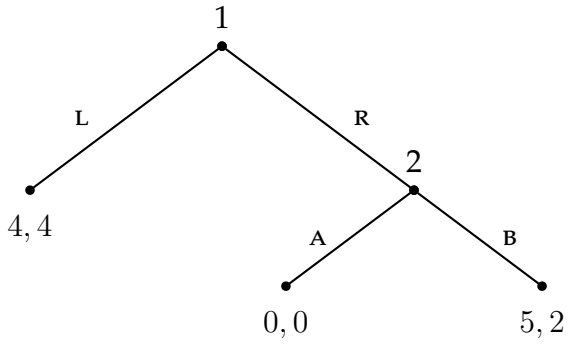
Before we move on to illustrate our ideas by use of a simple example, we should note that the SBE equilibrium exists. The result immediately follows from Battigalli and Dufwenberg [1].[7]

PROPOSITION 1. *Any blame game as described in Section 2, has a SBE in mixed strategies.*

4 A SIMPLE EXAMPLE

In this section we illustrate the notion of blame and the equilibrium analysis by using the simple sequential game displayed in Figure 1. Throughout the analysis, for ease of exposition, we will focus on pure strategy equilibrium. In this game, there are three outcomes (L), (R,A), and (R,B), and the material payoffs corresponding to them are given just below the terminal nodes.

FIGURE 1: A two-person extensive game with blame.



Let $\beta_i(\delta_i) := b_i - f(\delta_i)$ for some $1 \geq b_i \geq 0$ and a non-decreasing function of blame, f, such that $f(\delta_i) = 0$ for all $\delta_i \leq 0$. Also, we assume that $0 \leq f(\delta_i) \leq 1$ for all $\delta_i \leq 5$. Our

[7] Note that one can define "player i in player j's position" as a meta player. Note that in our formulation, this meta player's preferences do not depend on beliefs, and the optimality of his actions is guaranteed by the equilibrium condition.

choice of β_i suggests that if player i does not blame his opponent, he assigns a constant non-negative weight b_i to his opponent's material payoff. As player i blames more, the weight he assigns to player j's material payoff decreases and in fact when it is high enough (i.e. $f(\delta) \geq b_i$) player i becomes antagonistic to player j.

In order to illustrate how blame is computed let us consider the following scenario.

$$\sigma_{21}(\varnothing) = \text{L},\ \hat{\sigma}_{21}(\varnothing) = \text{R},\ \text{and}\ \hat{\sigma}_{212}(\text{R}) = \text{B}.$$

That is, (*i*) player 2's strategy, if he were in player 1's position, would be to play L, (*ii*) player 2 believes that player 1's strategy is to play R at the initial node, and (*iii*) player 2 believes that player 1's belief about his strategy is that he plays B at the history (R).

Note that if player 2 were in player 1's position with $\hat{\sigma}_{212}(\text{R}) = \text{B}$, his strategy $\sigma_{21}(\varnothing) = \text{L}$ would lead to a material payoff of $\pi_2(o_{(\sigma_{21},\hat{\sigma}_{212})}) = 4$ for a player in his position. On the other hand, player 2 believes that player 1 plays R under the belief that he will actually play B. As a result, he believes that player 1 expects to give him a material payoff of $\pi_2(o_{(\hat{\sigma}_{21},\hat{\sigma}_{212})}) = 2$. Therefore, while he would give a material payoff of 4 to a player in his position, he believes that player 1 expects to give him a payoff of 2. Hence, by Definition 1, he blames player 2 by $\delta_2(s_2, \mu_2) = 4 - 2 = 2$.

Note that in the above discussion we do not question why player 2 would play $\sigma_{21}(\varnothing) = \text{L}$ in player 1's position. We simply take it as given. If this scenario were to be part of an equilibrium however, we would require that σ_{21} is a maximizer of the utility u_{21} given the beliefs at the equilibrium.

Having determined how to compute a player's blame, we now turn to the analysis of SBE of the game. In the conventional case where players have no blame concerns ($b_i = 0$ and $f(\cdot) = 0$) the unique subgame perfect equilibrium predicts the outcome (R,B). In what follows, we will show how SBE differs from the subgame perfect equilibrium.

Since this is a game of compete information, players' b_i's are common knowledge. Hence, the equilibrium behavior of players will be potentially different depending on players' b_i's. In fact Figure 2 depicts different regions of the b_1, b_2 space each determining a different equilibrium. We will discuss an equilibrium that exists in the region where $0 \leq b_1 \leq 1$ and $1/2 \leq b_2 < \overline{b} := \max\{1/2, f(4) - 2/5\}$ and leave the analysis of the others to the next section.

Let us show that the following is the SBE of the game for our set of parameters (the rectangular region indicated with bold line in Figure 2):

$$\begin{array}{lll} \sigma_1(\varnothing) = \hat{\sigma}_{21}(\varnothing) = \hat{\sigma}_{121}(\varnothing) = \text{L} & \text{and} & \sigma_2(\text{R}) = \hat{\sigma}_{12}(\text{R}) = \hat{\sigma}_{212}(\text{R}) = \text{A} \\ \sigma_{12}(\text{R}) = \text{B} & & \sigma_{21}(\varnothing) = \text{L} \end{array}$$

Let us first show that the strategies σ_{12} and σ_{21} are optimal. Recall that in the equilibrium σ_{ij} is supposed to be a best response to $\hat{\sigma}^h_{iji}$ at any history h. The strategy $\sigma_{12}(\text{R}) = \text{B}$ is optimal because while it yields a utility u_{21} of $2 + 5b_1$, the action A yields 0. As a result, because $b_1 \geq 0$, player 1 finds it optimal to play B in player 2's position. Recall that the definition of u_{ij} assumes that player i does not blame himself when he puts himself in the position of player j, i.e. $\beta_i(0)$. This allow us to make the above comparison. Also, $\sigma_{21}(\varnothing) = \text{L}$ is optimal because it yields a utility u_{12} of $4 + 4b_2$ as opposed to 0, which the player expects to receive by taking action R given his belief $\hat{\sigma}_{212}(\text{R}) = \text{A}$.

In order to show the optimality of players' strategies σ_1 and σ_2 we should first understand how much they blame each other. When we look at the situation of player 1, based on his belief $\hat{\sigma}_{121}(\varnothing) = \text{L}$, when he puts himself in player 2's position he does not expect player 1 to choose R. Hence, since he does not expect (R) to be reached, his action in player 2's position is irrelevant and would not affect his payoff. Therefore, player 1 has no reason to blame player 2, i.e. $\delta_1(s_1, \mu_1) = 0$. Without blame, if he plays L his payoff is $4 + 4b_1$ and if he plays R his expected payoff is 0 since he believes that $\hat{\sigma}_{12}(\text{R}) = \text{A}$. Therefore, his best response is to choose L.

Now, let us show that player 2's strategy is also optimal. First, note that player 2's beliefs at (R) is updated to $\hat{\sigma}^{\text{R}}_{21}(\varnothing) = \text{R}$. Therefore, given his beliefs $\hat{\sigma}_{212}(\text{R}) = \text{A}$, and what he would do in player 1's position $\sigma_{21}(\varnothing) = \text{L}$, he blames player 1 by 4. In this case, his utility from A is 0 while his utility from B is $2 + 5\big(b_2 - f(4)\big)$. Since in the region of interest $b_2 \leq \bar{b} := \max\{1/2, f(4) - 2/5\}$ his optimal action is A. This equilibrium highlights the importance of blame in our formulation of player's notion of kindness. Note that player 2 has a relatively high b_2 and therefore when he puts himself in player 1's position he considers taking action L. When he believes (or finds out) that player 1 took action R, he blames him so much that he wants to carry out his threat of taking action A because, at that point, getting 0 is better for him than allowing player 1 to get a payoff of 5 and him a payoff of 2.

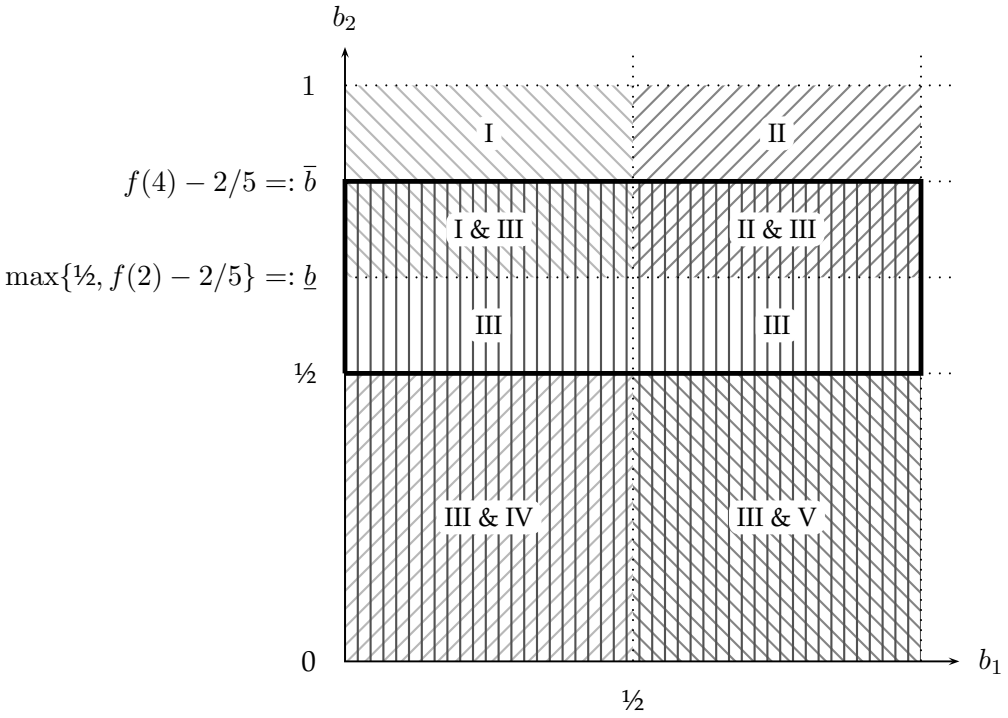
Figure 2 summarizes the equilibria of the game for different combinations of b_1 and b_2. The horizontal axis is b_1 and the vertical axis is b_2. We provide the full equilibrium characterization in Appendix B. There are five different equilibria labeled by Roman numerals. We list the equilibria as $\Big[\big(\sigma_1(\varnothing), \sigma_{12}(\text{R})\big); \big(\sigma_2(\text{R}), \sigma_{21}(\varnothing)\big)\Big]$. We omit the beliefs since $\sigma_i = \hat{\sigma}_{ji} = \hat{\sigma}_{iji}$.

$$\text{I: } \big[(\text{R,B});(\text{B,L})\big];\ \text{II: } \big[(\text{L,B});(\text{B,L})\big];\ \text{III: } \big[(\text{L,B});(\text{A,L})\big];\ \text{IV: } \big[(\text{R,B});(\text{B,R})\big];\ \text{V: } \big[(\text{L,B});(\text{B,R})\big].$$

The equilibrium that we analyze above is labeled as III, and the analysis is relevant for region (with heavy borderline) where, $1/2 < b_2 < \bar{b}$, and $0 < b_2 < 1$.

To give a flavor of what happens in other regions, consider the region where both b_1

FIGURE 2: Characterization of SBE of the game in Figure 1.



For each region, the type of equilibria that appear in the region are labeled in Roman numerals. This figure represents the case where $1/2 < f(2) - 2/5$ and $f(4) - 2/5 < 1$. We denote $\bar{b} := \max\{1/2, f(4) - 2/5\}$ and $\underline{b} := f(2) - 2/5$.

and b_2 are smaller than $1/2$ and let us describe equilibrium IV. Small values of b_i's imply that neither player cares greatly for the other's material payoff. As a result, if the game ever progressed to (R), player 2 would not blame player 1 for his action of R since he would have done the same thing. In addition, player 1 is willing to play R since he expects player 2 to play B. Hence in the SBE player 1 plays R and player 2 plays B. Let us have a look at the equilibrium I, which appears when b_1 is smaller than $1/2$ and b_2 is larger than $1/2$. This is a case where player 1 does not care much about player 2's material payoff, while player 2 does care about player 1's material payoff. In this equilibrium, at the history (R), player 2 blames player 1 for his action of R since he would have player L if he were in player 1's position. Even though player 2 blames player 1 for this action, he does not play A (i.e. punish player 1) because he cares about player 1's material payoff enough that it dominates his blame.

5 Experimental Design

The aim of our experiment was to test the elements of our theory of blame in a manner that focused on subjects' preferences and motivations for punishment. In order to do this we designed an experiment that did not involve any strategic interaction. Rather, the experiment was a direct test of preferences involving blame. This design had the obvious advantage of holding strategic considerations constant for our subjects and not having them confound observed behavior. Later in this section, we also present the analysis of an experiment—which involves a more strategic situation—run by Carpenter, Kariv and Schotter [4].

For our experiment sixty six subjects were recruited from the undergraduate population of New York University and asked to come to the experimental lab of the Center for Experimental Social Science. They engaged in one round of the experiment to be described below. A show up fee of $8 was given and subjects earned on average $15.06 for about 25 minutes.

The task that our subjects faced was extremely simple. It was composed of two stages. Subjects were not informed what the content of the second stage would be before they completed the first stage. But they were told that there would be a second stage.

In the first stage the subjects played a dictator game. That is, they were asked to split 10 tokens (convertible to dollars at the rate of 1 token = 1$) between themselves and an anonymous other person in the room. After all subjects made their choices they were randomly divided into two equally sized groups called Senders and Receivers and randomly matched in pairs. If a subject was chosen to be a Sender his payoff was equal to what he decided to keep for himself of the 10 tokens. If he was chosen to be a Receiver, his payoff was equal to the amount given to him by the Sender whom he was matched with. Subjects were not told their payoff from the first stage until after both stages of the experiment were completed.

After the first stage was completed subjects moved on to the second stage, where they were randomly matched with another subject in the lab. After they were matched, they were offered, as a Receiver, the amount that the subject whom they were matched with sent as a Sender in the first stage. In other words, if a subject was matched with a person who gave 3 tokens in the first stage, he was offered 3 tokens in the second stage as his payoff while the other subject got 7. Subjects did not have the option of rejecting proposals, but they could, if they wished, punish their pair member by reducing his payoff by 1 token at no cost to themselves. To elicit their response we used the strategy method in that rather than having the subjects punish directly, we asked them to set a cutoff before seeing the offer. If the offer was equal to or less than the cutoff, 1 token would be removed from the

other subject's payoff. For example, say that a subject chose a cutoff of 4. If the subject was matched with a subject who offered 3 tokens in the first stage, and hence kept 7 for himself, then the computer would reduce the pair member's payoff by 1 token from 7 to 6. If the Sender gave 5 tokens and kept 5 tokens, his tokens would not be reduced and will stay at 5.

The the second stage payoff was determined as follow. After subjects made their choices in the second stage, the computer randomly determined whether they are a Sender or a Receiver. If a subject was designated to be a Receiver the subject whom he was matched with was designated to be a Sender and vice versa. A Sender's payoff was equal to what he decided to keep for himself in the first stage minus any 1-token reduction by his pair member if there was a punishment.

The data of interest generated by the experiment are the Stage-1 offers and the Stage-2 cutoffs since those are the basic ingredients for both our theory and other theories.

The total payoff for the subjects in the full 2-stage experiment was equal to the sum of earnings in the first stage and the second stage plus the show up fee.

In the next section we will review the predictions of four theories that could be applied to make predictions in this experimental environment. We will then analyze the data in the light of the predictions of these theories.

6 Theoretical Predictions

If theories are to be of interest, they must generate predictions that are distinctly different from those of other, competing, theories and be testable using a simple and transparent experiment. The experiment outlined above, we feel, is an excellent testing ground for our theory of blame because it is simple and determines quite different types of predictions from all of the other leading theories that can be used to explain the data it generates—i.e., the Fehr and Schmidt [13]'s Inequity Aversion Model, Rabin's Theory of Kindness Reciprocity, Levine's Theory of Interdependent Preferences, and our Theory of Blame.

Our goal in this section of the paper is to describe the behavior prescribed by these theories of reciprocity in the particular experiment we have just discussed. Throughout this section we assume that a subject assigns a utility $u(x)$ to a material payoff x. For expositional ease let us assume that u is increasing, continuous, and concave.

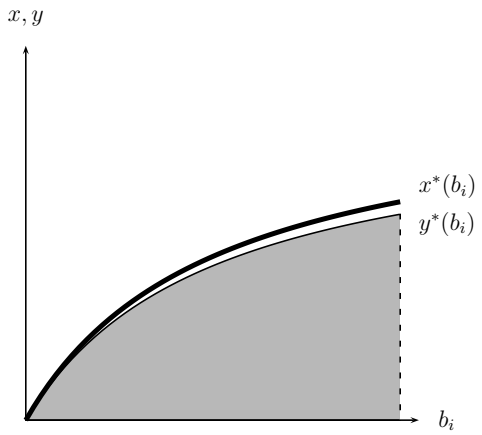
Since our experiment asks subjects to first divide 10 experimental tokens and to then decide on a punishment to be directed at those whose division they think is worthy of punishment, we will investigate each contending theory for its predictions on this division and punishment behavior. More precisely, we will discuss a subject's optimal allocation between himself, $10 - x^*$, and another subject, x^*, if he was given a material payoff of 10

tokens to divide and also determine whether the same individual is willing to punish a subject who gives him y when he is in the Receiver position. We will discuss each theory in turn.

Blame

Let us start with the predictions of our theory. Note that since the Sender's problem in the first stage is just to choose an optimal allocation, the subject chooses $x \in [0, 10]$ to maximize $u_i(10 - x) + \beta_i\big(\delta_i(s_i, \mu_i)\big)x$. Let us assume that $\beta_i\big(\delta_i(s_i, \mu_i)\big) := b_i - f\big(\delta_i(\mu_i, s_i)\big)$ where f is a non-negative and increasing function such that $f(0) = 0$ for all $\delta \leq 0$. Since the first stage is simply a dictator game, there is no room for blame; hence $\delta_i = 0$ for any subject i. It is therefore straightforward to see that the optimal amount x^* a subject will send is a non-decreasing function of b.

Figure 3: Prediction of the Theory of Blame



The optimal offer of a subject with b_i is $x^*(b_i)$. The subject can punish any offer below $y^*(b_i)$, i.e. shaded region.

The amount $x^*(b_i)$ is what subject i with a particular b_i will optimally offer to an anonymous subject when he is asked to allocate a material payoff of 10 as a Sender.

Now suppose that the subject is in the Receiver's position in the second stage and he is offered an amount y. Based on the premise of our theory, the subject evaluates this offer based on what he would have offered—in fact he did offer—as a Sender. Therefore, he updates his preferences by incorporating blame $\delta_i(s_i, \mu_i) = x^*(b_i) - y$. Therefore, the receiver may find it optimal to punish the sender for the offer y if $y < y^*(b_i) := x^*(b_i) - f^{-1}(b_i)$.

Figure 3 summarizes these observations. The bold line indicates the optimal offer of the Sender. Since the choice of f is arbitrary the punishment cutoff $y^*(b_i)$ is below $x^*(b_i)$ for any b_i. The important feature is that both the offer, $x^*(b_i)$, and the punishment cutoff, $y^*(b_i)$, are increasing in a subject's b_i with the cutoff everywhere below the offer.

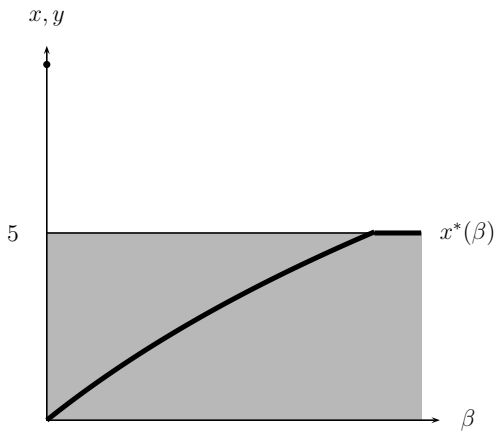
Inequity Aversion

The view that individuals can exhibit inequity aversion is studied by Fehr and Schmidt [13].[8] According to this theory, although a person likes more material payoff, he gets disutility from being far off from others. Based on Fehr and Schmidt [13]'s formulation of such preferences, in our problem, a subject chooses an offer x to maximize

$$u(10-x) - \alpha \max\left\{2x-10, 0\right\} - \beta \max\left\{10-2x, 0\right\}$$

where $\alpha > \beta \geq 0$. Note that the larger a subject's β parameter the "nicer" he is when functioning as a Sender since the more he dislikes having more than the Receiver. In that sense, it is directly comparable to our b_i. Also note that the α parameter is relevant for utility when the subject is receiving less that 5 (allocates more than 5 to the Receiver) while the β parameter is relevant when the subjects receives more than 5 (allocated less than 5 to the Receiver).

Figure 4: Prediction of the Theory of Inequity Aversion



The optimal offer of a subject with β is $x^*(\beta)$. Any subject can punish any offer below 5, i.e. shaded region.

[8] Bolton and Ockenfels [3] studies the same idea with a different formulation of preferences. Since the predictions of the two papers are the same we do not provide a separate discussion of that paper.

Since it is never optimal for a subject to allocate more than 5 to the Receiver (given $\alpha > 0$), it will be the β parameter that will determine a subject's allocation when functioning as a dictator. Hence, $x^*(\beta)$, which is the optimal amount a subject with parameter β, will offer increases in β. However, it never exceeds the most equitable allocation 5, because beyond 5 the subject bears a disutility from inequity aversion and receives less material payoff. The bold line in Figure 4 illustrates the optimal offer.

Now we turn to the question of when a subject punishes a Sender in the second stage. Let y be the offer that a Receiver gets. Any offer $y < 5$ creates a disutility since $\alpha > 0$. Therefore, since punishment reduces the existing inequality and does so at zero cost, the Receiver will find it optimal to punish any offer $y < 5$. Also note that a Receiver who offered $x^* > 0$ when he is in the Sender position must necessarily have $\beta > 0$. Therefore, for any offer $y \geq 5$, since punishment further increases inequality, the Receiver does not find it optimal to punish the Sender. Only when $x^* = 0$ it is possible that $\beta = 0$ and hence the agent is indifferent between punishing any offer y.

The figure depicts this behavior. The bold line indicates the optimal offer $x^*(\beta)$ of an agent as a Sender and the gray area indicates the offers that will be punished. Note that behavior under Fehr and Schmidt [13] differs qualitatively from that of our blame theory because, while in our theory the punishment cutoff is an increasing function of a subject's b_i, under Fehr and Schmidt [13], the punishment threshold is a constant of 5 and independent of a subject's α. This provides a very strong prediction for the Fehr and Schmidt [13]'s theory which is that all offers below 5 will be punished, a prediction that is distinctly deferent than our blame theory provides.

Kindness

Most reciprocity theories rely on an exogenous norm of kindness in order to determine how kind people are.[9] For instance, in the theory of Rabin [17], "[...] players have a shared notion of kindness and fairness and that they apply these standards symmetrically." The main postulate of this theory is that people return kindness with kind acts and unkindness with mean acts. To be precise, Rabin [17] defines kindness with reference to a predetermined allocation (e.g. equitable allocation) on the Pareto frontier of possible payoffs. If a player believes that his opponent's strategy leads to a payoff that is less than this predetermined allocation, then the player finds his opponent's act unkind. Although Rabin [17] shows that the results are valid for a large class of such reference points, whatever that reference point is, it is nevertheless exogenously determined.

Since strategic interaction in our experiment is eliminated it is relatively easy to deter-

[9] See [2, 8–11] for leading examples.

mine both giving and punishing behavior for all our competing theories. For Rabin, how much a subject is willing to give is determined by the marginal benefit he receives and how much he cares about being nice to an anonymous person. Therefore, generically, one would expect that the offers increase in b_i as in Figure 3. Rabin [17] assumes that if there is no opportunity for reciprocity, as in the dictator game, then since a person's utility is linear in his own material payoff and independent of anyone else's, $(b_i = 0)$, he would give zero when in the position of a dictator. If we allow $b_i > 0$, then as we said above we would expect more to be allocated to a Receiver as b_i increases.

On the other hand, in the second stage of the experiment, the punishment behavior is determined by how kind a player thinks the offer he receives. Precisely, the offer is unkind if it is below the predetermined norm and kind otherwise. Therefore, the theory predicts that a Sender is punished if the offer he makes is below the norm, which is arbitrary.

Interdependent Preferences

Levine [16] takes a novel approach in formulating altruism and reciprocity by using orthodox game theoretical tools.[10] In particular, in order to analyze experimental data, Levine [16] models the underlying game as a Bayesian game where the types determine how altruistic or kind a subject is towards another subject. A simplified version of the preferences are as follows.

$$u(10-x)+\frac{a_i+\lambda a_j}{1+\lambda}x.$$

That is, each subject i has a type $a_i \in (-1,1)$, which determines the weight they assign to the other player's material payoff. Players are also sensitive to who they are playing against. In particular, the utility function posits that the weight is higher if the opponent is nicer; i.e. higher a_j. Also the intensity of the sensitivity is captured by a parameter λ. That is, λ measures the intensity of reciprocity for player i. Clearly, if $\lambda = 0$, player i is not affected by whom he plays against. As $\lambda > 0$ increases, the agent becomes more reciprocal. Therefore, for $\lambda > 0$, player i is nicer against a nice player.

In our experimental setup, since a Sender does not know who he plays with, he chooses x to maximize

$$u(10-x)+\int \frac{a_i+\lambda a_j}{1+\lambda}dF(a_j)x$$

where F is the distribution of types.

It is easy to determine the optimal offer $x^*(a_i)$ of player i. Since the weight is an increasing function of a_i, the optimal offer is non-decreasing in a_i. Recall that in our experiment we keep the first and the second stages strategically independent. Therefore, the type of

[10] For a more general approach we refer the reader to Gul and Pesendorfer [15].

a player is revealed through his offer y as Sender in the first stage.

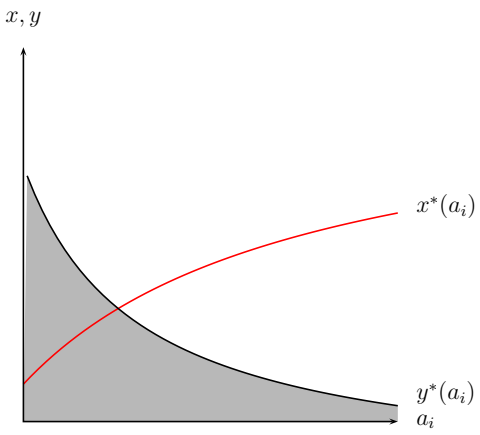
In the second stage, whether a Receiver punishes a Sender or not depends on his and the Sender's type. In fact, if $a_i < -\lambda a_j$ the Receiver punishes. This immediately suggests that there is an inverse relation between what an agent offers and his punishment threshold. The intuition is simple. According to Levine [16]'s formulation, how much a player cares about the material payoff of another player increases both in the other's as well as his own "kindness" (type). So a spiteful player may possibly punish another player despite his kind offer. Similarly, a very nice agent does not punish anybody except maybe very spiteful people.

Finally, note that Levine [16]'s theory differs from ours in that while for the theory of Blame as a player's type increases and he becomes nicer and nicer, he expects more and more from others since he is judging them by his own standards. For Levine [16], however, nice people are more and more tolerant of the nasty behavior of others because, in some sense, the two types are averaged in determining how much a subject cares about his opponent.

As a result, Levine [16]'s theory predicts punishment behavior that is almost directly opposite of ours. While our punishment cutoffs are increasing in how nice our subjects are, nice people demand nice behavior in a theory of blame, Levine's punishment cutoffs are decreasing in how nice a subject is.

Figure 5 illustrates the discussion.

FIGURE 5: Prediction of the Theory of Levine [16]



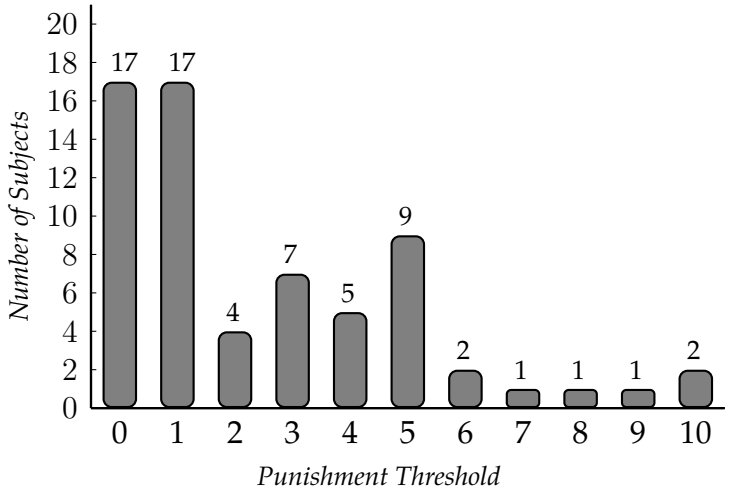
The optimal offer of a subject with a_i is $x^*(a_i)$. Any subject can punish any offer below $y^*(a_i)$, i.e. shaded region.

6.1 Preference Experiment

In light of the predictions discussed above we will present the results of our experiment by concentrating on the relationship between the offers made in the first stage and the cutoffs stated in the second stage. Remember, the different theories we will be comparing make stark and qualitatively different predictions. While the Fehr and Schmidt [13] (and other inequality averse models like Bolton and Ockenfels [3] predict that cutoffs will be 5 no matter what offer a subjects made, our blame theory predicts that cutoffs will be increasing in the offers made while Levine [16] suggests the opposite: cutoffs will be a decreasing function of offers. Hence, the only theory consistent with a positive relationship between offers and cutoffs is our theory of blame and we will demonstrate below that our data is consistent with such a positive relationship.

We begin by easily rejecting the hypothesis that our subjects acted in accordance with the Fehr and Schmidt [13]'s theory. Remember those theories predict that all subjects set a cutoff or approximately 5 no matter what offer they send. Hence, if our subjects acted according to the Fehr and Schmidt [13]'s model all cutoffs would be 5. Such a strong prediction is easily rejected since in fact only 9 of 66 observed cutoffs were equal to 5. Of those 9 subjects, however, 5 exhibited behavior that was also consistent with the theory of blame so only 4 subjects exhibited behavior that could only be explained by Fehr and Schmidt [13] (or Bolton and Ockenfels [3]). Hence, it is obvious that the cutoff of 5 was not one that was commonly used nor represented the central tendency of subjects since a median test rejects the hypothesis that 5 was the median of the distribution at less than the 1% level of significance using a binomial test.

Figure 6: Histogram of punishment thresholds



When it comes to our theory of blame the data is far more kind. As discussed in Section

6 our theory predicts that a subject sets his cutoffs below his offer and that cutoffs are increasing in the offers made in the first stage. These facts are easily established. For example, of our 66 subjects only 15 chose cutoffs that were above their offers and of these 15, 9 of them had made offers of zero and hence were at the corner of the feasible offer set. For subjects who made zero offers, however, any stated positive cutoff would violate our theory. For the 51 subjects who made interior offers, however, only 6 violated the predictions of our blame theory. To see this relationship graphically consider Figure 7 which shows the offers and cutoffs for each of our 66 subjects in the experiment arranged in ascending order of offers. As can be seen, for the overwhelming majority of subjects cutoffs stated were below offers made. The subjects who made zero offers and positive cutoffs constitute a strange subset since they revealed themselves to be totally selfish with respect to offers yet demanded positive amounts in order not to punish. In fact, some not only made zero offers but actually set cutoffs of as much as 9 and 10 indicating completely ego centric preferences. Unless these subjects also set a cutoff of zero, our predictions are bound to fail from them. It is interesting to note that the behavior of these subjects is consistent with the Rabin [17]'s theory since if $b_i = 0$, as is assumed by Rabin, then subjects are expected to offer 0 in the dictator game yet demand something positive in Stage 2. (Obviously, such subjects do not put themselves in the position of others and ask what they would have done since they would have given zero.) It is also consistent with the behavior of spiteful subjects ($b_i < 0$), since they would offer zero and punish for spite.

With respect to the relationship between offers and cutoffs a regression of cutoffs on offers indicates a positive relationship. When the regression is run on the full data set, while the relationship is positive the coefficient in front of the offer variable is not significant at the 5% level. This result, however, is dominated by the subjects who set zero as their cutoff. When we remove those subjects and concentrate only on those who made interior offers, the coefficient is highly significant and positive as shown in Table 1.

TABLE 1: Relationship between the cutoffs and offers

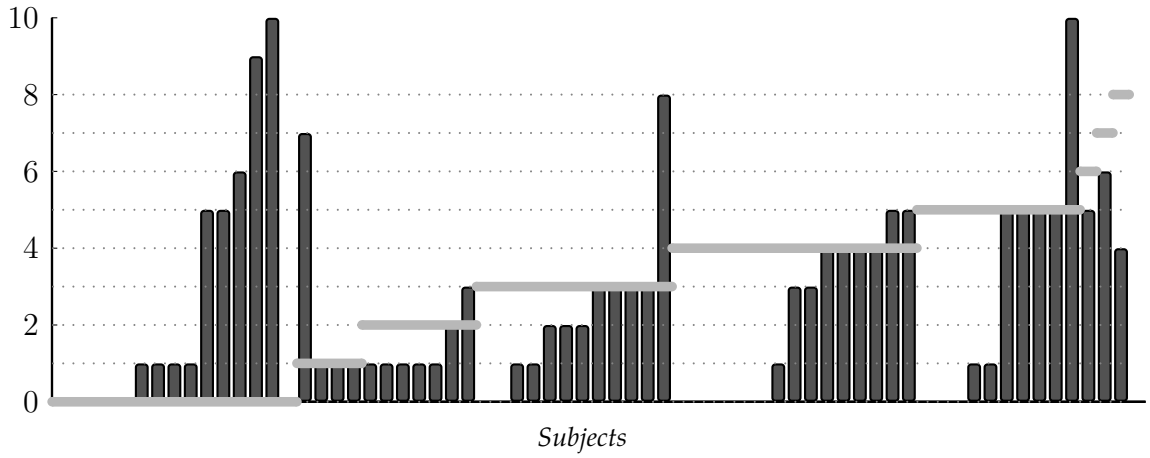
	Coeff	Std Err	t	$P > \|t\|$
offer	.447	.206	2.17	.035
constant	.872	.799	1.09	.280

$N = 52$, Adjusted $R^2 = .068$.

By demonstrating that the relationship between offers and cutoffs is positive we have rejected the hypothesis that the Levine [16]'s model does a good job at organizing our data since it posits that the relationship is negative.

Figure 8 presents the cumulative distribution of offers and cutoffs. As we can see, the

FIGURE 7: Offers vs cutoffs



Gray shaded bars indicate the punishment cutoff of each subject, while the light gray bar denotes a subject's offer.

distribution of cutoffs is significantly shifted to the left when compared to the distribution of offers. A Kolmogorov-Smirnov test run to compare these two distributions rejects the hypothesis of equality at the 5% ($D = 0.227, p < 0.045$) level in favor of the alternative that cutoffs are lower than offers. This shift to the left is predicted by our theory but not by any of the others.

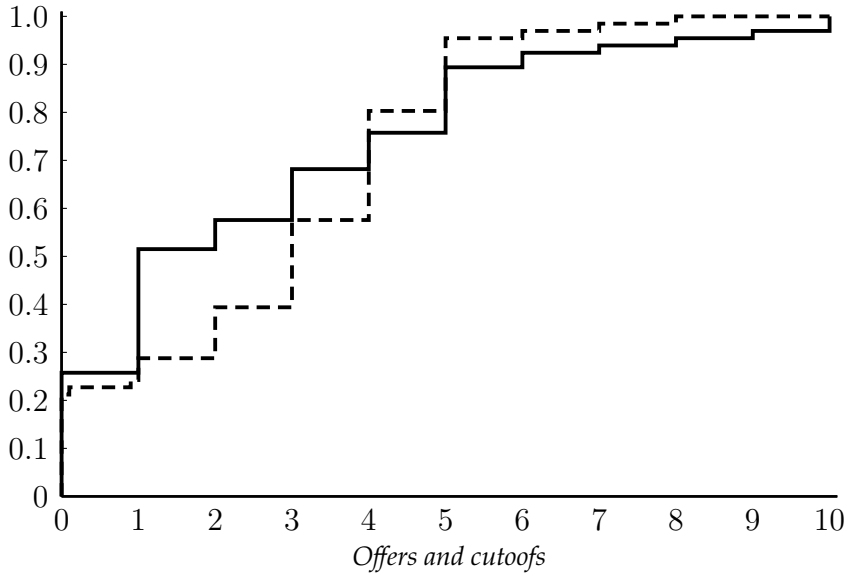
In summary, the data generated by our experiment lends support to the idea that subjects used elements of blame in their punishment behavior and rejects the idea that their behavior can be easily explained by any of the other theories we have discussed. The Fehr and Schmidt [13] and Bolton and Ockenfels [3]'s theories fails because cutoffs fail to be consistently equal to five while Levine's theory fails because the relationship between offers and cutoffs is positive in contrast to the prediction of Levine [16]'s theory. Totally selfish subjects, those who offer zero, are the exception to the rule and their behavior can be explained, perhaps, by Rabin [17] or Levine [16] since for Rabin [17] offering 0 is expected while for Levine [16] there is a inverse relationship between offers and cutoffs.

6.2 PUBLIC GOODS DATA

If our theory of blame is correct we should be able to detect it in other experiments as long as they have two main features: 1. That subjects be allowed to punish others (or reward them) and, 2. That subjects are placed in identical symmetric situations so they know how they would behave if they were in another person's shoes.

Such requirements are easily satisfied by public goods experiments with punishments

FIGURE 8: Cumulative distributions of offers and thresholds



The solid line indicates the offers while the dashed line indicates the cutoffs.

where, as in Fehr and Gaechter [12] for instance, subjects are allowed to punish others. This is true because in such games people can engage in the blame thought experiment since each person knows how much they contributed to the pubic good and can therefore compare their behavior to others and see if they want to blame or punish them.

Caution must be used when looking at such data, however, since most, if not all such experiments are embedded in a complete network where all subjects can see the contributions of all others and punish as many people as they wish. This network structure is not the best to use in assessing the motives for subject punishment, since it is steeped with free riding and other coordination problems which easily mask the motives for punishment. For example, say that I am in such a game with three others and observe the public goods contribution of all three. In addition, say that it is common knowledge that each person can observe and punish whomever he wants. Whether I punish then becomes a function of my beliefs about whether the others in my groups will punish and I might easily decide to free ride on their punishments in order to save money. In addition, who I punish is also a complex coordination game since I obviously would want to punish the person who has contributed the least but since such a person is likely to be punished by others, I might decide to punish the second lowest etc. The point is that my punishment behavior may be a poor indicator of my preference because it is confounded by these strategic considerations.

A better way to identify a subject's true preferences or motives for punishment would

be to look at behavior in a different network where such coordination and free riding problems are absent. Such an exercise was done by Carpenter et al. [4] where they look at four-person public goods games with punishments under a variety of network architectures. The best of these architectures for our purposes, however, is what they call the Direct Circle where subjects are arranged in a circle where Subject 1 can observe and punish Subjects 2, Subject 2 can observe and punish Subject 3, Subject 3 can observe and punish Subject 4 and finally Subject 4 can observe and punish Subject 1. In this case, each person has the sole responsibility to observe and punish just one person so there is no free riding or coordination problems and therefore the motivation for punishing is clearer.

In their experiment subjects first play a standard Voluntary Contribution Game in two stages. In the first stage they contribute and are told their payoffs and the mean contribution of the others in their group. In the second stage, after observing their first stage information, they decide whether to punish and by how much. Punishment points directed at another subject reduce the target's payoff by 10% for every one point used by the subject.

There are a number of theories one can construct to explain punishment in such public good games. One theory, investigated by De Quervain et al. [6] is a norm based theory that says that people punish others if they violate a group contribution norm. This norm is set exogenously, possibly based on some commonly held theory of fairness. For Fehr and Gaechter [12] the norm is that those who fail to contribute the mean amount or more are candidates for punishment. According to the theory of blame, however, the standard used to determine punishments is not an exogenously defined norm but rather a personal and individual standard based on how the person himself would behave in the situation being examined.

For the public goods game, subjects who subscribe to a theory of blame, should punish anyone who contributed less than they did and never punish those who contributed more, whether or not those people contributed more than the mean amount.

Such different theories of punishment can easily be investigated using the Carpenter et al. [4] data, and running a simple random effects probit regression where the probability of punishment is a function of the difference between a subject's contribution and that of his target's, the difference between the target's contribution and the group's mean, and the mean itself. Let pc_i be the public good contribution of subject i, pc_{-i} be the public good contribution of the person, subject i observes (the target), and m be the mean of the group's contribution in a given period. More formally the regression run is:

$$\Pr(pun) = \alpha + \beta_1 \Delta\_other^{+} \beta_2 \Delta\_other^{-} + \beta_3(pc_{-i} - m) + \beta_4(m) + \epsilon_i$$

where $\Delta\_other^{+} := pc_i - pc_{-i}$ if $pc_i - pc_{-i} > 0$ and 0 otherwise; and $\Delta\_other^{-} := pc_i - pc_{-i}$

if $pc_i - pc_{-i} < 0$ and 0 otherwise. If the theory of blame is responsible for punishment behavior we would expect that coefficient β_1 would be positive and significant while all other coefficients should be insignificantly different from zero since all that should matter for punishment is whether a subject's contribution was more than the contribution of the person he monitored, which is captured by Δ_other^+. The regression results are presented in Table 2.

TABLE 2: Punishment behavior in Carpenter et al. [4]: Directed circle

	Coeff	Std Err	Z	$P > \|Z\|$
Δ_other^+	.153	.032	4.86	.000
Δ_other^-	.046	.032	1.45	.148
$(pc_{-i} - m)$	-.038	.033	-1.16	.246
mean	-.014	.033	-.44	.663
constant	-1.422	.636	-2.24	.025

$N = 240$, Wald $\chi^2(4) = 45.92$, $\Pr > \chi^2 = .0000$.
Robust z-statistics are reported in the third column (clustering at the subject level).

As we see in Table 2, consistent with our expectations, the only variable with a statistically significant coefficient is Δ_other^+. The positive coefficient means that a subject is more likely to be punished the further his contribution falls below that of the person who observes him. The relationship of the target's contribution to the mean is insignificant. The coefficient in front of the mean variable itself is negative but again insignificant. The negativity of the coefficient makes sense since, as the mean increases, being below it is less of a transgression.

This simple regression lends support for the theory of blame and against the idea that punishment behavior is determined by the observed behavior of a subject in relation to some exogenously determined norm. What matters is how much a subject's contributions differed from that of the person who is watching him—a personal norm.

Other experiments and other data sets could be explored to investigate our theory of blame. It is our conjecture, however, that if one were to do so one would find that our theory has considerable support which would indicate that people, when judging others, are more likely to impose their own personal norms defined by what they would do if they were in the situation that others find themselves in rather than using some exogenous one-size-fits-all norm devised by others.

7 Conclusions

In this paper we have proposed and tested a theory of kindness that is an essential ingredient to any theory of reciprocity. Simply put, our theory of kindness states that in judging whether player i has been kind or unkind to player j, player j would have to put himself in the strategic position of player i and ask himself how he would have acted under identical circumstances. If j would have acted in a worse manner than i acted, then we say that j does not blame i for his behavior. If, however, j would have been nicer than i was, then we say that "j blames i" for his actions (i's actions were blameworthy.) After presenting a formal definition of this concept we investigated how it can be applied to the analysis of a dynamic psychological extensive form games and developed the notion of a Sequential Blame Equilibrium.

Using a simple modified dictator game experiment, we demonstrated that our theory has a substantial amount of explanatory power especially when compared to other competing models. The theory was then used to explain the observed punishments in a public goods game with networks run by Carpenter et al. [4].

It is our feeling that the theory of blame we present has some features that are both intuitively and empirically appealing. First, as we have argued, theories of kindness or fairness should be endogenous or at least based on the individual tastes and preferences of the person making an assessment. One-size-fits-all theories that impose the same norm (i.e., equity) on all decision makers fail to capture the individual nature of kindness assessments.

Finally, it is important to note that we do not claim that our theory describes the behavior of all potential subjects. Clearly, a significant portion of the population may behave according to other theories.

References

[1] Pierpaolo Battigalli and Martin Dufwenberg. Dynamic psychological games. *Journal of Economic Theory*, 144(1):1–35, 2009.

[2] Sally Blount. When social outcomes aren't fair: The effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes*, 63(2):131–144, 1995.

[3] Gary E. Bolton and Axel Ockenfels. Erc: A theory of equity, reciprocity, and competition. *American Economic Review*, 90(1):166–193, 2000.

[4] Jeffrey Carpenter, Shachar Kariv, and Andrew Schotter. Network architecture and mutual monitoring in public goods experiments. *Review of Economic Design*, 2012, forthcoming.

[5] Gary Charness and Matthew Rabin. Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117(3):817–869, 2002.

[6] Dominique J.-F. De Quervain, Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck, and Ernst Fehr. The neural basis of altruistic punishment. *Science*, 305(5688):1254–1258, 2004.

[7] Martin Dufwenberg and Georg Kirchsteiger. A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2):268–298, 2004.

[8] Armin Falk, Ernst Fehr, and Urs Fischbacher. On the nature of fair behavior. *Economic Inquiry*, 41(1):20–26, 2003.

[9] Armin Falk, Ernst Fehr, and Urs Fischbacher. Testing theories of fairness-intentions matter. *Games and Economic Behavior*, 62(1):287–303, 2008.

[10] Armin Falk, Ernst Fehr, and Christian Zehnder. Fairness perceptions and reservation wages—the behavioral effects of minimum wage laws. *Quarterly Journal of Economics*, 121(4):1347–1381, 2006.

[11] Armin Falk and Urs Fischbacher. A theory of reciprocity. *Games and Economic Behavior*, 54(2):293–315, 2006.

[12] Ernst Fehr and Simon Gächter. Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4):980–994, 2000.

[13] Ernst Fehr and Klaus M Schmidt. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3):817–868, 1999.

[14] John Geanakoplos, David Pearce, and Ennio Stacchetti. Psychological games and sequential rationality. *Games and Economic Behavior*, 1:60–79, 1989.

[15] Faruk Gul and Wolfgang Pesendorfer. Interdependent preference models as a theory of intentions. Princeton University, mimeographed, 2011.

[16] David K. Levine. Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics*, 1:593–622, 1998.

[17] Matthew Rabin. Incorporating fairness into game theory and economics. *American Economic Review*, 83(5):1281–1302, 1993.

[18] Andrew Schotter. *Free Market Economics: A Critical Appraisal*. Blackwell Publishers, New York, NY, 2nd edition, 1990.

[19] Joel Sobel. Interdependent preferences and reciprocity. *Journal of Economic Literature*, 43(2):392–436, 2005.

Appendices

Appendix A: Experimental Instructions

This is an experiment in economic decision-making. If you make good decisions you may be able to earn a good payment, which will be given to you at the end of the experiment.

Experimental Procedures

The experiment is composed of two stages. We will hand out instructions for Stage 2 after we are finished with Stage 1.

Stage 1

In Stage 1 of the experiment you will be given 10 tokens. Your task will be to divide this 10 tokens between you and an anonymous other person in this room. That means you will be asked to state how much of the 10 tokens you want to give to the other person, and therefore how much you would retain for yourself.

After you have made your choice, you will be randomly divided into two equally sized groups called *Senders* and *Receivers* and you will be matched in pairs. If you are chosen to be a *Sender* your payoff will be equal to what you have decided to keep for yourself of the 10 tokens. If you are chosen to be a *Receiver*, your payoff will be the amount given to you by the *Sender* you matched with.

You will not be told your payoff from Stage 1 until after the entire experiment is completed.

Note: The number you will be asked to enter in the screen will be the amount you want to give to another anonymous person in the room paired with you.

Stage 2

In Stage 2 of the experiment you will be randomly matched with another subject in the lab. This random matching will be *independent* of the one performed in Stage 1 so there is very little chance you will be matched with the same person. After you are matched you will be offered, as a *Receiver*, the amount your matched pair member sent as a *Sender* in Stage 1. In other words, if your matched pair member was a person who gave 3 tokens in Stage 1, you will be offered 3 tokens in Stage 2 as your payoff and your pair member will get 7.

Although you cannot change the amount you are given, you have the option to reduce the amount your match pair member keeps for himself/herself by deciding whether to reduce his/her payoff by 1 token. You can do this is a slightly indirect manner. Rather than stating whether you want to reduce your match pair member's tokens or not after you see the amount given to you, we will ask you **before** you see the offer to state an amount below which you will decide to reduce the pair member's payoff by one token. Call this

number your cutoff and assume for illustrative purposes that you entered a cutoff of 4 into the computer. If your matched pair member gave you 3 tokens and kept 7 tokens for himself/herself, since 3 is less than cutoff of 4, the computer will reduce your matched pair member's payoff by 1 token from 7 to 6. If the other person gave you 5 tokens and kept 5 tokens, his/her tokens will not be reduced and will stay 5.

Your final payoff will be determined as follow. After you have made your choice, the computer will randomly determine whether you are a *Sender* or a *Receiver*. Clearly, if you are a *Receiver* then your match pair member is a *Sender* and vice versa. If you are chosen to be a *Sender* your payoff will be equal to what you have decided to keep for yourself of the 10 tokens in Stage 1 minus any 1-token reduction by your pair member. If your math pair member decided not to take a token away, you will receive your payoff undiminished. Otherwise, you will receive one token less than what you kept for yourself. If you are chosen to be a *Receiver*, your payoff will be the amount given to you by your match.

Final Payoff Your final payoff from the two stages of the experiment will simply be the sum of your payoffs in each Stage. You will be paid $1 for each token you have at the end of the experiment.

Appendix B: Equilibrium Characterization of the Example

We can provide a similar analysis for the entire space of b_1, b_2. We will do this by determining player 1 and player 2's optimal behavior as we sweep b_1 and b_2 over their ranges. Once we determine players' optimal behavior independently, we can characterize the equilibrium behavior for the entire space of b_1, b_2.

Let us start with player 2. Player 2 is called upon to move at history (R). At the history (R), whatever his prior belief was, player 2 updates it to $\hat{\sigma}_{21}^{(\text{R})}(\varnothing) = \text{R}$. When player 2 is at (R), whether he will choose the action A or the action B, depends on what he would do when he puts himself in player 1's position since that will determine his blame at R. This, in turn depends on what he thinks player 1 believes he will do. There are two beliefs to consider: $\hat{\sigma}_{212}(\text{R}) = \text{A}$ and $\hat{\sigma}_{212}(\text{R}) = \text{B}$. Let us examine them one at a time.

First assume that $\hat{\sigma}_{212}(\text{R}) = \text{A}$. In this case, for any $b_2 \geq 0$, the best response of player 2 in player 1's position is L (i.e. $\sigma_{21}(\varnothing) = \text{L}$). This leads to a blame of $\delta_2^{(\text{R})} = 4$ if player 1 chooses R. In order for this scenario to be sustained as an equilibrium we need to have $\sigma_2(\text{R}) = \text{A}$ since $\hat{\sigma}_{212}(\text{R}) = \text{A}$. This is true only if player 2 blames player 1 enough that his payoff by choosing A is greater than that of choosing B, or equivalently $0 \geq 2 + \big(b_2 - f(4)\big)5$. This implies $b_2 \leq f(4) - 2/5$. So, at the history (R), for $b_2 \in [0, f(4) - 2/5]$, given the belief $\hat{\sigma}_{212}(\text{R}) = \text{A}$, player 2's optimal strategies are $\sigma_2(\text{R}) = \text{A}$ and $\sigma_{21}(\varnothing) = \text{L}$. If $b_2 > f(4) - 2/5$, the belief $\hat{\sigma}_{212}(\text{R}) = \text{A}$ is not consistent with the optimal strategy of player 2 since if $b_2 > f(4) - 2/5$ and $\hat{\sigma}_{212}(\text{R}) = \text{A}$ then player 2 would care sufficiently much about player 1's material payoff so as to choose B at history (R). This violates equilibrium consistency.

Second, let us assume that $\hat{\sigma}_{212}(\text{R}) = \text{B}$. Unlike the case above, best response of player 2 in player 1's position depends on b_2. Given the belief $\hat{\sigma}_{212}(\text{R}) = \text{B}$, player 2's utility in player 1's position, u_{21}, from action L is $4 + 4b_2$, and from action R it is $5 + 2b_2$. Therefore, for any $b_2 \leq 1/2$ it is optimal for him to play R, i.e. $\sigma_{21}(\varnothing) = \text{R}$. Since player 2 would have chosen R in player 1's position, he does not blame player 1 and optimally plays $\sigma_2(\text{R}) = \text{B}$.

When we look at the case $b_2 \geq 1/2$, player 2 cares so much about the material payoff of player 1 that his best response to the belief $\hat{\sigma}_{212}(\text{R}) = \text{B}$ in player 1's position is $\sigma_{21}(\varnothing) = \text{L}$. So, if player 2 were in player 1's position he would have given a payoff of 4 to the player in his position (to himself) but he believes player 1 expects to give him a payoff of 2 causing a blame of 2. Note that this scenario can be sustained as an equilibrium only if $\sigma_2(\text{R}) = \text{B}$ i.e. if $2 + 5\big(b_2 - f(2)\big) \geq 0$ or $b_2 \geq f(2) - 2/5$. As a result, in an equilibrium, at the history (R), given the belief $\hat{\sigma}_{212}(\text{R}) = \text{B}$,

- for any $b_2 \geq \max\{1/2, f(2) - 2/5\}$, player 2's optimal strategies are $\sigma_2(\text{R}) = \text{B}$ and $\sigma_{21}(\varnothing) = \text{L}$, and
- for any $b_2 \leq 1/2$, player 2's optimal strategies are $\sigma_2(\text{R}) = \text{B}$ and $\sigma_{21}(\varnothing) = \text{R}$.

The analysis of player 1's optimal action is relatively simple and can be summarized

as follows. If $b_1 \leq 1/2$ player 1 finds it optimal to play R as a best response to $\sigma_2(\text{R}) = \text{B}$; otherwise, if $\sigma_2(\text{R}) = \text{A}$, he plays L. If $b_1 \geq 1/2$, it is always optimal to play L for player 1. To demonstrate this first observe that for all b_1, if player 1 were in player 2's position, he would optimally play $\sigma_{12}(\text{R}) = \text{B}$. (Remember that player 1 does not blame himself for his own actions so if he arrives at history R in 2's position, he simply compares this payoff from action A, which is 0, to the payoff from action B, which is positive for all b_1. There are now two cases to consider: $\hat{\sigma}_{121}(\varnothing) = \text{L}$, and $\hat{\sigma}_{121}(\varnothing) = \text{R}$.

When $\hat{\sigma}_{121}(\varnothing) = \text{L}$ player 1's expected material payoff is 4 regardless of what he believes player 2's action is at the history (R). Similarly, when he puts himself in player 2's position, regardless of what he would do his expected payoff would be 4. Therefore, in this case, player 2 will not blame player 1. When, however, $\hat{\sigma}_{121}(\varnothing) = \text{R}$ and $\hat{\sigma}_{12}(\text{R}) = \text{A}$—i.e. player 1 believes that player 2 will play A even though he thinks player 1 plays R—then player 1 blames player 2 by $\delta_2 = 5$. But since $f(5) \leq 1$ by assumption, player 1 does not find it optimal to play A despite his blame. So, we do not observe this as an equilibrium. As a result, we simply observe that if $b_1 \leq 1/2$ player 1 finds it optimal to play R whenever $\sigma_2(\text{R}) = \text{B}$, otherwise if $\sigma_2(\text{R}) = \text{A}$, he plays L. If $b_1 \geq 1/2$, it is always optimal to play L for player 1.

Now that we have determined each player's optimal behavior consistent with equilibrium for all values of b_1 and b_2, we can classify equilibrium behavior as is done in 2. In what follows we will list the equilibria as $\Big[\big(\sigma_1(\varnothing), \sigma_{12}(\text{R})\big); \big(\sigma_2(\text{R}), \sigma_{21}(\varnothing)\big)\Big]$. We omit the beliefs since $\sigma_i = \hat{\sigma}_{ji} = \hat{\sigma}_{iji}$.

I: $\big[$(R,B);(B,L)$\big]$; II: $\big[$(L,B);(B,L)$\big]$; III: $\big[$(L,B);(A,L)$\big]$; IV: $\big[$(R,B);(B,R)$\big]$; V: $\big[$(L,B);(B,R)$\big]$.