Reciprocity-Induced Cooperation

By

Vincy Fon and Francesco Parisi

Abstract: The concept of reciprocity gains importance where there is no external authority to enforce agreements. Many legal systems foster meta-rules of reciprocity to facilitate cooperative outcomes. This paper considers the role of reciprocity rules in various strategic environments. We start by considering the effect of reciprocity constraints in a classic prisoner’s dilemma with two symmetric parties and linear payoffs. We extend the analysis to continuous strategies and then further extend the basic model of reciprocity to (a) asymmetric players, and (b) non-linear payoff functions. Then we examine the welfare properties of the reciprocity-induced equilibrium. In many game-theoretic situations, reciprocity facilitates the achievement of cooperative outcomes. Yet the reciprocity-induced equilibrium is not always socially optimal.

There are three important dimensions of reciprocity, namely the issues of emergence, enforcement, and effects of reciprocity. The first dimension considers the question of how norms of reciprocity emerge. Evolutionary psychologists have hypothesized that there is a behavioral foundation of reciprocity, and that humans have evolved mental algorithms for identifying and punishing defectors (see, e.g., HOFFMAN, McCABE AND SMITH [1998]). There is a considerable body of literature on the subject, as the issue of reciprocal behavior between individuals has become a matter of increasing interest in experimental and evolutionary economics (see, e.g. TRIVERS [1971]; BERG, DICKHAUT AND McCABE [1995]; SKYRMS [1996]). Further, contrary to the predictions based on assumptions of self-interested utility maximization, a significant body of literature has accumulated evidence that suggests that individuals are motivated by concerns of fairness and reciprocity (FEHR AND SCHMIDT [2000]).¹ The second dimension

¹ FEHR AND SCHMIDT [2000] note that the standard economics assumptions still work in the vast majority of cases in making predictions about behavior. Nonetheless, the empirical evidence on reciprocity and trust is robust enough that it cannot be dismissed as an aberration, and must be taken into account while modeling certain kinds of behavior, in both repeat and single-shot games.
considered in the literature concerns the problem of enforcement of reciprocity norms. The decentralized enforcement of reciprocity norms creates a second-order collective action problem (ELLIKSON [2001]). This problem results from the fact that the enforcement of reciprocity norms requires the punishment of non reciprocators. Such punishment creates a public benefit for the collectivity at large, while imposing a private cost for the individual enforcer. Thus, the second-order collective action problem adds a second dimension of strategic behavior for the sustainability of reciprocity.

In this paper we consider the third question of the actual effects of reciprocity on the solution of cooperation problems. For the purpose of analytical clarity, we separate the study of the effects of reciprocity from the study of the sources and enforcement of reciprocity studied in the existing literature. In doing so, we assume the existence of exogenous reciprocity constraints, such as the existence of some social or legal mechanisms that ensure an effective enforcement of reciprocity. This assumption is consistent with the recent findings of experimental and behavioral economics that show that humans exhibit a strong tendency towards reciprocity, as revealed by their tendency to cooperate and reciprocate, at a personal cost, even when there are no plausible future benefits from so behaving (GINTIS [2000, 262]).

The paper utilizes a game-theoretic model of reciprocity to revisit the conventional wisdom in which reciprocity facilitates the emergence of efficient cooperative outcomes. We first set a standard prisoner’s dilemma problem with discrete strategies and symmetric payoffs, looking at the effect of a reciprocity constraint on the parties’ choice. We then investigate whether the properties evidenced in the standard case
extend to the broader array of cooperation problems, where the parties are faced with continuous strategies, asymmetric payoffs, and nonlinear cost functions.

The economic model of reciprocity unveils the limits of the general intuition in which binding reciprocity constraints provide a viable solution to prisoner’s dilemma problems.

1 Reciprocity and Cooperation

The law and economics and constitutional political economy literature has given brief consideration to the effects of reciprocity constraints in Prisoners’ Dilemma games, in the context of constitutional design (Buchanan [1978]) and customary law formation (Parisi, [1995] and [2000]). In those studies, most of the illustrations of reciprocity in cooperation problems involve players with symmetric payoffs and discrete strategies. Our study extends the analysis beyond such initial conditions, considering the impact of exogenous reciprocity constraints on prisoner’s dilemma games with asymmetric linear and non-linear payoffs and continuous strategies.

In this paper, we consider the effects of reciprocity in the limit case of one-shot games. This is different from the form of reciprocity studied by Axelrod [1984] and Trivers [1971]. This extension is consistent with recent experimental evidence finding support for cooperation under full information even in single play experiments (McCabe, Rassenti and Smith [1996]).

We compare the Nash equilibria in such games to those obtainable if the parties are subject to a reciprocity constraint. We call such equilibrium “reciprocity-induced equilibrium.” Finally we compare the reciprocity-induced equilibrium to the ideal social
optimum, in order to verify the conditions under which a reciprocity constraint would generate socially efficient levels of cooperation.

1.1 A Taxonomy of Reciprocity

Consider these oft-quoted colloquialisms: ‘Do unto others as you would have done unto you.’ ‘If you scratch my back, I’ll scratch yours.’ ‘Tit for tat.’ These pieces of collective wisdom come to mind when many think of reciprocity. In the existing economic literature, the term reciprocity refers to a fairly broad range of concepts. PARISI [2000a] provides a taxonomy of reciprocity, which includes the concept of structural reciprocity, stochastic reciprocity, and induced reciprocity.

In an ideal world of structural reciprocity, the parties’ incentives are aligned perfectly, such that neither has an incentive to defect unilaterally. Such reciprocity exists, for example, in a world where the players are in a pure common interest game (SCHELLING [1980]). Cooperative outcomes in such cases generally do not require exogenous reciprocity constraints and do not rely on external enforcement mechanisms, such as a legal system, or a threat of retaliatory behavior, as the players will choose to cooperate in any case. When such perfect alignment of interests does not exist, which is more often than not, there is an incentive for opportunistic behavior. Here, reciprocity constraints may play an important role in two distinct settings.

One such setting, the stochastic reciprocity, considers reciprocity in the context of iterated games. Here reciprocity can be an effective constraint if the parties alternate roles in their iterated interactions. Much of the existing literature utilizes this concept of reciprocity. For example, AXELROD [1984], found that a tit-for-tat strategy outperformed
a ‘rational’ self-interested strategy in an iterated game; he suggests that cooperation is far more common and “normal” than expected, and the standard economic model of self-interest is not necessarily the best model for all circumstances. Joyce Berg and her co-authors found that reciprocity was an essential element of human behavior, and held that this accounted for trust extended to an anonymous counterpart (BERG, DICKHAUT AND McCABE [1995]). In the case of stochastic reciprocity, the players must undertake repeated transactions in a game. The source of randomness in the game could be role reversibility of the players, or a random distribution of asymmetric payoffs to the players over the repeated plays of the game. Stochastic reciprocity requires a pre-commitment by each player to a meta-strategy for the entire duration of the game. In such situations, cooperative strategies are likely to dominate, if there is a relatively high probability of future interaction, a relatively low discount rate of the players, and a sufficiently large fraction of reciprocators in the population. A higher probability of future interaction is more likely to increase the expected payoff from cooperation; a lower discount rate means that the future payoff is valued relatively highly in present value terms. Thus, both will increase the present value of cooperation.

In another distinct setting, reciprocity involves constraints that are exogenously induced on the parties’ strategies. This notion of reciprocity, which we term induced reciprocity, is the focus of the present paper. In this category, reciprocity is induced by

2 PARISI [2000b] terms this “silver reciprocity”

3 Stochastic reciprocity is similar to KEOHANE’S [1986] ‘diffuse reciprocity’, where an agent cooperates, not in expectation of a specific reciprocal reward, but some general reciprocal return in the future. This definition is readily applicable in international law, on which Keohane’s work is directly focused. The players are nations, who engage in repeated interactions with each other. Thus, the condition of a high probability of future interaction is fulfilled. It is reasonable to assume that states have low discount rates, since, in general, nations do have long lives and therefore long time horizons.
some external constraint. Real life examples of such constraint include biological and psychological tendencies to reciprocate, social norms, and institutional or legal constraints. Some of these reciprocity constraints rely on first-party enforcement (e.g., sense of guilt), while others rely on second-party or third-party enforcement (e.g., reputational costs and social or legal sanctions). In this paper we consider the effect of induced reciprocity in the case of one-shot games.\(^4\) The actual enforcement of reciprocity is thus assumed, be it by first, second, or third parties.

1.2 Reciprocity-Induced Strategies and Reciprocity-Induced Equilibrium

Below we consider the effect of exogenous reciprocity, assuming the existence of some of the above mentioned mechanisms to ensure an effective constraint on parties’ choices. This assumption may closely approximate reality in some environments. For example, observed human behavior reveals strong patterns of reciprocal behavior in one-shot as much as in repeated interactions. Evolutionary theorists have suggested that the tendency for humans to adopt norms of reciprocity in one-shot games may have evolved as a fitness-enhancing trait (GINTIS [2000, 262]). Likewise, in some other areas of human interaction, reciprocity is induced by social or legal constraints. For example, in customary law settings, norms of reciprocity acquire a fundamental importance (PARISI [1998] and [2000b]). Along similar lines, in public international law metarules of reciprocity are widely recognized and enforced.\(^5\)

\(^4\) This will help isolate the direct effect of the induced reciprocity constraint from the additional incentives created in a repeat interaction framework. The effects of a reciprocity constraint explored in this paper would obviously also apply to the case of repeat games. The tendency to cooperate in such settings would be reinforced by the pursuit of gains from future cooperation by the individual players.

\(^5\) See, e.g. Art. 21(1)(b) of 1969 Vienna Convention on the Law of Treaties. Reciprocity has effectively become a meta-rule for public international law, an essential principle for the practice of sovereign nations. See PARISI AND GHEI [2002] and PARISI AND SEVCENKO [2002]. See, however, Goldsmith and Posner
In this paper, we will build on these premises, assuming the existence of a reciprocity constraint and considering the effects of induced reciprocity constraints on the parties’ level of cooperation. We consider a simple form of reciprocity that allows the party who prefers a higher level of cooperation to revert to the lesser amount of cooperation chosen by the other party. That is to say, reciprocity rules only require parties to reciprocate the level of cooperation chosen by the other party, making such level of cooperation binding between the parties under our reciprocity rule. Put differently, if the parties’ reciprocity-induced levels of cooperation yield different levels of cooperation for the two players, the lesser of the two amounts of cooperation becomes the mutually binding level of cooperation for both individuals. This corresponds to a golden rule of reciprocity, which successfully binds each player’s strategy to that of his opponent (PARISI [2000a]). Automatic reciprocity of this type creates a symmetric constraint in the players’ strategies. Thus, for example, if one party’s level of reciprocity-induced cooperation equals "and the other party’s level equals $ and " < $, then the reciprocity-induced equilibrium will be (",")

We should note that the main virtue of this notion of reciprocity is that it encourages the truthful expression of preferences for both parties. When parties choose strategies under such reciprocity constraints they would do so without engaging in preference falsification, since neither party would have an incentive to withhold cooperation below the privately optimal level of reciprocal cooperation. As a consequence, the reciprocity mechanisms would trigger a level of cooperation equal to the level desired by the least cooperative player. In spite of such reversion towards the

[1999] and [2000], who express some skepticism (on empirical grounds) about reciprocity explanations of international cooperation involving more than two states.
least amount of voluntary cooperation, this mechanism yields a level of cooperation that always improves upon the Nash level of cooperation, absent reciprocity.  

Our results show that the reciprocity-induced outcomes always constitute a Pareto improvement with respect to the payoffs obtainable in the (unconstrained) Nash equilibrium. This may explain why parties may agree *ex ante* to a metarule of reciprocity, guaranteeing higher expected payoffs for both players.

2  *A Model of Reciprocity*

We start with a classic prisoner’s dilemma with two symmetric parties, as shown in Figure 1. Strategies available to both parties, \( s_1 \) and \( s_2 \), are either 0 or 1. We will refer to the strategy \( s_i = 0 \) as the no cooperation strategy, and \( s_i = 1 \) as the full cooperation strategy. As the convention goes, the first and second entries in each cell of the matrix refer to the payoff to the row player and the column player, respectively.

*Figure 1*

<table>
<thead>
<tr>
<th></th>
<th>( s_2 = 1 )</th>
<th>( s_2 = 0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( s_1 = 1 )</td>
<td>(-a + b, -a + b)</td>
<td>(-a, b)</td>
</tr>
<tr>
<td>( s_1 = 0 )</td>
<td>(b, -a)</td>
<td>(0, 0)</td>
</tr>
</tbody>
</table>

---

6 A hypothetical example often used in law school concerns the design of a contractual mechanism to induce parties to reveal truthful preference, when contributing to a public good. In order to avoid the traditional problems involved in gathering voluntary contribution to the provision of a public good, commitments to voluntary participation could be made contingent to the participation of all the other members, in such a way that each member of the community would be bound to contribute an amount equal to the lowest contribution received from the other members. In this way, each neighbor could truthfully reveal the highest amount that they would be willing to contribute, knowing that they would only be bound to pay the amount chosen by the least generous contributor in the group.
In a prisoner’s dilemma, both players face dominant defection strategies. Whatever the strategy of one party, the other party prefers no cooperation, or, \( s_i = 0 \). Hence, \(-a + b < b\) and \(-a < 0\) should hold. That is, the assumption \( a > 0 \) will ensure that no cooperation is the dominant strategy for either party. Meanwhile, the other party’s cooperation always produces a benefit for the receiving party, whatever its strategy. That is, \(-a + b > -a\) and \(b > 0\) should hold. This means that assuming that \( b > 0 \) will assure that a party is better off with the other party’s full cooperation. Lastly, to create the prisoner’s dilemma, the payoff to joint cooperation \((s_1 = 1, s_2 = 1)\) is greater than the payoff to joint defection \((s_1 = 0, s_2 = 0)\). Hence \(-a + b > 0\) should be assumed. To sum up, in order to capture the essence of the prisoner’s dilemma where the Nash equilibrium is inefficient, the following assumptions are imposed on the parameters throughout the paper: \(0 < a < b\).\(^7\)

In a prisoner’s dilemma with discrete strategies where the parties’ strategies are subject to a reciprocity constraint, the full cooperation strategies are adopted by the parties and the mutual cooperation outcome obtains in equilibrium (PARISI [1995]). In such equilibrium there is a definite improvement over the alternative Nash equilibrium. Indeed the reciprocity-induced equilibrium coincides with the social optimum.

We extend the case of discrete variables, where cooperation is either nil or full, to the case of continuous variables with infinitely many degrees of cooperation. In particular, we assume that \(s_1, s_2 \in [0, 1]\) and will refer to \(s_i\) as the degree of cooperation or the effort spent on cooperation. In order to investigate systematically the role of reciprocity in the prisoner’s dilemma problem, we first analyze situations with identical

\(^7\) Later, when we extend our analysis to the case where the two parties are asymmetric, a similar assumption will be imposed to preserve the prisoner’s dilemma problem for both parties.
players facing symmetric payoff functions. In Section 2.1, we consider the case of linear payoff functions and in Section 2.2 we extend the analysis to nonlinear cases. In Section 2.3 we further extend our analysis to the more general case where the two parties face different nonlinear payoff functions. In each of these cases, we study the impact of exogenous reciprocity constraints on the equilibrium of the game, comparing the reciprocity-induced outcome with the alternative Nash outcome (in the absence of reciprocity) and the ideal social optimum.

2.1 Symmetric Linear Payoffs

We begin our analysis of reciprocity by considering a strategic problem where the payoff functions are linear and symmetric for both players. Let the payoff function for party 1 be:

$$P_1(s_1, s_2) = -as_1 + bs_2.$$ 

Note that: $$P_1(1, 1) = -a + b,$$ $$P_1(1, 0) = -a,$$ $$P_1(0, 1) = b,$$ and $$P_1(0, 0) = 0.$$ Likewise, assume that the payoff function for party 2 is:

$$P_2(s_1, s_2) = -as_2 + bs_1.$$ 

Assuming $$0 < a < b,$$ we have a continuous extension of the prisoner’s dilemma considered above. It can readily be seen that the limits of the continuous variable $$s_i$$ generate the classic prisoner’s dilemma with discrete strategies, as illustrated in Figure 1.
(a) *Reciprocity-Induced Equilibrium*. We shall now consider the impact of a reciprocity constraint on the equilibrium strategies of the two players. In our world of induced reciprocity, each party knows that, within the limits of mutually agreeable level of cooperation, the other party will reciprocate. Thus the least level of cooperation chosen by the players will become the *de facto* level of cooperation for all. The payoff function to party 1, given reciprocity, thus becomes:

\[
\pi_1(s_1) = P_1(s_1, s_2)
\]

where

\[
\frac{d\pi_1}{ds_1} = \frac{\partial P_1}{\partial s_1} + \frac{\partial P_1}{\partial s_2} \cdot \frac{ds_2}{ds_1} = -a + b > 0.
\]

We should note that party 1’s payoff under reciprocity becomes an increasing function of his own level of cooperation \(s_1\), and full cooperation \(s_1 = 1\) becomes a dominant strategy for party 1 under reciprocity. Likewise, party 2’s payoff under reciprocity becomes an increasing function of his own level of cooperation \(s_2\), rendering full cooperation, \(s_2 = 1\), a dominant strategy for party 2 as well. Clearly, the reciprocity-induced equilibrium strategies are \((s_1^*, s_2^*) = (1, 1)\), and the reciprocity-induced equilibrium payoffs for the two parties become \(\pi_1^*(s_1^*) = \pi_2^*(s_2^*) = b - a > 0\). This allows us to contrast the reciprocity-induced outcomes to the Nash equilibrium which is obtained in the absence of reciprocity. Recall that, absent any constraint, the Nash equilibrium strategies were \(s_1 = s_2 = 0\) and the payoffs for the two parties were \(P_1(0,0) = P_2(0,0) = 0\). Thus, the reciprocity-induced payoff for each party represents a substantial improvement over the Nash equilibrium payoff.

---

8 We use superscript * to denote the reciprocity-induced strategies or outcomes, and a bar above the variables or the functions to denote the socially optimal strategies or outcomes.
(b) Social Optimum. We shall now consider the efficiency of the reciprocity-induced outcome in light of the Benthamite or Kaldor-Hicks criterion of welfare. Socially optimal strategies are those that maximize the aggregate payoff for the 2 parties. The social optimization problem thus becomes:

$$\max_{s_1, s_2} P(s_1, s_2) = \bar{P}_1(s_1, s_2) + \bar{P}_2(s_1, s_2) = (-as_1 + bs_2) + (-as_2 + bs_1) \quad s.t. \quad 0 \leq s_1 \leq 1, \quad 0 \leq s_2 \leq 1.$$ 

Since $0 < a < b$, the social welfare function $\bar{P}$ is an increasing function in both $s_1$ and $s_2$. That is to say, an increase in level of cooperation from either party increases the total payoff available to society. The social optimum then requires that both parties engage in full cooperation, with both strategies equal to 1, i.e. $(\bar{s}_1 = 1, \bar{s}_2 = 1)$. It can readily be seen that, in the present case of linear and symmetric payoffs, the reciprocity-induced equilibrium coincides with the social optimum. The presence of a reciprocity constraint solves the prisoner’s dilemma problem, allowing parties to undertake socially optimal strategies in equilibrium.

In this section, we found that the reciprocity-induced equilibrium yields a clear improvement over the alternative Nash equilibrium and matches the social optimum. These results should not be surprising since they amount to the most straightforward extensions of the discrete prisoner’s dilemma case under reciprocity in a continuous strategy space.

In order to obtain a better understanding of the parties’ incentives under a rule of reciprocity, we should consider the economic interpretation of the payoff function.
\( P_1(s_1, s_2) = -as_1 + bs_2 \). From the viewpoint of party 1, \( a \) is the marginal cost of his own cooperation effort while \( b \) is the marginal benefit derived from the other party’s cooperation effort. Under reciprocity, we thus have \( \pi_1(s_1) = -as_1 + bs_1 \), where \( a \) and \( b \) become the marginal cost and the marginal benefit of the party’s own cooperation effort.

Above, we considered the case of linear payoff functions, which was equivalent to assuming that both marginal cost and marginal benefit of increasing cooperation were constant. In the next two sections, we shall relax the linearity and the symmetry assumptions in turn to investigate what happens with increasing marginal cost and constant marginal benefit.

2.2 Symmetric Nonlinear Payoffs

We start by investigating what happens with increasing marginal cost and constant marginal benefit under symmetry. We shall see that the consideration of nonlinear payoffs is interesting in the present context, because it generates interior solutions that would not emerge from the linear case. Qualitatively similar interior solutions can be generated assuming nonlinearities in the cost or benefit functions (e.g. increasing marginal cost of cooperation or decreasing marginal benefit of cooperation, or both). For ease of exposition, in this paper we consider the case of increasing marginal cost and constant marginal benefit of cooperation.

Assume that the payoff functions for the two parties are:

\[ P_1(s_1, s_2) = -as_1^2 + bs_2 \quad \text{and} \quad P_2(s_1, s_2) = -as_2^2 + bs_1. \]
Again, we assume $0 < a < b$ to create the prisoner’s dilemma. In this case, the reciprocity-induced payoff function for party 1 is equal to $\pi_1(s_1) = -as_1^2 + bs_1$. The marginal cost of an increase in the level of cooperation for party 1 is $MC_1 = 2as_1$, and the corresponding marginal benefit for increased cooperation under reciprocity is $MB_1^R = b$. It is logical that the parties’ willingness to cooperate under reciprocity depends on whether the marginal benefit of reciprocal cooperation exceeds the rising marginal cost. Depending on whether $2a$, the marginal cost at full levels of cooperation, $s_i = 1$ exceeds $b$, the marginal benefit of cooperation, we need to distinguish two groups of cases. The first group encompasses situations where both parties, if subject to a reciprocity constraint, would choose a full level of cooperation. The second group of cases encompasses situations where, in spite of the existence of a binding reciprocity constraint, at least one party would choose less than full cooperation. We shall examine these two groups of cases in turn.

2.2.1 The case of full cooperation: $b \geq 2a$

In this case, $b$, the marginal benefit of cooperation is greater than or equal to $2a$, the marginal cost at full level of cooperation. Here, as will be shown below, both parties, if subject to a reciprocity constraint, would choose a full level of cooperation.

(a) Reciprocity-Induced Equilibrium. Since

$$\frac{d\pi_1}{ds_1} = -2as_1 + b \geq -2as_1 + 2a = 2a(1 - s_i) > 0 \text{ for all } s_i \in (0, 1), \pi_1 \text{ is an increasing}$$
function of $s_1$ for all $s_1$ in between 0 and 1. Hence $s_1 = 1$ is the dominant reciprocity-induced strategy for party 1.

With symmetric payoffs for party 2, $s_2 = 1$ will be chosen. Hence the reciprocity-induced equilibrium strategies are $(s_1^*, s_2^*) = (1, 1)$, and the corresponding reciprocity-induced equilibrium payoffs are $\pi_1^*(1) = \pi_2^*(1) = b - a > 0$. It is easy to see that the Nash equilibrium is $s_1 = 0, s_2 = 0$ and the payoffs under the Nash equilibrium for the two parties are $P_1(0, 0) = P_2(0, 0) = 0$. Hence once again, we see that reciprocity solves the prisoner’s dilemma and improves upon the Nash equilibrium.

(b) Social Optimum. The social problem for this case is:

$$\max_{s_1, s_2} P(s_1, s_2) = P_1(s_1, s_2) + P_2(s_1, s_2) = (-as_1^2 + bs_2) + (-as_2^2 + bs_1) \quad s.t. \quad 0 \leq s_1 \leq 1, \ 0 \leq s_2 \leq 1.$$ 

Since $\frac{\partial P}{\partial s_1} = -2as_1 + b > 0$, $P$ is an increasing function of both $s_1$ and $s_2$ when $b \geq 2a$.

Thus the social optimal strategies are $(\bar{s}_1 = 1, \bar{s}_2 = 1)$. Further, the social optimal payoffs are $P_1(1, 1) = P_2(1, 1) = b - a$ and the aggregate social optimal payoff is $\bar{P}(1, 1) = P_1(1, 1) + P_2(1, 1) = 2(b - a)$. These social optimal payoffs are identical to the reciprocity-induced equilibrium payoffs.

2.2.2 The case of partial cooperation: $b < 2a$

In this case, $b$, the marginal benefit of cooperation is less than $2a$, the marginal cost at full level of cooperation. Here, as shown below, in spite of the existence of a binding reciprocity constraint, at least one party would choose less than full cooperation.
(a) Reciprocity-Induced Equilibrium. When $b < 2a$, the payoff function with reciprocity attains maximum at $s_i = \frac{b}{2a}$ where $\frac{b}{2a}$ falls in between 0 and 1, with a maximal value of $\frac{b^2}{4a}$.

With symmetry, both parties choose the same strategies under reciprocity, $\frac{b}{2a}$. Hence the reciprocity-induced equilibrium strategies are: $(s_1^*, s_2^*) = \left( \frac{b}{2a}, \frac{b}{2a} \right)$. The reciprocity-induced equilibrium payoffs are $\pi_i^*(s_i^*) = \pi_i^*(s_j^*) = \frac{b^2}{4a} > 0$. Clearly, these payoffs are better than those obtained under Nash equilibrium prisoner’s dilemma, where both parties would face dominant defection strategies $s_i = 0$ and receives a payoff of $P_i(0, 0) = 0$.

(b) Social Optimum. We shall now turn to the social problem where $b < 2a$:

$$\max_{s_1, s_2} \overline{P}(s_1, s_2) = \overline{P}_1(s_1, s_2) + \overline{P}_2(s_1, s_2) = (-a s_1^2 + b s_2) + (-a s_2^2 + b s_1) \quad s.t. \quad 0 \leq s_1 \leq 1, 0 \leq s_2 \leq 1.$$

When the marginal cost at full cooperation exceeds the marginal benefit of reciprocal cooperation, $2a > b$, the social optimum does not require full cooperation: $(\overline{s}_1 = \frac{b}{2a}, \overline{s}_2 = \frac{b}{2a})$. The social optimal payoffs are $\overline{P}_1(\frac{b}{2a}, \frac{b}{2a}) = \overline{P}_2(\frac{b}{2a}, \frac{b}{2a}) = \frac{b^2}{4a}$. Again, these social optimal payoffs are identical to the reciprocity-induced equilibrium payoffs.
These results are quite interesting. Recall that in the case considered in Section 2.2.1 full cooperation would obtain in spite of the rising marginal cost, since the condition \( \frac{b}{2a} \geq 1 \) ensures that marginal benefit is greater than marginal cost in the relevant domain of feasible cooperation. In the case under examination, the parties continue to face symmetric payoff functions with increasing marginal cost to cooperation, but less than full cooperation is induced by the reciprocity constraint. The parties choose partial, rather than full, cooperation because, given \( \frac{b}{2a} < 1 \), the marginal cost of cooperation starts exceeding the benefit of reciprocal cooperation before the parties reach a full level of cooperation.

In this Section with symmetric nonlinear payoffs, the reciprocity-induced levels of cooperation always coincide with the social optimum. Partial, rather than full, cooperation can be privately and socially optimal. The possibility that the social optimum corresponds to a partial, rather than full, level of cooperation is due to the nonlinearities of the payoffs, since it would be inefficient to carry out cooperation beyond the point in which its marginal cost equals its marginal benefit. The reciprocity constraint allows the parties to obtain the highest achievable payoffs in equilibrium.

2.3 Asymmetric Nonlinear Payoffs

We now turn to the asymmetric cases where the parties face different marginal cost and benefit functions. We continue to assume that both parties face increasing marginal cost and constant marginal benefit. In particular, let the payoff function for party 1 be:
\[ P_1(s_1, s_2) = -as_1^2 + bs_2. \]

Also, let the payoff function for party 2 be:

\[ P_2(s_1, s_2) = -cs_2^2 + ds_1. \]

In this asymmetric setting, the prisoner’s dilemma requires that \( 0 < a < b \) and \( 0 < c < d \). Without loss of generality, we assume that party 1 is the “high-cost cooperator” (or, more precisely, the low net benefit cooperator), in the sense that \( \frac{b}{2a} \leq \frac{d}{2c} \). This follows from the fact that \( b \) and \( d \) are the reciprocity-induced marginal benefits for parties 1 and 2 respectively, while \( 2a \) and \( 2c \) are the marginal costs at full cooperation level \((s_i = 1)\) for the same parties. Before comparing the outcomes of such asymmetric case, we evaluate the different reciprocity-induced equilibria and the possible social optima.

\( (a) \) Reciprocity-Induced Equilibrium. In the case of asymmetric payoff functions, the desirable level of cooperation for the two players is likely to differ. In spite of the assurance of reciprocal cooperation, the parties may prefer different levels of cooperation. The strategy that one party may wish to adopt, assuming reciprocal cooperation, may not be matched by equal willingness from the other party. For this reason, we need to recall our definition of reciprocity-induced strategies (labeled \( s^* \)) and include an additional notation to characterize the individual agents’ maximum level of
agreeable cooperation (labeled \( s' \)). For party 1 with payoff function 
\[ P_1(s_1, s_2) = -as_1^2 + bs_2, \]
the desired level of cooperation, \( s'_1 \), assuming reciprocal cooperation, depends on whether the marginal cost at full cooperation exceeds the marginal benefit under reciprocity:

\[
\begin{align*}
\text{if } \frac{b}{2a} \geq 1, & \quad s'_1 = 1 \\
\text{if } \frac{b}{2a} < 1, & \quad s'_1 = \frac{b}{2a}
\end{align*}
\]

Likewise, for party 2 with payoff function \( P_2(s_1, s_2) = -cs_2^2 + ds_1 \), the desired level of cooperation, \( s'_2 \), depends on the relative magnitude of the marginal cost at full cooperation and the marginal benefit under reciprocity:

\[
\begin{align*}
\text{if } \frac{d}{2c} \geq 1, & \quad s'_2 = 1 \\
\text{if } \frac{d}{2c} < 1, & \quad s'_2 = \frac{d}{2c}
\end{align*}
\]

When \( 1 \leq \frac{b}{2a} \leq \frac{d}{2c} \), both parties choose full cooperation under reciprocity. Since the strategy each party chooses match the expected reciprocal strategy of the other party, full cooperation becomes the reciprocity-induced equilibrium strategies for both parties:
(s_1^* = 1, s_2^* = 1). As it will be seen below, such convergence of preferences will not likely be found in the case of partial cooperation.

Given the asymmetry, when \( \frac{b}{2a} < 1 \), the privately optimal values of cooperation may not coincide for the two players. Since party 1 is the high-cost cooperator, \( s_1' = \frac{b}{2a} \leq \frac{d}{2c} = s_2' \). Given our definition of reciprocity-induced cooperation, the two parties will only be bound to cooperate to the lesser level of cooperation desired by the two parties. The reciprocity-induced cooperation effort chosen by party 1, \( s_1' = \frac{b}{2a} \), thus becomes the binding strategy for both parties. Hence the reciprocity-induced equilibrium strategies for this asymmetric case are \( s_1^* = \frac{b}{2a}, s_2^* = \frac{b}{2a} \). Unlike the case of full cooperation examined above, the desired levels of cooperation for the two asymmetric players is unlikely to be equal (with the sole exception where the parties have asymmetric cost functions with identical marginal benefit-cost ratios).

We conclude the discussion on reciprocity-induced equilibrium by listing the equilibrium strategies and the corresponding equilibrium payoffs for different scenarios.

---

9 Whenever the privately optimal levels of reciprocal cooperation do not coincide for the two players, the reciprocity-induced equilibrium will only coincide with the preferred level for one of the parties. In our specific case, the cooperation choice of the high-cost cooperator would always dictate such equilibrium. It is easy to show that the equilibrium payoff for party 2 is \( \pi_2^*(s_2^*) = \frac{b(ad - bc)}{4a^2} > 0 \). If \( d/2c < 1 \), party 2 would have been happier if \( s_2' = d/2c \) were the reciprocal strategies chosen by both because \( \pi_2^*(s_2') < \pi_2^*(s_2^*) \). Likewise, if \( d/2c \geq 1 \), party 2 prefers the reciprocal strategy 1. In spite of these facts, party 2 is still better off under the reciprocity-induced equilibrium with a positive payoff, \( \pi_2^*(s_2^*) \), than under the Nash equilibrium where its payoff would have been zero.
1. If \( \frac{b}{2a} \leq \frac{d}{2c} \), then \((s_1^*, s_2^*) = (1, 1)\) with \(\pi_1^*(1) = b - a > 0\) and \(\pi_2^*(1) = d - c > 0\).

2. If \( \frac{b}{2a} < 1 \), then \((s_1^* = \frac{b}{2a}, s_2^* = \frac{b}{2a})\) with \(\pi_1^* (\frac{b}{2a}) = \frac{b^2}{4a} > 0\) and \\
\[
\pi_2^* (\frac{b}{2a}) = \frac{b(ad - bc)}{4a^2} > 0.
\]

The above results allow us to contrast the reciprocity-induced outcomes to the alternative Nash equilibrium. As shown in the previous symmetric cases, absent reciprocity, the Nash equilibrium strategies would be \(s_1 = s_2 = 0\) and the payoffs for the two parties would fall to \(P_1(0,0) = P_2(0,0) = 0\). Thus, in this case, the existence of a reciprocity constraint also leads to a substantial improvement over the Nash equilibrium payoff.

(b) Social Optimum. Next we turn our attention to the social optima. Similar to the previously examined cases, the social problem is:

\[
\max_{s_1, s_2} \bar{P}(s_1, s_2) = \bar{P}_1(s_1, s_2) + \bar{P}_2(s_1, s_2) = (-as_1^2 + bs_2) + (-cs_2^2 + ds_1) \\
s.t. \quad 0 \leq s_1 \leq 1, \quad 0 \leq s_2 \leq 1
\]

In order to find the social optimum, it is now important to compare the marginal cost of a strategy, say \(s_1\), with its social marginal benefit. That is, we need to compare the marginal cost of \(s_1\), \(MC_1 = 2as_1\), with its social marginal benefit, \(MB_1^S = d\), rather than comparing the marginal cost of \(s_1\), \(MC_1 = 2as_1\), with the reciprocity-induced marginal
benefit $MB_i^b = b$, as we did previously, when considering the private incentives.

Keeping this in mind, the different possible social optima are as follows:\(^{10}\)

1. If $\frac{d}{2a} < 1 \& \frac{b}{2c} < 1$, then $\bar{s}_1 = d < 1, \bar{s}_2 = b < 1$.

2. If $\frac{d}{2a} < 1 \leq \frac{b}{2c}$, then $\bar{s}_1 = d < 1, \bar{s}_2 = 1$.

3. If $\frac{b}{2c} < 1 \leq \frac{d}{2a}$, then $\bar{s}_1 = 1, \bar{s}_2 = \frac{b}{2c}$.

4. If $1 \leq \frac{d}{2a} \& 1 \leq \frac{b}{2c}$, then $\bar{s}_1 = 1, \bar{s}_2 = 1$.

It is noteworthy that, depending on the relative magnitude of the social marginal costs and benefits of cooperation, the social optimum may require partial cooperation from both parties, full cooperation from both parties, or a mixed combination of partial and full cooperation for the two parties. Given that alternative socially optimal strategies exist under different situations, the relationship between reciprocity-induced equilibrium and social optimum also varies, depending on the costs and benefits of cooperation for the parties. This result reveals the limits of reciprocity in inducing socially optimal strategies among heterogeneous players.

\(^{10}\) The payoffs for the individual players, as well as the aggregate payoffs for these social optima are listed in Appendix A.
3. Reciprocity and Socially Optimal Cooperation

The economic model of reciprocity verifies the general intuition according to which binding reciprocity provides a viable solution to prisoner’s dilemma problems. In the previous Section we have identified an important attribute of reciprocity in cooperation problems: reciprocity always induces the parties to adopt levels of cooperation higher than the alternative levels that they would otherwise adopt in a Nash equilibrium. We now explore the extent to which reciprocity constraints, while improving on the Nash, are also capable of generating socially optimal levels of cooperation. As it will be shown below, the ability of reciprocity constraints to generate socially optimal levels of cooperation depends in great measure on the homogeneity of the players.

In the case of symmetric players, the reciprocity-induced equilibrium is privately and socially optimal in both partial and full cooperation cases. With asymmetric players, however, the equilibrium level of cooperation will always be dictated by the high cost-cooperator. In the asymmetric case, therefore, there are several additional conditions for the reciprocity-induced equilibrium to be identical to the social optimum. In the following, we consider these conditions for socially optimal cooperation under reciprocity.

3.1 Symmetric versus asymmetric social optima

A first straightforward condition for the reciprocity-induced equilibrium to be identical to the social optimum requires that the social optimum lies along the symmetric strategy diagonal in the game’s strategy space. If the social optimum requires asymmetric
levels of cooperation for the two parties, the presence of a binding reciprocity constraint would obviously prevent the achievement of such ideal optimum. The above thus constitutes a necessary condition for social optimality.

Consider the following case in which the social optimum falls outside the reciprocity diagonal in the strategy space. We consider the simplest case where this could happen. Let the marginal benefits for the cooperation efforts of both parties be equal, and assume that $MC_1$ is greater than $MC_2$. That is, we assume that $b = d$ and $a > c$.\(^{11}\) In this case, first note that the desired cooperation effort under reciprocity for each party is equal to the corresponding socially optimal cooperation effort. That is, we have $s'_1 = \frac{b}{2a} = \frac{d}{2a} = \bar{s}_1$ and $s'_2 = \frac{d}{2c} = \frac{b}{2c} = \bar{s}_2$. Further, note that $s'_1 = s'_2 < s'_2 = \bar{s}_2$ because $a > c$.

Since the desired cooperation effort under reciprocity of party 1 is binding, we have $s'_1 = s'_2 = \bar{s}_1 < \bar{s}_2$. In this case, the reciprocity-induced equilibrium leads to an inadequate amount of cooperation effort.

Further conditions for optimality in the presence of asymmetric payoffs functions depend on whether the equilibrium strategies induce full or partial cooperation. We shall examine these two cases in turn.

### 3.2 The conditions for efficient reciprocity: the case of full cooperation

In the case of full cooperation, the reciprocity-induced equilibrium strategies $s'_1$, $s'_2$, and the social optimum strategies $\bar{s}_1$, $\bar{s}_2$ must all equal 1. In terms of the parameters

\(^{11}\) These assumptions on the parameters imply that party 1 is the high-cost cooperator.
of the model, this means the following must hold: \( \frac{b}{2a} \geq 1 \) (as \( s_1^* = 1 \)), \( \frac{d}{2c} \geq 1 \) (as \( s_2^* = 1 \)), \( \frac{d}{2a} \geq 1 \) (as \( s_1 = 1 \)), and \( \frac{b}{2c} \geq 1 \) (as \( s_2 = 1 \)). In other words, \( b \geq \max \{2a, 2c\} \) and \( d \geq \max \{2a, 2c\} \) must hold.

The conditions \( b \geq 2a \) and \( d \geq 2c \) indicate that the private marginal benefits of cooperation for the parties under reciprocity exceed the marginal costs at full levels of cooperation. The parties would extend cooperation beyond full cooperation, if they had an option to do so. The differences between the privately optimal levels of cooperation for the two parties are thus revealed only in the infeasible region of more-than-full cooperation, and are thus immaterial for the actual equilibrium at full cooperation. The parties converge to a full level of cooperation, not because they have identical preferences, but because full cooperation gives them the highest obtainable payoff in the feasible region.

The conditions \( b \geq 2c \) and \( d \geq 2a \) indicate that the social marginal benefits of cooperation for the parties exceed the marginal costs at full levels of cooperation. Thus the social optimum requires full cooperation for both parties.

The combined effect of these two sets of conditions thus guarantees that full cooperation will be both privately and socially optimal, in spite of the asymmetric cost and benefit functions for the two parties.
3.3 The conditions for efficient reciprocity: the case of partial cooperation

In the case of *partial cooperation*, in order for the reciprocity-induced equilibrium to coincide with the social optimum, two conditions must hold. First, party 1’s socially optimal strategy must be equal to its reciprocity-induced equilibrium strategy, since the high-cost cooperator’s chosen strategy is the binding strategy under reciprocity-induced equilibrium. That is,  \( s_1 = s_1^* < 1 \). Second, since the reciprocity-induced equilibrium requires equal levels of cooperation, socially optimal strategies must be equal for the two parties. That is,  \( \bar{s}_1 = \bar{s}_2 \). Translating these conditions to the parameters of the model, we have:

\[
\bar{s}_1 = s_1^* < 1 \quad \Rightarrow \quad \frac{d}{2a} = \frac{b}{2a} < 1 \quad \Rightarrow \quad b = d < 2a ,
\]

and

\[
\bar{s}_1 = \bar{s}_2 \quad \Rightarrow \quad \frac{d}{2a} = \frac{b}{2c} \quad \Rightarrow \quad a = c .
\]

Since  \( a = c, b = d < 2a \) must hold, we see that, whenever partial cooperation is socially optimal, reciprocity constraints will be able to deliver socially optimal equilibria only in the presence of identical parties with symmetric payoff functions.

This is an important result: if the social optimum requires partial levels of cooperation for the two parties, such equilibrium is obtainable under a reciprocity rule only if the players have symmetric payoff functions. In the case of asymmetric parties when the privately optimal level of cooperation for at least one party falls short of full cooperation (i.e., when the private marginal cost of cooperation comes to equal the benefit of reciprocal cooperation before full cooperation is achieved), the reciprocity-induced levels of cooperation do not coincide with the social optimum. Thus, unlike the
symmetric case in which the reciprocity-induced equilibrium is privately and socially optimal in both partial and full cooperation cases, with asymmetric players, partial cooperation will always be dictated by the high cost-cooperator and the resulting equilibrium level of cooperation will be privately optimal only for this party and would neither be privately optimal for the other party nor socially optimal.

4 Conclusion

Through genetic selection, cultural adaptation, or historical evolution, norms of reciprocity have emerged as behavioral, social or legal rules, governing a wide array of human, social and institutional settings. These reciprocity constraints often tend to become meta-rules in customary law and in other environments where there is no recognized rule of law or where enforcement mechanisms are absent.\textsuperscript{12}

In all the cases considered in this paper, the results reveal that reciprocity constraints always have a positive impact on the parties’ levels of cooperation. But, in the case of asymmetric players, reciprocity is not always capable of generating a global optimum. When looking at the aggregate payoffs of the players, we see that the reciprocity-induced outcome always improves upon the alternative Nash equilibrium in the absence of reciprocity. Such level of cooperation, however, may fall short (or occasionally exceed) the socially optimal level of cooperation for the two parties.

In this paper, starting from a standard discrete prisoner’s dilemma problem, we first extended the analysis to continuous strategic variables with symmetric linear payoff functions. In this case, we find that the reciprocity-induced equilibrium is one of full

\textsuperscript{12} See PARISI [2000b]. The classic on the evolution of norms, outside a formal legal system remains ELICKSON [2001].
cooperation and is always socially optimal. Next we considered symmetric nonlinear payoff functions for the two players. In this case, the reciprocity-induced equilibrium is also always identical to the social optimum. Here, it was found that the nonlinearity could induce either full cooperation or partial cooperation for both parties. The possibility of partial cooperation equilibrium was made possible by the fact that the marginal cost of cooperation of one party becomes equal to the marginal benefit of cooperation before it reaches the full cooperation level.

We next turn to the case of asymmetric nonlinear payoff functions. In general, the reciprocity-induced equilibrium is not identical to the social optimum. Further, with asymmetric nonlinear payoffs, the full-cooperation reciprocity-induced equilibrium and the social optimum may be the same when the marginal benefits are very high. On the other hand, partial-cooperation under reciprocity will be socially optimal only with symmetric parties.

Among the most important results, as mentioned above, is the fact that the reciprocity-induced equilibrium always improves upon the Nash equilibrium and hence will at least partially solve the prisoner's dilemma.

Another important implication of our study is that the payoffs corresponding to the socially efficient levels of cooperation for the two parties are never Pareto-superior to the payoffs of the reciprocity-induced equilibrium. This can be seen from the fact that, whenever the reciprocity-induced equilibrium and the socially optimal equilibrium do not coincide, the social optimum, albeit Kaldor-Hicks efficient, always worsens the position of one of the two players.
In such cases the reciprocity-induced equilibrium leaves some unexploited surplus for the parties: a move to the socially optimal levels of cooperation would increase the aggregate payoffs for the parties, thus allowing the gainers to fully compensate the losers for the additional cost of cooperation, yet capturing the unexploited surplus. This may explain why, in situations of asymmetry between the parties, reciprocity constraints may as such be unable to generate efficient outcomes. Different mechanisms of cooperation, such as explicit trading or contracting would instead allow the parties to converge towards global maxima.

It is expected that the greater the asymmetries between the players, the greater will the unexploited surplus from the reciprocity-induced equilibrium be. For example, when a high-cost cooperator and a low-cost cooperator are matched, the high-cost cooperator will pose a limit on the level of reciprocal cooperation chosen by the parties. The greater the gap between the parties’ cost of cooperation, the lower the ability of reciprocity to generate socially efficient outcomes. As a general result, we can thus anticipate reciprocity constraints to be quite effective devices in the context of symmetric players (or players with similar cost functions). On the contrary, reciprocity-induced outcomes are likely to fall short of global optimization, when the players face substantial asymmetries in their cost or benefit functions.

These considerations are in line with the findings of evolutionary socio-biology, showing that behavioral patterns of reciprocity tend to emerge and thrive in close-knit environments with homogeneous players. This further explains a well-known fact in social sciences. As shown in the literature on social norms and customary law, spontaneous cooperation under reciprocity is quite stable and effective in homogeneous
social groups. Reciprocity constraints fail to generate efficient levels of cooperation in heterogeneous groups (or, more generally, when there are substantial asymmetries in the cost-benefit ratios of the various players). The model of reciprocity presented in this paper may, in fact, help us understand the strength and the limitations of institutionally designed reciprocity constraints in solving prisoner’s dilemma problems in real life situations. Likewise, it would illuminate the practical limits of reciprocity in the formation of uniform social norms and egalitarian customs involving heterogeneous people. In this latter group of cases, we should thus expect greater reliance on other mechanisms of social and legal order, such as collective decision-making (e.g., political deliberation) or explicit contracting (e.g., treaty law formation among nation states).
Appendix A

Payoffs under social optimum with asymmetric nonlinear payoffs functions:

If \( \frac{d}{2a} < 1 \) \& \( \frac{b}{2c} < 1 \),

then \( \bar{s}_1 = \frac{d}{2a} < 1 \) \& \( \bar{s}_2 = \frac{b}{2c} < 1 \),

with \( \bar{P}_1(\bar{s}_1, \bar{s}_1) = \frac{-d^2}{4a} + \frac{b^2}{2c}, \quad \bar{P}_2(\bar{s}_1, \bar{s}_1) = \frac{-b^2}{4c} + \frac{d^2}{2a} \),

and the total social payoff is \( \bar{P}(\bar{s}_1, \bar{s}_2) = \frac{b^2}{4c} + \frac{d^2}{4a} \).

If \( \frac{d}{2a} < 1 \leq \frac{b}{2c} \),

then \( \bar{s}_1 = \frac{d}{2a} < 1 \) \& \( \bar{s}_2 = 1 \),

with \( \bar{P}_1(\frac{d}{2a}, 1) = \frac{-d^2}{4a} + b, \quad \bar{P}_2(\frac{d}{2a}, 1) = -c + \frac{d^2}{2a} \),

and the total social payoff is \( \bar{P}(\frac{d}{2a}, 1) = b - c + \frac{d^2}{4a} \).

If \( \frac{b}{2c} < 1 \leq \frac{d}{2a} \),

then \( \bar{s}_1 = 1 \) \& \( \bar{s}_2 = \frac{b}{2c} < 1 \),
with \( P_1(1, \frac{b}{2c}) = -a + \frac{b^2}{2c} \), \( P_2(1, \frac{b}{2c}) = -\frac{b^2}{4c} + d \),

and the total social payoff is \( P(1, \frac{b}{2c}) = \frac{b^2}{4c} + d - a \).

If \( 1 \leq \frac{d}{2a} \) & \( 1 \leq \frac{b}{2c} \),

then \( \bar{s}_1 = 1 \), \( \bar{s}_2 = 1 \),

with \( P_1(1,1) = -a + b \), \( P_2(1,1) = -c + d \),

and the total social payoff is \( P(1,1) = (b - a) + (d - c) + d - a \).
References


KOMTER, A.E. [1996], “Reciprocity as a Principle of Exclusion: Gift Giving in the
Netherlands,” *Sociology*, 30:2 (May), 299.

**LARGE, M.** [1999], “The Effectiveness of Gifts as Unilateral Initiatives in Bargaining,” *Sociological Perspectives*, 42:3 (September), 525.


**Vincy Fon**  
Assistant Professor of Economics  
George Washington University  
624 Funger Hall  
2201 G Street, N.W.  
Washington, D.C.  20052  
vfon@gwu.edu

**Francesco Parisi**  
Professor of Law  
George Mason University School of Law  
3401 North Fairfax Drive  
Arlington, VA  22201  
parisi@gmu.edu