

# Effecting Cooperation\*

Olivier Compte  
PSE, Paris

Andrew Postlewaite  
University of Pennsylvania

July 2007; this version May 2009

---

\*Compte: Paris School of Economics, 48 Bd Jourdan, 75014 Paris (e-mail: [compte@enpc.fr](mailto:compte@enpc.fr)); Postlewaite: Department of Economics, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6207 (e-mail: [apostlew@econ.upenn.edu](mailto:apostlew@econ.upenn.edu)). This paper is a substantially revised version of "Repeated Relationships with Limits on Information Processing". We thank Ted Bergstrom, Larry Samuelson, Alvaro Sandroni and the participants at many seminars at which this paper was presented for helpful comments. A large part of this work was done while Postlewaite visited the Paris School of Economics; their hospitality is gratefully acknowledged. The authors thank the Gould Foundation for financial support, and Postlewaite thanks the National Science Foundation for financial support.

**Abstract**

There is a large repeated games literature illustrating how future interactions provide incentives for cooperation. Much of this literature assumes public monitoring; players always observe precisely the same thing. Even slight deviations from public monitoring to private monitoring that incorporate differences in players' observations dramatically complicate coordination. Equilibria with private monitoring often seem unrealistically complex.

We set out a model in which players accomplish cooperation in an intuitively plausible fashion. Players process information via a mental system – a set of psychological states and a transition function between states depending on observations. Players restrict attention to a relatively small set of simple strategies, and consequently, might learn which perform well.

## 1. Introduction

Cooperation is ubiquitous in long-term interactions: we share driving responsibilities with our friends, we offer help to relatives when they are moving and we write joint papers with our colleagues. The particular circumstances of an agent’s interactions vary widely across the variety of our long-term relationships but the mechanics of cooperation are usually quite simple. When called upon, we do what the relationship requires, typically at some cost. We tend to be upset if our partner seems not to be doing his part and our willingness to cooperate diminishes. We may be forgiving for a time but stop cooperating if we become convinced the relationship is one-sided. We sometimes make overtures to renew the relationship when opportunities arise, hoping to rejuvenate cooperation. Incentives to cooperate stem from a concern that the relationship would temporarily break down, while incentives to be less cooperative when the relationship feels one-sided stem from the fear of being taken advantage of by a non-cooperative partner. Such simple behavior seems to be conducive to cooperation under a broad range of circumstances, including those in which we get only a noisy *private* signal about our partner’s efforts in the relationship, that is, when our partner does not always know if we are less than satisfied with their effort.

Despite the fundamental importance of cooperation in understanding human interaction in small or large groups, there is not currently a satisfactory theoretical foundation of such cooperation. The theory of repeated games has provided important insights about repeated interactions but does not capture the simple intuition in the paragraph above. When signals are private, the quest for “stable” rules of behavior (or equilibria) typically produces complex strategies that are finely tuned to the parameters of the game (payoffs, signal structure).<sup>1</sup> If the parameters of the game are changed slightly the rule fails to remain stable.<sup>2,3</sup>

The complexity of repeated game strategies with private monitoring begins with the fact that the number of possible histories increases exponentially with the number of interactions, and a strategy must specify what action is to be taken at each of these histories. If I and my spouse alternate cooking dinner and whoever cooks can either shirk or put in effort each time they cook, there will be approximately a billion possible histories after one month. For each of these billion histories, both I and my spouse will have gotten imperfect signals about the effort put in by the other on the nights they cooked, and for each of the histories, I must decide whether or not to put in effort the next time I cook. It is inconceivable that I recall the precise history after even a month let alone after several years.<sup>4</sup>

---

<sup>1</sup>See, e.g., Piccione (2002) and Ely and Valimaki (2002).

<sup>2</sup>In other words, equilibria are not strict.

<sup>3</sup>Fundamental to the standard approach to repeated games with private signals is the analysis of incentives of one party to convey to the other party information about the private signals he received, either directly (through actual communication), or indirectly (through the action played). Conveying such information is important to build punishments that generate incentives to cooperate in the first place.

Incentives to convey information however are typically provided by making each player indifferent between the various messages he may send, or the various actions he may play. There are exceptions, and some work such as Sekiguchi (1997) *do* have players provided with *strict* incentives to use their observation. But, these constructions rely on fine tuning some initial uncertainty about the opponent’s play (as shown in the work of Bagwell (1995)), and they typically produce strategies which depend in a complex way on past histories (as in Compte (2002)).

<sup>4</sup>The difficulty of choosing an action for each history is a problem not only for potential cooperators; even theorists analyzing repeated games find it prohibitive to do this. The recursive techniques of Abreu, Pearce and

A more realistic description is that I rely on some summary statistic in deciding whether or not to put in effort – the number of times it seemed effort was put in over the several times my spouse cooked, for example. In this way, histories are catalogued in a relatively small number of equivalence classes, and my action today depends only on the equivalence class containing the history. Our aim is to provide a *realistic* description of cooperation when players are strategic, and a central feature of the description is that players do not distinguish all histories in precise detail.

We ask more for a strategy to be realistic than that it can be represented with a small number of equivalence classes of histories. Repeated game strategies can often be represented in this way, but the classification of histories is for mathematical convenience, not on an *a priori* basis of how individuals pool histories.<sup>5</sup> Our aim is a positive model in which cooperation is attained via intuitively plausible behavior. A player’s pooling of histories is not simply a mathematical tool, but is meant to capture plausible cognitive limitations, and different histories in an equivalence class should have a natural similarity. Importantly, the categorization of histories should *not* depend on fine details of the problem at hand: small changes in the payoffs or the signal structure should not affect a player’s categorization of histories. Analogously, the actions taken for different equivalence classes of histories will be intuitively plausible in a satisfactory model of cooperation.

A realistic description of cooperative behavior should also include a notion of *how* players came to that behavior. In the standard approach to repeated games there is no realistic story of how players would arrive at the proposed equilibrium strategies. It seems extremely implausible that players could compute appropriate strategies through introspection in repeated games with private signals. Furthermore, equilibrium strategies in such a setting typically rely on my knowing not only the distribution of signals I receive conditional on the other player’s actions, but also on the distribution of *his* signals given my actions, something I never observe. Even if one entertains the possibility that players compute equilibrium strategies through introspection there is the question of how the players might know these signal distributions. One might posit that players could “learn” the equilibrium strategies, but the set of strategies is huge and it is difficult to see how a player might learn which strategies work well.<sup>6</sup> Learning equilibrium strategies is much more plausible when histories are catalogued into a small number of equivalence classes. When there is a finite number of actions and a finite number of equivalence classes, there is a finite number of pure strategies a player needs to consider, as opposed to the necessarily infinite number of pure strategies in the standard approach.

In our approach, players restrict attention to a relatively small set of simple strategies. Our goal is to find sets of strategies that have the following desirable properties: (i) the number of strategies in the set should be small enough that players might ultimately learn which perform

---

Stachetti (1990) are often used to characterize the set of equilibrium payoffs attainable as equilibria of repeated games without having to describe the strategies that attain those payoffs.

<sup>5</sup>In public monitoring games, a particular vector of continuation values may be assigned after different histories. Typically, histories might be categorized into equivalence classes, where the same vector of continuation values is assigned to all histories in an equivalence class, regardless of the actions and signals that constitute the histories.

<sup>6</sup>In the repeated game literature, allowing for public communication is often seen as a remedy to the problems caused by the fact that signals are private (see Compte (1998) and Kandori and Matsushima (1998)). However, while communication simplifies the analysis, all the comments above apply: when signals are conditionally independent, providing incentives to communicate truthfully is done by making players indifferent between the various messages they send.

well; (ii) the strategies should be based on a cataloging of histories that is intuitively plausible; and (iii) the sets of strategies should allow agents to cooperate under a broad set of circumstances.

One can think of the choice of actions as arising in a hierarchical behavior system in which the top level (how histories are cataloged) describes the ways in which a person might play in a broad set of games. The second level of the hierarchical system is activated when an individual is faced with a specific game. At this point, the player chooses from the set of feasible strategies given the cataloging of histories. While the set of possible strategies considered remains unchanged as the “fine details” of the games vary, the choice from that set may.

To repeat, the goal is a realistic description of cooperation when people are strategic. The games we play vary and our knowledge of the structure of the game may be limited. We need to adjust play through experimentation to learn which strategies perform well. Restrictions on the set of strategies allow players to identify best responses easily, and eventually adjust play as a function of the particular parameters of the game they are currently playing. This goal motivates the model that we set out below. We do not claim that this model is unique in achieving our goal, only that it is a plausible model that satisfies our stated desiderata.

### 1.1. Strategy restrictions

The restrictions on the strategies available to players are a crucial element of our approach.<sup>7</sup> We are not interested in arbitrary restrictions, but rather, on restrictions that might arise naturally. An individual’s action in any period is generally assumed to be a function of the history to that point. As we suggested above, we assume players do not recall all details of a history, but rather catalogue them into equivalence classes. We will refer to an equivalence class of histories as the *mental state* a player is in and limit the number of mental states available to a player. This restricts a player to behaving the same way for all histories of the game that lead to the same mental state.<sup>8</sup>

We think of the set of mental states not to be a choice variable, but rather a natural limitation of mental processing. We might feel cheated if we have put effort into a relationship and get signals that the other is not reciprocating. We can think of those histories in which one feels cheated as leading to a mental state (U)pset, and those histories in which one doesn’t feel cheated as leading to a mental state (N)ormal.<sup>9</sup> A *mental system* is the *set of mental states* one can be in along with a *transition function* that describes what combinations of initial mental state, actions and signals lead to specific updated mental states. We will assume in most of what we do that the transition function does not depend on the fine details of the game – the payoffs and the monitoring structure – but in principle it might. For example, in circumstances in which it is extremely costly for my partner to put in effort, I may not become upset if he does not seem to be doing so. However, a fundamental aspect of the transition function is that the individual does not have control over it.

---

<sup>7</sup>See Cherepanov, Feddersen and Sandroni (2008) for a very different way of restricting strategies people might employ.

<sup>8</sup>For many problems restricting a player to a finite number of mental states is natural. If there is a finite number of signals a player can receive following the play of a game in any period, and if players have bounded recall, the assumption that an individual has a finite number of mental states is without loss of generality.

<sup>9</sup>The case in which there are only two mental states is sufficient to illustrate our ideas. We do not suggest that people are necessarily limited to two mental states and our basic message holds when there are more mental states, as long as there is a finite number.

While the mental system may be the same across a variety of games, how one *responds* to being upset may be situational, that is, may depend on the particular game one is involved in, as well as on the behavior of the other player. If the cost of cooperation is very small, one might be hesitant to defect in state U and risk breaking a relationship that is generally cooperative, but not hesitate when the cost is large; whether in state U or state N, the individual may either cooperate or defect. Thus, while a player’s *available* strategies depend only on that player’s mental process - hence not necessarily on the fine details of the payoffs and informational structure of the game - the strategy he chooses will typically be sensitive to the specifics of the game at hand.

Our view is that there are limits to peoples’ cognitive abilities, and evolution and cultural indoctrination determine an individual’s mental system consisting of the states he can be in and the transition function that moves him from one state to another. Children experience a large number of diverse interactions, and how they interpret those experiences are affected by their parents and others they are (or have been) in contact with. A parent may tell his child that the failure of a partner to have reciprocated in an exchange is not a big deal and should be ignored, or the parent can tell the child that such selfish behavior is reprehensible and inexcusable. Repeated similar instances shape how the child interprets events of a particular type. Even in the absence of direct parental intervention, observing parental reactions to such problems shape the child’s interpretations.<sup>10</sup>

As stated above, our goal is to set out a model that allows cooperation between people using plausible strategies. We set out such a model in the next section and analyze the circumstances in which cooperation in a repeated prisoners dilemma problem with private monitoring is possible. In section 3 we compare players’ behavior with the behavior assumed in the standard model and analyze the robustness of the cooperative behavior that we identify. In this section we assume that there is occasionally a public signal that facilitates periodic synchronization. In section 4 we drop the assumption of such a public signal and show how synchronization can be accomplished without such a signal. Section 5 suggests evolutionary foundations of our model and concludes.

## 2. Model

*Gift exchange.*

There are two players who exchange gifts each period. Each has two possible actions available,  $\{D, C\}$ . Action  $D$  is not costly and can be thought of as no effort having been made in choosing a gift. In this case the gift will not necessarily be well perceived. Action  $C$  is costly, and can be interpreted as making substantial effort in choosing a gift; the gift is very likely to be well-received in this case. The expected payoffs to the players are as follows:

	$C$	$D$
$C$	1, 1	$-L, 1 + L$
$D$	$1 + L, -L$	0, 0

$L$  corresponds to the cost of effort in choosing the “thoughtful” gift: you save  $L$  when no effort is made in choosing the gift.

---

<sup>10</sup>It is quite possible that there is a conflict between the two; parents may want to indoctrinate their children to be different from themselves.

*Signal structure.*

We assume that there are two possible private signals that player  $i$  might receive,  $y_i \in Y_i = \{0, 1\}$ , where a signal corresponds to how well player  $i$  perceives the gift he received. We assume that if one doesn't put in effort in choosing a gift, then most likely, the person receiving the gift will not think highly of the gift. We will refer to  $y = 0$  as a "bad" signal and  $y = 1$  as "good".

Formally,

$$p = \Pr\{y_i = 0 \mid a_j = D\} = \Pr\{y_i = 1 \mid a_j = C\}.$$

We assume that  $p > 1/2$  and for most of the main text analysis we consider the case where  $p$  is close to 1.

In addition to this private signal, we assume that at the start of each period, players receive a public signal  $z \in Z = \{0, 1\}$ , and we let

$$q = \Pr\{z = 1\}.$$

The existence of a public signal  $z$  facilitates our exposition but can be dispensed with, as we demonstrate in section 4.

## 2.1. Strategies

As discussed above, players' behavior in any period will depend on the previous play of the game, but in a more restricted way than in general. There is a finite set of possible *mental states*,  $S_i$ , that player  $i$  can be in. Mental states capture the bounds on players' memories of the precise details of past play. For example, a player who has gotten one bad signal in the past twenty periods should not be assumed to remember this occurred in period sixteen or period seventeen. For simplicity, we assume that in the current example the players can be in one of two states  $U$ (pset) or  $N$ (ormal). The names are chosen to convey that at any time player  $i$  is called upon to play an action, he knows the mood he is in, which is a function of the history of (own) play and signals, but does not condition his action on finer details of the history.<sup>11</sup> One can interpret the restriction to strategies that are constant across the histories that lead to a particular mental state as being a limit on the player's memory.  $S_i$  is exogenously given, not a choice. Player  $i$ 's set of pure strategies is

$$\Sigma_i = \{\sigma_i, \sigma_i : S_i \longrightarrow A_i\}.$$

The particular state in  $S_i$  that player  $i$  is in at a given time depends on the previous play of the game. The *transition function* for player  $i$ , which we shall denote by  $T_i$ , is a function that determines the state player  $i$  will be in at the beginning of period  $t$  as a function of his state in period  $t - 1$ , his choice of action in period  $t - 1$ , and the outcome of that period – the signals  $y_i$  and  $z$ .<sup>12</sup> As is the set of states for player  $i$ , the transition function is exogenous. A player who

---

<sup>11</sup>For expository ease we assume that an individual's payoffs depend on outcomes, but not on the state he is in. The names that we use for the states suggest that the state itself could well be payoff relevant: whatever outcome arises, I will be less happy with that outcome if I'm upset. Our framework can easily accommodate state-dependence, and the qualitative nature of our conceptual points would be unchanged if we did so. We discuss this further in the discussion section below.

<sup>12</sup>In principle, the transition function could depend on more than the most recent signal, for example, whether two of the past three signals was "bad". For simplicity, in this paper we assume the given form.

has made an effort in his choice of gift but receives a bad signal may find it impossible not to be upset, that is, be in state  $U$ .

We assume the transition function for the example, which we will refer to as the *leading example* below, is as in the following figure.

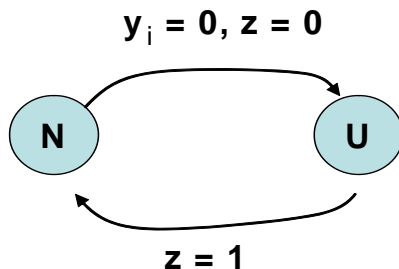


Figure 1: Transition

This figure shows which combinations of actions and signals will cause the player to move from one state to the other. If player  $i$  is in state  $N$ , he remains in that state unless he receives signals  $y = 0$  and  $z = 0$ , in which case he transits to state  $U$ . If  $i$  is in state  $U$ , he remains in that state until he receives signal  $z = 1$ , at which point he transits to state  $N$  regardless of the signal  $y$ .<sup>13</sup>

To summarize, a player is endowed with a *mental system* that consists of a set of mental states the player can be in and a transition function that describes what triggers moves from one state to another. Our interest is in finding stable patterns of behavior. Our structure requires that players' strategies are stationary: they do not depend on the period. This rules out strategies of the sort "Play  $D$  in prime number periods and play  $C$  otherwise", consistent with our focus on behavior that is a function of the information close to hand at the time choices are made.

Our candidate behavior for the players will be as follows. For player  $i$ ,

$$\begin{aligned}\sigma_i(N) &= C \\ \sigma_i(U) &= D.\end{aligned}$$

That is, player  $i$  plays  $C$  as long as he receives a gift that seems thoughtful, that is  $y_i = 1$ , or when  $z = 1$ . He plays  $D$  otherwise. Intuitively, player 1 triggers a "punishment phase" when he sees  $y_1 = 0$ , that is, when he didn't find the gift given to him appropriate. This punishment phase ends only when signal  $z = 1$  is received.

The public signal  $z$  gives the possibility of "resetting" the relationship to a cooperative mode. If the signal  $z$  is ignored and the mental process is defined by

---

<sup>13</sup>For this particular example, transitions depend only on the signals observed, and not on the individual's action. In general, it might also depend on the individual's action.

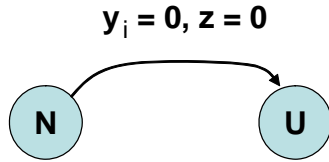


Figure 2: No “resetting”

then eventually, because signals are noisy, with probability 1 the players will get to state  $U$  under the proposed strategy and this will be absorbing: there would be nothing to change their behavior. The signal  $z$  allows for possible recoordination back to state  $N$  (and possibly cooperation).<sup>14</sup>

In our leading example, players stop being upset for exogenous reasons. Alternatively, in a two-state mental system the players could move from state  $U$  back to state  $N$  after seeing a good signal: you stop being upset as soon as you receive a nice gift.

Formally, players may be either in state  $N$  or in state  $U$ , but are endowed with the following transition function.

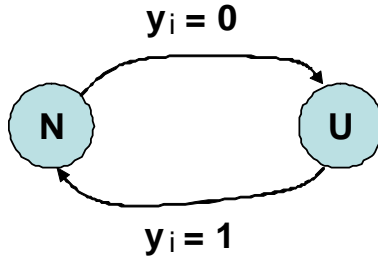


Figure 3: Forgiving transition

A player endowed with this alternative mental process, who would cooperate in  $N$  and defect in  $U$ , would be following a TIT for TAT strategy.<sup>15</sup>

## 2.2. An illustrative experiment

Before continuing with the formal description of our model, it is useful to give a real-world example to illustrate our idea of a mental system. Cohen *et al.* ran several experiments in which participants (students at the University of Michigan) were insulted by a confederate who would bump into the participant and call him an “asshole”. The experiment was designed to test the hypothesis that participants raised in the north reacted differently to the insult than did participants raised in the south. From the point of view of our model, what is most interesting is that the insult triggered a physical response in participants from the south. Southerners were upset by the insult, as shown by cortisol levels, and more physiologically primed for aggression, as shown by a rise in testosterone. We would interpret this as a transition from one mental state to another, evidenced by the physiological changes. This transition is plausibly not a choice on

<sup>14</sup>As mentioned above, we show the existence of a public signal  $z$  is not necessary for re-coordination below.

<sup>15</sup>We show below that cooperation is impossible if players have this mental process.

the participant’s part, but involuntary. The change in mental state that is a consequence of the insult was followed by a change in behavior: southerners were more likely to respond in an aggressive manner following the insult than were northerners. Moreover, Southerners who had been insulted were more than three times as likely to respond in an aggressive manner in a word completion test than were Southerners in a control group who were not insulted. There was no significant difference in the aggressiveness of Northerners who were insulted and those who were not.

The physiological reaction to an insult – what we would think of as a transition from one state to another – seems not to be “hard-wired”: physiological reactions to insults were substantially lower for northern students than for southern students. Indeed, the point of the Cohen *et al.* paper is to argue that there is a southern “culture of honor” that is inculcated in small boys from an early age. This culture emphasizes the importance of honor and the defense of it in the face of insults. This illustrates the view of our model expressed above that the transition function in our model can be thought of as culturally determined, but fixed from the point of view of an individual at the time decisions are taken.

### 2.3. Ergodic distributions and strategy valuation

For any pair of players’ strategies there will be an ergodic distribution over the pairs of actions played. While in general the ergodic distribution may depend on the initial conditions, we restrict attention to transition functions for which the distribution is unique. The ergodic distribution gives the probability distribution over payoffs in the stage game, and we take the payoff to the players to be the expected value of their payoffs given this distribution.

Formally, define a state profile  $s$  as a pair of states  $(s_1, s_2)$ . Each strategy profile  $\sigma$  induces transition probabilities over state profiles: by assumption each state profile  $s$  induces an action profile  $\sigma(s)$ , which in turn generates a probability distribution over signals, and hence, given the transition functions  $T_i$ , over next period states. We denote by  $Q_\sigma$  the transition matrix associated with  $\sigma$ , and by  $\phi_\sigma$  the ergodic distribution over states induced by  $\sigma$ . That is,  $\phi_\sigma(s)$  corresponds to the (long run) probability that players are in state  $s$ .<sup>16</sup>

We associate with each strategy profile  $\sigma$  the value induced by the ergodic distribution. This corresponds to computing discounted expected payoffs, and taking the discount to 1.<sup>17</sup> We denote by  $v(\sigma)$  this value (vector). Thus,

$$v(\sigma) = \sum_s g(\sigma(s))\phi_\sigma(s)$$

where  $g(\sigma(s))$  is the payoff vector induced by the strategy profile  $\sigma$  for state profile  $s$ .

#### *Equilibrium.*

*Definition:* We say that a profile  $\sigma \in \Sigma$  is an equilibrium if for any player  $i$  and any strategy  $\sigma'_i \in \Sigma_i$ ,

$$v_i(\sigma'_i, \sigma_{-i}) \leq v_i(\sigma).$$

<sup>16</sup>Formally,  $Q_\sigma(s', s)$  is the probability that next state profile is  $s'$  when the current state is  $s$ , and the vector  $\phi_\sigma$  solves  $\phi_\sigma(s') = \sum_s Q_\sigma(s', s) \phi_\sigma(s)$ .

<sup>17</sup>When discounting is not close to one, then a more complex valuation function must be defined: when  $\sigma$  is being played, and player  $i$  evaluates strategy  $\sigma'_i$  as compared to  $\sigma_i$ , the transitory phase from  $\phi_\sigma$  to  $\phi_{\sigma'_i, \sigma_{-i}}$  matters. Note however that the equilibria we will derive are strict equilibria, so they would remain equilibria under this alternative definition for discount factors sufficiently close to 1.

This is a weaker notion of equilibrium than traditionally used in repeated games because of the restriction on the set of strategies to be mappings from  $S_i$  to  $A_i$ .<sup>18</sup> Also note that  $\sigma_i$  as defined should not be viewed as a strategy of the repeated game.<sup>19</sup>

## 2.4. Successful Cooperation

Cooperation is possible for a range of gift exchange problems, that is, for a range of payoff values and signal distributions when the players have the mental system as shown in figure 1. The strategy profile  $\sigma$  in which each player plays  $C$  in the Normal state, and  $D$  in the Upset state will be an equilibrium for the case that  $q$  is .3 and  $(p, L)$  is in the shaded region:

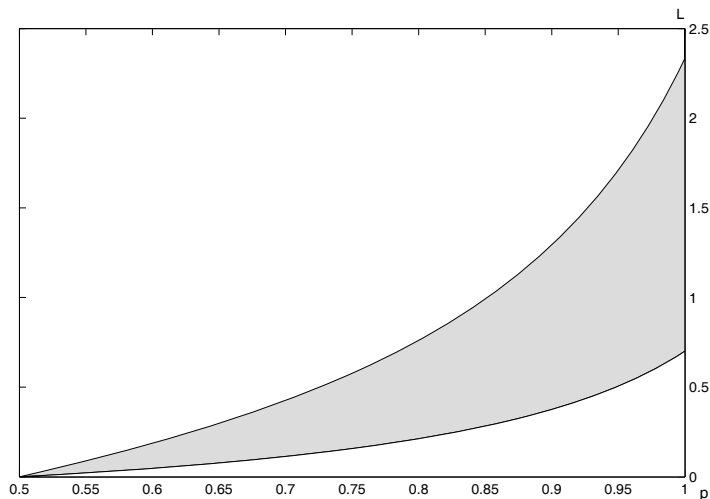


Figure 4:  $p - L$  combinations that allow cooperation

For each value of  $p$ ,  $\sigma$  is an equilibrium so long as  $L$  lies within *two bounds*. Intuitively,  $L$  cannot be too large or the incentive to deviate outweighs the possible cost of interrupting cooperation should the opponent get a bad signal.  $L$  cannot be too small or players will not have an incentive to punish (play  $D$ ) when they are in  $U$ . The first constraint is typical of all the repeated game literature. The latter constraint is specific to the private monitoring environment: incentives to trigger punishments are generally harder to provide.

We compute next these bounds for the case where  $p$  is close to 1. We consider the ergodic distribution induced by our candidate equilibrium strategy  $\sigma$  and examine in turn possible

<sup>18</sup>We restrict attention to pure strategies. However, our definitions can be easily generalized to accommodate mixed actions, by re-defining the set  $A_i$  appropriately, and having it include mixed actions. However, the spirit of our approach is that players should adjust play from experience, by checking from time to time the performance of alternative strategies. So if mixed actions are to be allowed, only few of them, rather than the whole set of mixed actions, should in our view be considered.

<sup>19</sup>A strategy of the repeated game is a mapping from histories to actions. The strategy  $\sigma_i$ , along with the mental system  $(S_i, T_i)$  would induce a repeated game strategy, once the initial state is specified.

deviations from that strategy. The transition over state profiles induced by  $\sigma$  is illustrated in figure 5.

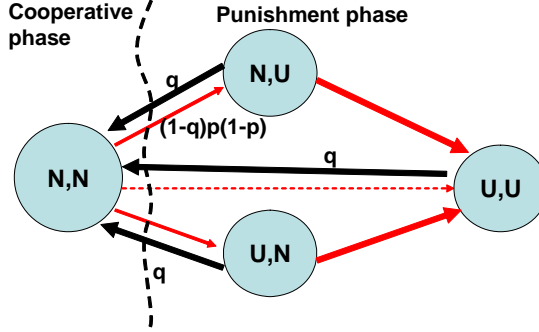


Figure 5: Transition over state profiles under  $\sigma$ .

When players follow our candidate equilibrium strategy, they alternate between cooperation and punishment phases. The probability of switching from cooperation to a punishment phase is  $\pi = (1 - q)(1 - p^2)$  (since switching occurs when either player receives a bad signal and  $z = 0$ ). The probability of switching from punishment to cooperation is  $q$ . Hence cooperative phases last on average  $1/\pi$  periods, while punishment phases last on average  $1/q$  periods.<sup>20</sup>

When player  $i$  plays  $D$  at both  $N$  and  $U$ , player  $j$  continues to alternate between phases of cooperation and defection. This is illustrated in figure 6.

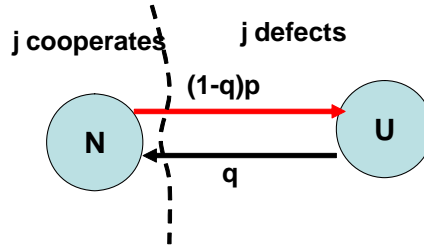


Figure 6:  $i$  defects always

Player  $i$  gets a higher payoff in cooperation phases; however, those phases are now much shorter, as his opponent switches to state  $U$  with probability  $(1 - q)p$ . For  $p$  close to 1, a defection at  $N$  almost certainly generates a bad signal, which triggers a punishment phase of expected length  $1/q$  with probability  $1 - q$ , hence an expected cost

$$\Delta = \frac{1 - q}{q}$$

<sup>20</sup>This is because punishment lasts  $1 + T$  periods with probability  $q(1 - q)^T$ , and because

$$1 + \sum_T Tq(1 - q)^T = 1/q.$$

corresponding to a per-period decrease in payoff of 1 for an expected number of period equal to  $\Delta$ .

Deviation is deterred if

$$L < \Delta. \tag{2.1}$$

When player  $i$  plays  $C$  at both  $N$  and  $U$ , he avoids triggering some punishment phases, illustrated in figure 7.

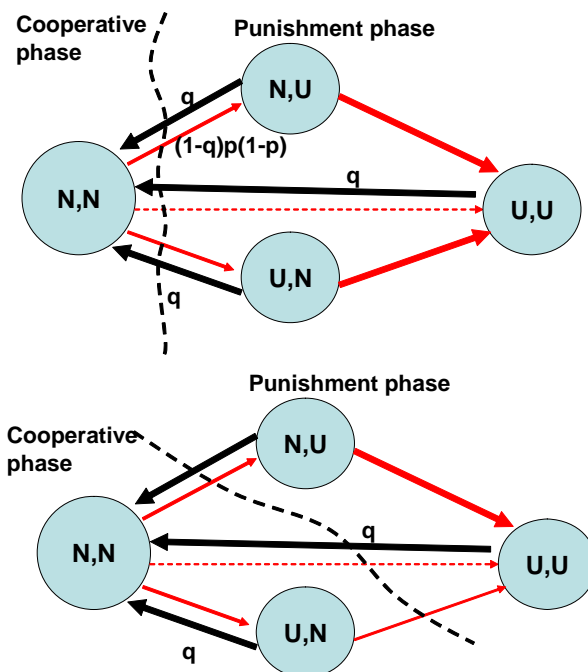


Figure 7:  $i$  cooperates always

However, he remains cooperative while player  $j$  is in a punishment phase. So  $L$  must be high enough for this option to be unattractive. More precisely, conditional on both players being in state  $N$ , there are events where only player  $i$  receives a bad signal, and events where only player  $j$  receives a bad signal.<sup>21</sup> Under the first event, player  $i$  initially gets 1 instead of  $1 + L$ , however he avoids the punishment phase, hence he makes a net gain of  $\Delta - L$ . Under the second event, nothing changes in the first period (because player  $i$  is still in state  $N$ ), but he then gets  $-L$  instead of 0 as long as the punishment phase continues,<sup>22</sup> hence an expected cost equal to  $L(\frac{1}{q} - 1) = L\Delta$ . Since these two events have equal probability, playing  $C$  at  $N$  and  $U$  is not a

<sup>21</sup>There are also events where both receive bad signal, but when  $p$  is close to 1, these are very unlikely events, and we can ignore them here. However, they would affect the computation in the general case where  $p$  is not close to 1.

<sup>22</sup>This is because when  $p$  is close to 1, player  $i$  switches to  $U$  with probability close to 1, hence he would have started playing  $D$  under the candidate equilibrium profile, while here, he does not.

profitable deviation if

$$\frac{1}{2}(\Delta - L) + \frac{1}{2}(-L\Delta) < 0, \text{ that is}$$

$$L > \frac{\Delta}{1 + \Delta}.$$

To summarize, for  $p$  close to 1, there is a range of parameters  $(q, L)$  for which the proposed strategy is an equilibrium strategy.<sup>23</sup> This range of parameters is such that  $\frac{\Delta}{1+\Delta} < L < \Delta$ , where  $\Delta = 1/q - 1$ . It is easy to check that outside this range, the only equilibrium entails both players defecting at both states.

Similar analysis can be performed for other values of  $p$ .

### 2.5. An Impossibility result

One might think that we obtain equilibrium cooperation in this example simply because we have severely restricted the set of strategies available to the players, thereby eliminating many potentially profitable deviations. We illustrate next that there are other seemingly reasonable two-state mental systems for which cooperation is (essentially) impossible in our environment.

Consider the forgiving transition described earlier, whereby a good signal makes a player switch back to the normal state.

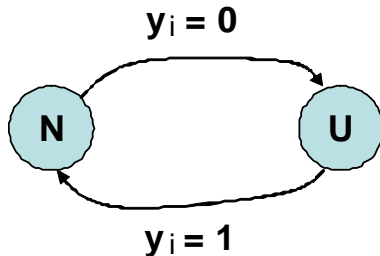


Figure 3: Forgiving transition

With such a mental process, for almost all values of  $p$  the only equilibrium entails both players defecting in both states.

**Proposition:** *If  $p \neq 1 - \frac{1}{2(1+L)}$ , and if each players' mental process is as defined above, then the only equilibrium entails defecting in both states.*

We provide a proof of this proposition in the appendix. The basic idea can be seen as follows. To fix ideas, consider the case where  $p$  is close to 1 and check that it cannot be an equilibrium that both players follow the strategy  $\sigma$  that plays  $C$  in  $N$  and  $D$  in  $U$ . If both players follow  $\sigma$ , then by symmetry, the induced ergodic distribution will put identical weight on  $(NN)$  and

<sup>23</sup>It is easy to check that playing  $D$  at  $N$  and  $C$  at  $U$  gives a lower value than always playing  $D$ .

( $UU$ ): the dynamic system has equal chances of exiting from ( $NN$ ) as it has of exiting from ( $UU$ ). As a result, players payoff will be bounded away from 1.

Consider next the case where player 1 deviates and plays the strategy  $\sigma^C$  that plays  $C$  at all states. There will be events where player 2 will switch to  $U$  and defect. However, since player 1 continues to cooperate, player 2 will soon switch back to  $N$  and cooperate. As a consequence, if player 1 plays  $\sigma^C$ , his payoff will remain arbitrarily close to 1. Hence it cannot be an equilibrium that both players play  $\sigma$ .

### 3. Discussion of example

Our leading example illustrates how cooperation can be achieved when strategies are constrained. Before going on, it is useful to compare this approach with the standard approach and discuss why cooperation is difficult when strategies are not constrained. We will then discuss richer mental processes.

#### 3.1. Comparison with standard approach

In the example, the strategies “play  $C$  when in  $N$  and play  $D$  when in  $U$ ” are in equilibrium for the parameters in the shaded region of figure 4: they are best responses to each other under the restriction that players use strategies in  $\Sigma_i$  that condition actions on states. These strategies however would not be best responses to each other without restrictions. Consider for example the first period following a punishment phase (players have just seen signal  $z = 1$  and have consequently just returned to state  $N$ ). Suppose player 1 receives a bad signal. Player 1 then transits to state  $U$ , and the equilibrium strategy in the example calls for him to play  $D$ , sending the play into the punishment phase again. However, at this first period after  $z = 1$ , player 2 was in state  $N$  as well, and most likely will remain in state  $N$  in the next period. So player 1 would be better off not triggering a punishment. In contrast, in our “constrained” approach, player 1 has incentives to trigger a punishment because he cannot condition his behavior on how he got to state  $U$ : all histories leading to  $U$  are pooled together, and on average across these histories, player 1 finds it optimal to play  $D$ .

To get a sense of how “un-pooling” histories complicates the analysis, let us try to construct a standard “unconstrained” equilibrium by considering a simple modification of our equilibrium strategy profile, allowing players to condition play on longer histories, and, for example, on whether they are just returning to  $N$ , or not. If equilibrium calls for cooperating following a return to state  $N$ , then a player would want to ignore signals in that first period following a return to state  $N$ . However, the best response to someone who behaves in that way is to play  $D$  in the first period following a return to state  $N$ . Equilibrium play in this first period following  $z = 1$  would thus have to involve mixed strategies, and later play would have to be finely tuned to the history of signals received subsequently. For example, a signal received in the first period following a return to  $N$  would have to be interpreted differently than a signal received in later periods, in a way that takes into account the fact that the opponent reacts differently to signals received shortly after a return to a cooperative phase than to signals received later. Thus, one cannot modify the strategies in our example in a simple way to effect a standard unconstrained equilibrium.

One difficulty in the search for strategies that support cooperation when strategies are not constrained (i.e., the standard approach) is precisely the possibility that beliefs are sensitive to the histories of signals received, and possibly the entire history of signals received. In general, there is then no natural way for players to be in one of a small number of states of mind in an equilibrium of the repeated game: distinct beliefs over the opponent’s continuation play are likely to generate distinct best responses, hence, according to our interpretation, distinct states of mind. But since the number of distinct beliefs typically grows over time, so does the number of states of mind.<sup>24</sup>

In contrast, our approach takes mental states and the transition across states as exogenous. In other words, we take as given the way players aggregate past information about play, that is, how they pool past histories. The strategic issue thus reduces to determining how one should play at each mental state, that is, choosing a mapping  $\sigma_i$  from states to actions, a much simpler task. In particular, our analysis is not based on a player forming beliefs about his opponent’s future play, and maximizing continuation payoffs given those beliefs.<sup>25</sup> Rather, we just compute the long-run value associated with any possible deviation.<sup>26</sup>

### 3.2. Adding more mental states

Players in our example had two possible mental states and two possible (pure) actions, which limited them to four pure strategies. This clearly limits them both in the existence of strategies that might lead to cooperation and in the possible profitable deviations. Adding a state can allow for more flexible reaction to signals that might permit cooperation which would have been impossible with only two states. For example, when signals are not very informative and  $q$  is small, the prescribed strategies in the example may not be an equilibrium. When there is substantial probability that I received a bad signal when the other player chose  $C$ , I might prefer not to trigger a punishment phase that could last for a long time. Hence the combination of relatively inaccurate signals of the opponent’s action and low probability of escaping the punishment regime may preclude an equilibrium in which the two cooperate.

Adding a state can allow cooperation that would be impossible with only two states for some parameters. Suppose there is a state  $M$  in addition to the states  $N$  and  $U$ , and define transitions as follows:

---

<sup>24</sup>See however the discussion of Mailath and Morris (2002), Phelan and Skrzypacz (2006) and Kandori and Obara (2007) in the related literature Section below. See also the literature on belief-free equilibria (e.g., Piccione (2002) and Ely and Valimaki (2002)).

<sup>25</sup>One can define rational beliefs for a player at a given informational state  $s_i$ , for a given profile  $\sigma$ , by considering the ergodic distribution over state profile  $\phi_\sigma(s_1, s_2)$  and by computing the conditional probabilities  $\Pr_{\phi_\sigma}\{s_j | s_i\}$ . Such a belief corresponds to the average of the beliefs given the possible histories leading to state  $s_i$  (the average taken with respect to the ergodic induced by  $\sigma$ ).

Note however that while beliefs can be defined, we don’t have in mind that players would know the mental process and strategy of the other player and use that knowledge to compute a belief. Beliefs should just be viewed as conditional probabilities, that is, as a way to understand correlations across each player’s state, under a given strategy profile.

<sup>26</sup>This long-run value does not depend on current beliefs because there is a unique ergodic distribution, and players are very patient.

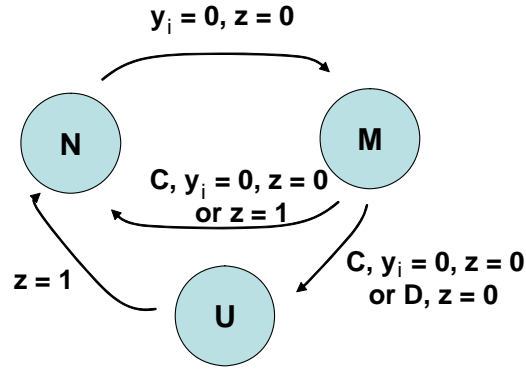


Figure 8: “Slow” transition to state  $U$

The interpretation of this is that if a player in state  $N$  gets a bad signal, he transits to state  $M$ , rather than state  $U$  as in the leading example. If the first signal a player in state  $M$  gets is bad, he transits to state  $U$ , and if it is positive, he returns back to state  $N$  as though the bad signal that sent him to  $M$  never occurred. Players remain in state  $U$  whatever action is played or private signal received, until signal  $z = 1$  occurs. Signal  $z = 1$  always sends players back to state  $N$ . The additional state, along with the amended transition function, allows for strategies that punish an opponent if there is some evidence he is not cooperating, but the evidence that triggers punishment can be either a single bad signal (e.g., the strategy plays  $C$  in  $N$  only, which we denote by  $\bar{\sigma}$ ), or two consecutive bad signals (e.g., the strategy plays  $C$  in  $N$  and  $M$ , and  $D$  in  $U$ , which we denote by  $\eta$ ). Computations as above show that the strategy profile  $(\eta, \eta)$  that entails both players cooperating in states  $N$  and  $M$  and punishing in state  $U$  is an equilibrium for some parameters for which cooperation is impossible with two states. The following figure illustrates this. We have set the value  $q$  to 0.3. For the set of parameters  $(p, L)$  in the shaded area,  $(\eta, \eta)$  is an equilibrium profile. For the parameters  $(p, L)$  between the top and bottom curved lines,  $(\sigma, \sigma)$  was an equilibrium profile in our simple two state case.

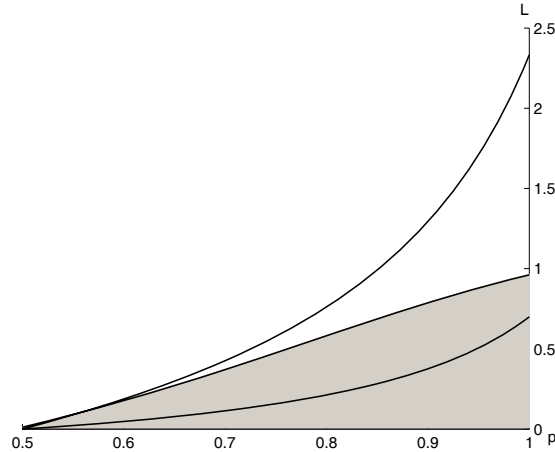


Figure 9: values of  $(p, L)$  for which  $(\eta, \eta)$  is an equilibrium when  $q = 0.3$ .

We make several comments about this example.

1. Intuitively, the additional state has two effects. Under the proposed strategy  $\eta$ , it allows for longer cooperation phases by reducing the probability that the relationship transits to a punishment phase (such transitions are inevitable even when players cooperate), and it allows players to be more cautious about deciding that their opponent has switched to  $U$  by waiting for two consecutive negative signals. This has two effects on incentives. It makes defecting in  $N$  more attractive because it takes two consecutive bad signals to make the other player switch to  $U$ . It also makes cooperating in  $U$  less attractive, because when a player eventually switches to  $U$ , it is very likely that the other player is already in  $U$ , hence incentives to play  $D$  in  $U$  are stronger. This is illustrated in figure 9: compared to the two-state case, the  $p - L$  region where cooperation is possible (under  $(\eta, \eta)$ ) shifts downward.

2. Allowing for a stochastic transition function is an alternative modification of the mental system in our leading example that could accomplish the two effects above. In the example, a player transits from state  $N$  to state  $U$  if he observes a bad signal (and  $z = 0$ ). Suppose the transition function is changed so that a player transits from state  $N$  to state  $U$  with probability  $1 - \mu$  after seeing a bad signal. This will make transition to  $U$  less likely, hence incentives to play  $D$  at  $N$  stronger. However, when a player finds himself in state  $U$ , it is more likely that his opponent is in state  $U$  as well, hence players will have stronger incentives to play  $D$  at  $U$ . Thus, for some parameters, this enables cooperation when it would not have been possible with the deterministic transition function in the example.

3. In addition to making cooperation possible for parameters under which it was impossible with only two states, a third state can increase the *value* of cooperation for parameters even if cooperation was possible with two states (i.e.,  $v(\eta, \eta) > v(\sigma, \sigma)$ ), as the punishment state is triggered less often.

4. Adding a state is not unambiguously good, however. As mentioned above, an additional state allows not only for more complex strategies to achieve cooperation, but more complex strategies for deviating as well. Despite the fact that both  $(\sigma, \sigma)$  and  $(\bar{\sigma}, \bar{\sigma})$  generate the same behavior (under both profiles, once a player observes a bad signal, he continues to defect until a signal  $z = 1$  arises), there are parameters for which  $(\sigma, \sigma)$  is an equilibrium, but  $(\bar{\sigma}, \bar{\sigma})$  is not. For  $(\bar{\sigma}, \bar{\sigma})$  to be an equilibrium, players must have an incentive to play  $D$  in  $M$ . This entails a strictly tighter constraint because playing  $C$  at  $M$  is less costly than playing  $C$  at  $U$ : in the event the opponent is already in  $U$ , the player is likely to get a negative signal and switch from  $M$  to  $U$  next period. For  $p$  close to 1 for example, the cost of playing  $C$  in  $M$  will be for only one period rather than for the duration of the punishment phase. Formally, the incentive constraint becomes:

$$\frac{1}{2}(\Delta - L) + \frac{1}{2}(1 - q)(-L) \geq 0.$$

Rearranging terms, we have

$$L \geq \frac{\Delta}{2 - q} = \frac{1 - q}{q(2 - q)},$$

a condition that is always more restrictive than  $L \geq \frac{\Delta}{1 + \Delta} = 1 - q$ . Figure 10 shows how the incentive condition changes for other values of  $p$ . (The  $p$  and  $L$  combinations for which cooperation was possible in the two-state case is again the region between the top and bottom curved lines.)

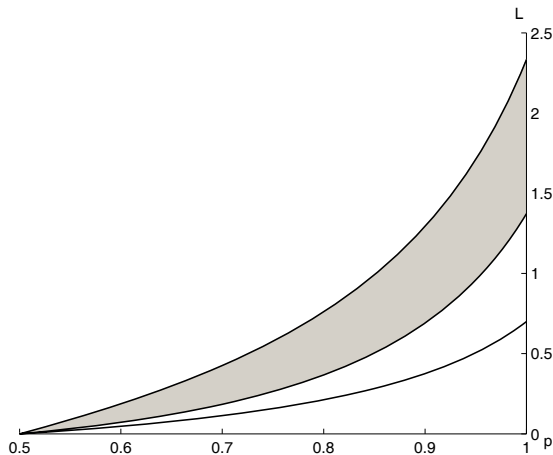


Figure 10: values of  $(p, L)$  for which  $(\bar{\sigma}, \bar{\sigma})$  is an equilibrium when  $q = 0.3$ .

### 3.3. Available strategies and One-shot deviations

We discussed above that if there were no restrictions on strategies, a player would like to deviate from the equilibrium strategy in the example by not playing  $D$  if he received a bad signal the first period after he returned to state  $N$ . Players in our framework are restricted to choosing the action they will play when in state  $N$  and the action they will play in state  $U$ . This restricts the set of strategies available, including the deviations available to them.

Our approach puts limits on the strategies available to players in various ways. First, as in standard Bayesian models, a player's state pools many histories of the game, and he is restricted to behave in the same way every time he finds himself in a particular state. For example, a player in our main example can choose to play  $C$  rather than  $D$  when he is in state  $U$ , but he cannot play  $C$  for some subset of the histories that lead to state  $U$  and play  $D$  for other histories that lead to  $U$ . If players are to behave differently for different subsets of histories, they must have different mental states that allow them to distinguish the sets. In the previous subsection, state  $M$  plays precisely that role. With this additional state, the player has available a strategy that allows him to play  $C$  after a bad signal (instead of going to state  $U$  and playing  $D$ ), and condition future play on the realization of next period's signal, i.e., reverting to state  $N$  after a good signal, or switching to state  $U$  after a second bad signal. For any expanded mental system, however, there will always be strategies for which deviations from that strategy cannot be accomplished within that system.

Second, some deviations are not available because current actions have consequences on the player's continuation state of mind. In the previous subsection, if a player play  $D$  in state  $M$ , he moves to state  $U$ . He cannot decide, upon seeing the signal  $y_i$ , whether he wants to continue to behave as though he were still in state  $N$ . He will be in state  $U$  regardless because in this example, a defection triggers a switch to state  $U$ . In other words, as in models with imperfect recall, after playing  $D$  in  $M$ , the player will be in state  $U$  and does not distinguish whether he is in  $U$  because he received two consecutive bad signals or because he played  $D$  in  $M$ .

Because of these restrictions, we may find equilibria for which there would be a profitable one-shot deviation; the player may simply not have available a strategy that would mimic the standard one-shot deviation.

Conversely, one might expect that a version of the one-shot deviation principle would hold in our framework: if it does not pay for me to deviate one time when I am in state  $U$  (and play  $C$  once), then it should not be the case that it is profitable for me to deviate each time I am in state  $U$ . This is not the case however. Returning to our main example, suppose that I am in state  $U$  and I consider deviating from my strategy one time only and playing  $C$ . Returning to my equilibrium strategy after this involves my playing  $D$  while in  $U$ . At best, this will lead to a delay in entering the punishment phase by one period. On the other hand, if I play  $C$  each time I am in  $U$  I avoid completely triggering the punishment phase any time only I receive a bad signal. Formally, deviating one time at  $U$  (and playing  $C$ ) is not profitable when

$$\Pr_{\sigma}(s_j = U \mid s_i = U)(-L) + \Pr_{\sigma}(s_j = N \mid s_i = U)(q(-L) + (1 - q)(1)) < 0$$

or equivalently,<sup>27</sup> when

$$L > \frac{q}{q^2 + 2(1 - q)}(1 - q).$$

This condition is less stringent than the equilibrium condition we found earlier.

*Comments.*

1. If a player has a profitable one-shot deviation in some state, he will have a profitable mixed strategy deviation. Indeed, choosing the action that would be a profitable one-shot deviation with small probability each time he is in that state will give a gain to the player (without the cost entailed in playing that action with probability one each time the player is in that state). If we did not restrict to pure strategies, or if we allowed players to play differently with small probability, we would derive stronger constraints, and in equilibrium, there would be no profitable one-shot deviations.

2. In our framework, one could define an alternative notion of a one-shot deviation whereby a player could decide to change state once (from  $U$  to  $N$ , say, and continue behaving as if he were now in state  $N$ ). Such deviations could be profitable, depending on whether it would be profitable for a player to add a small amount of noise to a particular transition.

### 3.4. Robustness

The example was kept simple in a number of ways to make clear how cooperation could be achieved when strategies are restricted. Some of the simplifications are not particularly realistic, but can be relaxed without affecting the basic point that cooperation is possible even when agents get private signals if strategies are restricted. In Compte and Postlewaite (2008) we show that the assumption that the players simultaneously chose actions in the basic stage game can be

---

<sup>27</sup>Note that when  $p$  is close to 1,

$$\phi_{\sigma}(UU) = (1 - q)[\phi_{\sigma}(UN) + \phi_{\sigma}(NU) + \phi_{\sigma}(UU)],$$

which implies, since  $\phi_{\sigma}(UN) = \phi_{\sigma}(NU)$  by symmetry,

$$\frac{\Pr_{\sigma}(s_j = U \mid s_i = U)}{\Pr_{\sigma}(s_j = N \mid s_i = U)} = \frac{\phi_{\sigma}(UU)}{\phi_{\sigma}(NU)} = \frac{2(1 - q)}{q}$$

relaxed so that their choices are sequential without altering the qualitative conclusions about cooperation. That paper also shows that it is straightforward to extend the analysis to problems in which agents are heterogeneous in costs and benefits, and in which agents are heterogeneous in their monitoring technologies. In general, agent heterogeneity typically restricts the set of parameters for which cooperation is possible.

The extensions examined in Compte and Postlewaite (2008) are not meant only as a robustness check though. As mentioned in the introduction, our goal is a realistic description of cooperation when people are strategic and the structure of the games they play varies. In the face of the variety of the games we play, players’ mental processes should be viewed as the linchpin of cooperation. These extensions are meant to capture the scope of a given mental process.

#### 4. When $z$ is not public: Resetting the relationship to cooperation

A central issue in relationships where monitoring is private is ensuring that players have incentives to trigger punishments. When monitoring is public, all players see the same signal, so there is no problem in coordinating a punishment phase: it is common knowledge what other players know. The issue in private monitoring is as in our example – when a player gets a bad signal it is equally likely that the other player may or may not have already received a bad signal, making it a nontrivial decision for a player to begin a punishment phase. The lack of common knowledge among players results in a second key issue – how do players get back to cooperation once a punishment has been triggered. We finessed the second issue in the analysis above by assuming the public signal  $z$  that facilitated recoordination after a punishment period. As we mentioned, a public signal  $z$  is a relatively simple way for the players to recoordinate, but not necessary. We demonstrate next how players can coordinate a move back to cooperation in the absence of a public signal: we first examine the case where  $z$  is “almost public”; then, illustrating the difficulties of recoordination, we examine the case where each player receives signals  $z_i$  that are independently distributed and show that the mental system in our leading example no longer allows for cooperation. Finally, we show how that simple mental system can be amended to support cooperation.

##### 4.1. Cooperation with “almost public” signal $z$

Because of the noisiness in the private signals the players get, the relationship will periodically fall into disrepair with neither exerting effort in choosing a gift. Without some way to “reset” the relationship, this would be an absorbing state with no further cooperation. The public signal  $z$  allows the players to recoordinate and start cooperation afresh. The limits on the probability of the public (or nearly public) signal are intuitive: if  $q$  is too low, players may prefer not to punish when they get a signal that the other has not put in effort, and if  $q$  is too high the punishment phase will not be sufficiently painful to deter deviation.<sup>28</sup>

It isn’t essential, however, that the signal be precisely public. If the parameters  $(p, q, L)$  are such that incentives are strict, then by continuity, incentives will continue to be satisfied if each

---

<sup>28</sup>A richer mental system can allow for cooperation even when the probability of the public signal is too large or too small for the two state mental system in the example. We discuss this below.

player receives a private signal  $z_i \in Z_i = \{0, 1\}$  with the property that  $\Pr(z_i = 1) = q$  and  $\Pr(z_1 = z_2)$  close enough to 1. As we show below however, the correlation between signals in this information structure cannot be too weak if cooperation is to be possible.

#### 4.2. Impossibility of cooperation with independent signals $z_1, z_2$

We next show that when  $p$  is close to 1, cooperation cannot be sustained in equilibrium if the signals  $z_1$  and  $z_2$  are independent.<sup>29</sup> Formally, we consider the mental process as before with the qualification that  $T_i$  is now defined over  $Z_i$  rather than  $Z$ . We investigate whether and when the strategy  $\sigma$  that plays  $C$  in  $N$  and  $D$  in  $U$  is an equilibrium.

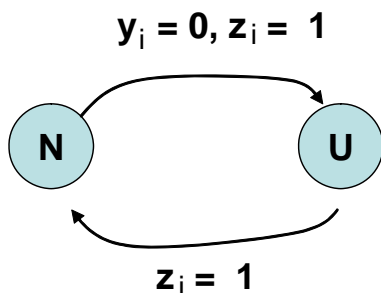


Figure 11: Independent “resetting”

**Proposition:** Fix the mental system as above. For any fixed  $q \in (0, 1)$ , for  $p$  close enough to 1, the strategy profile where each player cooperates in  $N$  and defects in  $U$  cannot be an equilibrium.

Assume  $p$  is close to 1. Under  $\sigma$ , cooperation phases last  $\frac{1}{2(1-p)(1-q)}$  on average (because each player has a chance  $(1-p)(1-q)$  of switching to  $U$  in each period), and punishment phases last  $\frac{1}{q^2}$  (since only in events where both players get signal  $z = 1$  at the same date that recoordination on cooperation is possible).<sup>30</sup> So for  $p$  close enough to 1, the value to following the proposed strategy profile is 1.

Compared to the case where  $z$  is public, the incentives to play  $C$  at  $N$  are unchanged: if player 1 plays  $D$  at both states, his opponent will be cooperative once every  $1/q$  periods on average, hence the condition

$$L < 1/q - 1 \tag{4.1}$$

still applies.

Incentives to defect at  $U$  however are much harder to provide. As before, by cooperating at  $U$ , player 1 ensures that a punishment phase is not triggered in the event state profile is  $UN$ . But there is another beneficial effect. In the event state profile  $UU$  occurs, the punishment phase

<sup>29</sup>Although we do not include the computations here, it can be shown that the result holds for all values of  $p \in (\frac{1}{2}, 1)$ .

<sup>30</sup>Omitting terms comparable to  $(1-p)$ , this is true whether the current state profile is  $(U, U)$ ,  $(U, N)$  or  $(N, U)$ .

that follows will last only  $1/q$  periods (as simultaneous transition to  $N$  is no longer required). So player 1 will only have incentives to defect at  $U$  if:

$$\frac{1}{2}L(1/q) > 1/q^2,$$

or equivalently

$$L > 2/q,$$

a condition which is incompatible with inequality (4.1).

Thus strategy  $\sigma$  cannot be an equilibrium: the length of punishment phases is substantially reduced when playing  $C$  at  $U$ , which makes playing  $C$  at  $U$  an attractive option.

### 4.3. Resetting without a public signal

In many situations, there may be no public signal that can facilitate the coordination back to cooperation. How can players then coordinate a move back to cooperation? We propose below a modification of our basic mental process that permits such coordination.

We consider a mental process with three states  $N, U$  and  $F$ . Transitions from  $N$  to  $U$  are as before. Transitions from  $U$  to  $F$  can be thought of as a player forgetting ( $F$ ) being upset over time. This is modelled as a probability of receiving a private signal  $z_i = 1$ , with  $z_i \in \{0, 1\}$  and  $\Pr\{z_i = 1\} = q$ . Signals are assumed to be drawn independently.

A key feature of our modified transition functions is what triggers a change from  $F$  to  $N$ ; what triggers the transition differs between the two players. For player 2, such a transition requires receiving a good signal,  $y_2 = 1$ . For player 1, such a transition is automatic. These transitions are summarized in Figure 12.

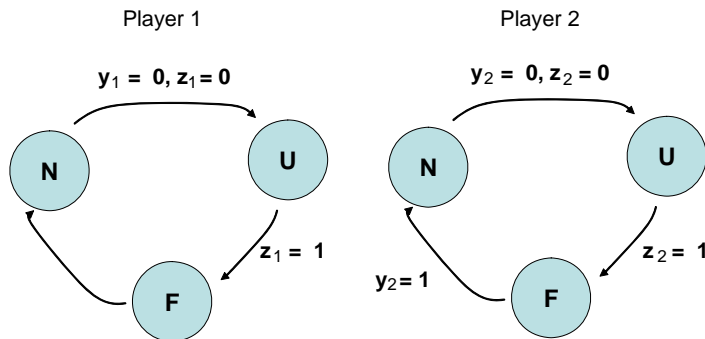


Figure 12: “Successful” independent resetting

We show in the appendix that there is a range of parameters for which it is an equilibrium strategy for player 1 to cooperate at  $N$  and  $F$ , and for player 2 to cooperate at  $N$  only.

In the case that  $p$  is close to 1, cooperation can be sustained for the  $q$  and  $L$  combinations in the shaded region in figure 13 below.

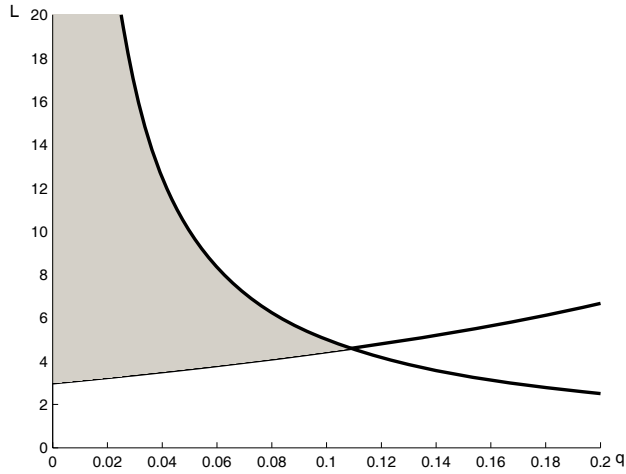


Figure 13: Values of  $q$  and  $L$  for which cooperation is possible (for  $p$  close to 1).

The key difference with the previous analysis is that players no longer simultaneously switch back to cooperation (because there is no public signal to allow that). Rather, the transition functions are as though one player acts as a leader in the relationship, and makes a gift (i.e., cooperate) as a signal that he understands that the relationship has broken down and needs to be restarted.

Intuitively, incentives to play  $C$  at  $F$  are easy to provide for player 1: When player 1 is in  $F$ , the other player has a non negligible chance (approximately equal to  $1/2$  if  $q$  is small) of being in  $F$  as well, hence playing  $C$  in  $F$ , though costly, generates a substantial chance of resetting cooperation. In contrast, incentives to play  $C$  at  $U$  are much weaker: playing  $C$  at  $U$  would also allow player 1 to reset cooperation in the event of a breakdown, but this would be a very costly strategy as it requires player 1 to possibly cooperate during many periods before player 2 switches back to  $N$ .

## 5. Further discussion

### 5.1. Learning to cooperate.

We emphasized above that one the desiderata for a satisfactory positive model of cooperation is that players should be able to *learn* how to cooperate. We describe here in more detail how players in our model might achieve that .

Our model of behavior has two levels: the mental process, which determines the set of strategies that a player might be using, and the strategy actually used, which is a mapping from mental states to actions. Our view is that mental processes, if not fixed, undergo changes at a slow pace while alternative strategies are sampled relatively often.<sup>31</sup>

<sup>31</sup>Note that we do not have in mind that players know everything about the structure of the interaction and the other player's mental process or strategy. Rather, we have in mind that, among the few strategies that are

Taking fixed the mental process (the possibility that mental processes evolve will be discussed below), we thus have in mind that players follow an experimental rule of thumb, whereby each starts with a particular strategy, and with small probability, randomly and independently samples alternative strategies to assess which perform best.<sup>32</sup>

We can distinguish two cases. First, players may currently be following strategies that are *not* in equilibrium, in which case experimentation will soon lead one of them to change strategy. Second, they are currently following strategies that *are* in equilibrium, and likely continue playing these strategies unless noise in evaluating performance, or simultaneous occurrence of experimentation, causes a player to switch strategies. These observations suggest that players will spend most of their time in an equilibrium of the game, jumping from one equilibrium to another depending on how noisy performance evaluation is.

Applied to our game, this implies that (i) for given parameters  $p, L$  and  $q$  that fall outside the region where  $(\sigma, \sigma)$  is an equilibrium, players will spend most of the time playing the strategy  $\sigma^D$  that plays  $D$  in both  $N$  and  $U$ , and that (ii) for given parameters  $p, L, q$  that fall inside the region where  $(\sigma, \sigma)$  is an equilibrium, players will spend most of the time alternating between  $(\sigma, \sigma)$  and  $(\sigma^D, \sigma^D)$ . This is a fairly standard model of experimentation for static games, and the prediction is consistent with that obtained in models such as Kandori, Mailath, and Rob (1993), with the qualification that there is no *a priori* reason to think that noise in evaluation would be small.

Our model is one of repeated games, but the simplification in the strategy space induced by a player's mental system allows us to analyze it as if it were a static game, and to obtain predictions as to whether players might learn to cooperate. It is unrealistic to assume that a player would be able to experiment with all the repeated game strategies (defined in the standard sense), and a mental system can be viewed as a way to guide the players as to which strategies he might sample.

>From an outsider's perspective, these constraints can be interpreted as a partial commitment, that is, a commitment to experiment only within a subset of the entire set of the strategies of the repeated game. The mental system puts some constraints on the feasible partial commitment, and our analysis shows that some partial commitments may allow for cooperation.

## 5.2. Evolution of mental system.

We have taken the mental system – the states and transition function – to be exogenously given. We did, however, suggest that one might think of these as having been formed by environmental factors. In the long run, evolution might influence both the set of mental states that are possible and the transition function. While beyond the scope of this paper, it would be interesting to understand how evolution shapes mental systems. We have pointed out above that it is *not* the case that evolution should necessarily favor more complicated mental systems; adding more states to a mental system that allows cooperation might make cooperation then impossible.

It is beyond the scope of this paper to analyze completely the evolutionary forces on mental processes, but we will illustrate why evolution does not necessarily lead to more mental states

---

available to them, they try them and learn which works best for them. No knowledge of the structure of the game or of the parameters is required for this.

<sup>32</sup>When experimentation occurs, it should last long enough for a player to get an accurate idea of which strategies perform well, and the probability of starting to experiment should be small enough that there is sufficient time to profit from the identification of good strategies.

with a simple example.<sup>33</sup> To assess evolutionary pressures away from the mental process in our leading example, we evaluate how a mutant, endowed with the process that has slow transition to the upset state, would fare against the mental process in our leading example. Using the notation of Section 3, there are two possible candidate equilibria:  $(\sigma, \bar{\sigma})$ , whereby each player defects after a single negative signal, and  $(\sigma, \eta)$ , whereby the mutant defects after two consecutive bad signals (player 2 is the mutant). In the first case, the conditions for which cooperation is possible are identical to those shown in Figure 10 (because  $\sigma$  and  $\bar{\sigma}$  generate the same behavior). In the second case, player 1 has a greater incentive to defect in  $N$  because it requires two negative signals for player 2 to switch to defect. Player 1 also has a greater incentive to cooperate in  $U$ : because it is more likely that player 1 switches to  $U$  first, there is a greater incentive for player 1 to avoid triggering a punishment, hence to play  $C$  in  $U$ . The consequence is that these incentive constraints may be incompatible, and hence, no parameters for which  $(\sigma, \eta)$  is an equilibrium. This is illustrated in Figure 14 for the case  $q = 0.3$ . The upper curve is the lower bound on  $L$  for which player 1 has an incentive to play  $D$  at  $U$  for any  $p$ . The lower curve is the upper bound on  $L$  for which player 1 has an incentive to play  $C$  at  $N$  for any  $p$ .

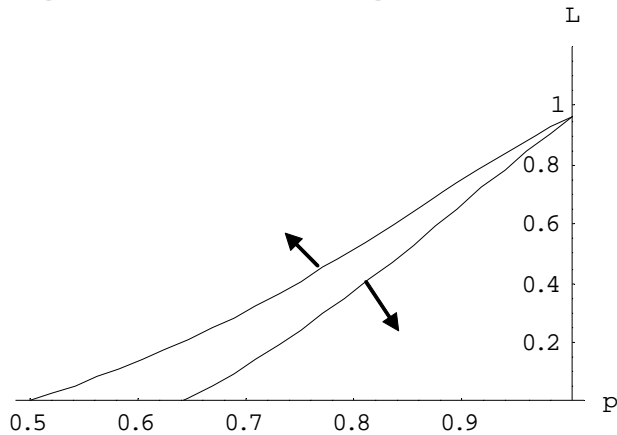


Figure 14: Incentive constraints for player 1 when  $(\sigma, \eta)$  is played

Overall, the mutant thus performs strictly worse than the regular player: because only  $(\sigma, \bar{\sigma})$  can be an equilibrium, his equilibrium payoff is no different than if he was a regular player when cooperation is achieved, but there are fewer  $(p, L)$  combinations for which cooperation can be supported.

While evolution might shape the mental systems one would expect to see in the long run, cultural transmission would logically be a way mental systems are shaped in the short run. We have illustrated the notion of mental systems with our leading example in which there is a single transition function of interest, but it is easy to see that it might be the case that there are several transition functions that allow cooperation in at least some circumstances. Models such as this might be useful in understanding why behaviors that are innocuous in some cultures trigger stronger reactions in others.

<sup>33</sup>Evolutionary pressures on mental systems are analyzed in more detail in Compte and Postlewaite (2009).

### 5.3. Extensions of the model

1. Our model has no explicit communication. We discussed how resetting might be accomplished without a public signal in section 5.3. Player 1 was exogenously designated as the leader in the relationship, and one can interpret the decision to play  $C$  in state  $G$  as an implicit communication that the relationship should restart. There, communication was noisy (because the other player does not receive a nice gift with probability one) and costly (it costs  $L$ , and  $L$  cannot be too small if cooperation is to be achieved).

One could mimic the implicit communication in this example with explicit communication by allowing player 1 to send a message at the end of each period, and by defining the transition induced by the message, if sent, on both players' mental states. We mention two examples of how one might define such transitions and how communication affects the possibility of cooperation.

Following the logic of the example in Section 4.3, we could assume that player 1 always sends a message when he is in state  $F$ , and if in state  $F$ , player 2 will transit from state  $F$  to state  $N$  after receiving a message. The discussion of that example suggests that communication can help here only if sending the message is sufficiently costly.

As a second example of how one could add the possibility of player 1 being able to explicitly send a message, consider the two state mental system without any resetting illustrated in Figure 2. Suppose that the transition function is such that when player 1 sends a message, player 2 will switch states; that is, the message makes a Normal player 2 Upset, and an Upset player 2 Normal. Communication here might be helpful even if there is no exogenous cost to sending the message since sending a message while in the normal state involves an endogenous cost: if player 2 is in  $N$ , sending a message will result in her being Upset next period.

2. There is no utility attached to mental states in our model; the states  $U$  and  $N$  are no more than collections of histories. It is straightforward to extend our model to the case in which utility is attached to states, or to particular transitions between states (going from upset to normal, for example). Strategies that constitute a strict equilibrium in our model would remain an equilibrium for small values attached to utilities in particular states or movements among states, but large values might affect the possibilities of cooperation. Suppose that in our main example, there is a negative utility attached to being in state  $U$  (relative to being in state  $N$ ). When player 1 is then in state  $N$ , deviating will be more costly than in the case we analyzed. Consequently, there will be a larger set of parameters for which players will not choose the strategy of playing  $D$  in both states. What if player 1 plays  $C$  in both states? It is easy to see that doing so will not decrease the proportion of time that 1 is in state  $U$ : whether he plays  $C$  or  $D$  in state  $U$ , he remains in that state until  $z = 1$ . Hence, playing  $C$  in both states will neither decrease the chance of transiting from state  $N$  to  $U$ , nor will it quicken his departure from  $U$ . Summarizing, modifying the example so that there is a decrease in utility when a player is in state  $U$  (or suffers a decrease in utility when he transits from  $N$  to  $U$ ) can only increase the set of parameters for which the cooperative strategies are an equilibrium.

3. One interpretation of mental states is that they are driven by emotions. A natural example of such a mental state is Guilt ( $G$ ). If our leading example is altered so that each player  $i$  observes not only his own signal  $y_i$  but also the other player's signal  $y_j$ , player  $i$  will know when player  $j$  did not like the gift, and player  $i$  might feel Guilty – an additional mental state ( $G$ ). If transitions to  $G$  from  $U$  or  $N$  never occur and transitions between  $U$  and  $N$  are as before (hence never take the other's signal into account), our analysis is unchanged: our

equilibrium would remain an equilibrium. However, if transitions to the state do  $G$  occur, then other outcomes may be sustainable as equilibria (equilibria similar to the public equilibria, for example).

More generally, our approach suggests thinking first of mental states, then of transitions that seem reasonable given the mental states, and ultimately of a class of games for which the corresponding mental system is useful.

4. Transitions over mental states have been defined as a function of a signal about how the game is played. If mental states are interpreted as emotional states, one can think of a player as receiving a signal of his opponent’s mental state, and that his own mental state as being affected by this. For example, if one perceives that the other is Upset, then one might feel Guilty (or maybe not). Of course, this transition to the guilty state may trigger a subsequent change in the opponent’s state, and so on. This suggests an extension where the actual signal structure depends not only on actions, as in standard repeated games, but also on the way histories are mapped into mental states.

5. In Compte and Postlewaite (2008) we extend the basic structure and ideas in this paper to the case of many agents who are randomly matched. As is intuitive, the range of parameters for which cooperation is possible is smaller than in the two-person case because there is a longer time between a player’s first defection and when he first meets opponents who do not cooperate as a result of his defection.

6. Compte and Postlewaite (2008) show how the apparatus in the paper can be employed to model social norms. There may exist a “norm” that prescribes acceptable behavior for a wide variety of problems, with agents receiving noisy signals about whether their partner has followed the norm or not. Two-state mental systems will allow support of the norm in a manner similar to the cooperation that is possible in the model we analyze in this paper. Agents will follow a norm’s prescriptions when they are in the “normal” state, and behave in their own self interest following observations that suggest their partner has violated the norm.

#### 5.4. Experiments

Our aim has been to set out a model that allows cooperation in ongoing relationships via behavior that can be represented by realistic strategies. The goal is to provide a richer conceptual framework than standard repeated games permit. While our goal was primarily conceptual, the model may suggest experiments. If people are allowed to play a repeated game with private monitoring similar to the one we analyze, in principle one could examine the actions people play and see whether they can be reasonably be structured as our model suggests. That is, one can ask what histories trigger a deviation from cooperation, whether those histories can be classified as a mental state, and whether the same histories continue to structure behavior as parameters of the game vary.

There have been experiments that investigate how players behave in repeated prisoner’s dilemma games with uncertainty of different kinds. In general, cooperation is less likely when actions are “noisy”, that is, there is randomness in what players observe about what their opponents have done<sup>34</sup>, and when the participants receive random payoffs. However, in these experiments, while players may not see exactly what their opponent has done, they know precisely what signals the opponent has seen. There has been relatively little work done on what

---

<sup>34</sup>See, e.g., Bendor et al. (1999) and Sainty (1999).

we have argued is the realistic setting in which players do not observe perfectly the actions taken by their opponents. Miller (1996) has run computer tournaments in which strategies for playing repeated prisoner’s dilemma games with uncertainty are pitted against each other, with more successful strategies “multiplying” more rapidly than less successful strategies. The form of the uncertainty that Miller considers is essentially as in our model. The action played in any period is seen with some noise by the opponent. Miller compares the strategies that succeed in this tournament when there is no noise in the observation of the opponent’s action and when the signals are 99% accurate and 95% accurate. While there are differences in this particular exercise and our model,<sup>35</sup> it is interesting to note two things. First, the successful strategies that emerge are simpler than they might have been in the sense that they use fewer states than was possible. This suggests that there is an “optimal” number of states in the sense that strategies that used more states did strictly worse. The second interesting thing about Miller’s results is that when the error rate in the signals of what an opponent has done increases from 1% to 5%, the number of states used by the most successful strategies decreases. While we don’t take this as compelling evidence for our general point of view, it is consistent with our framework.

### 5.5. Related literature

Although we have emphasized the difficulty in supporting cooperation when signals are private, there are monitoring structures for which cooperation is relatively easy to sustain. These are case when each player can be sure – or almost sure – of his opponent’s state of mind. Mailath and Morris (2002) analyze repeated games in which players get imperfect information about past play. Consider first games in which players get a public signal, that is, a signal that both see. Cooperation may be possible here, because if players choose pure strategies, they know exactly their opponents’ state of mind: the public signals and the actions prescribed by the equilibrium strategies will be common knowledge. Mailath and Morris then consider a perturbation of this information structure whereby each player gets the public signal with a small amount of noise. They focus on the case that players’ signals are *almost public*: for any signal a player receives, the probability that other players have received the same signal is close to one. Mailath and Morris show that if players’ strategies depend only on a finite number of past signals, the introduction of the small amount of noise into the players’ signals about past play doesn’t matter; strategies that comprise an equilibrium when there is no discrepancy in what players observe remain equilibrium strategies when players’ signals are almost public.<sup>36</sup>

This result provides sufficient conditions under which cooperation can be possible even when players receive private signals, but those conditions are quite stringent. In particular, when signals are almost public, I can predict very accurately what other players will next do. This is in sharp contrast to our example. First, the signals that players get are not helpful in predicting the signal received by the other player, and second, however accurate signals are, there are times (in state  $U$  for example) when a player may not be able to accurately predict what his opponent will do.<sup>37</sup>

<sup>35</sup>Miller considers finitely repeated games and allows transition functions to vary, for example.

<sup>36</sup>Phelan and Skrzypacz (2006) ask a related question. They consider strategies that depend on a finite number of states, but unlike Mailath and Morris, conditioning on past actions is allowed. Then, for any given monitoring structure, by looking at the set of beliefs generated after any history, they provide a general method for evaluating whether the candidate strategy is an equilibrium. See also Kandori and Obara (2007) for a related treatment.

<sup>37</sup>This is because even as  $p$  gets close to 1, the probability  $\Pr(s_2 = U \mid s_1 = U) = \phi_\sigma(UU)/(\phi_\sigma(UU) + \phi_\sigma(UN))$

It has long been understood that some Nash equilibria are sufficiently complicated that it is implausible that players will be able to identify the strategies and play them. One approach to taking complexity of strategies into account is to assume players use finite automata to implement their strategies. A finite automaton consists of a finite number of states and a transition function, as in our model. The complexity of a player’s strategy is defined to be the minimal size of a machine that can implement that strategy.<sup>38</sup> We differ from this literature in several respects. First, the literature using automata to implement strategies has players choosing both the transition function and the mapping from states to actions, taking fixed only the number of states available given the automaton’s size. In contrast, we take players’ transition functions as fixed with players’ choices being only the mapping from states to actions. Second, to our knowledge, this literature does not consider games with private monitoring. Third, the earlier literature used automata primarily as a tool to capture complexity; our modeling strategy takes more seriously mental systems as being a plausible, if crude, model of the process by which players may interact. There has also been work in single-person decision making problems that is analogous to the papers using automata to capture complexity costs.<sup>39</sup> While we take agents’ transition functions as fixed, the focus of this literature is on characterizing the optimal transition rule.

In our model a player must choose an action at a given date that depends only on which of the finite number of states that he is in at that time. The number of realized histories goes to infinity so a state is an information set that over time contains a large set of histories, and a player might prefer to choose different actions at different histories that led to a given state if he could distinguish the histories. This structure is analogous to the “absent-minded driver problem”<sup>40</sup> in which a driver who wishes to exit a limited access highway “at the second exit” must decide what to do when he arrives at an exit but cannot recall whether he has already passed an exit.

The model we study reduces to a stochastic game of a particular kind in which each player has his own state variable, and each player observes only his own state.<sup>41</sup> Most of the literature in stochastic games assumes that in each period, there is a state that is known to both players,<sup>42</sup> while our interest is in the case that players do *not* know their partner’s state. In addition, our interest is primarily in understanding the conditions under which cooperation is possible for specific classes of games.

Kandori (1992) and Okuno-Fujiwara and Postlewaite (1995) analyze the possibility of cooperation when players are repeatedly randomly matched to play a prisoner’s dilemma game with limited information about previous play. Those papers demonstrate that cooperation can be supported with limited information, but the nature of the limited information is different from that in this paper. In particular, there is no uncertainty about the action that an individual has taken in a given interaction unlike in our model.

Abdulkadiroglu and Bagwell (2007) analyze a two-person favor exchange model in which in each period one or the other, or neither, of the agents may be able to do a costly favor for the

---

remains bounded away from 0 and 1. (See also footnote 27.)

<sup>38</sup>See, e.g., Ben-Porath (1986) and Abreu and Rubinstein (1988).

<sup>39</sup>See Wilson (2004) and Cover and Hellman (1970) for such models of single-person decision problems and Monte (2007) for a strategic treatment of such models.

<sup>40</sup>See, e.g., Piccione and Rubinstein (1997).

<sup>41</sup>We thank Eilon Solan for this observation.

<sup>42</sup>However, see Altman et. al. (2005).

other agent.<sup>43</sup> When in a given period an agent has the opportunity to do a favor, this fact will be known to him but not to the potential recipient of the favor. If a favor is done, it is publicly observed by both agents. It is assumed that the benefit to the recipient of a favor is greater than the cost of doing the favor, it is efficient that all possible favors be done. It is straightforward to see that there is an incentive problem in inducing an agent to do a costly favor when his opponent will not know that such a favor was possible however. Abdulkadiroglu and Bagwell analyze the equilibria in this problem and demonstrate how relatively simple strategies can support significant cooperation. This problem differs from ours in the nature of the asymmetry of information. In the problem Abdulkadiroglu and Bagwell analyze, agents do not know precisely the game they face in any period. If an agent cannot do a favor for his opponent, he does not know whether or not his opponent can do a favor for him. But when an agent chooses to do a favor, both agents observe this with no uncertainty. In contrast, in our model agents know precisely the game in any period, but get an imperfect signal of the action taken by their opponent. It is this latter imperfection that makes coordination and cooperation in our framework so difficult.

## 6. Appendix

### No cooperation under the forgiving transition.

Consider first the case where player 2 follows the strategy  $\sigma$  that plays  $C$  in  $N$  and  $D$  in  $U$ . If player 1 adopts the same strategy, then by symmetry, the induced ergodic distribution puts identical weight on  $(NN)$  and  $(UU)$  on one hand, and on  $(NU)$  and  $(UN)$  on the other hand. (Intuitively, the dynamic system has equal chances of exiting from  $(NN)$  as it has of exiting from  $(UU)$ .)

$$\phi_{\sigma,\sigma}(NN) = \phi_{\sigma,\sigma}(UU) \text{ and } \phi_{\sigma,\sigma}(NU) = \phi_{\sigma,\sigma}(UN) \quad (6.1)$$

The value to player 1 from following that strategy is thus

$$\begin{aligned} v(\sigma, \sigma) &= \phi_{\sigma,\sigma}(NN) + (1+L)\phi_{\sigma,\sigma}(UN) - L\phi_{\sigma,\sigma}(NU) \\ &= \phi_{\sigma,\sigma}(NN) + \phi_{\sigma,\sigma}(UN) = \frac{1}{2}(\phi_{\sigma,\sigma}(NN) + \phi_{\sigma,\sigma}(UN) + \phi_{\sigma,\sigma}(NU) + \phi_{\sigma,\sigma}(UU)) \\ &= \frac{1}{2}. \end{aligned}$$

Now if player 1 cooperates in both states ( $\sigma^C$ ), player 2 will switch back and forth between states  $N$  and  $U$ , spending a fraction  $p$  of the time in state  $N$ . The value to player 1 from following that strategy is thus:

$$v(\sigma^C, \sigma) = p + (1-p)(-L)$$

and it exceeds 1/2 if

$$p > 1 - \frac{1}{2(1+L)}.$$

If player 1 defects in both states ( $\sigma^D$ ), player 2 will again switch back and forth between states  $N$  and  $U$ , but now spending a fraction  $1-p$  of the time in state  $N$ . The value to player 1 from

---

<sup>43</sup>Mobius (2001) and Hauser and Hopenhayn (2005) analyze similar models.

following that strategy is thus:

$$v(\sigma^D, \sigma) = (1 - p)(1 + L)$$

which exceeds  $1/2$  as soon as  $p < 1 - \frac{1}{2(1+L)}$ .

Finally, if player 1 follows the strategy  $\hat{\sigma}$  that plays  $D$  in  $N$  and  $C$  in  $U$ , then, as above, the dynamic system has equal chances of exiting from  $(NN)$  as it has of exiting from  $(UU)$ . Therefore, equalities (6.1) hold for the profile  $(\hat{\sigma}, \sigma)$ , and the value to player 1 from following  $\hat{\sigma}$  thus remains equal to  $1/2$ . It follows that unless  $p = 1 - \frac{1}{2(1+L)}$ , the strategy profiles  $(\sigma, \sigma)$  and  $(\hat{\sigma}, \sigma)$  cannot be equilibria. Similar considerations show that the strategy profile  $(\hat{\sigma}, \hat{\sigma})$  cannot be an equilibrium. As a result, only strategy profiles that are constant across states may be in equilibrium, hence the only equilibrium entails defecting in both states.

### Resetting without a public signal

Our candidate equilibrium strategy pair is as follows. For player 1,

$$\sigma_1(N) = C, \sigma_1(U) = D \text{ and } \sigma_1(F) = C$$

and for player 2,

$$\sigma_2(N) = C, \sigma_2(U) = D, \sigma_2(F) = D$$

Intuitively, when the state profile is  $(N, N)$ , both players cooperate until one player receives a bad signal and triggers a punishment phase. Once a punishment phase starts, two events may occur: Either player 2 moves to state  $F$  before player 1 moves to  $F$  (or at least no later than one period after player 2 does). In that case, the most likely event is that players will coordinate back to  $(N, N)$  (with probability close to 1).<sup>44</sup> Alternatively, player 1 moves to state  $F$  more than one period before player 2 moves to state  $F$ . In that case, the most likely event is that players switch to  $(N, F)$  or  $(N, U)$ , hence coordination back to  $(N, N)$  will take longer.

We show here the calculations of the set of  $q - L$  combinations that are consistent with cooperation when  $p$  is close to 1. We illustrate the main transitions for the state pairs for the case where  $p$  is close to 1 and  $q$  is small, but not too small:

$$0 < 1 - p \ll q \ll 1.$$

---

<sup>44</sup>This is because once in state profile  $(G, W)$ , player 1 plays  $C$  and moves to  $N$ , while player 2 receives (with probability close to 1) signal  $y_2 = 1$ , hence also moves to  $N$ .

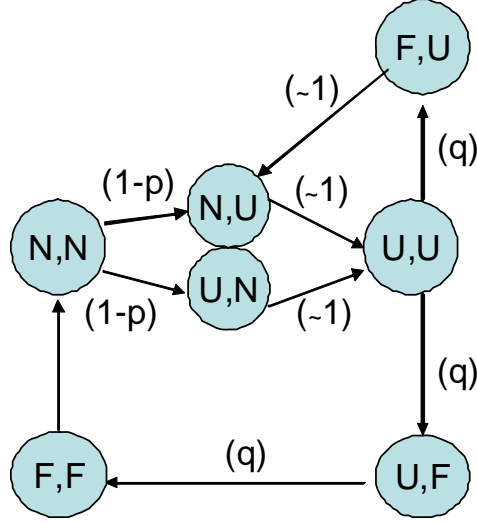


Figure 15: Transition of mental state

As mentioned above, we restrict attention in this example to this case; for the more general case where  $q$  is larger, tedious computations are required. We only report graphically the set of  $q - L$  combinations for which the proposed strategy profile is an equilibrium (as shown in Figure 14).

*Analysis:*

When players follow the proposed strategy profile, they alternate between long phases of cooperation (of length  $1/\pi$  with  $\pi = 2(1 - p)$ ), and relatively short punishment phases (of approximate length  $2/q$ ).

*Incentives for player 1 at U.* Under the proposed equilibrium strategy profile, the expected loss that player 1 incurs (compared to being in the cooperative phase) until coordination back to cooperation occurs is approximately  $2/q$ .<sup>45</sup>

When player 1 cooperates at  $U$ , he avoids triggering a punishment phase in the event  $(U, N)$ , so the occurrences of punishment phases are reduced by  $1/2$ . In addition, punishment phases are shorter, as coordination back to cooperation occurs as soon as player 2 transits to  $W$  (hence punishment length is reduced to  $1/q$ ), however they are more costly per period of punishment, as player 1 loses an additional  $L$  in each period (compared to the case where he would play  $D$ ).

<sup>45</sup>The exact cost is larger because there are a few periods in which player 1 cooperates while player 2 still defects. A better approximation of the cost is  $2/q + \frac{1}{2} + 3(1 + L)$ . To see this, compute, conditional on being in  $(U, U)$ , the expected loss that player 1 incurs (compared to being in the cooperative phase) until coordination back to cooperation occurs. We denote by  $\Delta$  that loss. We have:

$$\Delta = 1 + q[2(1 + L) + \Delta] + q[1/q + 1 + L] + (1 - 2q)\Delta$$

or equivalently:

$$\Delta = \frac{2}{q} + 3(1 + L)$$

Because there is equal chance of going to  $UU$  through  $NU$  or  $UN$ , the expected loss from a punishment phase is:

$$\frac{1}{2}(1 + L) + \frac{1}{2}(-L) + \Delta = \frac{1}{2} + \Delta$$

The condition is thus:

$$\frac{2}{q} < \frac{1}{2} \left( \frac{1}{q} \right) (1 + L)$$

or equivalently:

$$L > 3.$$

*Incentives of player 2 at N:* Defection generates short period of cooperation (cooperation lasts 2 periods), during which player 2 gains an additional payoff of  $L$ , and long periods of punishment (that last  $1/q$  periods) during which player 2 loses 1. Hence the condition

$$2L < \frac{1}{q}.$$

We omit the verifications of the other incentives, which are easier to check and automatically satisfied. **QED**

## 7. Bibliography

Abdulkadiroglu, Atila and Kyle Bagwell [2007], "Trust, Reciprocity and Favors in Cooperative Relationships," mimeo, Columbia University.

Abreu, D. and A. Rubinstein [1988], "The Structure of Nash Equilibrium in Repeated Games with Finite Automata," *Econometrica*, 56, No. 6, 1259-1281.

Altman, E., K. Avrachenkov, R. Marquez and G. Miller [2005], "Zero-Sum Constrained Stochastic Games with Independent State Processes," *Mathematical Methods of Operations Research* 62, 375-386.

Bagwell, K. [1995], "Commitment and observability in games", *Games and Economic Behavior*, 8, pp 271-280

Bendor, Jonathan, Roderick M. Kramer, and Suzanne Stout [1991], "When in Doubt ...Cooperation in a Noisy Prisoner's Dilemma," *Journal of Conflict Resolution* 35, pp 691-719.

Ben-Porath, E. [1986], "Repeated Games with Finite Automata," *Journal of Economic Theory*, 59, 17-32.

Cherepanov, Vadim, Timothy Feddersen and Alvaro Sandroni [2008], "Rationalization," mimeo, University of Pennsylvania.

Cohen, D., R. E. Nisbett, B. Bowdle and N. Schwartz [1996], "Insult, Aggression, and the Southern Culture of Honor: An 'Experimental Ethnography'," *Journal of Personality and Social Psychology* 70, 945-960.

Compte, O. [1998], "Communication in Repeated Games with Imperfect Private Monitoring", *Econometrica*, Vol. 66, pp. 597-626.

Compte, O. [2002], "On Sustaining Cooperation without Public Observations," *Journal of Economic Theory*, Vol. 102, pp. 106-150.

Compte, O. and A. Postlewaite [2008], "Repeated Relationships with Limits on Information Processing," mimeo, University of Pennsylvania.

Compte, O. and A. Postlewaite [2009], "The Extent of Cooperation," mimeo, University of Pennsylvania.

Cover, T. and M. Hellman [1970], "Learning with Finite Memory," *Annals of Mathematical Statistics* 41, 765-782.

- Ely, J. C. and J. Välimäki [2002], "A robust folk theorem for the prisoner's dilemma," *J. Economic Theory* 102, 84–105.
- Hauser, C. and H. Hopenhayn [2005], "Trading Favors: Optimal Exchange and Forgiveness," mimeo, UCLA.
- Kandori, M. [1992], "Social Norms and Community Enforcement," *Review of Economic Studies*, Vol. 59, pp. 63-80.
- Kandori, M., G. Mailath and R. Rob [1993], "Learning, Mutation and Long Run Equilibria in Games," *Econometrica*, Vol. 61, pp. 29-56.
- Kandori, M. and H. Matsushima [1998], "Private Observation, Communication and Collusion," *Econometrica*, Vol. 66, 627-652.
- Kandori, M., and I. Obara [2007], "Finite State Equilibria in Dynamic Games," mimeo.
- Mailath, G. and S. Morris [2002], "Repeated Games with Almost-Public Monitoring," *Journal of Economic Theory*, 102, 189-228.
- Miller, John H. [1996], "The Coevolution of Automata in the Repeated Prisoner's Dilemma," *Journal of Economic Behavior and Organization* 29, pp 87-112.
- Mobius, M. [2001], "Trading Favors," mimeo, Harvard University.
- Monte, D. [2007], "Reputation and Bounded Memory: A Theory of Inertia and Skepticism," mimeo.
- Okuno-Fujiwara, M. and A. Postlewaite [1995], "Social Norms in Matching Games," *Games and Economic Behavior*, Vol. 9, pp 79-109.
- Phelan, C. and A. Skrzypacz [2006], "Private Monitoring with Infinite Histories," mimeo.
- Piccione, M. [2002], "The repeated prisoner's dilemma with imperfect private monitoring," *Journal of Economic Theory* 102 , 70–83.
- Piccione, M. and A. Rubinstein [1997], "On the Interpretation of Decision Problems with Imperfect Recall," *Games and Economic Behavior* 20, 3-24.
- Sainty, Barbara [1999], "Achieving Greater Cooperation in a Noisy Prisoner's Dilemma: An Experimental Investigation." *Journal of Economic Behavior and Organization*, 39(4), pp. 421–35.
- Sekiguchi, T.[1997] "Efficiency in Repeated Prisoner's Dilemma with Private Monitoring", *Journal of Economic Theory*, 76, pp 345-361.
- Wilson, A. [2004], "Bounded Memory and Biases in Information Processing," mimeo, University of Chicago.