Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297
pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper 13-008

"Plausible Cooperation"

Third Version

by

Olivier Compte and Andrew Postlewaite

http://ssrn.com/abstract=2212687xxxxxxx

Plausible Cooperation*

Olivier Compte
PSE, Paris

Andrew Postlewaite
University of Pennsylvania

December 2010, This version December 2012

Abstract

There is a large repeated games literature illustrating how future interactions provide incentives for cooperation. Much of the earlier literature assumes public monitoring: players always observe precisely the same thing. Departures from public monitoring to private monitoring that incorporate differences in players' observations may dramatically complicate coordination and the provision of incentives, with the consequence that 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

*Compte: Paris School of Economics, 48 Bd Jourdan, 75014 Paris (e-mail: compte@enpc.fr); Postlewaite: Department of Economics, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6207 (e-mail: apostlew@econ.upenn.edu). This paper is a substantially revised version of two earlier working papers, "Repeated Relationships with Limits on Information Processing" and "Effecting Cooperation". 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.

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 restart 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 noncooperative 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, the theory of repeated games, while providing important insights about repeated interactions, 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),[1] or to the assumed sequencing/timing of actions and signals.[2] These "rules of behavior" fail to remain stable when the parameters of the game are changed slightly.[3,4] Their robustness to changes in timing is typically not addressed, nor is their plausibility. What we propose below is an alternative theory/description of how cooperation is accomplished when players are strategic, with the central concern that cooperation be attained via realistic and intuitively plausible behavior.

[1] See in particular the belief free literature in repeated games (Piccione (2002) and Ely and Valimaki (2002)). In belief free equilibria, each player is made indifferent between cooperating and defecting after any history of play and signals received.

[2] Repeated relationships are typically modeled as a stage game played repeatedly, with the players choosing actions simultaneously in the stage game. In reality, the players may be moving sequentially and the signals they get about others' actions may not arrive simultaneously. The choice to model a repeated relationship as simultaneous play is not based on a concern for realism, but for analytic convenience. A plausible theory of cooperation should not hinge on the fine details of the timing of actions: we should expect that behavior that is optimal when play is simultaneous to be optimal if players were to move sequentially.

[3] For example, in Compte and Postlewaite (2013), we show that the Ely-Valimaky construction is not robust to stochastic changes in the underlying monitoring structure.

[4] 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 necessary 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 (see Compte (1998) and Kandori and Matsushima (1998)), or the various actions he may play (belief free literature). There are exceptions, and some work such as Sekiguchi (1997) *does* 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 that depend in a complex way on past histories (as in Compte (2002)).

Finally, when public communication is allowed and signals are not conditionally independant, strict incentives to communicate past signals truthfully may be provided (Kandori and Matsushima (1998)). But the equilibrium construction relies on simultaneous communication protocols.

A theory of cooperative behavior faces at least two challenges:

A first challenge is a realistic *description* of strategies. A strategy is a complex object that specifies behavior after all possible histories, and the number of possible histories increases exponentially with the number of interactions. 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.

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 past 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. This concern is not novel. Aumann (1981) suggests focussing on stationary strategies, with the assumption that the agent only has a finite number of states of mind. The strategies we shall consider conform to Aumann's suggestion – endowing agents with a limited number of mental states.[5]

However we ask more for a strategy to be realistic than that it can be represented with a small number of states. Repeated game strategies can often be represented in this way, but then the classification of histories is for mathematical convenience, and not on an *a priori* basis of how individuals pool histories.[6]**.** Rather, our view is that a player's pooling of histories should be intuitively plausible, capturing plausible cognitive limitations (inability to distinguish finely between various signals), and/or reflecting how a player might plausibly interpret or react to observations.

A second challenge is a theory of cooperative behavior that is consistent with agents coming 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

[5] The path suggested has led to the study of repeated game equilibria in which players are ***constrained*** to using finite automata (Rubinstein 1986, Neymann (1985,1998), Ben Porath (1993)). It has also led to the study of repeated game equilibria in which strategies can be *implemented* by simple automata (Abreu (1986)), or approximated by finite automata (Kalai Stanford (1988)).

[6] When public signals are available or when monitoring is perfect, and one considers public equilibria (i.e., in which players condition behavior on public histories only), a vector of continuation values may be assigned after any public history, with histories leading to the identical continuation values being categorized into the same equivalence class. The vector of continuation values may thus be viewed as an endogenous summary statistic of past signals and play. Under perfect monitoring, Abreu (1988) derives a class of equilibria having a simple representation (with only few continuation values). A similar logic has been applied by Kalai and Stanford (1988), who get approximate equilibria with strategies using a finite number of states by partitioning the set of possible continuation values into a finite number of values.

that players could compute appropriate strategies through introspection in repeated games with private signals.[7] 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. Alternatively, 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: even if one restricted attention to strategies that are deterministic functions of histories, finding an optimal strategy amounts to finding an optimal partition of the histories among all possible partitions.

The path we propose addresses these two challenges. Our view is that the second challenge calls for a restriction on the strategies that agents choose among. The nature of the restriction one imposes is a critical issue. In this paper, we propose ***a priori*** constraints on how agents process signals, motivated by plausible psychological considerations or cognitive limitations, and ask when the restricted family of strategies so generated is conducive to cooperation.[8]

To summarize, 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 well; (ii) the strategies should be based on a cataloging of histories that is intuitively plausible; (iii) the sets of strategies allow agents to cooperate under a broad set of circumstances; and (iv) equilibrium cooperation obtains in a way that is robust to the parameters of the game and the timing of players' actions. 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.

Before going on we should emphasize two things. First, cooperation is not always possible in our framework, and second, even when equilibrium cooperation is possible, there is also an equilibrium in which cooperation doesn't obtain. We will have nothing to say about what determines whether the cooperative equilibrium or the uncooperative equilibrium arises when both exist. Although this is an interesting question, and although we believe that our framework

[7] In a different context (that of repeated games with perfect monitoring), Gilboa (1988) and Ben Porath (1990) have expressed a related concern, distinguishing between the complexity associated with *implementing* a repeated game strategy, and the complexity associated with *computing* best response automata. The concern that we express is not computational complexity per se, but rather the ability to perform relevant computations when precise knowledge of distributions is lacking.

[8] Although Aumann (1981) is not motivated by learning considerations, he mentions that assuming a bound on the number of states of mind would "put a bound on the complexity a strategy can have, and enable an analysis in the framework of finite games." In particular, in the context of a repeated prisoners' dilemma with perfect observations, he reports an example in which only few strategies are compared. Although Kalai, Samet and Stanford (1988) argue that the example lacks robustness, the path we follow is in the spirit of Aumann's example.

makes it more plausible that players would learn to cooperate, this is beyond the scope of this paper.

1.1. Strategy restrictions

The restrictions on the strategies available to players are a crucial element of our approach. We are not interested in arbitrary restrictions, but rather, on restrictions that might arise naturally.

As suggested above, we shall first endow each player with a finite set of *mental states*, restricting a player to behaving the same way for all histories of the game that lead to the same mental state. In addition, we shall endow agents with *transition functions* that describe what combinations of initial mental state, actions and signals lead to specific updated mental states. Mental states and transitions jointly define what we call a *mental system*. Contrasting with traditional approaches to bounded rationality in repeated games (i.e., the automaton literature), we do not attempt to endogenize the mental system that would be optimal given the parameters of the problem and behavior of other players. We do not think of a mental system as a choice variable, but rather as a natural limitation of mental processing. For example, 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. In principle, transitions could be a complex function of past histories, with evidence accumulating in the background, up to the point where, suddenly, one feels cheated. In principle, whether one feels cheated could also depend on some aspects of the game being played: 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 in our model is that it is exogenous**:** the individual does not have control over it.

A mental system characterizes how signals are interpreted and processed. Modeling mental systems as exogenous reflects our view that there are limits to peoples' cognitive abilities, and that evolution and cultural indoctrination should have more influence on one's mental system than the particular circumstances in which a specific game is played. 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 shapes the child's interpretations.

Taking the mental system as exogenous has one obvious consequence: this restricts the strategies *available* to agents. The restriction may be drastic when there are only few mental states. But there is still scope for strategic choices: even if upset, we assume that the agent still has a choice of what action to choose (either defect or cooperate in the case of a prisoner's dilemma). In other words, while the mental system is assumed to be the same across a variety of games, how one *responds* to being upset is assumed to be situational, and 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 when upset, and risk breaking a relationship that is generally cooperative, but not hesitate when the cost is large.

In proposing restrictions on strategies available to agents, we are aware that other types of restrictions could be proposed. One could for example argue that one's mental state may be more adequately represented by a continuous variable, rather than a discrete variable taking only two values: Upset or Normal. One could also argue that how one plays is determined by the mental state, and that agents have some control over what makes them Upset. Whatever description one finds plausible, one central issue remains: *what is the scope of strategic choice*? What aspects of choice do agents have control over? Our perspective is that this scope is limited, or at the very least, the scope should capture the agent's sophistication, i.e., his ability to tailor his behavior to the specific environment that he faces. If one finds it plausible to describe strategies using a continuum of mental states, then, one should also prescribe reasonable restrictions on the mappings from states to actions that agents would plausibly compare.[9]

It is beyond the scope of this paper to analyze how the nature of the restriction imposed affects the ability of agents to cooperate. But the spirit of the exercise we are proposing is this: to understand which type of restrictions are consistent with cooperation, and for what type of strategic environment.

Plan. We now turn to the formal model, which follows the steps described above. In Sections 2 and 3, our model assumes that there is occasionally a public signal that facilitates periodic synchronization. We analyze the circumstances in which cooperation is possible and discuss the robustness of our result. Next, in section 4 we drop the assumption of such a public signal and show how synchronization can be accomplished without such a signal. In Section 5 we discuss the results and possible extensions.

From a positive perspective, our model provides a simple and plausible theory of how players manage to cooperate despite the fact that signals are private. Strategic choices are limited yet they are sufficient to highlight (and isolate) the two key differences between games with public

[9] Strategy restrictions constitute a standard modelling device. One restricts actions or signals to few, thereby capturing the agents' inability to finely adjust behavior to the environment. The restrictions proposed here have a similar objective, that of limiting an agent's ability to adjust too finely the underlying structure of the problem he faces. The novelty lies in considering more general classes of restrictions, i.e., not necessarily arising from measurability conditions (See Compte and Postlewaite (2012) for a discussion of these issues and applications to other strategic environments).

and private monitoring: (a) the incentives to trigger punishments, which require that the cost incurred from remaining cooperative while the other defects be sufficiently large (see section 2); (b) the difficulty in recoordinating to cooperation once a punishment phase has been triggered. As explained in Section 4, one solution is that players have asymmetric roles, with one being a leader and the other a follower. When players have symmetric roles, sustaining cooperation will require asymmetric responses to good and bad signals respectively, with good signals triggering a stronger forgiving/redemption effect, as compared with the deterioration that bad signals produce.

From a broader theoretical perspective, constraints on strategies act as partial commitment (to using only strategies in that subset). Sources of partial commitment may be informational, cognitive, or psychological. These constraints may stem from informational, cognitive, or psychological considerations. They may also be interpreted as characterizing the agent's degree of sophistication. As explored in Section 2.6, the path proposed illustrates a tension between sophistication and robustness: sophistication (i.e. a richer set of strategies) may be conducive to higher cooperation level for some parameters of the games and specific strategies, but it may decrease the chance of cooperation, i.e., the range of parameters for which these strategies remain in equilibrium.

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 received. 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.

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". We restrict attention to two signals in the example, but discuss how the analysis can be extended to a continuum of signals in Section 3.5.

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 traditional models. A player is endowed with a *mental system* that consists of a finite set S_i of *mental states* the player can be in, and a *transition function* T_i that describes what triggers moves from one state to another: the function T_i determines the mental 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 outcomes of that period and possibly previous periods. The restriction we impose is that a player may only condition his behavior on his mental state, and not on finer details of the history.[10] Given this restriction, all mappings from states to actions are assumed admissible. Player i's set of pure strategies is:

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

We will illustrate the basic ideas with an example in which the players can be in one of two states U(pset) or N(ormal).[11] The names of the two states 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

[10] Note that our structure requires that players' strategies be stationary: they do not depend on **calendar time**. This rules out strategies of the sort "Play D in prime number periods and play C otherwise", consistent with our focus on behavior that does not depend on fine details of the history.

[11] The restriction to two mental states is for expository purposes. The basic insights that cooperation can be sustained via intuitively plausible strategies continue to hold when agents have more finely delineated mental states; we discuss this at the end of this section.

history of (own) play and signals.[12] Both S_i and T_i are exogenously given, not a choice: a player who 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.

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

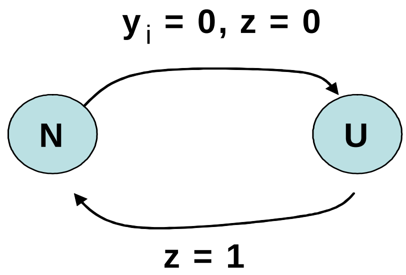


Figure 1: Transition

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.[13] The mental system thus determines how observations are aggregated over time, hence how histories are pooled: some histories lead to state N, others lead to state U.

The simple form of the transition function – that given the previous state and the action taken, it depends only on the most recent signal – is for simplicity. In principle, the transition function could depend on more than this most recent signal, for example, whether two of the past three signals was "bad", or it could also be stochastic. We discuss this at the end of this section and consider stochastic transitions in Section 4.

Given the mental system above, our candidate behavior for each player i will be as follows,

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

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

[12] 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-dependent payoffs, and the qualitative nature of our conceptual points would be unchanged if we did so.

[13] 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.

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

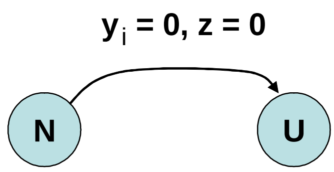


Figure 2: No "resetting"

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).

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.

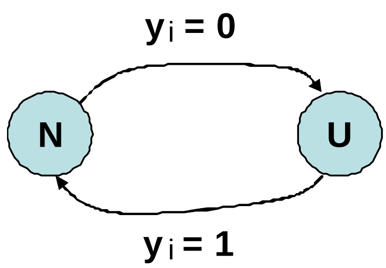


Figure 3: Forgiving transition

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.[14]

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.* (1996) ran several experiments

[14] We show below that cooperation is essentially impossible if players have this mental process.

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 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 culturally driven: physiological reactions to insults were substantially lower for northern students than for southern students.

Indeed, the point of the Cohen *et al.* (1996) 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 expressed above that the transition function in our model can be thought of as culturally determined.

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.[15] 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 σ 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 ϕ_σ the ergodic distribution over states induced by σ. That is, $\phi_\sigma(s)$ corresponds to the (long run) probability that players are

[15] While in general the ergodic distribution may depend on the initial conditions, we restrict attention to transition functions for which the distribution is unique.

in state s.[16]

We associate with each strategy profile σ the value induced by the ergodic distribution. This corresponds to computing discounted expected payoffs, and taking the discount to 1.[17] 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 σ 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 $\sigma'_i \in \Sigma_i$,

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

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.[18,19]

2.4. Successful cooperation

We are interested in equilibria in which the players cooperate at least some of the time asymptotically. This *requires* players playing the strategy "play C in N and D in U".[20] We shall call σ_i^* this strategy for player i.

The strategy $\sigma^* \equiv (\sigma_1^*, \sigma_2^*)$ cannot be an equilibrium for all parameters p, q and L. For example fix p and q. Then L *cannot be too large.* If L is sufficiently large it will pay a player

[16] Formally, define $Q_\sigma(s', s)$ as the probability that next state profile is s' when the current state is s. That is, Q_σ is the transition matrix over state profiles induced by σ. The vector ϕ_σ solves $\phi_\sigma(s') = \sum_s Q_\sigma(s', s)\ \phi_\sigma(s)$.

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

[18] We restrict attention to pure strategies. However, our definitions can be easily generalized to accomodate 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. We discuss the issue of mixed strategies in detail in the discussion section below.

[19] Note that σ_i as defined should not be viewed as a strategy of the repeated game. A strategy of the repeated game is a mapping from histories to actions. The strategy σ_i, along with the mental system (S_i, T_i) would induce a repeated game strategy, once the initial state is specified.

[20] This is because there cannot be an equilibrium where a player cooperates always, or when he plays D in N and C in U . Indeed, when player 2 plays D in N and C in U, then, defecting always is the best response for player 1 because it maximes the chance that player 2 cooperates.

to deviate to "always defect": the occasional reward of L to the deviating player will be more than enough to compensate for causing the opponent to very likely switch from N to U.

Also, L *cannot be too small.* When a player gets a bad signal he is not sure if his opponent is now in U and is playing D, or if the signal is a "mistake". If it's the latter case, playing D will likely lead to a spell of noncooperation. If L is small, there is little cost to playing C to avoid this; thus there is a lower bound on L that is consistent with the strategies above being an equilibrium.

There is thus an upper bound and a lower bound on L. The upper bound reflects a standard consideration, that gains from deviating from cooperation should not be too large; otherwise they cannot be offset by punishment phases. The lower bound is specific to private monitoring games: private signals can always be ignored, so incentives to trigger punishment have to be provided; and such incentives obtain when the cost of remaining cooperative while the other defects is too large. That basic intuition will be unaltered when we move to more mental states.

We now turn to a formal derivation of these two bounds, showing that they can be compatible. We refer to ϕ_{ij} as the long-run probability that player 1 is in state $i \in \{U, N\}$ while player 2 is in state j when both players follow σ^*. By definition we have

$$v_1(\sigma_1^*, \sigma_2^*) = \phi_{NN} - L\phi_{NU} + (1 + L)\phi_{UN}.$$

By symmetry, $\phi_{NU} = \phi_{UN}$, so this expression reduces to

$$v_1(\sigma_1^*, \sigma_2^*) = \phi_{NN} + \phi_{UN} = \Pr_{\sigma^*}(s_2 = N) \equiv \phi_N.$$

Consider now the alternative strategy σ^D (respectively σ^C) where player 1 plays D (respectively C) in both states U and N. Also call ϕ_j^D (respectively ϕ_j^C) the long-run probability that player 2 is in state $j \in \{U, N\}$ when player 1 plays the strategy σ^D (σ^C) and player 2 plays σ_2^*. We have:

$$v_1(\sigma^D, \sigma_2^*) = (1 + L)\phi_N^D.$$

This expression reflects the fact that playing σ^D induces additional gains when the other is in the normal state; but this has a cost because of an adverse effect on the chance that player 2 is in the normal state ($\phi_N^D < \phi_N$ when $p > 1/2$). The expressions above imply that the deviation to σ^D is not profitable when

$$L \leq \bar{L} = \frac{\phi_N}{\phi_N^D} - 1. \tag{2.1}$$

When player 1 plays σ^C, he obtains:

$$v_1(\sigma^C, \sigma_2^*) = \phi_N^C - L\phi_U^C.$$

This expression reflects the fact that playing σ^C changes the probability that player 2 is in N ($\phi_N^C > \phi_N$ when $p > 1/2$) in a way that benefits player 1: player 2 is more likely to be in N when player 1 always cooperates than when he follows σ^*, because he avoids triggering a punishment/Upset phase when he receives bad signals by mistake. But this has a cost: he loses L whenever player 2 is in U. The deviation to σ^C is thus not profitable when

$$L \geq \underline{L} \equiv \frac{\phi_N^C - \phi_N}{\phi_U^C} = \frac{\phi_U - \phi_U^C}{\phi_U^C} = \frac{\phi_U}{\phi_U^C} - 1 \tag{2.2}$$

We turn to checking when these two bounds are compatible. First we consider $p = 1/2$ and p close to 1.

(i) $p = 1/2$.

The distribution over player 2's state is then independent of player 1's strategy, so $\phi_N^C = \phi_N = \phi_N^D$, hence the two bounds coincide (and they are equal to 0).

(ii) p is close to 1.

Then mistakes are rare, so $\phi_N \simeq 1$. When player 1 always defects, player 2 essentially receives only bad signals. If in state N, this signal triggers a change to U with probability $(1-q)$. Since it takes on average $1/q$ periods to return to N, the fraction of the time player 2 spends in N when player 1 always plays D is $\frac{1}{1+(1-q)/q} = q$, hence $\phi_N^D \simeq q$. This gives us the upper bound

$$\bar{L} = \frac{1}{q} - 1.$$

Now assume player 1 follows σ^C. Mistakes occur with probability $(1-p)$, leading with probability $(1-q)$ to an Upset phase of length $1/q$, hence $\phi_U^C \simeq (1-p)(1-q)/q$. When player 1 is following σ^*, he too reacts to bad signals, and in turn induces player 2 to switch to U (though with one period delay). So ϕ_U is larger than ϕ_U^C, but no larger than $2\phi_U^C$. Thus

$$\underline{L} = \frac{\phi_U}{\phi_U^C} - 1 \leq 1$$

and the two bounds are compatible when $q < 1/2$.[21]

The following Proposition, which is obtained by using exact expressions for the long run

[21] This is a rough approximation of the upperbound. Exact computation for p close to 1 gives $\underline{L} = 1 - q$.

distributions, shows that the inequalities are compatible for any $p > 1/2$ and any $q \in (0, 1)$.

Proposition 1: *For any $p > 1/2$ and any $q \in (0, 1)$, we have $0 < \underline{L} < \bar{L}$. Besides, for any q, both $\bar{L}$ and $\underline{L}$ are increasing with p.*

Details of the computation are in the Appendix.[22] We note here that the long-run distributions are easily derived. For example, the long run probability ϕ_{NN} satisfies

$$\phi_{NN} = q + (1 - q)p^2\phi_{NN}$$

(both players are in state N either because resetting to cooperation occurred, or because players were already in state N and no mistake occurred). Solving, this gives $\phi_{NN} = \frac{q}{1-(1-q)p^2}$.

Finally, there is still a deviation that has not been checked: the possibility that player 1 chooses to play C in U and D in N. We show in the appendix that this deviation can only be profitable when L is large enough. This thus imposes another upper bound on L, possibly tighter than $\bar{L}$. We check that for q not too large, that constraint is not tighter.

Proposition 2: *There exists $\bar{q}$ such that for any $q < \bar{q}$, $p \in (1/2, 1)$ and any $L \in (\underline{L}, \bar{L})$, it is an equilibrium for both agents to play the strategy C in N and D in U.*

Our proof of Proposition 2 proposes a formal argument that does not pin down a particular value for $\bar{q}$. Direct (but tedious) computations however show that Proposition 2 holds with $\bar{q}$ as large as 0.4.

The shaded region in the graph below shows how the range of L for which cooperation is possible varies as a function of p for the particular value of q equal .3.

[22] We thank an anonymous referee for suggesting the outline of the calculation.

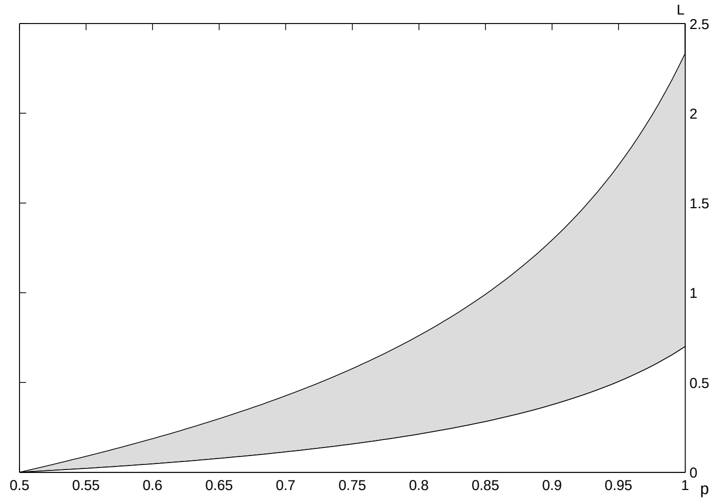


Figure 4: $p - L$ combinations that allow cooperation for $q = .3$

Note that as long as the signal about the opponent's effort is informative, there are some values of L for which cooperation is possible. Both the lower bound and upper bound on such L's are increasing as p increases, as is the size of the interval of such L's.

2.5. Tit-for-tat

A mental system generates restrictions on the set of strategies available to players and these restrictions eliminate potentially profitable deviations. It is not the case, however, that seemingly reasonable mental systems necessarily make cooperation possible. The Forgiving mental system described above causes a player to be responsive to good signals: good signals make him switch back to the normal state.

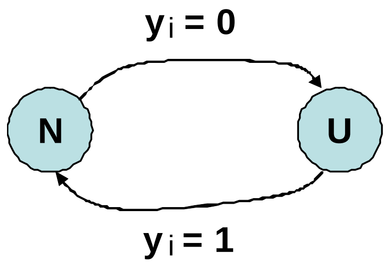


Figure 3: Forgiving transition

Along with the strategy of playing C in N and D in U, this mental system induces a tit-for-tat strategy. With such a mental system however, for almost all values of p the only equilibrium entails both players defecting in both states.

> **Proposition 3:** *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 leave the proof of this proposition in the appendix, but give the basic idea here. 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 σ that plays C in N and D in U. If both players follow σ, then by symmetry, the induced ergodic distribution will put identical weight on (NN) and (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 σ^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 σ^C, his payoff will remain arbitrarily close to 1. Hence it cannot be an equilibrium that both players play σ.

2.6. More mental states

The example illustrates how cooperation can be sustained via intuitively plausible strategies when the strategies available to players are restricted. It also illustrates that in spite of the restriction, and so long as the constant strategies are included in the set, it is not trivial to support cooperation, highlighting two relevant economic constraints. This of interest however only to the degree that the constraints imposed have some claim to realism. We find it compelling that people use a coarse perception of the past. What is less compelling perhaps is that the player is restricted to two mental states. The restriction to two mental states in the example, however, is not crucial for the insight from the example.

Consider a possibly large set of mental states $S_i = (s_i^1, ..., s_i^n)$, each state s_i^k reflecting an increasing level of annoyance as k increases. The set of possible strategies expands dramatically when the number of states rise, so it is natural in this case to focus on a limited set of strategies – monotone (or threshold) strategies: if one cooperates at one level of annoyance, then one also cooperates at mental states where one is less annoyed. Assume that when a player cooperates, receiving a good signal moves him to a lower state if possible (from s_i^k to s_i^{k-1} for $k \geq 2$) with probability h, while a bad signal moves him to a higher state (from s_k to s_{k+1}) with probability h'. In addition, assume that when a player defects, his state does not change. Signal $z = 1$ moves the players back to state s^1.[23]

[23] This assumption is not essential. Signal $z = 1$ could move players back to some intermediate state n_0.

We denote by σ_{k_0} the strategy that cooperates at all states s^k for $k \leq k_0$, and by σ_0 the strategy that defects at all states. The parameter k_0 thus characterizes the threshold or leniency level associated with strategy σ_{k_0}. We are interested in whether σ_{k_0} can be supported as equilibrium behavior for some k_0.

In the appendix, we show that when players restrict attention to threshold strategies the central qualitative insights in the two state example above continue to hold. When a player deviates by being *more* lenient, he triggers a punishment phase less often, but this is at the expense of having to bear the cost of continuing to cooperate longer in the event the other has already started a punishment phase. Hence L should not be too small; otherwise that deviation will be profitable. And if a player deviates by being *less* lenient, he reacts more quickly to bad signals (and saves L when the other has already started a punishment phase), but he triggers a punishment phase more often (which is costly unless L is large). So L should not be too large otherwise that deviation will be profitable.

In general, when there are many mental states, as we increase the set of strategies that players are endowed with, the set of parameters for which any given strategy can be an equilibrium is reduced (because there are more feasible deviations), but there are potentially more equilibrium strategies.

As the cooperative threshold increases, there typically will be a level beyond which cooperation is infeasible for most parameters. However, as long as threshold strategies constitute an equilibrium, higher thresholds reduce the probability of entering a punishment phase, and hence give higher welfare.

2.7. More complicated transition functions

Our leading example analyzes the case in which there are two states and transitions that only depend on the most recent signal. The latter feature simplifies the computations of the ergodic distributions, but our analysis and incentive conditions (2.1) and (2.2) would be unchanged regardless of how complicated the transition functions are. Of course it is not the case that any transition function is consistent with cooperation, as Proposition 3 illustrates. In general, however, cooperation will be consistent with more complicated transition functions than that in the leading example, with the set of parameters for which cooperation is possible depending on the transition function.

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. We will then discuss extensions.

3.1. Ex ante and interim incentives

The distinction between ex ante and interim incentives is irrelevant in our framework. When a strategy profile $\sigma = (\sigma_1, \sigma_2)$ is played, the value that player i obtains is $v_i(\sigma)$ and it is computed by considering the ergodic distribution over state profiles induced by σ. Neither the date at which this computation takes place, nor the beliefs that player i might have about the other player's current mental state are specified.

We do not specify beliefs because we do not think of players as having common knowledge over the signal structure nor **over** the mental systems. However as a modeler, we could very well consider a particular history h_i of the game for player i, define the belief that a player would have after that history, and compute the continuation payoff $v_i^{h_i}(\sigma)$ that player i obtains if he follows σ from then on. Because we are considering arbitrarily patient players, however, $v_i^{h_i}(\sigma)$ is arbitrarily close to $v_i(\sigma)$. So if σ is a strict equilibrium in our sense, it will also be a sequential equilibrium in the usual sense given the assumed strategy restrictions: after any history h_i, conforming to σ is optimal among all the strategies available to player i.

3.2. Pure versus mixed strategies

We have restricted our analysis to pure strategies. A legitimate question is whether our equilibrium strategy profile σ^* remains an equilibrium when the set of strategies is expanded to allow for mixed strategies. We allow below for mixed strategies in which a player puts weight at least $1 - \varepsilon$ on a pure action, with ε small.[24]

The following Proposition shows that the equilibrium is indeed robust.

> **Proposition 4:** *Consider the set of strategies $\Sigma_i^\varepsilon : S_i \to A_i^\varepsilon$ where A_i^ε is the set of mixed strategies that put weight at least $1 - \varepsilon$ on a pure strategy. Call P^ε the set of parameters (q, p, L) for which σ^* is a strict equilibrium. If $(q, p, L) \in P^0$ then $(q, p, L) \in P^\varepsilon$ for ε small enough.*

The proof is in the appendix. It is interesting to note that the proof of this Proposition shows that looking at one-shot deviations generates strictly weaker restrictions than the ones we consider. The reason is that checking for one-shot deviations only does not ensure that multiple deviations are deterred.[25]

[24] One motivation for doing so is that this is a way to allow a player to deviate for **a single** period, very infrequently so, thereby allowing us to effectively check for one-shot deviations.

[25] The reason for this non-equivalence is that a mental states pools many possible histories, so the structure of each player's decision problem becomes analogous to one with imperfect recall, as in the "absent-minded driver"'s problem examined in Piccione and Rubinstein (1997).

3.3. Robustness

The example was kept simple in a number of ways to make clear how cooperation could be achieved when the set of strategies is 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. We pointed out in the introduction that we are often unsure about the precise timing of actions and signals in repeated relationships that we study. In Compte and Postlewaite (2008) we show that the assumption in this paper that the players simultaneously choose actions in the basic stage game can be relaxed so that their choices are sequential without altering the qualitative conclusions about cooperation. Thus the equilibrium behavior that we derive is robust not only to changes in payoffs or monitoring structure, but also to changes in the timing of decisions. 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.

In the initial example that we analyzed an agent received a binary signal about her partner's effort. This assumption makes the analysis more transparent but can easily be replaced with a signal structure in which agents might receive signals in the unit interval, with higher signals being stronger evidence that the partner put in effort. One may then modify each agent's mental system by defining a threshold in the signal set with signals below the threshold treated as "bad" and signals above treated as "good". With a mental system modified in this way, the qualitative features of the equilibria we study are essentially unchanged with this signal structure.[26]

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. 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

[26] The difference is that we would no longer have $p = \Pr(y = 1 \mid C) = \Pr(y = 0 \mid D)$. This assumption was made for convenience, however (so that a single parameter describes the monitoring structure). In the more general setup suggested here, we would have two parameters: $p = \Pr(y = 1 \mid C)$ and $r = \Pr(y = 0 \mid D)$. Also, even for a priori symmetric monitoring structures, we could lose the symmetry of the distribution over binary outcomes if players were to use a different threshold.

is no problem in coordinating a punishment phase. 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.

Another key issue is how do players get back to cooperation once a punishment has been triggered. We finessed that second issue by assuming the public signal z that facilitated recoordination after a punishment period. We demonstrate next how players can coordinate a move back to cooperation in the absence of a public signal. We first illustrate that the issue of recoordination is not trivial. We examine the case where the players receive private signals z_i, $i = 1, 2$ (instead of a public signal z) and where players' mental systems are as in our leading example. With this simple mental system, cooperation can be supported if z_1 and z_2 are highly correlated. We show however that when the signals z_i are independently distributed cooperation can no longer be supported. Finally, we show two alterations of that simple mental system that allow cooperation and recoordination.

4.1. The difficulties of recoordination with private signals: An illustration.

Assume that each player i receives a private signal $z_i \in Z_i = \{0, 1\}$ and consider the mental process as before with the qualification that T_i is now defined over Z_i rather than Z. (See figure 5 below.)

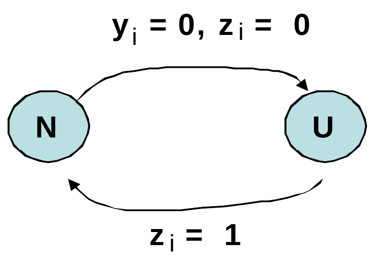


Figure 5: Independent "resetting"

If the parameters (p, q, L) are such that incentives are strict in the public resetting signal case, then by continuity, incentives will continue to be satisfied if the signals z_i have the property that $\Pr(z_i = 1) = q$ and $\Pr(z_1 = z_2)$ close enough to 1.

The correlation between signals in this information structure cannot be too weak if cooperation is to be possible however: if the two signals z_1 and z_2 are independent, cooperation cannot be sustained in equilibrium when p is close to 1.

> **Proposition 5:** *Fix the mental system as above with z_1 and z_2 independent. 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.

We leave the proof to the Appendix.

In essence, recoordination is less likely because it requires two simultaneous signals $z_1 = 1$ and $z_2 = 1$, so incentives to trigger punishments are more difficult to provide: the upper and lower bound on L are no longer compatible.

4.2. Resetting without a public signal

Public and nearly public signals can facilitate the coordination back to cooperation, but they are not necessary. Slightly more complicated, but still plausible, mental systems can support cooperation when there is no correlation in players' signals. We mention here two such possibilities that are analyzed in detail in the appendix.

Modified Tit-for-Tat. The first example is a stochastic modification of the Tit-for-Tat example discussed in Section 2.5. In that example, bad signals caused a transition from N to U and a good signal caused a transition back to N. Cooperation was essentially impossible with that mental system, but cooperation may be possible if the transitions are made stochastic. Suppose the transitions from N to U are modified so that a bad signal only causes a transition to U with probability h.[27] Also suppose that the transitions from U to N are modified so that (i) with probability b, player i forgets and transits to N *independently* of the signal received; and (i) if still in U, a good signal triggers a transition back to N with probability k.[28]

Transitions are summarized in Figure 6:

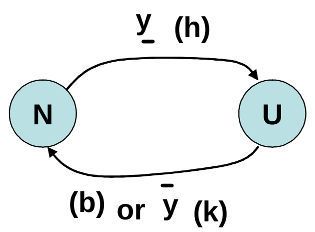


Figure 6: Modified Tit-forTat

For some configurations of parameters b, h and k, cooperation (play C in N and D in U) will be an equilibrium for a broad set of values of p and L, demonstrating that robust cooperation can be achieved without public signals.

Asymmetric mental systems. In the second example the two players have different mental systems, each with three states N, U and F. Transitions from N to U are as before. Transitions

[27] Good signals do not cause such transitions.

[28] Thus, the mental system combines a feature of our initial mental system (exogenous transition to N, or forgetfulness) and a feature of our Forgiving mental system (transition to N triggered by good signals).

from U to F are stochastic, depending on independent private signals z_1 and z_2, $z_i \in \{0, 1\}$ and $\Pr\{z_i = 1\} = q$. (the transitions away from U can be thought of as a player forgetting being upset).

A key feature of the modified transition functions is that what triggers a change from F to N 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 7.

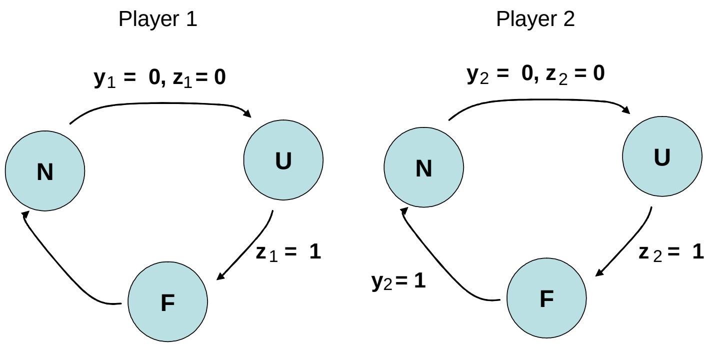


Figure 7: "Successful" independent resetting

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.[29]

The key difference with the analysis in section 2 is that players no longer simultaneously switch back to cooperation (because there is no public signal to support that). Rather, the transition functions are as though one player acts as a leader in the relationship, and makes an effort in choosing 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.

In this example both players have three state mental systems. It is relatively simple to demonstrate that one can obtain a similar outcome if player 1 has a two state mental system

[29] It can also be demonstrated that the analysis is robust to changes in timing.

of the forgetting kind, but with slower reactions.[30] When player 2 has the mental system in the current example, re-coordination is solved: when player 1 comes back to N, there is a high chance that player 2 is already in F (because player 1 is slower to forget). Cooperation by player 1 then makes player 2 transit to N, and because player 1 does not react with probability 1 to bad signals, there is a substantial chance that both stay in N.

While we do not attempt a characterization of the pairs of mental systems that can support cooperation in this paper, we point this out to demonstrate that robust cooperation may be possible even when a public signal is absent, and to suggest that asymmetric mental systems could help recoordination.

5. Further discussion

Evolution of mental systems. 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. 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, or reduce the set of parameters under which cooperation is possible. The question of evolution is discussed in more detail in Compte and Postlewaite (2009).

Extensions of the model.[31] Our model has no explicit communication. In the three state asymmetric example, Player 1 was exogenously designated as the leader in the relationship, and one can interpret the decision to play C in state F as an implicit communication that the relationship should restart. There, communication was noisy (because the other player does not receive a good signal 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.

Direct utility from being in a mental state. 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).

[30] By slower reactions, we mean that he reacts to bad signals with probability smaller than 1, say $1/2$, and that he forgets (and transit back to N independently of the signal received) with a probability smaller than q, say $q/2$.

[31] These extensions are discussed in more detail in Compte and Postlewaite (2009).

Cooperation in larger groups. The basic structure and ideas in this paper can be extended 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.[32]

Social norms. We have restricted attention to play in a prisoner's dilemma game to focus attention on the central ingredients of our model. It is straightforward to extend the basic idea to more general games, including asymmetric games. 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.

5.1. 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. This can be the case when 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 (2002) then 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 doesn't matter, because each can predict very accurately what other players will next do.[33] 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.[34]

As mentioned in the introduction, one branch of the repeated game literature aims at taking into account the complexity of strategies into account, assuming players use finite automata to

[32] There is a literature that analyzes the possibility of cooperation when players are repeatedly randomly matched (Kandori (1992) and Okuno-Fujiwara and Postlewaite (1995)), but in these models, some public information is or becomes available to players.

[33] When one moves away from almost public signal monitoring, beliefs about each other's past histories may begin to matter, and even checking that a candidate strategy profile is an equilibrium becomes difficult, even when the strategy can be represented as an automaton with a small number of states. Phelan and Skrzypacz (2006) address this issue. See also Kandori and Obara (2007) for a related treatment.

[34] This is because even as p gets close to 1, the probability $\Pr(s_2 = U \mid s_1 = U) = \phi^{\sigma}_{UU}/(\phi^{\sigma}_{UU} + \phi^{\sigma}_{UN})$ remains bounded away from 0 and 1.

implement their strategies.[35] The strategies that we consider can be represented as automata, yet we differ from this literature in several respects. First, in this literature players choose 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.[36] Second, to our knowledge, this literature does not consider games with private monitoring. Third, this literature used automata primarily as a tool to capture the complexity of a given strategy, while we emphasize the restriction on the *set* of strategies as a modelling device to capture the players' limited ability to tailor their behavior to the underlying parameters of the game. Fourth, our modeling strategy takes more seriously mental systems as being a plausible, if crude, model of the process by which players may interact, thereby shaping the restriction that we consider.

The model we study reduces to a dynamic game of a particular kind in which the underlying state is the profile of mental states.[37] Most of the literature in dynamic games assumes that in each period, there is a state that is known to both players,[38] while in our model the state is partially known: each player only knows his own mental state.

Following Mobius (2001), there is an active literature on favor/gift exchange between two players. The payoff structure of the stage game is analogous to the gift exchange model that we consider, and in some of these models players may receive relevant private information (for example about the opportunity to do favors as in Abdulkadiroglu and Bagwell (2007)).[39] These papers, however, assume public observations (i.e. whether a favor has been made) that allow players to coordinate future play.

Finally, in a recent paper, Romero (2010) provides an interesting example which, at least at a formal level, bears some resemblance to ours: there is a restriction to a limited set of automata that each individual may consider using (hence a limited set of strategies), and one of these automata (Win Stay Lose Shift) is an equilibrium in this restricted class.[40] This strategy facilitates recoordination when moves are simultaneous, but does poorly when players move in sequence: it would typically generate long sequences of miscoordinations after a bad signal.[41]

[35] The complexity of a player's strategy is defined to be the minimal size of a machine that can implement that strategy. See, e.g., Abreu and Rubinstein (1988) and Ben-Porath (1993)

[36] There has also been work in single-person decision making problems that is analogous to the papers using automata to capture complexity costs. 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. While we take agents' transition functions as fixed, the focus of this literature is on characterizing the optimal transition rule.

[37] We thank Eilon Solan for this observation. Dynamic games have the property that the structure of the continuation game varies with some underlying state. The dynamic nature of the game would be reinforced if we attached payoffs to being in a particular mental state.

[38] However see Altman et al. (2005).

[39] Hauser and Hopenhayn (2008) analyze a similar model in continuous time.

[40] Win Stay Lose Shift is an automaton that plays C after $C\overline{y}$ or $D\underline{y}$, and D otherwise.

[41] After an initial bad signal $\underline{y}$ (unfortunately) received by player 1, say, the most likely sequence of signals

6. Appendix

Proof of Proposition 1

In what follows, we define $\mu = (1-q)/q$.

Recall that ϕ_j^D denotes the long run probability that player 2 is in state j when 1 defects at all states and 2 plays σ_2^*. We have: $\phi_N^D = q + (1-q)(1-p)\phi_N^D$, implying that

$$\phi_N^D = \frac{q}{1-(1-q)(1-p)} = \frac{1}{1+\mu p}$$

Similarly we have $\phi_N^C = q + (1-q)p\phi_N^C$, implying that

$$\phi_N^C = \frac{q}{1-(1-q)p} = \frac{1}{1+\mu(1-p)}$$

Recall that ϕ_{ij}denote the long-run probability that player 1 is in state $i \in \{U, N\}$ while player 2 is in state j, under the candidate equilibrium σ^*. As we already explained, we have $\phi_{NN} = q + (1-q)p^2\phi_{NN}$ implying that

$$\phi_{NN} = \frac{q}{1-(1-q)p^2} = \frac{1}{1+\mu(1-p^2)}$$

Next, we similarly have $\phi_{UN} = (1-q)(1-p)\phi_{UN} + (1-q)p(1-p)\phi_{NN}$; implying that

$$\phi_{UN} = \frac{\mu p(1-p)}{1+\mu p}\phi_{NN},$$

hence:

$$\begin{aligned} \phi_N &= (1+\frac{\mu p(1-p)}{1+\mu p})\phi_{NN} = (\frac{1+\mu(1-p^2)+\mu(2p-1)}{1+\mu p})\phi_{NN} \\ &= \phi_N^D(1+\mu(2p-1)\phi_{NN}) \end{aligned}$$

These equalities allow us to derive the bounds $\bar{L}$ and $\underline{L}$ as a function of μ and p (and to simplify computations, ϕ_{NN}). Specifically,

We have:

$$\bar{L} = \frac{\phi_N}{\phi_N^D} - 1 = \mu(2p-1)\phi_{NN}.$$

would lead to persistent miscoordination:

player 1's play and signals	C	$\underline{y}$	D	$\underline{y}$	C	$\underline{y}$	D	$\overline{y}$	D	$\underline{y}$	C	$\underline{y}$	D
player 2's play and signals		C	$\underline{y}$	D	$\overline{y}$	D	$\underline{y}$	C	$\underline{y}$	D	$\overline{y}$	D	$\underline{y}$

We then use $\phi_N = \phi_N^D(1+\bar{L})$ and $\underline{L} = \frac{1-\phi_N}{1-\phi_N^C} - 1$ to obtain, after some algebra:

$$\underline{L} = \bar{L}p\phi_N^D = \frac{p}{1+p\mu}\bar{L}$$

This shows that $\underline{L} < \bar{L}$. Besides, since ϕ_{NN} increases with p, and since $\frac{p}{1+p\mu}$ increases with p, both $\bar{L}$ and $\underline{L}$ are increasing functions of p. Q.E.D.

Proof of Proposition 2

We already checked that for any $L \in (\underline{L}, \bar{L})$, neither σ^C nor σ^D are profitable deviations. To complete the proof, we need to check that the strategy $\widetilde{\sigma}$ that plays D in N and C in U is not profitable. Call $\widetilde{\phi}$ the long run distribution over state profiles. We have:

$$v_1(\widetilde{\sigma}, \sigma_2^*) = (1+L)\widetilde{\phi}_{NN} + \widetilde{\phi}_{UN} - L\widetilde{\phi}_{UU}.$$

Since $v_1(\sigma_1^*, \sigma_2^*) = \phi_N$, the deviation is not profitable whenever (letting $\widetilde{\phi}_N = \widetilde{\phi}_{NN} + \widetilde{\phi}_{UN}$):

$$(\widetilde{\phi}_{NN} - \widetilde{\phi}_{UU})L < \phi_N - \widetilde{\phi}_N$$

Lemma 1: $\phi_N > \widetilde{\phi}_N$

Intuitively, under $\widetilde{\sigma}$, player 1 plays D in N, triggering player 2 to exit from N. There is a small counterbalancing effect because with probability $(1-p)^2$ players may end up in UN and play CC. However this is not enough to increase the long-run probability that 2 is in N above ϕ_N.

Lemma 2: There exists $\bar{q}$ such that for all $q < \bar{q}$ and $p \in (1/2, 1)$, $\widetilde{\phi}_{NN} < \widetilde{\phi}_{UU}$

Intuitively, since player 1 plays D in N, it is relatively easy to exit from NN to NU and then to UU (the probability of such exit is bounded away from 0, uniformly on $p \in (1/2, 1)$, while exit from UU requires an occurrence of the resetting signal z.

Combining these two Lemmas implies that under the conditions of Lemma 2, $\widetilde{\sigma}$ cannot be a profitable deviation.

To check Lemma 1 formally, let us compute explicitly the long-run probabilities $\widetilde{\phi}_{ij}$. The long run probability $\widetilde{\phi}_{NN}$ satisfies: $\widetilde{\phi}_{NN} = q + (1-q)p(1-p)\widetilde{\phi}_{NN}$ implying that

$$\widetilde{\phi}_{NN} = \frac{q}{1-(1-q)p(1-p)} = \frac{1}{1+\mu(1-p(1-p))}$$

where $\mu = (1-q)/q$. We similarly have $\widetilde{\phi}_{NU} = (1-q)(1-p)\widetilde{\phi}_{NU} + (1-q)p^2\widetilde{\phi}_{NN}$ implying that

$$\widetilde{\phi}_{NU} = \frac{\mu p^2}{1+\mu p}\widetilde{\phi}_{NN},$$

and $\widetilde{\phi}_{UN} = (1-q)p\widetilde{\phi}_{UN} + (1-q)p(1-p)\widetilde{\phi}_{NN}$, implying that

$$\widetilde{\phi}_{UN} = \frac{\mu p(1-p)}{1+\mu(1-p)}\widetilde{\phi}_{NN},$$

Simple algebra permits to conclude that $\widetilde{\phi}_N = \widetilde{\phi}_{NN} + \widetilde{\phi}_{NU} < \phi_N$. Q.E.D.

Proof of Proposition 3

Consider first the case where each player i follows the strategy σ_i^* that plays C in N and D in U, and let ϕ denote the long-run distribution over states induced by $\sigma^* = (\sigma_1^*, \sigma_2^*)$. By symmetry, and since the dynamic system has equal chances of exiting from NN and of exiting from UU, we have:

$$\phi_{NN} = \phi_{UU} \text{ and } \phi_{NU} = \phi_{UN} \tag{6.1}$$

The value to player 1 from following that strategy is thus

$$\begin{aligned} v_1(\sigma_1^*, \sigma_2^*) &= \phi_{NN} + (1+L)\phi_{UN} - L\phi_{NU} \\ &= \phi_{NN} + \phi_{UN} = \frac{1}{2}(\phi_{NN} + \phi_{UN} + \phi_{NU} + \phi_{UU}) \\ &= \frac{1}{2}. \end{aligned}$$

Now if player 1 cooperates in both states (σ^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 (σ^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 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 $\widehat{\sigma}$ 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 $(\widehat{\sigma}, \sigma)$, and the value to player 1 from following $\widehat{\sigma}$ thus remains equal to 1/2. It follows that unless $p = 1 - \frac{1}{2(1+L)}$, the strategy profiles (σ, σ) and $(\widehat{\sigma}, \sigma)$ cannot be equilibria. Similar considerations show that the strategy profile $(\widehat{\sigma}, \widehat{\sigma})$ 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.

Many states with threshold strategies

Consider a possibly large set $S_i = (s^1, ..., s^n)$, and to fix ideas, define the following transition over states. Assume that when a player cooperates, receiving a good signal moves him to a lower state (from s_k to s_{k-1}) with probability h, while a bad signal moves him to a higher state (from s_k to s_{k+1}) with probability h'. When a player defects, his state does not change. Signal $z = 1$ moves the players back to state s^1.[42]

Denote by σ_{k_0} the strategy that cooperates at all states s^k for $k \leq k_0$, and by σ_0 the strategy that defects at all states. The parameter k_0 thus characterizes the leniency level associated with strategy σ_{k_0}. Note that the strategies σ_0 and σ_n coincide respectively with σ^D and σ^C, and that, given our assumed transition, the strategy σ_1 could be derived with two states as before. We are interested in whether σ_{k_0} can be supported as equilibrium behavior for some k_0 not small.

Fix $k_0 \geq 1$. For any k, the strategy profile (σ_k, σ_{k_0}) induces an ergodic distribution over actions profiles $a \in A = \{CC, CD, DC, DD\}$, which we denote by $\phi_k = \{\phi_k^a\}_{a \in A}$. It will also be convenient to denote by $\phi_k^{2,C}$ the probability that 2 cooperates when (σ_k, σ_{k_0}) is played: $\phi_k^{2,C} = \phi_k^{CC} + \phi_k^{DC}$. We shall say that players are in a cooperative phase when CC is being played, and in a punishment phase otherwise. When CD or DC is being played, we shall say that players are in a transition phase.[43]

We have:

$$\begin{aligned} v_1(\sigma_k, \sigma_{k_0}) &= \phi_k^{CC} + (1+L)\phi_k^{DC} - L\phi_k^{CD} \\ &= \phi_k^{2,C} - L(\phi_k^{CD} - \phi_k^{DC}) \end{aligned}$$

By symmetry, we have $\phi_{k_0}^{DC} = \phi_{k_0}^{CD}$, so $v_1(\sigma_{k_0}, \sigma_{k_0}) = \phi_{k_0}^{2,C}$ so equilibrium conditions write as

$$\phi_k^{2,C} - L(\phi_k^{CD} - \phi_k^{DC}) \leq \phi_{k_0}^{2,C} \text{ for all } k.$$

[42] This assumption is not essential. Signal $z = 1$ could move players back to some state n_0.

[43] Thus a punishment phase starts by a transition phase (unless both start defecting simultaneously).

From these conditions we thus obtain either lower or upper constraint on L. For example, if player 1 adopts a more lenient strategy ($k > k_0$), he triggers a punishment phase less often so $\phi_k^{2,C} > \phi_{k_0}^{2,C}$, the equilibrium condition requires $\phi_k^{CD} > \phi_k^{DC}$ (transition phases should be on average more costly to the more lenient player) and it imposes a lower bound on L.

Proof of Proposition 4

Consider parameters (p, q, L) for which σ^* is a strict equilibrium. We investigate whether player 1 has a profitable deviation when he puts weight at least $1 - \varepsilon$ on a pure strategy. If player 1 deviates to a strategy nearby σ^C, he obtains a payoff nearby $v_1(\sigma^C, \sigma^*)$, so since σ^* is strict, these deviations cannot be profitable for ε small enough. The same argument applies to. σ^D or $\widetilde{\sigma}$ (the strategy that plays D in N and C in U).

So we only need to investigate the case where player 1 would put weight $1 - \varepsilon$ on σ^*. Note that because ε is arbitrarily small (so that deviations occur arbitrarily infrequently).

Case 1: deviations to σ^D.

We consider first deviations that put weight ε on σ^D. We call $\sigma_1^{\varepsilon,D}$ that strategy, and $\phi^{\varepsilon,d}$ the long run distribution over state profiles induced by $(\sigma_1^{\varepsilon,D}, \sigma_2^*)$, so that $\phi^{0,d} = \phi$ corresponds to the long-run distribution induced by σ^*. The condition that $\sigma_1^{\varepsilon,D}$ is not profitable for ε small enough writes as:

$$\frac{\partial}{\partial \alpha}[\phi_{NN}^{\alpha,d} - L\phi_{NU}^{\alpha,d} + (1 + L)\phi_{UN}^{\alpha,d}] \mid_{\alpha=0} + L\phi_N^{0,d} < 0, \tag{6.2}$$

where the first term in this expression characterize the (adverse) effect on long-run probabilities induced by the deviation, while the second characterizes the extra gain player 1 obtains from playing D (these gains happen when player 2 is in the normal state.

Define $p(\alpha) = p(1 - \alpha) + \alpha(1 - p) = p - \alpha(2p - 1)$. We have $\phi_{NN}^{\alpha,d} = q + (1 - q)pp(\alpha)\phi_{NN}^{\alpha,d}$, implying:

$$\phi_{NN}^{\alpha,d} = \frac{1}{1 + \mu pp(\alpha)}$$

We also have $\phi_{NU}^{\alpha,d} = (1 - q)[(1 - p)\phi_{NU}^{\alpha,d} + (1 - p(\alpha))p\phi_{NN}^{\alpha,d}]$, and $\phi_{UN}^{\alpha,d} = (1 - q)[(1 - p)\phi_{UN}^{\alpha,d} + (p(\alpha))(1 - p)\phi_{NN}^{\alpha,d}]$, thus implying:

$$\phi_{NU}^{\alpha,d} = \frac{\mu p(1 - p(\alpha))}{1 + \mu p}\phi_{NN}^{\alpha,d} \text{ and } \phi_{UN}^{\alpha,d} = \frac{\mu p(\alpha)(1 - p)}{1 + \mu p}\phi_{NN}^{\alpha,d}.$$

It is immediate to see that $\frac{\partial}{\partial \alpha}[\phi_{UN}^{\alpha,d} - \phi_{NU}^{\alpha,d}] \mid_{\alpha=0} < 0$, so for (6.2) to hold, it is sufficient that

$$L\phi_N^{0,d} < -\frac{\partial \phi_N^{\alpha,d}}{\partial \alpha} \mid_{\alpha=0} \tag{6.3}$$

Letting $h(\alpha) = \frac{\mu p(\alpha)(1-p)}{1+\mu p}$, so that $\phi_N^{\alpha,d} = (1 + h(\alpha))\phi_{NN}^{\alpha,d}$, and $h = h(0)$, we get

$$-\frac{\partial \phi_N^{\alpha,d}}{\partial \alpha} \mid_{\alpha=0} = -(1+h)\frac{\partial \phi_{NN}^{\alpha,d}}{\partial \alpha} \mid_{\alpha=0} -h'(0)\phi_{NN}^{\alpha,d}$$

Using $-h'(0) = \frac{\mu(1-p)(2p-1)}{1+\mu p}$ and $-\frac{\partial \phi_{NN}^{\alpha,d}}{\partial \alpha} = \mu p(2p-1)(\phi_{NN}^{\alpha,d})^2$, inequality (6.3) holds when

$$L < p\mu(2p-1)\phi_{NN} + (1-p)\mu(2p-1)\frac{1}{1+\mu p}$$

Since $\frac{1}{1+\mu p} > \frac{1}{1+\mu p^2} = \phi_{NN}$, expression (6.3) holds a fortiori when

$$L < p(\mu(2p-1)\phi_{NN} = \bar{L}$$

Case 2: Deviations to σ^C.

Using similar notations, the condition that $\sigma_1^{\varepsilon,C}$ is not profitable for ε small enough writes as:

$$\frac{\partial}{\partial \alpha}[\phi_{NN}^{\alpha,c} - L\phi_{NU}^{\alpha,c} + (1+L)\phi_{UN}^{\alpha,c}] \mid_{\alpha=0} -L\phi_U < 0, \tag{6.4}$$

where the first term in this expression characterizes the (now positive) effect on long-run probabilities induced by the deviation, while the second characterizes the extra loss player 1 suffers from playing C in events where player 2 is in state U.

It is immediate to check that

$$\phi_{NN}^{\alpha,c} = \phi_{NN}^{0,c} = \phi_{NN} \qquad \phi_{NU}^{\alpha,c} = \phi_{NU}^{0,c} = \phi_{NU}, \text{ and } \phi_{UN}^{\alpha,c} = \frac{\mu p(1-p)}{1+\mu(1-p(\alpha))}\phi_{NN}$$

Condition (6.4) is thus equivalent to

$$(1+L)\frac{\partial \phi_{UN}^{\alpha,c}}{\partial \alpha} < L\phi_U$$

Since $\frac{\partial \phi_{UN}^{\alpha,c}}{\partial \alpha} = \frac{\mu(2p-1)}{1+\mu p}\phi_{UN}^{0,c}$, since $\phi_{UN}^{0,c} = \frac{\mu p(1-p)}{1+\mu p}\phi_{NN}^{0,c}$, since $\phi_U = 1 - \phi_{NN} - \phi_{UN}$ and since $1 - \phi_{NN} = \mu(1-p^2)\phi_{NN}$, one obtains

$$L > \underline{\underline{L}} \equiv \frac{\mu p(2p-1)}{(1+\mu p)^2 + \mu p(1+\mu p^2)}$$

Simple algebra shows that for any $p > 1/2$ and $\mu > 0$, $\underline{L} > \underline{\underline{L}}$.[44]

[44] Recall that $\underline{L} = \frac{\mu p(2p-1)}{(1+\mu p)(1+\mu(1-p^2))}$.

Finally, observe that for deviations to strategies that would put weight on all σ^C, σ^D and $\widetilde{\sigma}$, the first order effect for ε small is a combination of the effect characterized above, so these deviations are not profitable either.

Proof of Proposition 5

Assume p is close to 1. Under σ, 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).[45] 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

$$L < 1/q - 1 \tag{6.5}$$

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 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 that is incompatible with inequality (6.5).

Thus strategy σ 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.

Recoordination with stochastic tit-for-tat

The analysis in Section 2.4 applies: following the candidate equilibrium strategy σ^* induces a long run payoff equal to ϕ_N; deviating to σ^D generates a payoff equal to $(1+L)\phi_N^D$, and deviating

[45] Omitting terms comparable to $(1-p)$, this is true whether the current state profile is (U,U), (U,N) or (N,U).

to σ^C generates a payoff equal to $\phi_N^C - L\phi_U^C$. Deterring these deviations thus, as before, requires:

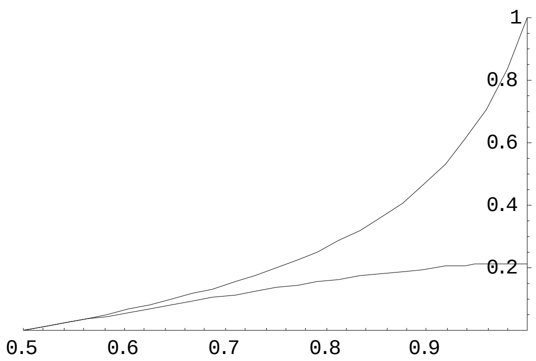
$$\underline{L} = \frac{\phi_U}{\phi_U^C} - 1 < L < \bar{L} = \frac{\phi_N}{\phi_N^D} - 1.$$

The difference with Section 2.4 however is that the long run probabilities take different values (they are functions of p, b, k and h), and these long-run probabilities are more difficult to compute. We state here our main result:

Proposition 6: $\underline{L} < \bar{L}$ if and only if $h < (1-b)k$.

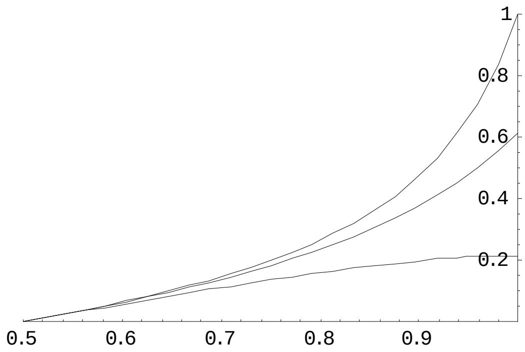
Intuitively, the higher k and the smaller h, the easier it is to recoordinate on state NN from state UU. Indeed, once one player, say player 1, switches to N and cooperates, the other player will be likely to switch to N in the next period if k is large enough. If h is high as well however, then it is quite possible that the player who has initiated a move back to cooperation (i.e. player 1) returns to state U before player 2 switches to N, and recoordination does not occur.

In the following graph, the (p, L) for which the two inequalities above are compatible when $h = b = 0.1$ and $k = 0.9$ are those between the two curves:



Note that this graph shows the parameters for which a player does not gain by deviating to the strategy σ^C that plays C at all states, or to the strategy σ^D that plays D at all states. For these parameters, σ^* is an equilibrium if only threshold strategies are considered. If we include the possibility that a player deviates to the strategy σ^{DC} that plays C at U and D at N, then another (possibly tighter) upper constraint must hold. Intuitively, σ^{DC} is a "manipulative" strategy in which a player cooperates to induce the other cooperate, and then defects, hoping that the other would not react too quickly: this strategy may be profitable precisely when the other is quickly forgiving (k high), and slow to react to bad signals (h small). The following figure adds this new upper constraint to the set of parameters for which σ^* is an equilibrium.

The figure is again drawn for $b = h = 0.1$ and $k = 0.9$.



The middle curve is the new upper constraint that must be satisfied to ensure that σ^{DC} is not a profitable deviation.

Computations: Proposition 6 is an immediate corollary of a stronger result which we now state, and which applies to any mental system where transitions depend only on the signal received. Any mental system induces probabilities that a given player will switch from one state to the other, as a function of the action (C or D) played by the other player. For transitions from N to U, we let:

$$p_C = \Pr(N \to U \mid C) \text{ and } p_D = \Pr(N \to U \mid D)$$

and for transitions from U to N:

$$q_C = \Pr(U \to N \mid C) \text{ and } q_D = \Pr(U \to N \mid D).$$

The following proposition holds:

Proposition 7. *Assume $p_D > p_C$ and $q_C > q_D$. Then $\underline{L} < \bar{L}$ if and only if $p_C + q_C > p_D + q_D$.*

When $p_D > p_C$ and $q_C > q_D$, the mental system has a Tit-for-Tat flavor: a player tends to become upset when his opponent defects, and forgiving when his opponent cooperates. Proposition 7 says that a necessary condition for such a mental system to support cooperation is that forgiveness induced by cooperation ($q_C - q_D$) is stronger than the deterioration induced by defection ($p_D - p_C$). Proposition 6 is a corollary because given our assumption about the mental system, we have:

$$\begin{aligned} p_C &= (1-p)h \text{ and } p_D = ph \\ q_C &= b + (1-b)pk \text{ and } q_D = b + (1-b)(1-p)k \end{aligned}$$

so the conditions of Proposition 7 are satisfied.

Proof: We need to compute ϕ_N^D, ϕ_N^C and ϕ_N and check whether and when $\Delta \equiv \phi_N(1 - \phi_N^C) - (1-\phi_N)\phi_N^D$ is positive. The long-run probabilities ϕ_N^C and ϕ_N^D are easy to compute. We have $\phi_N^C = (1-p_C)\phi_N^C + q_C(1-\phi_N^C)$, which yields:

$$\phi_N^C = \frac{q_C}{q_C + p_C}.$$

Similarly, we have $\phi_N^D = (1-p_D)\phi_N^D + q_D(1-\phi_N^D)$, implying that:

$$\phi_N^D = \frac{q_D}{q_D + p_D}.$$

To compute $\phi_N = \phi_{NN} + \phi_{UN}$, we have to find a probability vector $\phi = (\phi_{NN}, \phi_{NU}, \phi_{UN}, \phi_{UU})$ which is fixed point of:

$$\phi = \phi.M \text{ where } M = \begin{pmatrix} (1-p_C)^2 & (1-p_D)q_C & q_C(1-p_D) & (q_D)^2 \\ (1-p_C)p_C & (1-p_D)(1-q_C) & p_D q_C & q_D(1-q_D) \\ p_C(1-p_C) & p_D q_C & (1-p_D)(1-q_C) & (1-q_D)q_D \\ (p_C)^2 & p_D(1-q_C) & p_D(1-q_C) & (1-q_D)^2 \end{pmatrix}.$$

It can be verified that for all (p_C, p_D, q_C, q_D) satisfying the conditions of the proposition, Δ has the same sign as

$$(p_C + q_C - p_D - q_D)[(p_D - p_C)q_C + p_D(q_C - q_D)].$$

Recoordination with asymmetric mental systems

In this example the two players have different mental systems, each with three states N, U and F. Transitions from N to U are as before. Transitions from U to F are stochastic, depending on independent private signals z_1 and z_2, $z_i \in \{0,1\}$ and $\Pr\{z_i = 1\} = q$. For player 2, a transition from F to N requires receiving a good signal, $y_2 = 1$ while for player 1, such a transition is automatic. These transitions are summarized in Figure 7.

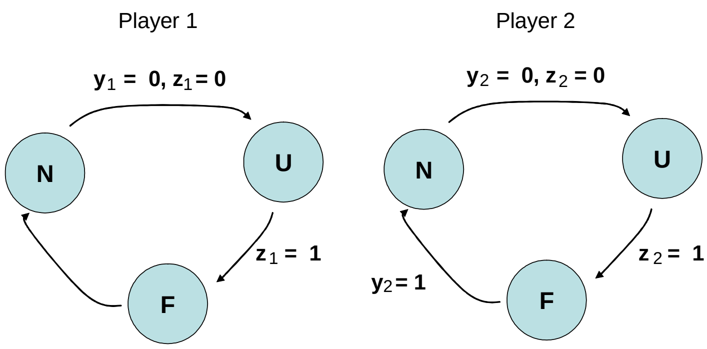


Figure 7: “Successful” independent resetting

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 or at the same time player 1 moves to F. In that case, the most likely event is that players will coordinate back to (N, N) (with probability close to 1).[46] Alternatively, player 1 moves to state F before player 2 moves to state F. In that case, the most likely event is that players switch to (N, F) or (N, U), and then back to state U for player 1, 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:

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

[46] This is because once in state profile (F, F), 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.

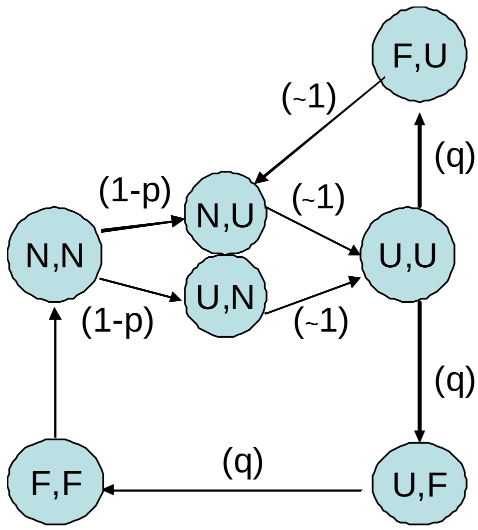


Figure 8: Transition of mental state

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 8).

Analysis:

When players follow the proposed strategy profile, they alternate between long phases of cooperation (of length $1/\pi$ with $\pi \simeq 2(1-p)$), 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 $2/q$.[47]

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 F (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). The condition is thus:

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

[47] The exact cost is larger because before recoordination actually occurs, there is at least one period (and possibly more periods in case of failed attempts) in which player 1 cooperates while player 2 still defects. It can be shown that a better approximation of the cost is $2/q + \frac{1}{2} + 3(1+L)$.

or equivalently:

$$L > 3.$$

Incentives of player 2 at N: Defection generates short periods 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 looses 1. Hence the condition

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

We omit the verification of the other incentives, which are easily checked and automatically satisfied. QED

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 9.

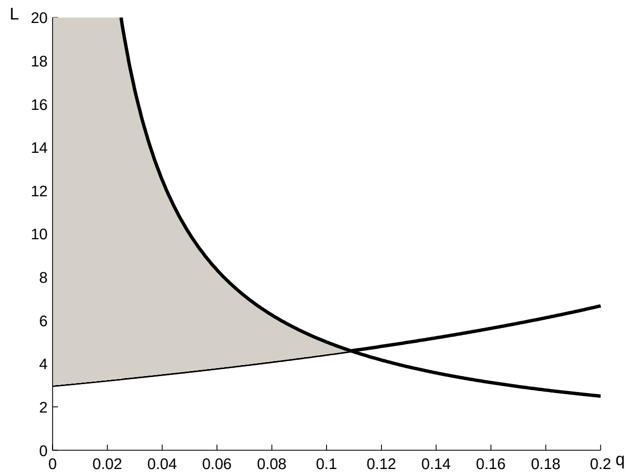


Figure 9: Values of q and L for which cooperation is possible (for p close to 1).

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). 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.

7. Bibliography

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

Abreu, D. [1986], "Extremal Equilibria of Oligopolistic Supergames", *Journal of Economic Theory*, 39, 191-225.

Abreu, D. [1988], "On the Theory of Infinitely Repeated Games with Discounting", *Econometrica* , 56, 383-396.

Abreu, Dilip, David Pearce and Ennio Stacchetti [1990], "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica* 58, 1041-1063.

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.

Aumann, R., [1981], *Survey of repeated games*, in "Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern," pp. 11-42, Bibliograph. Inst., Zurich.

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

Ben-Porath, E. [1990], "The Complexity of Computing a Best Response Automaton in Repeated Games with Mixed Strategies", *Games and Economic Behavior* 2 1-12.

Ben-Porath, E. [1993], "Repeated Games with Finite Automata," *Journal of Economic Theory*, 59, 17-32.

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 [2009], "The Extent of Cooperation," mimeo, University of Pennsylvania.

Compte, O. and A. Postlewaite [2012], "Strategy restrictions and Limited knowledge", mimeo

Compte, O. and A. Postlewaite [2013], "Belief Free Equilibria," 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.

Gilboa, I. [1988]. "The Complexity of Computing Best-Response Automata in Repeated Games," *Journal of Economic Theory*, 45, 342-352.

Kalai, E. and W. Stanford [1988], "Finite Rationality and Interpersonal Complexity in Repeated Games", *Econometrica* , 56, 397-410

Kalai, E., D. Samet and W. Stanford [1988], "A Note on Reactive Equilibria in the Discounted Prisoner's Dilemma and Associated Games", *International Journal of Game Theory*, 17, 177-186

Hauser, C. and H. Hopenhayn [2008], "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. 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.

Mobius, M. [2001], "Trading Favors," mimeo, Harvard University.

Monte, D. [2007], "Reputation and Bounded Memory: A Theory of Inertia and Skepticism," mimeo.

Neyman, A. [1985], "Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoner's Dilemma," *Economics Letters*. 19:227–229.

Neyman, A. [1998], "Finitely Repeated Games with Finite Automata," *Mathematics of Operations Research* , 23, 513-552

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.

Romero, J. [2010], "Bounded Rationality in Repeated Games," mimeo, Caltech.

Rubinstein,A. [1986], "Finite automata play the repeated prisoner's dilemma," *Journal of Economic Theory*, 39, 83-96

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.