Reputation and Bounded Memory in Repeated Games with Incomplete Information

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Daniel Monte

Dissertation Director: Professor Benjamin Polak

December 2007

Contents

List of Tables

List of Figures

ABSTRACT

Reputation and Bounded Memory in Repeated Games with Incomplete Information

Daniel Monte

2007

This dissertation is a study of bounded memory and reputation in games. In the first chapter we show that the optimal use of a finite memory may induce inertia and infrequent updating on a player's behavior, sometimes for strategic reasons. The setting is a repeated cheap-talk game with incomplete information on the sender's type. The receiver has only a fixed number of memory states available. He knows that he is forgetful and his strategy is to choose an action rule, which is a map from each memory state to the set of actions, and a transition rule from state to state. Unlike in most models of bounded memory, we view memory as a conscious process: in equilibrium strategies must be consistent with beliefs. First, we show that the equilibrium transition rule will be monotonic. Second, we show that when memory constraints are severe, the player's transition rule will involve randomization before he reaches the extreme states. In a game, randomization has two interpretations: it is used as a memory-saving device and as a screening device (to test the opponent before updating).

The second chapter shows that bounded memory with sequential rationality constraints can explain long-run reputation even in the case of parties with opposite interests. Memory is modeled as a finite set of states. The strategy of the player is an action rule, which is

a map from memory to actions, and a transition rule from state to state. In equilibrium, strategies and beliefs must be consistent. The setting is an infinitely repeated game with one-sided incomplete information. The informed player is either a zero-sum normal type or a commitment type playing a mixed strategy. Under full memory, types are revealed asymptotically. Bounded memory in the form of a finite automata cannot, by itself, explain long-run reputations, as we show in the paper. However, if memory is modeled as a conscious process with sequential rationality constraints, then long-run reputation will be sustained in any Markovian equilibrium.

In the third chapter we study a repeated adverse selection model in which the best contract for the principal is to reproduce the outcome with sequential rationality constraints. The model suggests that contracting will only have an impact when the adverse selection problem is *less* severe.

Acknowledgments

I am grateful to my dissertation committee: Dirk Bergemann, Stephen Morris and Ben Polak. I am also thankful for conversations with Attila Ambrus, Eduardo Faingold, Hanming Fang, John Geanakoplos, Itzhak Gilboa, Ehud Kalai, Pei-yu Lo, George Mailath, Wolfgang Pesendorfer, Maher Said, Lones Smith, Peter Sorensen and Joel Watson.

Chapter 1

Introduction

We study the implications of bounded memory in games. A typical assumption in economics is that people have a perfect memory. In models of long-term relationships, players condition their strategies on the entire history of the game, irrespective of how long and complicated that history may be. Yet, in reality, most people forget things. They categorize. They often ignore information and their updating may be infrequent. We study a model of bounded memory that captures these features.

Throughout this dissertation, memory is modeled as a finite set of states. All the agent knows about the history of the game is her current memory state. The player's strategy is to choose an action rule, which is a map from each memory state to the set of actions, and a transition rule from state to state. First, we characterize the equilibrium memory rule in a reputation game and show that it may induce inertia and infrequent updating, sometimes for strategic reasons. Second, we show that in a long-term relationship, learning (or type separation) is never complete, in contrast to recent results in reputation games.

We view both action choice and updating rule as a conscious process, unlike the literature on finite automata that assumes commitment to the ex-ante strategy. Therefore, the player in our model is subject to sequential rationality constraints. The action chosen at each state must be optimal given the beliefs at that state. And, the transition rule from each state must be optimal given the beliefs at that memory state and taking as given the strategy—both action and transition rules—at all her states. The reason for taking the strategy at all states as given when deciding on an action or on which state to move is that if an agent deviates today, she will not remember it tomorrow. This idea of sequential rationality constraints in bounded memory comes from Piccione and Rubinstein (1997) and Wilson (2003), but these authors studied single-person decision problems. Here we study games, where the inability to commit matters.

In the first chapter, "Reputation and Bounded Memory", we focus on reputation games, where one player is trying to learn her opponent's type (for example, repeated games with incomplete information). In these games memory plays a central role, since remembering the exact history of the game is important for learning. The setting is a repeated zero-sum game with two players, one of whom has bounded memory. She faces a player who, with some exogenous prior probability, is committed to a pure strategy. This game is based on Sobel's (1985) credible advice model, in which a policy maker is uncertain about her adviser's preferences.

We characterize the equilibrium action and updating rule in this game. We show that the bounded memory player must have beliefs about her opponent's type that are one and zero in her two “extreme” memory states. The transition rule from state to state must satisfy a weak monotonicity property; hence, the resulting endogenous updating rule will resemble Bayesian updating whenever possible.

When the number of memory states is not "large enough," the player will use randomization in her transition rule before reaching the extreme states. After a signal, the player randomizes between updating and remaining on the same state. This leads to infrequent updating and "inertia" on the player's behavior. Similar to single-person games, randomization can be interpreted as a memory-saving device: with no capacity to store all the informative signals, the player optimally decides to discard some of them. However, we show that, in games, there is an additional strategic role for randomization. It is used as a screening device: the player is "testing" the opponent before updating.

The second chapter, "Bounded Memory and Limits on Learning," contributes to the literature on reputation and repeated games with incomplete information. A celebrated recent result in this literature is that, asymptotically, the play of the game converges to the play of a complete information game (see Cripps et al. (2004), for example). This means that players can profit from a "false" reputation only in the short-run. Constant opportunistic behavior will lead to statistical revelation of the actual type, which means no long-run reputation.

The setting is very similar to that in the previous chapter. Here, the bounded memory agent faces a player who, with some exogenous probability, might be committed to a mixed strategy, instead of a pure strategy. This difference is analogous to the distinction between games with perfect monitoring versus imperfect monitoring.

We show that under bounded memory, we will not have full learning (or type separation), even in the long run. The main intuition for this result is that with bounded memory the agent can hold only a finite number of beliefs in equilibrium. And, these beliefs cannot be too far apart from each other, or else the sequential rationality constraints would not be satisfied. Therefore, with initial uncertainty about types and bounded memory on the uninformed player, long-term reputations can be sustained even in the extreme case in which agents have opposite preferences.

Finally, the third chapter of this dissertation, studies a credibility model in which the receiver has the ability to commit to a strategy before the first stage game. We show that this ability to commit will not improve the receiver's payoff. In other words, the optimal contract in this credibility model is to reproduce the perfect Bayesian equilibrium outcome of the game without the ability of commitment. This suggests that the ability to commit will only benefit a player in cases in which the utility of the receiver is not directly opposed to that of the sender.

Chapter 2

Reputation and Bounded Memory

2.1 Introduction

An implicit assumption in most economic models is that people have a perfect memory and update their beliefs using Bayes' rule. In models of long-term relationships, players condition their strategies on the entire history of the game, irrespective of how long and complicated that history may be. Yet most people forget things. They categorize. They often ignore information and update infrequently. This chapter studies a model of bounded memory that captures these memory imperfections.

In our model, the bounded memory player has only a fixed number of memory states available. He knows the information in the current period, but he is forgetful between periods. All he knows about the history of the game is his current memory state. He is aware of his memory constraints and chooses the best strategy to deal with them. He chooses both an action rule, which is a map from each memory state to the set of actions, and a transition rule from state to state.

In this chapter we focus on reputation games. These are games in which one player is learning about his opponent's type. We characterize the equilibrium memory rule in a reputation game and show the implications on the agent's behavior. We show that the transition rule must satisfy an intuitive weak monotonicity property; loosely speaking, good news leads the bounded-memory player to move to a memory state associated with a weakly higher reputation. But we also show that when memory constraints are severe, the player will use randomization in his updating rule, which may induce "inertia" in his behavior and infrequent updating. Thus, unresponsiveness to new information is, in fact, optimal for the player when his memory is small.

A key innovation in this chapter is that we view both action choice and memory constraints as a conscious process. At all points in the game, the player is aware of his memory constraints and consciously optimizes given what he knows. Thus, the player is subject to sequential rationality constraints. The action he chooses at each memory state and the transition rule from each state must be optimal given his beliefs at that state and taking as given the strategy - both action and transition rules - at all his states. The reason the player takes his own strategy at all states as given when deciding on an action or on which state to move is that if he deviates today, he will not remember it tomorrow.

Conscious memory distinguishes our model from the standard finite automata in the literature. Like ours, the standard automaton has a fixed set of states, a transition rule and an action rule. But standard automata can be committed to a strategy ex ante, and hence does not face sequential rationality constraints. It is as if the standard automaton unconsciously follows the pre-scribed action and transition rules chosen at the start of the game. The idea of sequential rationality in bounded memory was introduced by Piccione and Rubinstein (1997) and Wilson (2003), but these authors studied single-person decision problems. Here we study games, where the inability to commit matters.[1]

The setting of this chapter is a repeated cheap-talk game with incomplete information. It is based on Sobel's (1985) credible advice model, where a policy maker is uncertain about his adviser's preferences. On every period the adviser, or sender, knows the true state of the world and reports it to the policy maker, the receiver. However, the sender need not report truthfully; his reporting strategy will depend on his preferences. The sender is either a commitment type, someone who always tells the truth, or a strategic type, someone with opposite preferences to those of the receiver.

Once the receiver observes the report from the sender, he takes an action and the payoffs are realized. Payoffs depend only on the state of the world in the current period and on the action taken by the receiver. At the end of the period, the state of the world is verified, and the receiver knows whether the sender has lied to him or not. The receiver then updates his beliefs (Bayesian updating in Sobel's model) concerning the sender's type. Thus, the receiver acts based on the report of his adviser at the same time that he is learning about his opponent's type. In our model, the receiver has bounded memory; instead of updating using Bayes' rule, he adheres to broadly defined categories. For example, if the receiver had only three memory states, he might categorize the sender as "a friend", "an enemy", or "still unclear".

We show in propositions 2.2 and 2.3 necessary conditions for equilibria. In particular, we show that the updating rule must be weakly increasing as long as the receiver obtains

[1] Rubinstein (1986) and Kalai and Neme (1992) also study automata models with a perfection requirement. The solution concept used in this paper is substantially different, though, since it requires consistent beliefs, as will be discussed later.

truthful reports from the sender. This implies that the receiver's belief that the sender is committed to the truth is stochastically higher after a truthful report. Moreover, because a false report leaves no uncertainty in the mind of the receiver, he moves to his 'lowest' memory state after this signal. We show that, even when the bounded player has very few memory states and hence can not keep track of large amounts of information, in his "lowest" state his belief on the sender being honest is zero, and in his "highest" state the belief is one. Surprisingly, this result holds even for the minimal case of only two memory states, or one-bit memory. This is shown in proposition 2.2.

Propositions 2.4 and 2.5 show that if the prior on the sender's type is higher than a particular threshold, the receiver will use deterministic transition rules. If this condition is not met, then the equilibrium transition rule will require randomization. Informally, this means that when the receiver does not have enough memory to keep track of the truthful reports, he will use randomization to overcome the memory problem and test the sender before updating.

The role of random transition rules in the optimal finite memory has been studied in single person decision problems. Hellman and Cover (1970) studied the two-hypothesis testing problem with a finite automaton (with ex-ante commitment to the strategy). A decision maker has to make a decision after a very long sequence of signals. However, the decision maker cannot recall all the sequence and has, instead, to choose the best way to store information given his finite set of memory states. A key result of the paper is that, for a discrete signal case, the transition rule is random in the extreme states. The authors concluded that, perhaps counter intuitively, the decision maker uses randomization as a memory-saving device.

The benefits of random transition rules for a decision-maker were also shown by Kalai and Solan (2003). They showed that randomization is necessary in a single person decision problem when the decision maker is restricted to automata. Their paper also showed the advantages of randomization in the transition rule versus randomization in the choice of action, a subject also discussed in this chapter.

Wilson (2003) studied a problem similar to Hellman and Cover (1970). In her model the decision maker was subject to sequential rationality constraints. The optimal memory rule obtained is similar to Hellman and Cover's and includes randomization in the extreme states. Moreover, she showed that modeling human memory as an optimal finite automaton can explain several biases in information processing described in the literature (see Rabin (1998) for a survey on behavioral biases).

Our results suggest that in an incomplete information game randomization in the transition rule is needed as a memory-saving device in much the same way as in Hellman and Cover (1970 and 1971), Kalai and Solan (2003) and Wilson (2003). However, unlike these single player models, this chapter shows that in games there is an additional strategic role for randomization. In the incomplete information game, randomization is used as a screening device: to test the opponent and give incentives for the opponent's type to be revealed early in the game.

In the second chapter we will consider an extension of the model studied here. In that chapter the commitment type of sender plays all actions with positive probability. With full memory, types are revealed asymptotically. However, if the uninformed player has bounded memory, we show that reputation will be sustained in any Markovian equilibrium. I.e., types are never fully separated. This finding contrasts with recent results on reputation games where the strategic effects of reputation eventually washes off, as in Benabou and Laroque (1992), Jackson and Kalai (1999) and Cripps et al (2004).

A player with bounded memory can hold only a finite number of beliefs in equilibrium. In Monte (2006) the commitment type plays a mixed strategy and the actions do not reveal as much information as it does in the present chapter. Thus, the beliefs that the bounded memory player holds in equilibrium cannot be too far apart, or else the incentive compatibility constraints wouldn't be satisfied: there would not be an incentive to move from one state to another regardless of the action observed. This imposes a maximum difference between the lowest and the highest beliefs. Thus, we can calculate a bound on learning, which is given by the extreme beliefs. In the present chapter, on the other hand, the beliefs can be far apart since one of the actions is fully revealing and thus, induces a substantive change on the bounded memory player's belief.

The study of the implications of an imperfect memory has taken two different modeling strategies in the literature. One approach is to make explicit assumptions about the memory process, while assuming that the agent is not aware of these limitations. This memory process could be, for example, bounded recall, where the agent is able to recall only the information of the last k periods.[2] Or, it could be based on memory decay, such as studied by Mullainathan (2002) and Sarafidis (2007). There are also the papers by Mullainathan (2001) and Fryer and Jackson (2003), where agents are restricted to hold a finite set of

[2] There are several papers on multi-player games with bounded recall, for example, Kalai and Stanford (1988), Lehrer (1988) and, more recently, Huck and Sarin (2004).

posteriors. In these papers the updating rule (categorization) is given exogenously; it is not part of the player's strategy.

The second approach in modeling memory restrictions is to assume constraints on the agent's memory, but such that he is fully aware of these limitations. The agent then decides on the optimal strategy given this constraint. The memory rule itself becomes part of the player's strategy.

This second approach includes the automata models, such as Hellman and Cover (1970). These models have also been studied to capture bounded rationality in implementing a strategy. For some of the early papers modeling economic agents as automata, see Neyman (1985), Rubinstein (1986) and Kalai and Stanford (1988).

The bounded memory model with sequential rationality constraints suggests that there is an alternative interpretation for the player, modeling him as a collection of agents.[3] These agents act with the same interests and do not communicate with each other except through the use of a finite set of messages (the memory states). Thus, this model is in many ways similar to dynastic repeated games as in Lagunoff and Matsui (2004) and Anderlini et al (2006). Each generation does not remember the past, but receives a message (from a finite set) from the previous generation. The current generation's memory about the game must be contained in the message received. In this sense, it is also similar to modeling a player as a team.[4]

This chapter is organized as follows. Section 2.2 consists of the description of the model

---

[3] Modeling a player as an organization of multiple selves was done earlier by Stroz (1956) and Isbell (1957).

[4] See Radner (1962) for a model of decisions with teams.

and the definitions of memory, strategies, as well as the equilibrium concept. The case of two memory states is shown in section 2.3. Section 2.4 gives the main result of the chapter: the characterization of the memory rule and the condition for the receiver to have deterministic transition rules, given a memory with n states. We show the example of 3 memory states in section 2.5. In section 2.6 we present a discussion of the incentive compatibility concept and a comparison with an automaton model. Section 2.7 concludes the chapter. Most of the proofs are in the appendix.

2.2 Model

The setting of our study, a model based on Sobel (1985), is a repeated cheap-talk game in which the receiver has incomplete information on the sender's type. Before the first stage game, Nature draws one of two possible types for the sender, about which the receiver is uninformed. With probability ρ the sender is a behavioral type committed to a pure strategy: he always tells the truth (truth and lie will be defined below). This behavioral type will be denoted B. With probability $(1-\rho)$ the sender is a "strategic type" S, with utility opposite to the receiver's.[5]

The timing of every stage game is the following. Nature draws a state of the world in every period, $\omega_t \in \Omega = \{0,1\}$, each happening with probability $\frac{1}{2}$. The sender observes ω_t and sends a message $m_t \in \{0,1\}$ to the receiver. This message has no direct influence on the player's payoffs. We will say that the sender tells the *truth* when $m_t = \omega_t$. Otherwise, he *lies*.

[5] Sobel (1985) calls the honest type the "Friend" and the strategic type, the "Enemy".

The receiver observes the message and takes an action a_t in the interval $[0, 1]$. After he takes the action, the payoffs are realized and the states are verified. At this point, the receiver can tell whether the sender has lied to him. Based on this information, the receiver updates his belief on the sender's type.

The game is repeated, but after every period there is an exogenous stopping probability η. This variable is capturing an exogenous probability that the relationship will end. We will focus on the case where this probability η is very small so that the players expect the game to go on for a very long horizon. The players discount their repeated game payoff using this stopping probability and also using a discount factor $\delta \leq 1$.

The receiver maximizes his goal when he takes an action that matches the state of the world. He is worse off when his action is 'far' from the true state. The particular functional form of utility considered in this chapter is a quadratic loss function. Thus, the stage game payoff of the receiver is: $u_R = -\left(a_t - \omega_t\right)^2$. The strategic sender has preferences completely opposite to those of the receiver, $u_S = \left(a_t - \omega_t\right)^2$.

Under full memory, the trade-off for the strategic sender is between building reputation or revealing himself. He might want to mimic the behavioral type and build reputation for the following stage game. Or, he might want to lie and reveal himself. Once he lies, he plays a zero-sum game with the receiver, and the unique equilibrium of this subgame is babbling, which means that the receiver ignores the sender's message when taking an action. We will later see that this trade-off is still present in the game with a bounded memory receiver.

Memory and Strategies A history in this game is defined as Nature's choice of the actual type, the sequence of action profiles, states of the world, Nature's choice about the repeated game ending or continuing, and the memory states of the receiver. The set of histories in the game is denoted by H. The sender, who is unconstrained, will condition his strategy on the observed history of the game.

Since the names of the states are irrelevant, we will define the action space for the sender to be $\{T, L\}$ where T is a "truth" and L is a "lie". We define the strategic sender's strategy as:

$$q : H \to \Delta \{T, L\} .$$

With slight abuse of notation, we will refer to $q(h)$ as the probability of telling the truth given the history h.

To simplify the analysis, we assume that at every period of the game the sender knows the receiver's current memory state. This assumption will leave out the sender's inference problem. We discuss this assumption further in section 2.4.2. We focus on equilibria in which the probability that the sender will tell the truth or lie will vary only across states, but not across time. Thus, we look only at equilibria with Markovian strategies.

The memory of the receiver is defined as a finite set of states $\mathcal{M} = \{1, 2, ..., n\}$. A typical element of $\mathcal{M}$ is denoted by s_i or s_j, or simply i or j.

At the start of each period, the receiver must decide on an action based on his current memory state, which is all the information that he has about the history of the game. We can write his action rule as:

$$a : \mathcal{M} \to [0, 1] , \tag{2.1}$$

interpreted as the probability (at the current memory state) that the receiver will follow the sender's advice.

At the end of each period, the receiver must decide which memory state to move to next based on his current memory state and whether that period's message was true or false. Allowing for the possibility of randomization, we can write the transition rule as a map

$$\varphi : \mathcal{M} \times \{T, L\} \rightarrow \Delta(\mathcal{M}). \tag{2.2}$$

We denote $\varphi_{_T}(i, j)$ as the probability of moving from state i to state j given that the sender told the truth. This transition rule will determine how the receiver updates beliefs.

One way to think of this is that the bounded memory player's knowledge about the history of the game is summarized by an n–valued statistic s_i, which is updated according to the map φ.

Finally, it is also part of the receiver's strategy to decide, before the first stage game, his initial distribution over the memory states $\varphi_0 \in \Delta(\mathcal{M})$.

The strategy for the receiver is the pair (φ, a) and we denote the strategy profile by $\sigma = (\varphi, a, q)$.

Beliefs As described, we view memory as a conscious process. Players know that they are forgetful. At every memory state they will hold a distribution of beliefs over the set of histories. Given a strategy profile $\sigma = (\varphi, a, q)$, the memory states form a partition of the possible histories H, so we can write h element of s_i for a history that would result in the receiver being at state s_i. Let $\mu(h|s_i, \sigma)$ denote the belief of the receiver in state s_i and given the strategy profile σ that the correct history is h. As usual, at any information set the beliefs about all histories must sum up to one

$$\sum_{h \in s_i} \mu(h|s_i, \sigma) = 1.$$

We need to define how the bounded memory player forms these beliefs.[6] Following Piccione and Rubinstein (1997), we assume that the beliefs correspond to "relative frequencies" as follows.

Let $f(h|\sigma)$ be the probability that a particular play of the game passes through the history h given the strategy profile σ. For each history h and memory state s_i, let the receiver's belief be given by the relative frequency as defined below.

Definition 2.1. ***(Consistency)***

A strategy profile σ is consistent with the beliefs μ if, for every memory state s_i and for every history $h \in s_i$, we have that the beliefs are computed as follows:

$$\mu(h|s_i, \sigma) = \frac{f(h|\sigma)}{\sum_{h' \in s_i} f(h'|\sigma)}. \tag{2.3}$$

Notice that denominator in expression (2.3) can be greater than one. The underlying reason for this is that the receiver only keeps track of the time (the period of the game) insofar as his transition rule allows. Thus, for example, depending on the transition rule, a t-period history and its parent $t-1$-period sub-history could place the receiver in the same memory state. This contrasts with what would be the receiver's information sets in

[6] Since the player is not forgetful within the period, but only across periods, we only have to define how he computes beliefs at the beginning of a stage game. At the end of the stage the player updates his beliefs using Bayes' rule.

the standard game without bounded memory. In the extreme case of one memory state, all histories must be in the same state and the denominator in (2.3) would be $\frac{1}{\eta}$, where recall that η is the exogenous stopping probability. Even in this case, however, the exogenous stopping probability ensures that beliefs are well defined; the bounded memory player will have well defined priors over the time periods.

Let H_B be the set of histories where the actual type is B. Similarly, H_S is the set of histories for which the actual type is S; hence, $H_B \cup H_S = H$. At the beginning of a stage game, given some memory state s_i, the *prior* belief that the opponent is a behavioral type is denoted by:

$$\rho_i \equiv \Pr\left(B|s_i, \sigma\right) = \sum_{h \in s_i \cap H_B} \mu\left(h|s_i, \sigma\right). \tag{2.4}$$

At the beginning of every stage game, we denote $\pi_i \equiv \Pr\left(T|s_i, \sigma\right)$ as the probability that the sender will tell the truth in that stage game, given the current memory state s_i. Since the sender is using a Markovian strategy, we can write the probability of truth as:

$$\pi_i = \rho_i + \left(1 - \rho_i\right) q_i. \tag{2.5}$$

After observing whether the signal was true or false, the receiver updates his belief concerning the probability that the sender is a behavioral type. We denote this *posterior* after a truth as $p_i^B \equiv \Pr\left(B|T, s_i\right)$. These beliefs are computed using (2.4) and (2.5).

$$p_i^B = \frac{\rho_i}{\pi_i}. \tag{2.6}$$

After a lie, the posterior on the sender being a behavioral type is zero, for a behavioral type always tells the truth.

In a game with full memory, the player's posterior in the end of a stage game is also his prior in the next stage game. This is not true in general for games with bounded memory players. In any stage game, the player does not necessarily know which was the previous stage game; or the belief he held in the last period. Upon reaching a memory state s_i, the receiver will hold a belief about his opponent given by (2.4), regardless of the actual history. Since all his knowledge about the history of the game is given by the statistic s_i, the belief he holds in s_i must depend only on this information.

Imperfect Recall and Incentive Compatibility[7] In our concept of optimality, we use the notion of incentive compatibility as described by Piccione and Rubinstein (1997)[8] and Wilson (2003). The assumption that we make is that at every information set the player holds beliefs induced by the strategy profile σ. If there is a deviation in the play of the game, the agent will not remember it, and his future beliefs will still be the ones induced by the strategy σ. Thus, a player might decide to deviate at a particular time, but he cannot trigger a sequence of deviations.

We say that a pair (μ, σ) is *incentive compatible* when it satisfies two conditions: one for the sender and another for the receiver.

First, the strategy of the strategic sender is a best response for him given the strategy of the bounded memory player (φ, a). Since the sender is unconstrained and conditions

[7] Absentmindedness as defined in Piccione and Rubinstein (1997) is a special case of imperfect recall. In this paper the bounded memory player is in fact absentminded. The issues of games with absentminded players discussed in this section applies more generally to games with imperfect recall as well.

[8] They refer to this condition as "modified multiself consistency".

his strategy on the entire history of the game, the incentive compatibility condition for the sender is the usual best response.

Second, the strategy of the bounded memory player is a best response for him at every point in time, taking as given the strategy for the sender and his own strategy at all memory states. The strategy (φ, a) is incentive compatible if at any information set s_i, there are no incentives to deviate given the beliefs at s_i and taking the strategy σ fixed. Again, the reason for taking his own strategy as given when deciding on which action to take or what state to move is that a deviation is not remembered in future periods and the beliefs in the following periods will given by the strategy $\sigma = (\varphi, a, q)$.

Given a strategy profile σ, every memory state will have an associated expected continuation payoff conditional on the actual type of the sender. The expected continuation payoff for the receiver at memory state i, given that the sender is a behavioral type, is denoted by v_i^B. This expected continuation payoff is the stage game payoff and the continuation payoff induced by the strategy profile. Formally, the expected continuation payoff v_i^B can be written as:

$$v_i^B = -\left(1-\pi_i\right)^2 + \left(1-\eta\right)\delta \sum_{j\in\mathcal{M}} \varphi_T\left(i,j\right) v_j^B. \tag{2.7}$$

The first term on the right of (2.7) is the payoff of the receiver in the stage game. This payoff is given by the equilibrium action a_i and given the strategy of the behavioral type of sender, which is to tell the truth with probability 1. The second term is the expected continuation payoff of the continuation game. This depends on the transition rule and on the associated continuation payoffs of all states reached with positive probability given the transition rule φ. The expected continuation payoff for the receiver given a strategic sender is denoted by v_i^S. Under the Markovian assumption, we can write this expected payoff as:

$$v_i^S = -q_i (1 - \pi_i)^2 - (1 - q_i) \pi_i^2 + (1 - \eta) \delta \left(q_i \sum_{j \in \mathcal{M}} \varphi_T (i, j) v_j^S + (1 - q_i) \sum_{j \in \mathcal{M}} \varphi_L (i, j) v_j^S \right). \tag{2.8}$$

When deciding on an action to take, and on which state to move, the bounded memory player makes his decisions based on the expected continuation payoffs associated with his decisions. Thus, in the context of this game, the incentive compatibility constraint can be written as two separate conditions: one condition for the transition rule and another one for the action rule.

The condition for incentive compatibility on the action rule of the receiver requires that he takes the myopic best action at all stage games. For suppose not: at some memory state i the specified action is different then the myopic best one. If the receiver deviates to the best current action he will not remember it in the following period. Since histories are private, the sender will only punish the receiver for this deviation if this punishment was profitable even in the case of no deviations. This implies that it must not be profitable, and thus, the receiver should deviate and play the myopic best one.

The incentive compatibility condition for the transition rule requires that the receiver moves to the memory state that gives him the highest expected payoff given his beliefs. Thus, if his transition rule assigns positive probability to move from state i to state j after a truth, for example, then given his beliefs at state i, it must be optimal for him to do so. We state the definition of incentive compatibility in the receiver's transition rule.

Definition 2.2. ***(Incentive Compatibility: Transition Rule)***

If a strategy $\sigma = (\varphi, a, q)$ *is* incentive compatible*, then the transition rule* φ *satisfies the following condition. For any states* $i, j,$ *and* $j' \in \mathcal{M}$*:*

$$\begin{aligned}\varphi_T(i,j) &> 0 \Rightarrow p_i^B v_j^B + \left(1-p_i^B\right) v_j^S \geq p_i^B v_{j'}^B + \left(1-p_i^B\right) v_{j'}^S, \\ \varphi_L(i,j) &> 0 \Rightarrow v_j^S \geq v_{j'}^S.\end{aligned}$$

We can interpret the bounded player as a collection of different selves; each self acting at a different point. Under this multi-self interpretation, we say that a strategy is incentive compatible if one self cannot gain by deviating from his equilibrium strategy, given the beliefs induced by this strategy and *assuming* that all other selves are playing the equilibrium strategy. The assumption in this definition is that the interim player can remember the equilibrium strategy, but cannot remember deviations during the game.

For a further discussion of imperfect recall, time consistency and incentive compatibility, see Aumann et al (1997), Gilboa (1997) and Piccione and Rubinstein (1997).

Equilibrium We define equilibrium using the notion of incentive compatibility. An equilibrium in this game is such that the strategies and beliefs are consistent, and the strategies are incentive compatible. The strategy of the sender is a best response for him given the strategy of the receiver and the strategy of the receiver is incentive compatible as in definition 2.2.

Definition 2.3. ***(Incentive Compatible Equilibrium)***

The strategy profile $\sigma = (\varphi, a, q)$ *is an* incentive compatible equilibrium *if there exists a belief* μ *such that the pair* (μ, σ) *is* ***consistent*** *and the strategy* σ *is* ***incentive compatible****.*

Sequential equilibrium is not the appropriate solution concept for games with absentmindedness, as was pointed out by Piccione and Rubinstein (1997, p18). The formal notion of sequential equilibrium requires the strategy of the player to be optimal at every information set, given the beliefs induced by this strategy. In games with absentmindedness the continuation strategy need not be optimal, since the player cannot revise his entire strategy during the play of the game (as described in section 2.2). In other words, the player might be "trapped" in bad equilibria.

In games with imperfect recall there are typically multiple equilibria (even in one person games). That the ex-ante decision maker will coordinate his actions in the most profitable equilibrium is an assumption of this model. We take the view that there are compelling reasons to assume that, ex-ante, the receiver can coordinate on the most profitable equilibrium, as was suggested by Aumann et al (1997). The memory rule will describe the agents' heuristics on updating beliefs, and in our view, Nature will play the role of coordinating on the first best for the bounded memory player. Thus, one way to think about this problem is as a mechanism design. The principal is the ex-ante player and the agents are the unbounded opponent and all the interim selves of the bounded memory player. The principal must choose the optimal mechanism given the set of equilibria between the interim agents and the unbounded player.

2.3 Two Memory States

In this section we restrict attention to the two-memory state case. This is a very special case, since the memory is minimal: one bit only. It will be very useful to our purposes since

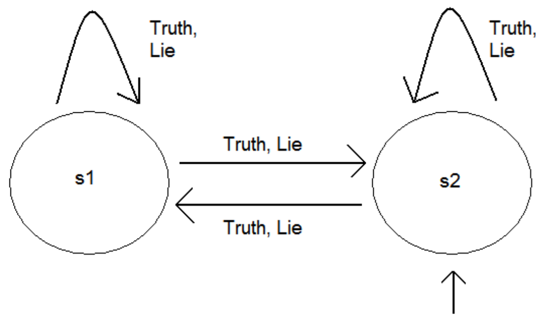

Figure 2.1: Updating Rule

the resulting equilibrium in this two-state world will show us the outcome on the extreme states of more general memories ($n > 2$).

An updating rule for the two-memory state case is a probability of switching from state 1 to state 2 and vice-versa, after receiving a truthful signal or a lie. A general updating rule is depicted in figure 1.

We can interpret this situation as a person that thinks only through two categories; he either thinks of his opponent as a "bad person" or as a "good person".

There are multiple equilibria in this two-state case when the prior on the behavioral type of sender is not very small.[9] Among these equilibria, the one that gives the receiver the highest ex-ante expected payoff is depicted below.

With the rule of figure 2, the receiver starts at some memory state, say memory state 2, and remains there as long as he keeps receiving truthful signals. After the first lie he moves to the other state, which is absorbing.

To construct this equilibrium, lets consider the case where the expected continuation

[9] For a small prior about the behavioral type of sender, babbling in both states is the only possible equilibrium. Babbling is characterized by a belief of $\frac{1}{2}$ in both states and with the strategic sender telling the truth with a probability that is just enough to make the receiver indifferent between believing the message or not.

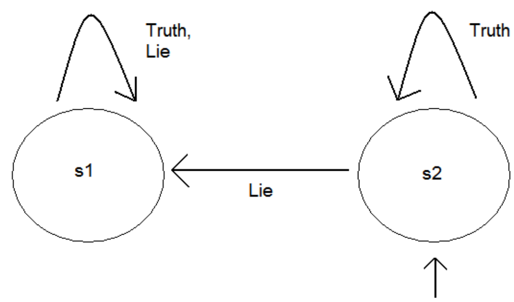

Figure 2.2: Rule that Separates after a Lie

payoff given a strategic type is higher in state 1, $v_1^S > v_2^S$. We know that a lie completely reveals the type of the sender. The receiver will then find optimal to move to state 1 whenever he observes a lie, regardless of his current state. Thus, the transition rule must assign probability one after a lie to state 1 $\varphi_L(i,1) = 1$, for $i = 1, 2$.

The strategic type of sender strictly prefers to lie in state 2. The intuition for this is that the trade-off between current payoff and reputation incentives does not exist in this highest state. The reputation concerns disappear, since this last state is the highest belief that the receiver can hold. The current payoff from telling the truth is worse than the babbling payoff (otherwise, telling the truth would be profitable even in the current period and this would be a contradiction in equilibrium). Thus, even though after the sender lies he is moved to the absorbing state 1, it is still strictly better for him to lie right away in state 2.

Therefore, in state 2, the sender tells the truth with probability zero, $q_2 = 0$. This implies that after a truth in state 2 the receiver's posterior is one $p_2^T = 1$. Because strategies must be incentive compatible, the receiver moves to the memory state with highest expected continuation payoff given a behavioral type. In this two-memory state case, if both states are reached in equilibrium, it must be that $v_2^B > v_1^B$. Thus, the receiver prefers to remain at memory state 2 after a truth: $\varphi_T(2,2) = 1$.

We analyze only the case where the exogenous probability of ending the game is very small $\eta \to 0$. Before we state the main result of this section, note that when the stopping probability η goes to zero, the expected length of the game increases.

In this example, the long-run probability of having a behavioral type in state 1 is zero, since, given a behavioral type, the receiver eventually reaches state 2 and stays there forever. Thus, by incentive compatibility, the receiver will assign $\varphi_T(1,1) = 1$. Similarly, given a behavioral type, the receiver eventually reaches state 2 and remains there until the end of the game, whereas the strategic type visits that state at most once. Thus, the belief in state 2 approaches one.

Therefore the equilibrium transition rule in the two-state case where both states are reached in equilibrium is given by figure 2. Since the transition rule completely separates the liars, whenever the receiver reaches memory state 1 he can be sure that he is dealing with a strategic type of sender. Thus the only possible belief that the truth is being told in that state is the one associated with babbling: $\pi_2 = \frac{1}{2}$.

To compute the belief in memory state 2 we have to compute the beliefs about the time periods. Our first result is that the receiver will hold "extreme" beliefs, that completely separate the types, even in this case of a minimal memory (2 states). This result will generalize for the case where the receiver is has more than two memory states: his two extreme states will have reputations zero and one.

Proposition 2.1. ***(Extreme Beliefs)***

For the two memory state game, the unique non trivial equilibrium is such that: $\lim_{\eta \to 0} \pi_1 = \frac{1}{2}$ *and* $\lim_{\eta \to 0} \pi_2 = 1$.

Proof. The sender will lie with probability 1 in state 2, thus $q_2 = 0$. This is true because the sender strictly prefers to lie in that state, $U_s(L|s_2) > U_s(T|s_2)$, regardless of the transition probability. Thus, if state 2 is the initial state, then it must be that:

$$\pi_2 = \Pr(t = 1|s_2)\,\rho + \Pr(t = 2|s_2) + \Pr(t = 3|s_2) \ldots$$

Where the probabilities of time periods are given by:

$$\Pr(t = 1|s_2) = \frac{1}{1 + \sum_{t=1}^{\infty} (1-\eta)^t \rho} = \frac{\eta}{\eta + (1-\eta)\,\rho}.$$

Thus: $\pi_2 = \frac{\rho}{\eta + \rho - \rho\eta}$, which leads us to:

$$\lim_{\eta \to 0} \pi_2 = \frac{\rho}{\eta + \rho(1-\eta)} = 1. \tag{2.9}$$

Moreover, since only behavioral types tell the truth in this state, the posterior on this type after a truth is one. By incentive compatibility, it must be that the receiver does not move to another memory state after a truth and, thus $\varphi_T(2,2) = 1$. If the transition in state 1 is positive, i.e. $\varphi_T(1,2) > 0$ then eventually the behavioral type gets "locked" in state 2 forever. This implies that in state 1 the belief about the behavioral type goes to zero. Thus, babbling is the unique outcome in this state and $\lim_{\eta \to 0} \pi_1 = \frac{1}{2}$. □

We conclude that the receiver, having a very small memory, will start the game with "long run beliefs". Another interesting property of the equilibrium is that the receiver keeps track of the liars. The strategic sender will gain not because the receiver will forget in case he lies, but because the receiver doesn't know the period that he is in when he starts the game. In other words, the receiver is confused about the time period when he is in state 2, so he doesn't know if he has already separated all the liars. This inflates the belief in state 2 and gives the sender a high payoff in the initial period.

2.4 n Memory States

Consider now the general case where the bounded memory player is restricted to n memory states, where $n > 2$.

Designing the best response for players with a bounded number of states has been shown to be an NP-complete problem, even for the simple case of a repeated prisoner's dilemma with complete information.[10] In our setting, for every state reached with positive probability by the equilibrium updating rule, the incentive compatibility constraints must be satisfied. Computing a best-response automaton and checking whether the incentive compatibility constraints are satisfied seems to be a computationally infeasible task.

Fortunately, though, we can show necessary conditions for equilibria. We can then characterize the equilibrium transition rule of the bounded memory player. We show that the equilibrium transition rule must satisfy a weak monotonicity condition, and hence the resulting updating rule resembles Bayesian updating whenever possible.

From what follows, we label the states in increasing order of continuation payoffs given

[10] See Papadimitriou (1992).

a behavioral type. Thus, if $i > j$ then $v_i^B > v_j^B$.

As has been pointed out in the literature,[11] there are typically multiple equilibria in games with imperfect recall. In this game, there are many equilibrium memory rules in which the receiver has redundant states. All the results in the appendix allow for these "bad equilibria". The most intuitive way to think of the memory rule, though, is to have in mind a rule without these redundant states. I.e., with n different memory states (holding different beliefs in equilibrium).

Throughout the chapter we consider only strategies in which all states are reached with positive probability in equilibrium.[12] Suppose, for now, that all states have different v_i^B and, consequently, different v_i^S (the cases of states with $v_i^B = v_j^B$ are considered in the appendix).

2.4.1 Equilibrium Updating Rule

Our main result is shown in the proposition below. We show that any equilibrium memory rule will satisfy a weakly increasing property. The equilibrium updating rule is such that the receiver separates the liars. Since only the strategic type can play this action, this signal is completely revealing. Thus, the receiver's posterior belief after a lie is zero. He will then move to his lowest state, and therefore $\varphi_L(i, 1) = 1$ for any memory state i. The same intuition holds for the case where the strategic sender strictly prefers to lie. In this case, a truth is completely revealing, since it is played only by a behavioral type in equilibrium. The receiver then moves to his highest state with probability one.

---

[11] See Piccione and Rubinstein (1997) and Aumann et al. (1997).

[12] States not reached in equilibrium do not play any role, not even as a threat (since, as we will show, there will be an absorbing babbling state). Thus, we can ignore these states without loss of generality.

While the receiver might ignore true signals, by not updating after receiving them, he will never update to a worse belief after a truth. The receiver will get a better payoff from staying in the same state rather than moving to a lower state. One interpretation of this result is that the receiver might not pay attention (update) to some signals, but he will never forget the information that he already holds.

Finally, the extreme states must have beliefs about the opponent's type that are zero and one. The intuition is that at state s_n there are no reputation incentives, thus the bad type of sender will lie right away. If the receiver is at this memory state, the only chance that the sender is the strategic type is if this is the first stage game being played at this memory state. In other words, the strategic sender will stay in this state for at most one period. On the other hand, if the sender is an honest type, the state is absorbing and this type will be in state s_n forever. The probability of being at state s_n for the first time goes to zero as the stopping probability gets smaller. The same argument holds for what happens at state s_1. If this is not the initial state, then only the strategic type of sender can reach this state. In this case, the result is obvious. If this is the initial state, the probability of having a strategic sender at that state goes to one as the death rate goes to zero. Note that since this state is absorbing, in equilibrium it will not be the initial state.

We state the result below for the case where the stopping probability is very small, $\eta \to 0$. In the appendix we show a more general version of the proposition, which holds for any stopping probability η, and which allows for redundant states.

Proposition 2.2. ***(Increasing Property)***

If the strategy profile $\sigma = (\varphi, a, q)$ is an equilibrium, then:

1. *After observing a lie move to an absorbing "babbling" state:* $\varphi_L(j,1) = 1$.
2. *Never go back after observing a true signal:* $\pi_j > \pi_i \Rightarrow \varphi_T(j,i) = 0$.
3. *Initial state is the lowest one after the "babbling state"* $\varphi_0(2) = 1$.
4. *The lowest belief approaches zero:* $\lim_{\eta \to 0} \rho_1 = 0$.
5. *The highest belief approaches one:* $\lim_{\eta \to 0} \rho_n = 1$.

At this point, we have ruled out some memory rules that could never be played in equilibrium–in particular, rules with loops and rules that don't separate the liars.

Although we have shown that the equilibrium updating rule must satisfy a weakly increasing property, we still want to understand how the updating happens after true signals. The proposition below tells us part of the story. All the results depend on a condition that the posteriors about the sender's type are different on the states. To weaken this restriction, in the appendix we prove the following lemma: $\pi_j > \pi_i \Rightarrow p_j^B \geq p_i^B$.[13]

Proposition 2.3. ***(Weak Monotonicity)***

Consider only memory rules with states with different posteriors, i.e., states where $p_i^B \neq p_j^B$. *Then, for any two states* $i, j \in \mathcal{M}$, *we have that:*

1. *(Single crossing)* $\varphi_T(i,k) > 0$, $\varphi_T(i,m) > 0$ *and* $\varphi_T(j,k) > 0 \Rightarrow \varphi_T(j,m) = 0$,

[13] It can be easily shown that for $n \leq 4$, i.e. for memories with less than or equal to four states, we must have that $\pi_j > \pi_i \Rightarrow p_j^H > p_i^H$. And, thus, the properties in proposition 2.3 hold without any additional restrictions.

for $\forall k, m$ *such that* $\pi_k \neq \pi_m$.

2. *(No Jumps)* $\varphi_T(i, k') > 0$,$\varphi_T(i, k'') > 0 \ \Rightarrow \varphi_T(j, k) = 0$,

   for $\forall k' < k < k''$.

3. *(Monotonicity)* *If* $\varphi_T(i, m) > 0 \Rightarrow \varphi_T(j, m') = 0$,

   for $\forall\ m' < m$.

Proof. We first prove the single crossing property. Suppose that $\varphi_T(i, k) > 0$ and also that $\varphi_T(i, m) > 0$. This implies that:

$$p_i^B\left(v_k^B - v_m^B\right) + p_i^S\left(v_k^S - v_m^S\right) = 0. \tag{2.10}$$

Suppose now that $\varphi_T(j, k) > 0$ and $\varphi_T(j, m) > 0$, then

$$p_j^B\left(v_k^B - v_m^B\right) + p_j^S\left(v_k^S - v_m^S\right) = 0. \tag{2.11}$$

If $p_i^B \neq p_j^B$ then (2.10) and (2.11) cannot hold at the same time. Thus, two states must have at most one state in common in their transition rules.

The next step is to show a "no jump" result for states where p_i^B and p_j^B are different. Suppose that $\varphi_T(i, k+1) > 0$ and $\varphi_T(i, k-1) > 0$. This implies that:

$$p_i^B\left(v_{k+1}^B - v_k^B\right) + p_i^S\left(v_{k+1}^S - v_k^S\right) \ \geq \ 0, \tag{2.12}$$

$$p_i^B\left(v_k^B - v_{k-1}^B\right) + p_i^S\left(v_k^S - v_{k-1}^S\right) \ \leq \ 0. \tag{2.13}$$

If in addition we also have that $\varphi_T(j,k) > 0$. Then it must be true that :

$$p_j^B\left(v_{k+1}^B - v_k^B\right) + p_j^S\left(v_{k+1}^S - v_k^S\right) \leq 0, \tag{2.14}$$

$$p_j^B\left(v_k^B - v_{k-1}^B\right) + p_j^S\left(v_k^S - v_{k-1}^S\right) \geq 0. \tag{2.15}$$

The equations above cannot hold for $\pi_{k+1} > \pi_k > \pi_{k-1}$ and $p_i^B \neq p_j^B$.

Finally, to prove the monotonicity condition, first note that by incentive compatibility we must have that:

$$\varphi_T(j,m) > 0 \Rightarrow p_j^B v_m^B + p_j^S v_m^S \geq p_j^B v_{m\prime}^B + p_j^S v_{m\prime}^S,$$

which means that:

$$p_j^B\left(v_m^B - v_{m\prime}^B\right) + p_j^S\left(v_m^S - v_{m\prime}^S\right) \geq 0. \tag{2.16}$$

Note that $\left(v_m^B - v_{m\prime}^B\right) \geq 0$ and $\left(v_m^S - v_{m\prime}^S\right) \leq 0$.Thus, since $p_i^B > p_j^B$ (and consequently $\left(p_j^S > p_i^S\right)$), we have that:

$$p_i^B\left(v_m^B - v_{m\prime}^B\right) + p_i^S\left(v_m^S - v_{m\prime}^S\right) > 0, \tag{2.17}$$

which proves our last condition. □

This monotonicity result tells us that for any two states with different posteriors, the transition rule of both states might have at most one state in common, and this is the highest point on the support of the transition rule of the lower posterior state. Moreover, the lower posterior state does not move to any state in the higher posterior state's support, except for this first point.

As we argued before, there are compelling reasons to focus only on the equilibria that give the receiver the highest payoff. Lemma 2.1 below shows that we can ignore the redundant states without loss of generality. This result tells us that any equilibrium in which the receiver is using a redundant state can be reproduced with a memory without redundant states. Therefore, when searching for the equilibrium that gives the receiver the highest expected payoff, we can focus only on rules where all states have different beliefs.

Lemma 2.1. ***(Redundant States)***

Consider a receiver with memory $\mathcal{M}$ that has only n states. The strategy $\sigma = (\varphi, a, q)$ gives the receiver a payoff of U_R^*. Now suppose that $\pi_i = \pi_j$, for some $i, j \in \mathcal{M}$. Then, there $\exists$ $(\varphi, a, q)'$ for memory some other memory $\mathcal{M}'$ with $n-1$ states and that gives the receiver utility the same payoff U_R^*.

Proof. Let $\pi_i = \pi_j$. From proposition 2.2 this implies that $v_i^S = v_j^S$. Thus, if both states are reached in equilibrium it must be that $v_i^B = v_j^B$. The receiver is always completely indifferent between the two states i and j after a truth or lie. If $p_i^B = p_j^B$, then the states are identical and we can consider them as being a single state (just rewrite the transition rules). If $p_i^B > p_j^B$, then they must have the same transition rules, or else $v_i^B = v_j^B$ would not hold. But, if they have the same transition rules then again they are identical and we can group them as one. □

A class of memory rules that satisfies propositions 2.2 and 2.3 is depicted in figure 3 below.

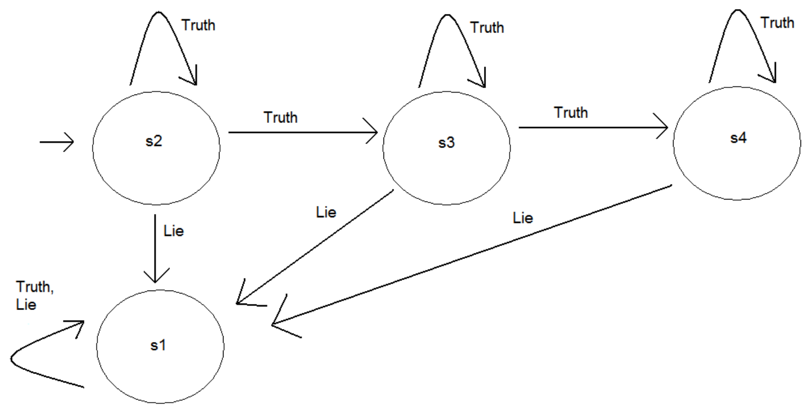

Figure 2.3: A Class of Equilibrium Memory Rules

The results suggest that this is the class of memory rules that the receiver will use. I.e., a strategy in which the transition rule has only positive probability in staying in the same state or moving to the next one.

In the following section, we will show the conditions under which the updating rule is deterministic, meaning that $\varphi_T(i, i+1) = 1$ for all states $i < n$. For $n = 3$, or 4 we can show that the memory rule must be the one shown in figure 3, allowing for the possibility that $\varphi_T(i, i) = 0$. It is possible, though, that for $n > 4$ the equilibrium transition is stochastic, but not exactly like the one depicted above. This case would suggest that the receiver is wasting resources by not fully using his memory states. Such a rule might exist in equilibrium, as long as the memory rule satisfies the conditions in propositions 2.2 and 2.3.

2.4.2 Deterministic Updating Rule

We have characterized the equilibrium transition rule. In this section, we show under what conditions this transition will be deterministic. We say that the receiver's memory is not binding when he updates his beliefs using Bayes' rule, with no bias whatsoever. There are cases, though, in which the transition rule is deterministic, but the updating differs from Bayesian in the last state. The memory of the receiver will confuse him in this extreme state and there will be biases in information processing. In this section we show the conditions on the parameters under which the receiver will use deterministic transition rules (the algorithm uses the same reasoning whether one wants to compute the threshold for Bayesian updating or for deterministic rules only).

We present the result in two propositions. The first one shows that, given a memory of size n, there is a threshold in the prior space such that if the prior is smaller than this threshold, the receiver will not use deterministic transition rules. We then prove another result showing that this is in fact also sufficient for equilibrium with deterministic transition rules. This sufficient condition is in fact a strong result by itself; thus, if the sender is using a best response and the transition rules are not random, the receiver will find it in his best interest to follow the specified transition rules. Given this result, one can relate it to Bayesian updating: if we describe Bayesian updating as an updating rule with an infinite number of memory states and deterministic transition rules, the player will find it in his best interest to keep playing this strategy, i.e., it will be incentive compatible as well. Thus, in this context, Bayesian updating is consistent with a large enough number of memory states.

When the condition of the threshold described below is not met, there are no equilibria with deterministic transition rules (besides the trivial one, where all states have the same expected continuation payoff). Thus, randomization is needed.

Proposition 2.4. ***(Deterministic Transition Rule: Necessary Condition)*** *Given any number of memory states $n > 2$, there exists a threshold on the prior about the behavioral type ρ_n^* such that if the actual prior is smaller than this threshold $\rho < \rho_n^*$ then there is no equilibrium with deterministic transition rules.*

The proof of the proposition above is by induction (shown formally in the appendix). The first step is to note that the last state will have belief 1, following the intuition of the two state case. The receiver will use pure strategy only if the belief in state s_{n-1} is at least as high as some threshold π_{n-1}^*, which depends on the parameters δ, n and η. If the belief is lower than this threshold, the sender will prefer to tell the truth and be updated with probability one to the highest state. Moreover, by incentive compatibility there is a lower bound on the posterior state s_{n-1}. That is, if the posterior on the sender's type is lower than this lower bound, the receiver will find it in his best interest to remain in that state after a true signal. Together, this implies that at every stage game there is a lower bound on the prior on the sender's type at that stage game. However, the prior on state s_{n-1} is the posterior of state s_{n-2}. Using the same reasoning backwards we find that there must be a lower bound on the prior for the receiver to play pure strategy. In the appendix we show how to compute this lower bound given the parameters δ, n and η.

The next proposition shows a sufficient condition for deterministic transition rules.

Proposition 2.5. ***(Deterministic Transition Rule: Incentive Compatibility)*** *Let the transition rules be deterministic: $\varphi_T(i, i+1) = 1$, and the strategy for the sender be a best response for him. Then it will be incentive compatible for the receiver to move only* *to the next state after a true signal:*

$$p_{i-1}^B v_i^B + \left(1 - p_{i-1}^B\right) v_i^S \geq p_{i-1}^B v_s^B + \left(1 - p_{i-1}^B\right) v_s^S, \qquad \forall s > 0.$$

Therefore, given a memory of size n, as long as the prior ρ is higher than the threshold ρ_n^*, which is shown in the appendix, the receiver will be able to reproduce Bayesian updating and there will be no information loss.

The following result shows that there is at most one equilibrium in which the receiver is using a pure strategy.

Proposition 2.6. ***(Deterministic Transition Rule: Uniqueness)***
Fix the number of memory states n *and the initial prior* ρ. *There is at most one equilibrium with deterministic transition rule without redundant states.*

Note that when memory states are unobservable by the strategic sender, the deterministic equilibria would still hold. In equilibrium the sender would know the current memory state. In the case where there are no equilibria with only deterministic transition rules, the structure of the transition rule would also be the same, since the results in proposition 2.3 would still hold.

2.5 Example: Three Memory States

In this section we show the equilibria for the case involving three memory states. The main goal of this section is to exemplify the mechanics of the model and to show how to compute the equilibria in a bounded memory game.

We use the results of proposition 2.2. The lowest state is equivalent to a babbling state,

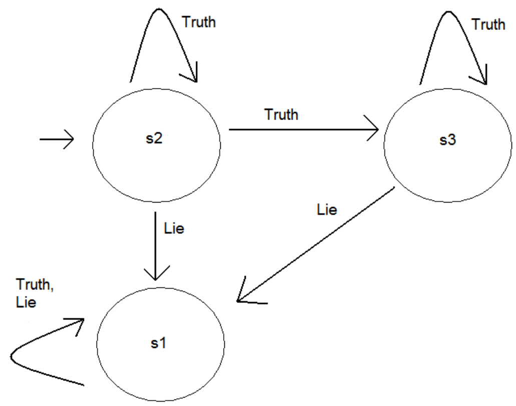

Figure 2.4: Three Memory States

where the probability of truth is $\pi_1 = \frac{1}{2}$; moreover, this lowest state is absorbing. Also, the belief in the highest state is one $\pi_3 = 1$. Finally, the receiver will start at the intermediate state $\varphi_0(2) = 1$. It remains for us to calculate the belief in state 2 π_2, the transition probability from state 2 to state 3 $\varphi_T(2,3)$, as well as the strategy of the sender. We focus on Markovian equilibria only, i.e., equilibria in which the strategy of the sender depends only on the current memory state. Figure 4 below depicts the equilibrium transition rule.

We know from the previous section that there is a threshold on the prior of the behavioral type such that the equilibrium involves only a deterministic transition rule. Thus, if this prior is higher than the threshold, $\rho > \bar{\rho}_3$, then the equilibrium with three non-redundant states involves deterministic transition from state 2 to state 3. This means that after a truth, the receiver updates to state 3 with probability one.

In this section we characterize the equilibria when $\rho < \bar{\rho}_3$. We already know that the equilibria must be such that $\varphi_T(2,2) > 0$. In equilibrium, the sender must be mixing between telling the truth and lying in state 2, or else the receiver would not mix himself, by incentive compatibility. The indifference condition of the sender is that lying in state 2 gives the same expected continuation payoff for him as telling the truth in this same sate, which means that:

$$\pi_2^2 + \delta\frac{1}{4} = (1-\pi_2)^2 + \delta\left(\varphi_T(2,2)\,\pi_2^2 + (1-\varphi_T(2,2))\right). \tag{2.18}$$

This gives us the following quadratic equation:

$$\delta\varphi_T(2,2)\,\pi_2^2 - 2\pi_2 + 1 + \delta(1-\varphi_T(2,2)) - \delta\frac{1}{4} = 0. \tag{2.19}$$

Solving for the belief π_2 gives us:

$$\pi_2 = \frac{1-\sqrt{1-\delta\varphi_T(2,2)\left(1+\delta(1-\varphi_T(2,2))-\delta\frac{1}{4}\right)}}{\delta\varphi_T(2,2)}. \tag{2.20}$$

We interpret the transition probability $\varphi_T(2,2)$ as a testing parameter, since it is capturing the probability that the sender will not be upgraded, even though the signal was truthful. We show that in the three state case, there is a trade-off between action rule and transition rule, or between actions and testing.

Lemma 2.2. ***(Testing)***

In equilibrium, the lower the receiver's belief about the truth, the more he will test the sender before updating.

Proof. Differentiating the indifference condition of the sender, which is given by (2.19), gives us:

$$\delta\varphi_T(2,2)\,2\pi_1 d\pi_1 - 2d\pi_1 + \delta\pi_1^2 d\varphi_T(2,2) - \delta d\varphi_T(2,2) = 0,$$

and finally,

$$\frac{d\varphi_T(2,2)}{d\pi_1} = \frac{2}{\delta}\frac{\delta\pi_1\varphi_T(2,2) - 1}{\left(1-\pi_1^2\right)} < 0.$$

□

Condition (2.20) is necessary for equilibrium in this three state case. Another necessary condition is that the receiver must be indifferent between updating to state 3 or staying in state 2 after a truth. For this indifference condition we have that:

$$p_2\left(v_3^H - v_2^H\right) + (1-p_2)\left(v_3^S - v_2^S\right) = 0.$$

Substituting the continuation values gives us:

$$p_2\frac{(1-\pi_2)^2}{1-\delta\varphi_T(2,2)} + (1-p_2)\left(\pi_2^2 - 1\right) = 0.$$

Solving for the posterior p_2 and knowing that $\pi_2 < 1$ this implies:

$$p_2 = \frac{(\pi_2+1)(1-\delta\varphi_T(2,2))}{1-\pi_2+(\pi_2+1)(1-\delta\varphi_T(2,2))}. \tag{2.21}$$

Now the two conditions missing are that these beliefs π_2 and p_2^T must be consistent in equilibrium, according to 2.3. The posterior p_2^T can be written as:

$$p_2^T = \gamma_1\frac{\rho}{\rho+(1-\rho)\,q_1} + \gamma_2\frac{\rho_2}{\rho_2+(1-\rho_2)\,q_2} + \gamma_3\frac{\rho_3}{\rho_3+(1-\rho_3)\,q_3} + ... \tag{2.22}$$

where the $\gamma' s$ indicate the Bayesian updating of time periods. First note that: $\rho_2 + (1 - \rho_2)\, q_2 = \frac{\rho + (1-\rho) q_1 q_2}{\rho + (1-\rho) q_1}$.

In general, we will have:

$$\rho_t + (1 - \rho_t)\, q_t = \frac{\rho + (1 - \rho)\, q_1 q_2 ... q_t}{\rho + (1 - \rho)\, q_1 ... q_{t-1}}. \tag{2.23}$$

Let $\Delta = f_1 + f_2 + f_3 + ...$ Where f_i is the frequency of period i. Then, $\gamma_i = \frac{f_i}{\Delta}$, and, in general we must have that:

$$\gamma_t = \frac{(1 - \eta)^{t-1} (\varphi_T (2,2))^{t-1} (\rho + (1 - \rho)\, q_1 q_2 \times ... \times q_t)}{\Delta}. \tag{2.24}$$

Before we compute what (2.22) should be, let's calculate each term of the equation. But first, also note that:

$$\rho_t = \frac{\rho}{\rho + (1 - \rho)\, q_1 q_2 \times ... \times q_{t-1}}.$$

The individual beliefs of the time periods can be written as:

$$\gamma_t \frac{\rho_t}{\rho_t + (1 - \rho_t)\, q_t} = \frac{(1 - \eta)^{t-1} (\varphi_T (2,2))^{t-1}}{\Delta} \rho.$$

Thus, (2.22) can be simplified:

$$p_2^T = \frac{\rho}{\Delta} + \frac{(1 - \eta)(\varphi_T (2,2))\, \rho}{\Delta} + \frac{(1 - \eta)^2 (\varphi_T (2,2))^2 \rho}{\Delta} + ...,$$

which in turn can be written as:

$$p_2^T = \frac{\rho}{\Delta} \frac{1}{(1-(1-\eta)\,\sigma)}. \tag{2.25}$$

And, for the Markovian case, the term Δ can be calculated as:

$$\Delta = (\rho + (1-\rho)\,q) + (1-\eta)\,(\varphi_T\,(2,2))\,\left(\rho + (1-\rho)\,q^2\right) + (1-\eta)^2\,(\varphi_T\,(2,2))^2\,\left(\rho + (1-\rho)\,q^3\right) + ...$$

This term can be simplified further to obtain the following expression:

$$\Delta = \frac{\rho + (1-\rho)\,q - (1-\eta)\,(\varphi_T\,(2,2))\,q}{(1-(1-\eta)\,(\varphi_T\,(2,2)))\,(1-(1-\eta)\,(\varphi_T\,(2,2))\,q)}. \tag{2.26}$$

Thus, substituting (2.26) in (2.25) gives us the following expression for the receiver's posterior:

$$p_2^T = \frac{\rho\,(1-(1-\eta)\,(\varphi_T\,(2,2))\,q)}{\rho + (1-\rho)\,q - (1-\eta)\,(\varphi_T\,(2,2))\,q}. \tag{2.27}$$

Similarly, for the belief π_2 we have that:

$$\pi_2 = \begin{bmatrix} (1-(1-\eta)\,(\varphi_T\,(2,2)))\,(1-(1-\eta)\,(\varphi_T\,(2,2))\,q) + \\ +(1-\eta)\,(\varphi_T\,(2,2))\,(\rho + (1-\rho)\,q - (1-\eta)\,(\varphi_T\,(2,2))\,q) \end{bmatrix}^{-1} \\ [\rho\,(1-(1-\eta)\,(\varphi_T\,(2,2))\,q) + (1-\rho)\,q\,(1-(1-\eta)\,(\varphi_T\,(2,2)))] \tag{2.28}$$

If the beliefs and strategies $\left(p_2^H, \pi_2, \varphi, a, q\right)$ satisfy the system of equations described by (2.20), (2.21), (2.27) and (2.28) above, then we have an equilibrium in the 3 memory state case.

2.6 Standard Automata

The automata models are in many ways similar to a bounded memory player. An automaton, like a bounded memory player, is a finite set of states with a transition rule and an action rule. When we model the memory of the player as an automaton, we ignore incentive compatibility constraints and the memory is designed to be the ex-ante optimal one. As it turns out, however, in single player games with no discounting, this distinction is nonexistent: Piccione and Rubinstein (1997) show that the ex-ante optimal strategy will also be incentive compatible.

In a game, there are two reasons that an equilibrium with automata could differ from one with a bounded memory player. The first one is the same as in a single player game with discounting. Think of a very impatient decision maker. Ex-ante, this player will design a strategy to achieve a higher payoff in the initial periods. As the game starts, the player might think that he is not in the initial periods any more and will take in consideration the payoffs of future periods. This distorts the incentives between the initial period and the period where the game has already started. An automaton would allow an individual to commit to actions and avoid the 'temptations' to deviate that his future selves would confront.

The second reason is the ability to commit against an opponent. Thus, modeling the player's memory as an automaton would require a further assumption–namely, that the player can credibly commit to his strategy.[14]

We take the view that both approaches have their own interest, but this chapter focuses

[14]Since in this paper the strategic sender is playing a zero-sum game with the receiver, it is not clear whether commitment would increase the receiver's payoff absent discounting effects.

	Bounded	Memory		Automata		
ρ	π_2	$\varphi_T(2,2)$	U_R	π_2^A	$\varphi_T(2,2)$	U_R^A
0.1	0.5	1	-1.25	0.5313	0.9358	-1.2415
0.2	0.6220	0.7256	-1.1777	0.5919	0.8007	-1.1730
0.3	0.6593	0.6219	-1.0736	0.6385	0.6814	-1.0716
0.4	0.6931	0.5142	-0.9523	0.6791	0.5607	-0.9514
0.5	0.7255	0.3931	-0.8181	0.7168	0.4274	-0.8178
0.6	0.7581	0.2459	-0.6736	0.7540	0.2663	-0.6735
0.7	0.7926	0.0492	-0.5198	0.7920	0.0532	-0.5198
0.8	0.8	0	-0.36	0.8	0	-0.36
0.9	0.9	0	-0.19	0.9	0	-0.19

Table 2.1: Automata: more testing than Bounded Memory

only on the case where incentive compatibility is indeed an issue. We show that in some situations the automaton can do better than the bounded memory player, while in others the automaton does just as well (obviously, automata can never do worse, since the set of incentive compatible memory rules is a subset of the memory rules described by an automaton). In fact, we show some results for the three state case, where the automaton does better than the bounded memory player.

One thing to note in table 1 below is that the lower the prior on the sender's type, the higher $\varphi_T(2,2)$ which means that the receiver will test the sender more. These are the equilibria for which the receiver is mixing on his updating rule. If ρ is very high (in this case the threshold is 0.72), there will be no randomization. All these equilibria were computed for $\eta = 10^{-60}$ and $\delta = 0.8$.

The comparison between the automaton and the bounded memory player is shown in the table below.[15]

[15] We present examples of automata that do better, but do not explicitly solve for the optimal automata. This is an interesting open question that we leave to future work.

The results in table 1 show that the three state automaton does better than a bounded memory receiver with the same number of states. Most importantly, it does so through more testing. The transition from state 2 to state 3 is higher with the bounded memory player than it is with an automaton. In fact, if the bounded memory player used the same transition as the automaton, after a truth the bounded memory player would find it interim optimal to move to state 3 and not to randomize. The incentives to move to state 3 would break down the equilibrium.

To summarize, an automaton will perform better than a bounded memory player by committing to test more.

2.7 Conclusion

This chapter is a study of bounded memory in a reputation game. It differs from the existing literature on imperfect memory by considering a game in which the memory rule is chosen by the player and satisfies incentive compatibility constraints. Equilibrium with bounded memory and incentive compatibility constraints was already studied in single player games, but this is the first time it has been done in a multi-player game.

Our view is that, although forgetful, players have some ability to determine what to remember and what to forget. A player might think that a fact is particularly important and, knowing that he will likely forget it, he will rehearse the fact and increase his chances of remembering it. Most models of bounded memory assume that, during the play of a game, people have no control whatsoever over what to remember or what to forget.[16]

[16]In models of optimal finite memory, such as the automata models or Dow's (1991) search model, the player decides on the memory rule before the game starts. Once in the game, he has no control over his

We showed that in this game the updating rule is rather simple: monotonic and weakly increasing. In particular, given the memory size n, if the prior on the behavioral type is high enough, the bounded memory player will use deterministic transition rules. In fact, he might do just as well as if he used Bayes' rule. Or, if the prior is higher than a particular threshold, but not "high enough," he will suffer loss (as compared to a Bayesian player) in the extreme state, when he gets confused about the time periods.

The second contribution of this chapter is to show the updating rule when memory constraints are severe. In these cases the receiver will use random transition rules in the initial states. Despite the multiplicity of equilibria that games with bounded memory have, there are necessary conditions on the updating rule for all equilibria. These conditions suggest a particular updating rule (stay put or go forward), when the receiver can coordinate on the equilibrium that gives him the highest payoff. This randomization in the transition rule is used for two different reasons. First, it is used to overcome the memory problem by not storing all the signals. This intuition was also present in single player games. Most importantly, however, in a two player game, randomization will be used as a strategic element: to test the opponents before updating.

In a broader sense, this chapter is part of an emerging literature on restricted capacity to deal with information. Players fail to use Bayes rule due to some constraint on their technology. This departure from Bayes' rule could result from a cost on updating new information (Reis (2007)), a restriction on acquiring new information (Sims (2003)), a cost to thinking through the implications of a particular action (Bolton and Faure-Grimald

---

memory.

(2005)), or memory constraints. In a repeated interaction, this ability to sort information is very important because of the substantial amount of data that some equilibria require, combined with possible cognitive restrictions of the agents.

The results that we see in the recent papers suggest that these constraints lead to inertia and inattention. Due to a restricted capacity in dealing with information, players cannot execute Bayes rule and will choose the information to memorize, and to acquire. In other words, they will sort the information received and ignore part of it. This chapter confirms this intuition in the context of a two player game, showing that the agents will ignore information and update only sporadically when their memory is constrained.

In the model presented, the strategic sender and the receiver had opposite preferences. The zero-sum nature of this relationship did not leave any room for cooperation when the bad type of sender was caught. Still unclear are the implications of bounded memory in sustaining cooperation in repeated interactions. The study of the role of bounded memory and reputation in a more general environment, without this zero-sum nature, is an open road of research.

Finally, in this chapter we modeled human memory as a finite set of states with sequential rationality constraints. One is tempted to apply what was learned here to other situations involving limited storage capacity, for example, to apply this model to the context of an organization that keeps track of signals about their clients. The imperfect communication between workers within a firm suggests this analogy.

2.8 Appendix

2.8.1 n Memory States

This section is divided as follows. First, we show a general version for proposition 2.2 in the text. This theorem is true regardless if the transition rule is deterministic (in which case it is trivially true) or not. Then we show in which cases the receiver will use deterministic transition rules.

We need extra notation for this section. In general, we denote the sender's expected continuation payoff in some state s_i as $U_S(s_i)$. His expected continuation payoff from telling the truth in that state is $U_S(T|s_i)$ and from lying it is $U_S(L|s_i)$. This utility is given by a current payoff of telling the truth (or lying) and an expected continuation payoff that depends on the transition rule φ as well as on $U_S(s_j)$ for all $j \in \mathcal{M}$.

2.8.2 Random Transition Rules

Define l as the state with highest expected continuation payoff if the receiver is facing a strategic sender. Formally: $\mathcal{D} \equiv \{l \in \mathcal{M} | v_l^S \geq v_i^S, \forall i \in \mathcal{M}\}$, similarly define: $\mathcal{U} \equiv \{u \in \mathcal{M} | v_u^B \geq v_i^B, \forall i \in \mathcal{M}\}$.

Proposition 2.7. ***(Increasing Updating Rule: General version of Proposition 2.2)***

If the strategy profile $\sigma = (\varphi, a, q)$ is an equilibrium, then:

1. *After Lie: $\varphi_L(j, l') = 0$ where $l' \notin \{l|\pi_l = \min_i \pi_i\}$.*

2. *If $U_S(L|i) > U_S(T|i) \Rightarrow \varphi_T(i, h\prime) = 0$ where $h' \notin \{h|\pi_h = \max_i \pi_i\}$.*

3. *After True:* $\pi_j > \pi_i \Rightarrow \varphi_T(j,i) = 0$ *(don't go back after a True signal).*

4. $\varphi_0(i) = 0, \forall \pi_i > \pi_{(2)}$.

5. $\lim_{\eta \to 0} \rho_l = 0,\ \forall l \in \mathcal{D}$.

6. $\lim_{\eta \to 0} \rho_u = 1,\ \forall u \in \mathcal{U}$.

We show the proof of this proposition through several different lemmas.

Our first result comes from incentive compatibility. If $\Pr(B|i,L) = 0,\ \forall i$,we must have that after a lie, the receiver moves to a state with the highest expected continuation payoff given that the sender is strategic. As defined above, the receiver moves to a state where the expected continuation payoff for the receiver conditional on the bad type of sender is equal to v_l^S (and for the sender is $U_S(l)$).

Before we state the first lemma, denote

$$j^* \in \mathcal{M}(j) \equiv \left\{ \begin{array}{c} j \in \mathcal{M} |\text{ after a true } p_j^B v_{j^*}^B + p_j^S v_{j^*}^S \geq p_j^B v_{j\prime}^B + p_j^S v_{j\prime}^S; \\ \text{after a lie: } v_{j^*}^S \geq v_{j\prime}^S, \forall j' \in \mathcal{M} \end{array} \right\}.$$

Thus, the payoff of the sender after lying is:

$$U_S(L|i) = \pi_i^2 + (1-\eta)\,\delta \sum_{i^*} \varphi_L(i,i^*)\, U_S(l).$$

Similarly, the payoff of the sender after telling the truth is:

$$U_S(T|i) = (1-\pi_i)^2 + (1-\eta)\,\delta \sum_{j^*} \varphi_T(i,j^*)\, U_S(j^*)$$

Lemma 2.3. $j \notin \mathcal{D} \Rightarrow \varphi_L(i,j) = 0, \forall i \in \mathcal{M}$.

Proof. By incentive compatibility, $\varphi_L(i,j) > 0 \Rightarrow v_j^S \geq v_{j'}^S, \forall j' \in \mathcal{M}$. Therefore we can write the payoff of the sender after lying as:

$$U_S(L|i) = \pi_i^2 + (1-\eta)\,\delta U_S(l)\,.$$

□

We now show a lemma that will be very helpful in subsequent results. The lemma is that whenever the sender reaches a state where $\pi_i = 1$, i.e., the highest possible belief, then the sender will strictly prefer to lie. This is because by lying the sender gets the highest possible current payoff and is then placed on the lowest state l. However, lying or telling the truth in l is strictly better for the sender than telling the truth in a state with belief higher than $\frac{1}{2}$.

Lemma 2.4. *In the highest state the strategic sender lies with probability one (except for the trivial equilibrium where all the states are the same):*

$$U_S(L|n) > U_S(T|n)\,.$$

Proof. We can write the utility of the strategic sender as:

$$\begin{aligned} U_S(L|i) &= \pi_n^2 + (1-\eta)\,\delta U_S(l)\,, \\ U_S(T|i) &= (1-\pi_n)^2 + (1-\eta)\,\delta \sum_{j\in\mathcal{M}} \varphi_T(i,j)\,U_S(j)\,. \end{aligned}$$

We can write the expected continuation payoff of the sender as:

$$\begin{aligned} U_S(j) &= (1-\pi_j)^2 + (1-\eta)\,\delta \sum_{s\in\mathcal{M}} \varphi_T(j,s)\,(1-\pi_s)^2 + \ldots \\ &\quad + (1-\eta)^t\,\delta^t \pi_k^2 + (1-\eta)^{t+1}\,\delta^{T+1} U_S(l)\,. \end{aligned} \tag{2.29}$$

Note also that telling the truth in any state gives the strategic sender a lower current payoff than the babbling payoff, and lying at state n gives the strategic sender the highest current payoff among all other states. Also, for $\forall j$ it must be true that $(1-\pi_j)^2 \leq \pi_l^2$, and also that $\pi_j^2 \leq \pi_n^2$. In state s_n we can write the utility for the sender as:

$$U_S(L|n) = \pi_n^2 + (1-\eta)^t\,\delta^t \pi_l^2 + (1-\eta)\,\delta\pi_l^2 + \ldots + (1-\eta)^{t+1}\,\delta^{t+1} U_S(l)\,, \tag{2.30}$$

and since we have that:

$$\begin{aligned} (1-\pi_j)^2 + (1-\eta)^t\,\delta^t \pi_k^2 &< \frac{1}{4} + (1-\eta)^t\,\delta^t \pi_n^2 \\ &< \pi_n^2 + (1-\eta)^t\,\delta^t \frac{1}{4} \leq \pi_n^2 + (1-\eta)^t\,\delta^t \pi_l^2. \end{aligned}$$

We can substitute in (2.29) and (2.30) to get that: $U_S(j) \leq U_S(L|n)\,, \forall j$. In particular this holds for $j = n$. □

Corollary 2.1. *If the state has belief 1 then the sender strictly prefers to lie:*

$$\pi_i = 1 \Rightarrow U_S(L|i) > U_S(T|i)\,.$$

Lemma 2.5. *The sender weakly prefers to lie in all the states:*

$$U_S(L|i) \geq U_S(T|i), \forall i.$$

Proof. Suppose $U_S(T|i) > U_S(L|i) \Rightarrow q_i = 1 \Rightarrow \pi_i = 1$.By the corollary above, we have a contradiction. □

We show that the best states to move once the type of sender is identified as strategic are those with lowest beliefs. In other words, that $\pi_l = \pi_1$. The proof is by showing that placing a strategic sender on state s_1 gives the receiver a higher payoff than if the sender is placed on state s_l $(l > 1)$. Remember that after a lie, the receiver knows with probability 1 that the sender is strategic.

From now on, we write q_i independently of the particular history h. We do this w.l.o.g. because the argument holds following any history for which the current state is s_i.

Sending the bad sender to v_l^S gives the receiver the following payoff:

$$\begin{aligned} v_l^S &= q_l \left\{ -(1-\pi_l)^2 + (1-\eta)\,\delta\Sigma_{j^*}\varphi_T(l,j^*)\,v_{j^*}^S \right\} + \\ &\quad + (1-q_l)\left\{-\pi_l^2 + (1-\eta)\,\delta v_l^S\right\}. \end{aligned} \tag{2.31}$$

However, in this state i the strategic sender weakly prefers lying to telling the truth. For if is this not the case, $q_i = 1 \Rightarrow \pi_i = 1$, which implies that lying is actually better for the sender. So we have to consider only the case where

$$(1-\pi_i)^2 + (1-\eta)\,\delta\Sigma_{j^*}\varphi_T(i,j^*)\,U_S(j^*) \leq \pi_i^2 + (1-\eta)\,\delta U_S(i)\,.$$

Thus equation (2.31) can be written as:

$$v_l^S = -\pi_l^2 + (1-\eta)\,\delta v_l^S. \tag{2.32}$$

Now consider a deviation where the receiver receives a lie and decides to place the sender in the lowest belief state instead of moving to the state where the expected continuation payoff is v_l^S. This deviation gives the receiver a payoff of:

$$v_1^S = q_1\left\{-(1-\pi_1)^2 + (1-\eta)\,\delta\Sigma_{j^*}\varphi_T(1,j^*)\,v_{j^*}^S\right\} + (1-q_1)\left\{-\pi_1^2 + (1-\eta)\,\delta\bar{v}_i^S\right\}.$$

Again, we have only to consider the case where:

$$(1-\pi_i)^2 + (1-\eta)\,\delta\Sigma_{j^*}\varphi_T(i,j^*)\,U_S(j^*) \leq \pi_i^2 + (1-\eta)\,\delta U_S(i)\,.$$

For if this is not true then $q_1 = 1$ and state 1 would not be the lowest belief state. Thus, again we can write:

$$v_1^S = -\pi_1^2 + (1-\eta)\,\delta v_l^S. \tag{2.33}$$

However we can compare the expected payoff on equations (2.32) and (2.33) to see that:

$v_1^S \geq v_l^S$, since :

$$-\pi_1^2 + (1-\eta)\,\delta v_l^S \geq -\pi_l^2 + (1-\eta)\,\delta v_l^S.$$

This means that after a lie, the receiver always prefers to place the bad sender on state 1. $\varphi_L(i,1) = 1, \forall i$.

Lemma 2.6. *Memory state 1 has highest expected payoff given a strategic sender:* $1 \in \mathcal{D}$.

Proof. The expected payoff of the receiver given a strategic type of sender is given by:

$$v_l^S = q_l \left\{ -(1-\pi_l)^2 + (1-\eta)\,\delta\Sigma_{j^*}\varphi_T(l,j^*)\,v_{j^*}^S \right\} + (1-q_l)\left\{-\pi_l^2 + (1-\eta)\,\delta v_l^S\right\}.$$

However,

$$(1-\pi_l)^2 + (1-\eta)\,\delta\Sigma_{j^*}\varphi_T(l,j^*)\,U_S(j^*) \leq \pi_l^2 + (1-\eta)\,\delta U_S(l),$$

for if the sender strictly prefers to tell the truth in state l, then we would have that $\pi_l = 1$. And lying would be strictly preferred as we saw in corollary (1). This would be a contradiction.

Thus we can write v_l^S as :

$$v_l^S = -\pi_l^2 + (1-\eta)\,\delta v_l^S.$$

Now consider the expected continuation payoff of placing a strategic sender in state 1. Again, we need only to consider the case where

$$(1-\pi_1)^2 + (1-\eta)\,\delta\Sigma_{j^*}\varphi_T(1,j^*)\,U_S(j^*) \leq \pi_1^2 + (1-\eta)\,\delta U_S(1).$$

Thus, we can write v_1^S as:

$$v_1^S = -\pi_1^2 + (1-\eta)\,\delta v_l^S.$$

However, $\pi_1 \leq \pi_l \Rightarrow -\pi_1^2 \geq -\pi_l^2$, and finally:

$$-\pi_1^2 + (1-\eta)\,\delta v_l^S \geq -\pi_l^2 + (1-\eta)\,\delta v_l^S.$$

Thus, $v_1^S \geq v_l^S$. Since by definition of v_l^S, $v_1^S \leq v_l^S$, we proved this lemma. □

The corollary below shows an immediate consequence of this lemma is that unless there is a state π_2 such that $\pi_2 = \pi_1$ and $v_2^S = v_1^S$, we must have that $\varphi_L(i,1) = 1$.

Corollary 2.2. *All the states with lowest expected continuation payoff for the sender must have the same belief:*

$$i \in \mathcal{D} \Rightarrow \pi_i = \pi_1.$$

Proof. Since we ordered the states by π_i, by definition $\pi_1 \leq \pi_l$. Suppose $\pi_l > \pi_1$. As shown in the lemma above:

$$\begin{aligned} v_l^S &= -\pi_l^2 + (1-\eta)\,\delta v_l^S, \\ v_1^S &= -\pi_1^2 + (1-\eta)\,\delta v_l^S. \end{aligned}$$

If $\pi_l > \pi_1 \Rightarrow v_l^S < v_1^S$. This is a contradiction. □

Corollary 2.3. *For any state j such that $\pi_j > \pi_1$ then by incentive compatibility it must be true that $\varphi_L(i,j) = 0$.*

Proof. After a lie we have that $\Pr(B|i,L) = 0, \forall i$. Then by incentive compatibility it must be that $v_1^S > v_j^S$ which implies that $\varphi_L(i,j) = 0$. □

In the following lemma we show that, in equilibrium, the order of the states is exactly the opposite of the order by v_i^S. This means that a state with higher belief has lower expected continuation payoff given that the sender is strategic. The proof relies on the fact that after lying the sender is placed in a state where his expected payoff is v_1^S. Again, this lemma relies on the first result of this section, which says that lying is always weakly preferred by the sender.

Lemma 2.7. *π_i and v_i^S have the exact opposite ordering.*

Proof. Consider any state s_i. The expected payoff conditional on the type of sender being strategic can be written as $v_i^S = -\pi_i^2 + (1-\eta)\,\delta v_1^S$. Consider two states $s_i$and s_j such that $\pi_j > \pi_i$. Then it must be that:

$$-\pi_i^2 + (1-\eta)\,\delta v_1^S > -\pi_j^2 + (1-\eta)\,\delta v_1^S, \tag{2.34}$$

but (2.34) implies that $v_i^S > v_j^S$. □

This lemma leads us to the following result: the order of states will be the same as the order by v_i^B. This means that states with higher beliefs have higher expected continuation payoff for the receiver given that the sender is a behavioral type. The proof of this corollary relies on incentive compatibility. If a state is reached with positive probability, then there must not exist another state that has higher expected continuation payoff for the receiver regardless of the types of sender (i.e. higher v_i^S and v_i^B). Since a state with lower belief has higher v_i^S it must be that this state with lower belief has lower v_i^B. Otherwise for whatever posterior the receiver holds, it is always strictly better to move to this lower belief state than to the original state.

Lemma 2.8. *For states reached with positive probability, π and v^B have the exact same ordering.*

Proof. Suppose $\pi_k > \pi_j$, and $v_j^B \geq v_k^B$. If j is reached with positive probability, then $\exists\ i^*$ such that:

$$p_{i^*}^B v_j^B + \left(1 - p_{i^*}^B\right) v_j^S \geq p_{i^*}^B v_{j'}^B + \left(1 - p_{i^*}^B\right) v_{j'}^S, \qquad \forall j'.$$

Since $\pi_k > \pi_j$, we already know that $v_j^S > v_k^S$. Thus,

$$p_{i'}^B v_j^B + \left(1 - p_{i'}^B\right) v_j^S \geq p_{i'}^B v_k^B + \left(1 - p_{i'}^B\right) v_k^S, \qquad \forall i'.$$

In particular, for $i' = i^*$. Thus, it must be that k is never reached with positive probability.

□

Lemma 2.9. *If the receiver knows with probability one that the sender is behavioral type, he will update to the state with highest expected continuation payoff given a behavioral type of sender:*

$$U_S\left(L|i\right) > U_S\left(T|i\right) \Rightarrow \varphi_T\left(i, h\right) = 1.$$

Proof. Since the strategic type strictly prefers to lie at state s_i it must be true that after any history h we have that $q_i = 0$. This in turn implies that $\Pr\left(B|i, T\right) = 1$. Since we know that $v_h^B \geq v_{i'}^B, \forall i'$ and also that v_i^B and π_i have the same ordering, we must have that:

$$n = \arg\max_{i'} p_i^B v_{i'}^B + \left(1 - p_i^B\right) v_{i'}^S = \arg\max_{i'} v_{i'}^B.$$

Thus, $\varphi_T(i, n) = 1$. □

Lemma 2.10. *$n \in \mathcal{U}$ and $\pi_u = \pi_n$,$\forall u \in \mathcal{U}$.*

Proof. First we show that $v_n^B = v_u^B$, $u \in \mathcal{U}$. We also know that $q_n = 0$. Suppose $v_u^B > v_n^B$ then, we have that the transition to state u has probability one $\varphi_T(n, u) = 1$ (since $q_n = 0$).

$$\begin{aligned} v_u^B &= -(1 - \pi_u)^2 + (1 - \eta)\,\delta \sum_{u^*} \varphi_T(u, u^*)\, v_{u^*}^B \\ &\leq -(1 - \pi_n)^2 + (1 - \eta)\,\delta \sum_{u^*} \varphi_T(u, u^*)\, v_{u^*}^B \\ &\leq -(1 - \pi_n)^2 + (1 - \eta)\,\delta v_u^B = v_n^B. \end{aligned}$$

Thus, $v_u^B > v_n^B$ cannot happen. The proof that $\pi_u = \pi_n$ is analogous to corollary 2. □

The next lemma will be important in order to show that the receiver will not move to a lower state after a true signal.

Lemma 2.11. *If the sender strictly prefers to lie on state i and is indifferent in state j, then $\pi_i > \pi_j$:*

$$U_S(L|i) > U_S(T|i) \text{ and } U_S(L|j) = U_S(T|j) \Rightarrow \pi_i > \pi_j.$$

Proof. Suppose $U_S(L|i) > U_S(T|i)$, $U_S(L|j) = U_S(T|j)$ and $\pi_i \leq \pi_j$.

$$\begin{aligned} U_S(L|i) &= \pi_i^2 + (1-\eta)\,\delta U_S(1), \\ U_S(T|i) &= (1-\pi_i)^2 + (1-\eta)\,\delta U_S(B), \end{aligned}$$

$$\pi_i^2 + (1-\eta)\,\delta U_S(1) > (1-\pi_i)^2 + (1-\eta)\,\delta U_S(B). \tag{2.35}$$

But, we also have that:

$$\pi_j^2 + (1-\eta)\,\delta U_S(1) = (1-\pi_j)^2 + (1-\eta)\,\delta \Sigma_{j^*} \varphi_T(i, j^*)\, U_S(j^*). \tag{2.36}$$

Since, $\pi_i \leq \pi_j$, we have that:

$$\pi_j^2 + (1-\eta)\,\delta U_S(1) \geq \pi_i^2 + (1-\eta)\,\delta U_S(1) > (1-\pi_i)^2 + (1-\eta)\,\delta U_S(h).$$

However: $U_S(h) \geq U_S(i), \forall i$ and $(1-\pi_i)^2 > (1-\pi_j)^2$ .Thus,

$$(1-\pi_i)^2 + (1-\eta)\,\delta U_S(h) > (1-\pi_j)^2 + (1-\eta)\,\delta \Sigma_{j^*} \varphi_T(i, j^*)\, U_S(j^*).$$

Finally, from (2.35) and (2.36) we have that:

$$\pi_j^2 + (1-\eta)\,\delta U_S(1) > (1-\pi_j)^2 + (1-\eta)\,\delta \Sigma_{j^*} \varphi_T(i, j^*)\, U_S(j^*).$$

This is a contradiction with equation (2.36). □

The lemma below shows that the receiver will not walk backwards after receiving a true signal. This is true because after receiving this true signal, the receiver does better staying in the same place rather than degrading the sender. Both the current and the future payoff are higher.

Lemma 2.12. *After a true signal the transition rule is weakly increasing:*

$$\pi_j > \pi_i \Rightarrow \varphi_T(j,i) = 0.$$

Proof. Suppose $\pi_j > \pi_i$ and $\varphi_T(j,i) > 0$. First note that by incentive compatibility it must be true that:

$$p_j^B v_i^B + \left(1 - p_j^B\right) v_i^S \geq p_j^B v_j^B + \left(1 - p_j^B\right) v_j^S.$$

However, it can also be written as:

$$p_j^B v_i^B + \left(1 - p_j^B\right) v_i^S = p_j^B \left(-(1-\pi_i)^2 + (1-\eta)\,\delta \sum_{i^*} \varphi_T(i,i^*)\, v_{i^*}^B \right) + \left(1 - p_j^B\right) v_i^S.$$

But $v_i^S = -U_S(i) = U_S(L|i) \geq U_S(T|i)$, with strict inequality only if $q_i = 0$.

If $U_S(L|i) > U_S(T|i) \Rightarrow \varphi_T(i,n) = 1$, implying that $\pi_i > \pi_j$ (see lemma (20) that implies that if $q_i = 0$ and $q_j > 0 \Rightarrow \pi_i > \pi_j$). Thus, we conclude that $U_S(L|i) = U_S(T|i)$.

Therefore, v_i^S can be written as:

$$v_i^S = -(1-\pi_i)^2 + \sum_{i^*} \varphi^T(i,i^*)\, v_{i^*}^S. \tag{2.37}$$

The expected continuation payoff of moving to state s_i after observing the truth in state s_j can be written using (2.37) as:

$$p_j^B v_i^B + \left(1 - p_j^B\right) v_i^S = -(1-\pi_i)^2 + \sum_{i^*} \varphi^T(i,i^*) \left(p_j^B v_{i^*}^B + \left(1 - p_j^B\right) v_{i^*}^S\right). \tag{2.38}$$

If, instead of going to state i after a truth, the receiver decides to stay in state j for one more period, he gains from that:

$$p_j^B v_j^B + \left(1 - p_j^B\right) v_j^S = -\left(1 - \pi_j\right)^2 + \sum_{j^*} \varphi_T\left(j, j^*\right)\left(p_j^B v_{j^*}^B + \left(1 - p_j^B\right) v_{j^*}^S\right). \tag{2.39}$$

By incentive compatibility and definition of $j^*$and i^* we have that

$$p_j^B v_{j^*}^B + \left(1 - p_j^B\right) v_{j^*}^S \geq p_j^B v_{i^*}^B + \left(1 - p_j^B\right) v_{i^*}^S. \tag{2.40}$$

Using (2.40) in (2.38) and (2.39) gives us:

$$p_j^B v_j^B + \left(1 - p_j^B\right) v_j^S \geq p_j^B v_i^B + \left(1 - p_j^B\right) v_i^S.$$

□

Lemma 2.13. *The receiver always starts either at the lowest memory state or at the lowest after the babbling state:*

$$\varphi_0\left(i\right) = 0, \forall \pi_i > \pi_{(2)}.$$

Proof. The ex-ante receiver chooses in which memory state to start the game. He will start the game at state i_0 such that $i_0 = \arg\max_i \rho v_i^B + (1 - \rho) v_i^S$. Given the results 1 and 3 in proposition (2.7), we have that $\rho < p_j^B, \forall j > 1$. Thus, if $\varphi_0\left(i'\right) > 0$,for some $\pi_{i'} > \pi_{(2)}$, then state $s_{i'}$ is not reached with positive probability in the game, except for time $t = 0$. □

This concludes the proof of proposition 2.7. To relate this proposition with the one presented in the text, we need two additional results:

Lemma 2.14. *The beliefs are extreme:*

$$\begin{aligned}\lim_{\eta\to 0}\rho_l &= 0,\ \textit{for}\ \forall l\in\mathcal{D},\\ \lim_{\eta\to 0}\rho_u &= 1,\ \textit{for}\ \forall u\in\mathcal{U}.\end{aligned}$$

Proof. We can calculate the posterior of the sender's type on any state $l \in \mathcal{D}$ as:

$$p_l^B = \sum_{h\in s_l\cap H_B} \mu\left((h,T)\,|s_l\right) \tag{2.41}$$

However, given the results on 2.7.1 and 2.7.3 from proposition 2.7 together with the fact that the strategic senders will either remain on one of the states in $\mathcal{D}$ forever or will visit it infinitely often, this state, call it l,will be such that i holds. For this, note in this case we have that as $\eta \to 0$, $\Pr\left(h_1|s_l\right) \to 0$ where h_1 means that the time period is 1 and therefore $\Pr\left(B|s_l\right) \to 0$. By incentive compatibility it will then imply that $\varphi_T\left(l,l\right) = 1$ and consequently $\pi_l = 0.5$.

Eventually all the strategic types will have lied. In particular, since states are observable, $q_u = 0$, for $\forall u \in \mathcal{U}$. There are no reputation incentives on the last state, and all strategic senders lie when they reach that state.

In other words, as $\eta \to 0$ we have that the strategic senders will be locked in the lowest state and also that $U_S\left(L|u\right) > U_S\left(T|u\right), \forall u \in \mathcal{U}$ since in the highest states there are no reputation incentives. Thus, eventually only behavioral types remain in the last state, and they stay in the state forever. We then have that $\lim_{\eta\to 0}\rho_u = 1$, for $\forall u \in \mathcal{U}$. □

We now show that the order of beliefs is the same as the order of posteriors. Consider two states π_i and π_j ,$\forall i,j$ such that: $\pi_j > \pi_i$ but also such that in equilibrium the posteriors have different order: $p_j^B < p_i^B$. We will show that this is a contradiction.

Using the monotonicity lemma, we can prove our result. The intuition is that if you have a state s_i with lower belief π and at the same time higher posterior than another state s_j, then the sender can't be indifferent between lying and telling the truth in states s_i and s_j.

Lemma 2.15. *The beliefs of the states are weakly ordered according to the posteriors:*

$$\pi_j > \pi_i \Rightarrow p_j^B \geq p_i^B.$$

Proof. Consider any two states i and j such that: $\pi_j > \pi_i$ and $p_j^B < p_i^B$. This implies that $U_S(T|i) = U_S(L|i)$ and $U_S(T|j) = U_S(L|j)$ cannot hold at the same time. Recall that

$$\begin{aligned} U_S(T|i) &= (1-\pi_i)^2 + (1-\eta)\,\delta \sum_{i^*} \varphi_T(i,i^*)\,U_S(i^*)\,, \\ U_S(L|i) &= \pi_i^2 + (1-\eta)\,\delta U_S(1)\,. \end{aligned}$$

Since the beliefs (π_i) have the same order as $U_S(i)$, from the monotonicity lemma we have that $\sum_{i^*} \varphi_T(i,i^*)\,U_S(i^*) \geq \sum_{j^*} \varphi_T(j,j^*)\,U_S(j^*)$ .Thus:

$$\begin{aligned} U_S(T|i) &= (1-\pi_i)^2 + (1-\eta)\,\delta \sum_{i^*} \varphi_T(i,i^*)\,U_S(i^*) \\ &> (1-\pi_j)^2 + (1-\eta)\,\delta \sum_{j^*} \varphi_T(j,j^*)\,U_S(j^*) \\ &= U_S(T|j)\,. \end{aligned}$$

At the same time we have that:

$$\begin{aligned} U_S(L|i) &= \pi_i^2 + (1-\eta)\,\delta U_S(1) \\ &< \pi_j^2 + (1-\eta)\,\delta U_S(1) \\ &= U_S(L|j)\,. \end{aligned}$$

We have that $U_S(T|i) > U_S(T|j)$ and also that $U_S(L|i) < U_S(L|j)$. Thus, $U_S(T|i) = U_S(L|i)$ which, in turn, implies that $U_S(L|j) > U_S(T|j)$. However,

$$U_S(L|j) > U_S(T|j) \Rightarrow q_j = 0,$$

which implies that $p_j^B = 1$. This is a contradiction. Thus, the only possibility is if:

$$U_S(T|j) = U_S(L|j) \Rightarrow U_S(T|i) > U_S(L|i) \Rightarrow \pi_i = 1,$$

but again we have a contradiction. □

2.8.3 Deterministic transition rules

This section shows necessary and sufficient conditions for the bounded memory player to use non random transition rules. The result below shows a necessary condition on the prior, given a memory size n.

Proof of Proposition 2.4. This shows the lower bound on the priors so that the receiver plays a pure strategy. The proof is by induction. Consider first the two last states, $n-1$ and n. We want to compute a threshold on the prior of that memory state such that the receiver will use $\varphi_T(n-1,n)=1$.

We know that $\pi_n = 1$, if $\pi_{n-1}^2 + (1-\eta)\,\delta\frac{1}{4} > (1-\pi_{n-1})^2 + (1-\eta)\,\delta 1$. Then lying is better than telling the truth and $q_{n-1}=0$, implying that $\pi_{n-1}=\rho_{n-1}$. But if the equation above holds with equality $\pi_{n-1}^2 + (1-\eta)\,\delta\frac{1}{4} = (1-\pi_{n-1})^2 + (1-\eta)\,\delta 1$, then the sender is indifferent between lying and telling the truth. Rearranging the incentive compatibility of the sender we have that:

$$\pi_{n-1} = \frac{1}{2} + (1-\eta)\,\delta\frac{3}{8}. \tag{2.42}$$

Thus, we need to find the lower bound on prior or, equivalently, the highest q that can support (2.42). The intuition is that if q is too high, the posterior will be low and the receiver will not want to move forward, so we need to consider the receiver's incentive compatibility constraint as well.

To compute the incentive compatibility of the receiver, note that: $v_n^B = 0$; $v_{n-1}^B = -(1-\pi_{n-1})^2$; $v_n^S = -1 - \frac{(1-\eta)\delta}{1-(1-\eta)\delta}\frac{1}{4}$; and $v_{n-1}^S = -\pi_{n-1}^2 - \frac{(1-\eta)\delta}{1-(1-\eta)\delta}\frac{1}{4}$.

For the receiver's incentive compatibility to hold, we need that:

$$p_{n-1}^B\left(v_n^B - v_{n-1}^B\right) + \left(1-p_{n-1}^B\right)\left(v_n^S - v_{n-1}^S\right) \geq 0.$$

In this context, rearranging terms and substituting the posteriors and the expected continuation payoffs we have that:

$$\frac{\rho_{n-1}}{\rho_{n-1} + \left(1-\rho_{n-1}\right)q_{n-1}}\left(v_n^B - v_{n-1}^B\right) + \left(1 - \frac{\rho_{n-1}}{\rho_{n-1} + \left(1-\rho_{n-1}\right)q_{n-1}}\right)\left(v_n^S - v_{n-1}^S\right) \geq 0,$$

which happens if and only if:

$$\rho_{n-1} \geq \frac{\pi_{n-1} + \pi_{n-1}^2}{2}. \tag{2.43}$$

For any ρ_{n-1} that is smaller than the threshold above, we need more q to induce the π needed for (2.42) and this would mean that the posterior is too low for the receiver to want to go up. If, on the other hand, the prior is strictly higher than (2.43) then we need a lower q and (2.42) is maintained. We showed that $\varphi_T(n-1, n-1) = 0$, moving forward is better for the receiver.

The conclusion of this result is that if we arrive at state s_{n-1} with a "prior" $\rho_{n-1} < \frac{\pi_{n-1}+\pi_{n-1}^2}{2}$ then we can't have a pure strategy, and it must be that $\varphi_T(n-1, n-1) > 0$. If we arrive at state s_{n-1} with a "prior" $\rho_{n-1} \geq \frac{\pi_{n-1}+\pi_{n-1}^2}{2}$ then using pure strategy is best response for the receiver.

Now let's look at state s_{n-2} and generalize the argument for states $i = n-2, n-3, ...1$. The necessary conditions for $\varphi_T(n-2, n-1) = 1$ are the following.

Suppose (2.42) and (2.43) so that the last two states the receiver plays pure strategy. We want to find conditions for $\varphi_T(n-2, n-1) = 1$.

If (2.42) does not hold with equality, i.e., if it is better for the sender to lie in state s_{n-1}, then the lower bound is higher. Thus we focus on the case where (2.42) holds with equality. More on this appears later. We use the equation:

$$\pi_{n-2} = \frac{1}{2} + (1-\eta)\frac{\delta}{2}\left(\pi_{n-1}^2 - \frac{1}{4}\right), \tag{2.44}$$

together with $\rho_{n-1} \geq \frac{\pi_{n-1}+\pi_{n-1}^2}{2}$ which is the same as $\frac{\rho_{n-2}}{\pi_{n-2}} \geq \frac{\pi_{n-1}+\pi_{n-1}^2}{2}$, in order to write this condition as:

$$\rho_{n-2} \geq \left(\frac{\pi_{n-1} + \pi_{n-1}^2}{2}\right)\pi_{n-2}. \tag{2.45}$$

If ρ_{n-2} is smaller than in equation (2.45) then when we get to state s_{n-1} the receiver will rather stay put than go forward.

We can now generalize the argument and we'll have that for all $i \leq n-2$:

$$\rho_i \geq \frac{\pi_{n-1} + \pi_{n-1}^2}{2} \prod_{k=i}^{n-2} \pi_k. \tag{2.46}$$

■

Corollary 2.4. *As the number of memory states increase $n \to \infty$, the threshold computed in (2.46) goes to zero:*$\rho_n^* \to 0$.

The result above guarantees that moving from state $n-1$ to state n, $\varphi_T(n-1,n) = 1$, is incentive compatible. But what guarantees that $\varphi_T(i-1,i) = 1, \forall i < n$? In other words, what guarantees that there will not be any incentive to deviate from the specified deterministic transition rule? The next lemma answers these questions. I show that if the receiver is playing pure strategy $\varphi_T(i,i^*) = 1$, the beliefs are computed through Bayesian updating and are such that the sender is playing a best response. Then it will be incentive compatible for the receiver not to deviate from the pure strategies. First we check for a deviation from moving forward to staying put. Then we generalize this result to any deviation of going backwards. The second step is to show that going forward one state (equilibrium) is better than jumping.

Lemma 2.16. $-\pi_j(1-\pi_j) > -\pi_j\left(1-\pi_{j'}\right)^2 - (1-\pi_j)\pi_{j'}^2$, $\forall j, j' \in \mathcal{M}$.

Proof.

$$\begin{aligned}
\pi_j(1-\pi_j) &< \pi_j\left(1-\pi_{j'}\right)^2+(1-\pi_j)\,\pi_{j'}^2 \iff \\
\pi_j-\pi_j^2 &< \pi_j-2\pi_j\pi_{j'}+\pi_j\pi_{j'}^2+\pi_{j'}^2-\pi_j\pi_{j'}^2 \iff \\
-\pi_j^2 &< -2\pi_j\pi_{j'}+\pi_{j'}^2 \iff \\
\pi_j^2-2\pi_j\pi_{j'}+\pi_{j'}^2 &> 0 \iff \left(\pi_j-\pi_{j'}\right)^2>0
\end{aligned}$$

This holds for any $\pi_j, \pi_{j'}$. □

To prove proposition 2.5 in the text, we show two lemmas.

Lemma 2.17. *Suppose that the transition rule is deterministic,* $\varphi_T(i,i+1)=1$, *and the strategy for the sender is a best response for him. Then it must be true that:*

$$p_{i-1}^B v_i^B+\left(1-p_{i-1}^B\right)v_i^S \geq p_{i-1}^B v_{i-s}^B+\left(1-p_{i-1}^B\right)v_{i-s}^S, \forall s>0.$$

Proof. We need to show that deviating to state s_{i+1-s} will not be a best reply for the receiver after a true signal is received in state s_i. Note that we can write the equilibrium payoff using the q and the discount factors.

$$\Pi_{eq}=-\rho_i\left(\sum_{k=i}^{n}(1-\pi_i)^2\right)-(1-\rho_i)\left\{q_i\left((1-\pi_i)^2+\delta U_S(i+1)\right)+(1-q_i)\left(\pi_i^2+\delta\frac{1}{4}\frac{1}{1-\delta}\right)\right\}. \tag{2.47}$$

We want an appropriate way to write (2.47) so that we can compare with the payoff from a deviation. Note that we can write $\rho_i+(1-\rho_i)\,q_iq_{i+1}=\pi_{i+1}\pi_i$, $\rho_i+(1-\rho_i)\,q_iq_{i+1}q_{i+2}=\pi_{i+2}\pi_{i+1}\pi_i$, and so on. However, $(1-\rho_i)\,q_i\,(1-q_{i+1})=(1-\pi_{i+1})\,\pi_i$; $(1-\rho_i)\,q_iq_{i+1}\,(1-q_{i+2})=$

$(1-\pi_{i+2})\,\pi_{i+1}\pi_i$ and so on. We can then write (2.47) as:

$$\begin{aligned}\Pi_{eq} &= -\pi_i\,(1-\pi_i)-\delta\left(\pi_i\pi_{i+1}\,(1-\pi_{i+1})+(1-\pi_i)\,\frac{1}{4}\frac{1}{1-\delta}\right)- \\ &\quad -\delta^2\left(\pi_i\pi_{i+1}\pi_{i+2}\,(1-\pi_{i+2})+\pi_i\,(1-\pi_{i+1})\,\frac{1}{4}\frac{1}{1-\delta}\right)+...\end{aligned} \tag{2.48}$$

The deviation payoff can be written in the same way, but with q^{dev} as being the best response for the sender after a deviation. Note however, that $U_S\,(L|i-1)=U_S\,(T|i-1)\,,$ thus $(1-\pi_{i-1})^2+\delta U_S\,(i)=\pi_{i-1}^2+\delta\frac{1}{4}\frac{1}{1-\delta}$ and therefore, any $q_i^{dev}\in[0,1]$ will not change equation (2.48). In particular, consider $\tilde{q}_i=q_i^{eq}$. In fact, consider the same modification for the entire strategy for the sender, i.e., $\tilde{q}_j=q_j^{eq},\forall j\geq i$.

Let's rewrite the deviation payoff replacing the qs in the way suggested above. We want to compare the payoffs period by period. At all periods before reaching state s_{n-s} lemma (23) tells us that the equilibrium payoff is higher. It remains for us to show what happens at state s_{n-s}. The payoff in this case is

$$-\rho_i\,(1-\pi_{n-s})^2-(1-\rho_i)\prod_{k=i}^{n-1}q_k\pi_{n-s}^2,$$

which can be written as:

$$-\prod_{k=i}^{n-1}\pi_k\left[\pi_n^*\,(1-\pi_{n-s})^2+(1-\pi_n^*)\,\pi_{n-s}^2\right].$$

However, we have that:

$$\begin{aligned}\Pi_{eq}\,(n-s+1) &> \Pi_{dev}\,(n-s+1) \iff \\ \pi_n^*\,(1-\pi_n)^2+(1-\pi_n^*)\,\pi_n^2 &< \pi_n^*\,(1-\pi_{n-s})^2+(1-\pi_n^*)\,\pi_{n-s}^2\end{aligned}$$

Note that this will happen if and only if:

$$
\begin{aligned}
1-\pi_n^* &< \pi_n^* - 2\pi_n^*\pi_{n-s} + \pi_n^*\pi_{n-s}^2 + \pi_{n-s}^2 - \pi_n^*\pi_{n-s}^2 \iff \\
0 &< (1-\pi_{n-s})\left\{2\pi_n^* - (1+\pi_{n-s})\right\}.
\end{aligned}
$$

Finally, this happens if and only if:

$$2\pi_n^* > 1 + \pi_{n-s}.$$

However, a necessary condition for equilibrium in pure strategy was that it should be incentive compatible for the receiver to update in state s_{n-1}. This condition is that $\rho_{n-1} \geq \frac{\pi_{n-1}+\pi_{n-1}^2}{2}$,knowing that we have that $\pi_n^* = p_{n-1}^B = \frac{\rho_{n-1}}{\pi_{n-1}}$, but $\rho_{n-1} \geq \frac{\pi_{n-1}+\pi_{n-1}^2}{2}$, thus $\pi_n^* \geq \frac{1+\pi_{n-1}}{2} > \frac{1+\pi_{n-s}}{2}$. Thus, we showed that the equilibrium payoff is greater than the deviation payoff at every period. □

Lemma 2.18. *Under deterministic transition rules we must have that:*

$$\rho_i v_i^B + (1-\rho_i)\, v_i^S \geq \rho_i v_{i+s}^B + (1-\rho_i)\, v_{i+s}^S.$$

Proof. The equilibrium payoff is again given by (2.47), and again we can write as in equation (2.48). We can further change the q and write (2.48) with $q_{n-s} = 0$,instead. This change in q_{n-s} will not change the value of Π_{eq} since $U_S\,(L|n-s) = U_S\,(T|n-s)$ or, $(1-\pi_{n-s})^2 +$

$\delta U_S(n-s+1) = \pi_{n-s}^2 + \delta$.The deviation payoff is:

$$\Pi_{dev} = -\rho_i \left(\sum_{k=i+s}^{n} (1-\pi_k)^2 \right) - (1-\rho_i) \left\{ q_i^{dev} \left((1-\pi_{i+s})^2 + \delta U_S(i+s+1) \right) + \left(1 - q_i^{dev} \right) \left(\pi_{i+s}^2 + \delta \frac{1}{4} \frac{1}{1-\delta} \right) \right\}. \tag{2.49}$$

Replace q_j^{dev} for $\tilde{q}_j$ for all $j \in \{i+s, ..., n-1\}$. Consider $\tilde{q}_i = q_i^{eq}$. In fact, consider the same modification for the entire strategy of the sender, i.e., $\tilde{q}_{j+1} = q_j^{eq}$. We first show that a deviation to the immediately higher state is not profitable. Then, we extend the argument to all other states. There is also an alternative proof through induction. Even if the bounded memory player could choose his beliefs satisfying only the incentive compatibility of the sender, he would still choose the same beliefs induced by the deterministic transition rules.

Once we use $\tilde{q}$ as the deviation probabilities for the sender, then (2.49) can be written as:

$$\begin{aligned} \Pi_{dev} = & -\left[\pi_i (1-\pi_{i+1})^2 + \pi_i \pi_{i+1}^2 \right] - \\ & -\delta \left(\pi_i \left[\pi_{i+1} (1-\pi_{i+2})^2 + (1-\pi_{i+1}) \pi_{i+2}^2 \right] + (1-\pi_i) \frac{1}{4} \frac{1}{1-\delta} \right) - \\ & -\delta^2 \left(\pi_i \pi_{i+1} \left[\pi_{i+2} (1-\pi_{i+3})^2 + (1-\pi_{i+2}) \pi_{i+3}^2 \right] + \pi_i (1-\pi_{i+1}) \frac{1}{4} \frac{1}{1-\delta} \right) + ... \end{aligned} \tag{2.50}$$

We now want to compare the payoffs in (2.48) but with $q_{n-1} = 0$ and (2.50) *period by period*. Note that according to lemma (9) we have that the payoff in (2.48) is greater than the payoff in (2.50) in every period before $n-i$. At this period, $\tilde{q}_{n-1} = 0$. Period $n-i$ we have that $-\rho_i (1-\pi_{n-1})^2 - (1-\rho_i) \left(\prod_{k=i}^{n-2} q_k \right) \pi_{n-1}^2$ whereas in the deviation we have that:$-\rho_i (1-\pi_n)^2 - (1-\rho_i) \left(\prod_{k=i}^{n-2} q_k \right) \pi_n^2$.

We want to show that:

$$-\rho_i (1-\pi_{n-s})^2 - (1-\rho_i)\left(\prod_{k=i}^{n-s+1} q_k\right)\pi_{n-s}^2 > -\rho_i (1-\pi_n)^2 - (1-\rho_i)\left(\prod_{k=i}^{n-2} q_k\right)\pi_n^2,$$

but this happens if and only if:

$$-\rho_i (1-\pi_{n-s})^2 - (1-\rho_i)\left(\prod_{k=i}^{n-s+1} q_k\right)\pi_{n-s}^2 > -(1-\rho_i)\left(\prod_{k=i}^{n-2} q_k\right).$$

This can be written as

$$\begin{aligned} (1-\pi_{n-s})\left\{(1-\rho_i)\left(\prod_{k=i}^{n-s+1} q_k\right)(1+\pi_{n-s}) - \rho_i (1-\pi_{n-s})\right\} &> 0, \\ 1+\pi_{n-1} - \rho_i\left(\prod_{k=i}^{n-s+1} q_k\right) - \rho_i\left(\prod_{k=i}^{n-s+1} q_k\right)\pi_{n-s} + \rho_i \pi_{n-s} &> 0. \end{aligned}$$

Finally, this implies that

$$1-\rho_i\left(\prod_{k=i}^{n-s+1} q_k\right) + \pi_{n-1} + \rho_i \pi_{n-1}\left(1-\left(\prod_{k=i}^{n-s+1} q_k\right)\right) > 0,$$

which is always true. This argument can be extended to all states with higher beliefs. I.e., deviating to state $i+2$ is worse than $i+1$ and so on. □

In the lemma below we show that there is at most one equilibrium in pure strategies when there are no identical states.

Proof of Proposition 2.6. Let π and π' be the vectors of beliefs associated to two different equilibria in pure strategies (if the beliefs are identical, then we must have that the equilibrium is in fact unique). Assume w.l.o.g. that $\pi_i > \pi'_i$ for some $i \in \mathcal{M}$. This implies that $\pi_{i+1} > \pi'_{i+1}$, for $\forall i < n-1$. This result is true because of the incentive compatibility of the sender, for if $\pi_i > \pi'_i$ and $\pi_{i+1} \leq \pi'_{i+1}$ then it must be that either the receiver is not playing a pure strategy or that the sender is not indifferent between telling the truth or lying in state i in one of the two equilibria. This would imply that the sender is a deterministic transition rule in state i in one of the two equilibria. Given this result, now let's examine two possibilities:

It could be that $\pi_{n-1} = \pi'_{n-1}$ implies that $\pi_{n-2} = \pi'_{n-2}$; also $\pi_{n-3} = \pi'_{n-3}$; and so on, which is a contradiction.

It could also be that $\pi_{n-1} > \pi'_{n-1}$. However, by incentive compatibility of the sender we will have that $\pi_{n-2} > \pi'_{n-2}$ and so on. Thus, $\pi_1 > \pi'_1 \Rightarrow q_1 > q'_1$, which in turn implies that $p_1^B < p_1'^B$. We know that $\pi_2 > \pi'_2$ hence $q_2 > q'_2$. Following the argument we get that $p_{n-2}^B < p_{n-2}'^B$, but $\pi_{n-1} > \pi'_{n-1}$. This is a contradiction since in this case it must be that $\pi_{n-1} = p_{n-2}^B$ and $\pi'_{n-1} \geq p_{n-2}'^B$. ■

Chapter 3

Bounded Memory and Limits on Learning

3.1 Introduction

Many economic interactions have been going on for a very long time. Interestingly, in some cases it is not clear whether both parties benefit from the relationship. Notorious examples are of countries that have been negotiating and engaging in disputes for centuries, sometimes with no clear benefits for one of the parties. The question that we ask in this chapter is the following: can we explain long-term relationships even when parties have opposite preferences?

If agents have opposite interests, a repeated interaction will not suffice to sustain a long-term relationship between them. Including prior uncertainty about the opponent's motives can only sustain reputation in the short run. In a repeated game with incomplete information and zero-sum normal type, the actual type will be asymptotically revealed, as we will show in the chapter. This convergence result implies that the two parties might cooperate in the short-run, but not in the long-run.

We show that incorporating bounded memory on the uninformed player will impact the possibility of long-run reputations. Thus, bounded memory explain long-term relationships *even* in the extreme case where parties have opposite interests.

Memory is modeled as a finite set of states in which the strategy of the player is to choose a transition rule and an action rule. At every stage game the only information that the player has about the history of the game is her current memory state. She can then compute a best response based on the beliefs about the actual history at that point, knowing that she is forgetful across periods. Thus, unlike a non-deterministic finite automaton, the equilibrium strategy of the bounded memory player must satisfy incentive compatibility constraints.[1]

The setting is a repeated zero-sum game with two-players, one of which has bounded memory. He faces a player that, with some exogenous probability, is committed to a specified mixed strategy. In a repeated game with incomplete information, a player with a bounded number of states faces two constraints: bounded complexity on implementing a strategy and bounded ability on updating beliefs about the actual type. The literature on automata has focused on the first issue, while we focus on the second.[2] In the setting of this chapter, the bounded memory player is uninformed about the type of his opponent. Moreover, the complete information game has a unique equilibrium in the repeated game. Thus, the complexity of implementing a strategy is simple, and the issue is on updating beliefs and

---

[1] Memory modeled in such a way was already studied in Monte (2006), where we also discuss the comparison between the bounded memory models and the finite automata models.

[2] See, for example, Neyman (1985), Rubinstein (1986), and Kalai and Stanford (1988) for repeated games with automata.

learning.

Our main result is shown in proposition 3.3 in the text. We compute an upper bound on the learning ability of the uninformed player as a function of his memory size. Although this upper bound decreases exponentially as memory increases, it is always bounded away from zero. This implies that learning is never perfect in a world with bounded memory.

The impossibility of learning is due to the incentive compatibility constraints, and not to the memory restriction itself. In fact, we show that, without these incentive compatibility constraints, the player is able to learn even with a two state-automaton. Thus, if the players could credibly commit to a memory rule, learning would be possible.

The main intuition for our result is that with bounded memory the agent can hold only one of a finite number of beliefs in equilibrium. And, these beliefs cannot be too far apart from each other, or else the sequential rationality constraints would not be satisfied.

Earlier papers have worked on the question of learning in the long run. Under different settings, Hart (1985), Kalai and Lehrer (1993), and more recently Sorin (1999), for example, showed that under some conditions the beliefs over an underlying stochastic process will eventually converge.

In reputation games, or repeated games with one-sided incomplete information, there is an analogous result. Either the uninformed player's beliefs about the informed player's type converges to the "correct one" in the limit or the types will eventually pool in some equilibria of the complete information game. Aumann and Maschler (1995), Benabou and Laroque (1992), Jackson and Kalai (1999) and Cripps, Mailath and Samuelson (2004) showed that this convergence result holds with an impressive generality, robust to different monitoring technologies and different underlying games. Essentially, their main result asserts that in an incomplete information game, after an arbitrarily long history, any equilibrium of the continuation game must be an equilibrium of the complete information game. The strategic use of reputation will be eventually washed off.

The benchmark model in this chapter is closer to Benabou and Laroque's (1992) model. They showed that the Markov perfect equilibria in an imperfect monitoring game where parties have opposite interests will exhibit this asymptotic revelation of types.[3] We use some of their arguments in section 3.3.

Explaining permanent reputations with imperfect memory is also currently being studied by other authors. Ekmecki (2005) showed that if the memory of the uninformed player is restricted (in the form of a finite set of ratings) then there exists a rating system (set of ratings and a transition rule) that can explain permanent reputation. The main difference between the two papers is that here memory is endogenous. It is part of the uninformed player's strategy and has to respect incentive compatibility requirements. In Ekmecki (2005) the memory process is exogenous: designed by a third party.

Other authors have worked on alternative explanations for permanent reputations. For example, in a game where types are continuously changing, permanent reputation can be sustained as shown by Holmstrom (1999), Cole, Dow and English (1995) and Mailath and Samuelson (2001). In a related study, Bar-Isaac (2004) showed that a model of reputation in teams can endogenously introduce this type uncertainty and thus sustain reputation.

[3] There are two main differences between our benchmark model and Benabou and Laroque's (1992). First, we have a perfect monitoring game with a commitment type playing a mixed strategy, whereas they have an imperfect monitoring game with the equivalent of an action type (in the honest equilibrium). Second, the payoff matrix for both players is different in the two papers.

Finally, in the first chapter we studied a similar game. There, we considered the case where the commitment type was playing a pure strategy. As a result, types were revealed in finite time with probability one. In this chapter we consider a behavioral type that is committed to a mixed strategy. In particular, a strategy that plays all actions with positive probability after any history. The difference is analogous to the distinction between perfect monitoring and imperfect monitoring games. This modification allows us to study the long run effects of limited memory.

The reason for why we had full learning in the first chapter, but not here is that, in the former, the equilibrium beliefs associated to every state did not impose a bound on the posteriors. This happened because one of the actions was fully revealing (never played by the commitment type). Here, the commitment type is playing a mixed strategy that assigns positive probability to every action. Thus, in equilibrium, the beliefs can't be too far away from each other, since this would violate the sequential rationality constraints as we show in section 3.6.1. This implies that we cannot have full learning.

Section 3.2 describes the model. In section 3 we solve the model for the full memory case. We present a definition of learning with bounded memory in section 3.4. Learning with a two-state automaton is shown in section 3.5. Section 3.6 is the main part of the chapter where we show that under bounded memory reputation will always be sustained. We conclude in section 3.7.

		player1: Informed	
		Head	Tail
player 2:	Head	$1, -1$	$-1, 1$
Uninformed	Tail	$-1, 1$	$1, -1$

Figure 3.1: Payoffs: Matching Pennies

3.2 Model

3.2.1 Repeated Game with Incomplete Information

The game is an infinitely repeated zero-sum game with incomplete information. We will consider the two-player matching pennies case. There are two players: one informed, called player 1, and one uninformed, player 2. In the beginning of the game, nature draws one of two possible types for the informed player: $k \in \{B, S\}$. Either a behavioral type (B) or a strategic type (S). The uninformed player is not aware of nature's choice. The payoffs of the stage game are shown in figure 1. The repeated game payoff is discounted by a rate $\delta < 1$.

The commitment type is playing a given mixed strategy known by both players. The normal type maximizes his payoffs, given by the payoff matrix in figure 1.

One interpretation for this game is the following. A policy maker is uninformed about his adviser's motives. With some exogenous probability, this adviser is playing a known action, which can be thought of as giving the correct advice about some issue. This loyal adviser makes mistakes with a fixed probability (committed to a mixed strategy). There is also a chance, though, that the adviser has opposite preferences and his goal is to make the policy maker worse-off. He might pretend to be loyal, but his ultimate goal is to mislead his employer.

There are two main reasons for considering this particular game. First, the question of learning under bounded memory is more clear cut here. In a repeated game with incomplete information, the uninformed player needs memory states for two different reasons: to be able to play complex strategies and to learn about his opponent's type. Thus, modeling a player with bounded memory is capturing bounded rationality in the way a player implements his strategy as well as in his ability to learn. [4] We want to 'isolate' the effects of bounded memory on the ability to learn, so we chose a game in which the strategies are not complex, even under full memory. In the complete information game there is a unique equilibrium in the repeated game. This implies that the only issue in this game is on learning the opponent's type.

The second for choosing the repeated matching pennies game is that we want to stress that an infinitely renewable reputation can happen even in a world in which parties have completely opposite interests. In a general two-player game (not zero-sum) there might be a possibility of types pooling in the same equilibrium of the repeated game. In this case, the question of learning looses its byte. The zero-sum nature of this game, though, will ensure that the types will not play the same equilibrium continuation strategy in any subgame. In other words, the strategic type will not mimic the behavioral type forever. If he did so, he would get negative payoffs indefinitely, which is lower than zero, which he can always guarantee himself with.

---

[4] Modeling bounded rationality in implementing a strategy has been widely studied in the automata literature–see Neyman (1985), Rubinstein (1986), and Kalai and Stanford (1988), for example. On the other hand, linking bounded rationality with learning is one of the main contributions of Monte (2006).

3.2.2 Memory and Strategies

In the reputation game we make no restrictions on the strategic type's memory. He can recall the exact history of the game. We denote H to be the set of histories, where a typical element of this set is h. A history h is the sequence of action profiles and the sequence of memory states of the uninformed player.

We will assume throughout the chapter that the memory state of player 2 is observed by player 1 at every point in time. Thus, the history observed by the informed player is the actual history of the game. We define the strategy of the strategic type of the informed player as

$$q : H \rightarrow \Delta \{Head, Tail\}.$$

Where the set of actions of the stage game is $\{Head, Tail\}$. Moreover, we will restrict attention to strategies that are Markovian in the memory states.

The behavioral type is committed to a simple mixed strategy. After every history of the game he plays heads with the same probability $\bar{q}$, i.e. $\bar{q} = \Pr(Head|h)$, $\forall h \in H$. Where $\frac{1}{2} < \bar{q} < 1$ and $\bar{q}$ is known by both players.

The uninformed player has bounded memory. Memory is defined as a finite set of states $\mathcal{M} = \{1, 2, ..., n\}$, and the typical elements of this set are s_i or s_j, or simply i or j.

The strategy of the bounded memory player is to choose the map from states to action, which we call the action rule,

$$a : \mathcal{M} \rightarrow \Delta \{Head, Tail\}.$$

Also, the bounded memory player chooses a transition from state to state

$$\tau : \mathcal{M} \times \{Head, Tail\}^2 \to \Delta(\mathcal{M}),$$

which determines how he update beliefs. Finally, he decides on an initial state $\tau_0 \in \Delta(\mathcal{M})$, which is decided before he enters the first stage game. We denote $\varphi_H(i,j)$ as the probability of moving from state i to state j given that the opponent has played heads, regardless of his own action.[5]

3.2.3 Beliefs

We will *assume* that the beliefs of the bounded memory player are computed using the invariant distribution (when it exists) of the types over the memory states. I.e. conditional on the actual type of player 1, the strategy profile $\sigma = (\tau, a, q)$ will induce a distribution over the memory states. If this distribution converges to some invariant distribution, we refer to this distribution as f^k where $f^k(i)$ is the ergodic probability of being at state s_i given that the actual type of player 1 is $k \in \{B, S\}$.

When this distribution does not converge, we will use the limit of the average of the distributions. Let $\Pr(s_i|k,\sigma,t)$ be the probability that at time t the bounded memory player will be at memory state s_i given that the actual type of player 1 is k and that the strategy profile is σ. The long-run average is given by $\bar{f}^k(i) := \lim_{t\to\infty} \frac{1}{t}\sum_{j=0}^{t-1} \Pr(s_i|k,\sigma,j)$ for $k \in \{B, S\}$.

As we said, though, we will consider only the class of equilibria in Markov strategies

[5]The player ignores his own action in the transition rule because it does not reveal anything about his opponent's type. This is due to the simultaneous nature of the stage game.

(Markovian on the memory states). This is possible given the assumption that memory states are observable by the informed player. In the equilibria that we look at, the strategic type of player 1 will play a stationary strategy that is conditioned only on the current memory state of the uninformed player. The reason for why we focus on this particular class of equilibria is that we will have an invariant distribution over the memory states that will illustrate nicely the intuition for our main result.[6]

Given this class of strategies, the strategy profile $\sigma = (\tau, a, q)$ will induce a transition matrix for each type of the informed player. Let $P^B = \left[\tilde{\varphi}^B(i,j)\right]$ and $P^S = \left[\tilde{\varphi}^S(i,j)\right]$ be the stochastic matrices that determine the evolution of both types of the informed player over the memory states. For the behavioral type, the i^{th} row and j^{th} column element of P^B is given by:

$$\tilde{\varphi}^B(i,j) := \bar{q}\varphi_H(i,j) + (1-\bar{q})\varphi_T(i,j).$$

For the strategic type,we have:

$$\tilde{\varphi}^S(i,j) := q_i\varphi_H(i,j) + (1-q_i)\varphi_T(i,j).$$

The stochastic matrices P^B and P^S define transition probabilities on $\mathcal{M}$ which guarantees us that the sequence $\left\{\frac{1}{n}\sum_{i=0}^{n-1}\left(P^k\right)^i\right\}_{n=1}^{\infty}$ converges to some stochastic matrix R. This means that $\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n-1}\Pr(s_i|k,\sigma,j) = \Pr(s_i|k,\sigma,t=0)R$, for any sequence

[6] The assumption of observable memory states and the Markovian restriction is not crucial for our results. We made these assumptions only to have well defined Markov matrices. We could, instead, have used stopping probabilities over the time periods and compute beliefs using a "frequentist" approach as suggested in Piccione and Rubinstein (1997) and used in Wilson (2006), both papers study a decision problem. This was done in a multi-player game in Monte (2006).

$\{\Pr(\cdot|k,\sigma,t)\}_t = \left\{\Pr(\cdot|k,\sigma,0)(T)^t\right\}_t$.[7] The product $\Pr(\cdot|k,\sigma,t=0)\,R$ is an invariant distribution, which we denote $f^k$and use it to compute the beliefs.

Long-Run Distributions

First, lets consider only the case where either f_i^B or f_i^S is positive. This means that all states are visited infinitely often given the strategy profile σ. When both f_i^B and f_i^S are zero, we assume that the bounded memory player will hold beliefs according to a modified and very long game. We will define and discuss this refinement later.

For the case of positive ergodic distribution of at least one of the two types in all the memory states, we assume that beliefs are computed in the following way.

Whenever the receiver enters a memory state, he holds a belief about his opponent's type. In other words, there is an equilibrium reputation level associated to every memory state. At the beginning of the stage game, this prior on the opponent's type is calculated using Bayes rule given the ergodic distribution. Formally, we define the prior on the behavioral type as follows.

Assumption (Prior Belief in a Memory State)

When f_i^B or f_i^S is positive, the prior *belief on the opponent being a behavioral type at the beginning of a stage game when the current memory state is s_i is given by:*

$$\rho_i := \Pr(B|s_i) = \frac{\rho f_i^B}{\rho f_i^B + (1-\rho) f_i^S}. \tag{3.1}$$

This is the crucial assumption of the section, perhaps of the chapter. All the other

[7] See Stokey and Lucas (1985), theorem 11.1

beliefs will be obtained by using Bayes' rule and the belief at the beginning of the stage game, ρ_i. For example, the belief that the action played in a stage game at memory state s_i is head, is equal to the probability of a behavioral type times the probability that this type plays heads, plus the probability of a strategic type multiplied by his probability of playing heads. Formally, we have that the belief associated to the action played in each memory state is denoted π_i and computed according to the following equation:

$$\pi_i := \Pr\left(Head|s_i\right) = \rho_i \bar{q} + \left(1-\rho_i\right) q_i. \tag{3.2}$$

At the end of the stage game in memory state s_i, and after either action, $Head$ or $Tail$, was played by player 1, the uninformed player will hold a *posterior* belief on the actual type of player 1 being behavioral. We denote this posterior belief by p_i^H or p_i^T. This posterior is computed using the prior ρ_i and Bayesian updating given the strategy profile σ and memory state s_i. It is given by:

$$p_i^H := \Pr\left(B|Head, s_i\right) = \frac{\rho f_i^B \bar{q}}{\rho f_i^B \bar{q} + \left(1-\rho\right) f_i^B q_i}, \tag{3.3}$$

$$p_i^T := \Pr\left(B|Tail, s_i\right) = \frac{\rho f_i^B \left(1-\bar{q}\right)}{\rho f_i^B \left(1-\bar{q}\right) + \left(1-\rho\right) f_i^S \left(1-q_i\right)}. \tag{3.4}$$

We also know that $\Pr\left(S|Head, s_i\right) = 1 - p_i^H$ and similarly for action tail.

Refinement: Reducible Memories

For the case where both f_i^B and f_i^S are zero, we will consider the following refinement. The uninformed player has probability zero of being in a state where both f_i^B and f_i^S are zero, as the number of time periods increase. However, if he finds himself in such a state, he will hold beliefs about the underlying true type of player 1. We take the view that the beliefs in such a state are not "free". I.e., even in states where the ergodic distribution assigns probability zero the beliefs must be consistent.

The way we define beliefs in the case of states not reached with positive probability by the distribution f^B and f^S is to think of very long, but finite games. The uninformed player forms beliefs by considering an alternative game Γ_N of length N, where N is "sufficiently large". In this modified game the strategy profile used is exactly the same as the one in the original game, but truncated at the N^{th} stage game. We define the frequency of a type in memory state s_i during the execution of the strategy profile σ in the game Γ_N as:

$$f_{Ni}^{B} = \Pr\left(s_i|B,\sigma,\Gamma_N\right) = \lim_{N\to\infty}\sum_{h\in s_i}\Pr\left(h|B,\sigma,\Gamma_N\right), \tag{3.5}$$

where $\Pr\left(h|\sigma,B,\Gamma_N\right)$ is the probability of reaching a history h during a play of game Γ_N and the execution of the strategy profile σ, given that the actual type is the behavioral type. Similarly, we define

$$f_{Ni}^{S} = \Pr\left(s_i|S,\sigma,\Gamma_N\right) = \lim_{N\to\infty}\sum_{h\in s_i}\Pr\left(h|S,\sigma,\Gamma_N\right). \tag{3.6}$$

We use this refinement only for states not reached with positive probability by the ergodic distribution, thus, the sum in (3.5) and (3.6) will converge. Again, the prior in a stage game is defined as:

$$\rho_{Ni} := \Pr\left(B|s_i,\sigma,\Gamma_N\right) = \frac{\rho f_{Ni}^{B}}{\rho f_{Ni}^{B} + (1-\rho)\, f_{Ni}^{S}}. \tag{3.7}$$

Again, the bounded memory player's prior belief on the action of the stage game and on his posterior on player 1's type in memory state s_i are given by Bayes' rule, using (3.7). We obtain similar expressions to (3.2), (3.3), and (3.4), but substituting f_i^k by the distribution induced by this modified game: f_{Ni}^k.

We are now ready to define consistent beliefs. Let μ denote the beliefs of the bounded memory player: $\mu = (\rho, p, \pi)$.

Definition 3.1. ***(Consistency)***

A pair (σ, μ) is consistent if given a strategy profile σ, the beliefs μ are computed according to (3.1)-(3.2). Whenever the memory states have probability zero in the long run, beliefs are computed using (3.7).

3.2.4 Equilibrium

In games with forgetfulness, agents might return to the same information set. They forget that they were already in that situation. In particular, they forget what they *did* when they were in that situation. The beliefs that the player holds are the ones induced by the strategy profile σ. If the player deviates at some point in time, he will not remember it, since he is forgetful and does not recall his own actions.

However, we view memory as a conscious process. The player makes conscious decisions on what to remember and on what to forget. Given the beliefs induced by the strategy profile, he will decide on the best action to take, knowing that he will not remember these actions in the future.

For these reasons, the solution concept for games with bounded memory is not the sequential equilibrium. In a game without forgetfulness, sequential equilibrium implies that at any information set, the continuation strategy is optimal for the player given his opponent's strategy and given his beliefs at that information set. In a game with bounded memory, the continuation strategy at an information set need not be optimal. The player is not able to revise his entire strategy, since he does not remember actions, or 'revised plans'. Informally, the player is playing a coordination game with all his different selves, and he might be trapped in a bad coordination equilibrium.

Our concept of optimality involves only optimal actions and transitions, given the beliefs induced by the strategy profile and taking as given the player's own behavior in future nodes. We refer to this concept of optimality as incentive compatibility. This concept was first suggested by Piccione and Rubinstein (1997) and also used by Wilson (2003), both for decision problems.[8] In this dissertation it is the first time studied in a game.

Given a behavioral strategy profile $\sigma = (\tau, a, q)$, the beliefs that the bounded memory player holds in every information set are the beliefs induced by σ. Should the player decide to deviate at some point and use a transition τ' or an action a' different than the one specified by (τ, a), then in the following stage game, the beliefs that this player holds are still the beliefs induced by $\sigma = (\tau, a, q)$.

When deciding on an action to take, and on which state to move, the bounded memory player takes his decisions based on expected continuation payoffs associated with his decisions. He takes his strategy–both action rule and transition rule–as given. The reason for doing this is that if he deviates today, he will not remember it tomorrow.

We will say that the pair (σ, μ) ,is *incentive compatible* if two conditions hold. First, for the informed player we must have that his strategy is a best response for him given a

---

[8] Piccione and Rubinstein (1997) denoted this concept as modified multiself consistent. We call it incentive compatibility, following Wilson (2003).

memory rule (τ, a). Second, for the uninformed player there are no incentives from deviating in any time t given the initial strategy σ and given the action rule a.

For the first condition, we can write:

$$U_S(\tau, a, q) \geq U_S(\tau, a, q') \qquad \forall q' \in Q. \tag{3.8}$$

Where $U_S(\tau, a, q)$ is the expected repeated game payoff for the strategic type of player 1, given the strategy profile σ. Q is the strategy space for the strategic type.

To define the second condition formally, we need extra notation. For every strategy profile $\sigma = (\tau, a, q)$ each memory state has an associated expected continuation payoff. We denote v_i^k as the expected continuation payoff for player 2 at memory state s_i, given that the actual type of player 1 is $k \in \{B, S\}$. We can write these payoffs as a sum of two terms. The first term of v_i^B and v_i^S corresponds to the expected payoff in the stage game given that the memory state is s_i. The second term corresponds to the expected continuation payoff after the first stage game at memory state s_i. This term depends on the associated v_i of all states and on the transition rule τ. Formally we write:

$$v_i^B = (2a_i - 1)(2\bar{q} - 1) + \delta \left(\bar{q} \sum_{j \in \mathcal{M}} \varphi_H(i,j)\, v_j^B + (1 - \bar{q}) \sum_{j \in \mathcal{M}} \varphi_T(i,j)\, v_j^B \right),$$

$$v_i^S = (2a_i - 1)(2q_i - 1) + \delta \left(q_i \sum_{j \in \mathcal{M}} \varphi_H(i,j)\, v_j^S + (1 - q_i) \sum_{j \in \mathcal{M}} \varphi_T(i,j)\, v_j^S \right).$$

Thus, we define an incentive compatible strategy for the bounded memory player as follows.

Definition 3.2. ***(Incentive Compatibility: Memory Rule)***

If a strategy $\sigma = (\tau, a, q)$ is an incentive compatible equilibrium, then it must be true that the memory process (τ, a) satisfies the following conditions for $\forall i, j, j' \in \mathcal{M} \quad \forall k \in \{H, T\}$:

$$\tau_{ij}^{k} > 0 \Rightarrow p_i^k v_j^B + \left(1 - p_i^k\right) v_j^S \geq p_i^k v_{j'}^B + \left(1 - p_i^k\right) v_{j'}^S, \tag{IC1}$$

$$a_i^* = \arg\max_{a \in [0,1]} (2a - 1) \{\pi_i - (1 - \pi_i)\}. \tag{IC2}$$

The first condition, says that when taking the decision of to which memory state to move, player 2 chooses the optimal state, with its associated expected payoff, given his beliefs about the opponent's type p.

The second condition, (IC2), says that if $\sigma = (\tau, a, q)$ is an incentive compatible equilibrium, then the action rule implies taking the myopic best action every stage game. Suppose $a \neq a^*$ where a^* is the optimal myopic action. Suppose also that player 2 deviates and play a^*, and then transition to state s_j, for example. Player 1 knows about this deviation and might want to punish player 2. The fact that player 2 is forgetful implies (by assumption) that whenever he reaches state s_j he will assume that on-equilibrium path actions were taken. Now suppose that there is a way to indicate (through actions of player 1 that he is punishing player 2). Then, if this deviation phase is profitable for player 1, he might as well do it even if no deviation had taken place. Player 1 evaluates what gives him a higher payoff and plays according to it. This implies that player 2 cannot be punished for a deviation of the equilibrium action rule.

Under the multi-self interpretation, a strategy is incentive compatible if no interim self wants to deviate from the equilibrium strategy assuming that all future selves are following it, and all past selves have been following the equilibrium strategy as well. The bounded memory player can deviate from his equilibrium strategy, but he cannot revise his entire strategy. In other words, he cannot trigger a sequence of deviations once the game has started.

We are now ready to define the equilibrium concept of this game. The main part of the concept is to understand the beliefs and the incentive compatibility constraints. The equilibrium of this game must be such that the strategies are incentive compatible and consistent with the beliefs. We formalize the equilibrium concept for this bounded memory game below.

Definition 3.3. ***(Equilibrium)***

The strategy profile $\sigma = (\tau, a, q)$ *is an* incentive compatible equilibrium *if there exists a belief* μ *such that the pair* (μ, σ) *is* consistent *and* incentive compatible*.*

3.3 Equilibrium with Full Memory

In this section we analyze the benchmark model and look at the equilibrium under full memory. Our main purpose is to show that although types are not revealed in finite time, as we show in the proposition below, they will be so in the limit.

This is a model that explains short-run relationships between two parties with opposite preferences. The uninformed player has to act every period and is trying to learn about his opponent's motives at the same time. The strategic type of the informed player has a current incentive for misleading the uninformed player. However, he might find it profitable to 'pretend' to be the commitment type, and explore the benefits of reputation in a future period. We will show in the proposition below that reputation evolves in a realistic fashion, with ups and downs, without ever reaching complete certainty.

We will focus on Markovian equilibrium only, since they suffice for our purposes.[9] Under full memory, the posteriors of the uninformed player are denoted by $\rho^H_{t+1} := \Pr(B|\rho_t, Head)$ and $\rho^T_{t+1} := \Pr(B|\rho_t, Tail)$. The player updates using Bayes' rule.

$$\rho^H_{t+1} = \frac{\rho\bar{q}}{\rho\bar{q} + (1-\rho)\,q(\rho)}, \tag{3.9a}$$

$$\rho^T_{t+1} = \frac{\rho(1-\bar{q})}{\rho(1-\bar{q}) + (1-\rho)(1-q(\rho))}. \tag{3.9b}$$

In equilibrium, it must be true that the strategic type is mixing after every history. If this was not true, than after some history h one of the two actions $\{Head, Tail\}$ would only be played by the behavioral type. This would imply that deviating from the equilibrium strategy by playing this action would give a reputation of one to the strategic type forever. Thus, for such an equilibrium to exist we need that the current payoff must be larger than the continuation payoff with reputation one. A discount factor not too low rules out this possibility. We state the proposition below, and prove it in the appendix.

Proposition 3.1. ***(Reputation Incentives under Full Memory)***

In equilibrium it must be true that

$$U_S(Head|h) = U_S(Tail|h), \text{ and } 0 < q(h) < 1, \ \forall h \in H.$$

[9] In this game Markovian equilibrium does exist and is unique (see Benabou and Laroque (1992) for a proof in a slightly different context). We suspect that it is indeed the unique equilibrium in the game.

At any reputation level the strategic type of the informed player is indifferent between playing heads and tails. Unlike Benabou and Laroque (1992), for example, where a high enough reputation induces this strategic type to strictly prefer to play tail and "milk" down his reputation, here the perfect monitoring nature of the game implies that if in the on-equilibrium path the player strictly prefers one action to the other, this would create the possibility of jumps in the reputation, which in turn would give even higher incentives for the sender to play the action and get the reputation jump. This will ruin the possibility of "equilibrium with reputation jumps".

The reputation $\{\rho_t\}_{t\in\mathbb{N}}$ is a Markov process (as shown above in (3.9a) and (3.9b)). It depends on the sequence of actions, on the equilibrium q and on the actual type of the informed player. From the uninformed player's point of view the reputation of the opponent follows a *martingale*:

$$\begin{aligned}\mathbb{E}\left[\rho_{t+1}|\rho_t=\rho\right] &= \pi\left(\rho\right)\rho^H_{t+1}+\rho\left(1-\bar{q}\right)+\left(1-\pi\left(\rho\right)\right)\rho^T_{t+1} \\ &= \rho_t.\end{aligned}$$

However, conditional on the actual type of the opponent, the reputation evolves differently. When the actual type is a behavioral type, we have that:

$$\mathbb{E}\left[\rho_{t+1}|\rho,B\right]=\bar{q}\rho^H_{t+1}+\left(1-\bar{q}\right)\rho^T_{t+1}.$$

We use equations (3.9a), (3.9b) and also the fact that:

$$\frac{\bar{q}^2}{\pi\left(\rho\right)}+\frac{\left(1-\bar{q}\right)^2}{1-\pi\left(\rho\right)}>1.$$

This implies that conditional on the actual type being behavioral, the reputation of this type will evolve according to:

$$\mathbb{E}\left[\rho_{t+1}|B,\rho\right] = \rho\left\{\frac{\bar{q}^2}{\pi\left(\rho\right)} + \frac{\left(1-\bar{q}\right)^2}{1-\pi\left(\rho\right)}\right\} > \rho. \tag{3.10}$$

The reputation tends to increase every period and is a *strict submartingale.*

We show that the evolution of beliefs on the informed player's type given that the actual type is a strategic type follows, instead, a supermartingale. Before we can show this result, we first prove the following lemma.

Lemma 3.1. *The strategic type plays head less often than the behavioral type:*

$$q_i < \bar{q}, \quad \forall m_i \in \mathcal{M}.$$

Proof. Suppose that there exists a reputation level ρ such that $q\left(\rho\right) > \bar{q}$. This implies that $\pi > \bar{q} > \frac{1}{2}$. Also it must be that the updating is given by: $\rho^H = \frac{\rho\bar{q}}{\pi} < \rho$ whereas $\rho^T = \frac{\rho(1-\bar{q})}{(1-\pi)} > \rho$. Thus, playing tail gives the sender a better current payoff (since $a \geq \frac{1}{2}$) and also a better reputation level. □

Conditional on the type of player 1 being strategic, the reputation of this strategic type will evolve according to:

$$\mathbb{E}\left[\rho_{t+1}|\rho,S\right] = q\left(\rho\right)\rho^H_{t+1} + \left(1-q\left(\rho\right)\right)\rho^T_{t+1}.$$

Using equations (3.9a) and (3.9b) we can write this expected increase in the belief on the type being behavioral given that it is actually the strategic type is given by:

$$\mathbb{E}\left[\rho_{t+1}|\rho, S\right] = \rho \left\{ \frac{\bar{q}q(\rho)}{\pi(\rho)} + \frac{(1-\bar{q})(1-q(\rho))}{1-\pi(\rho)} \right\} < \rho, \tag{3.11}$$

thus the belief ρ_t follows a *supermartingale* when the actual type is strategic.

The proposition below shows that the beliefs will converge conditional on the type of your opponent. Using the Martingale Convergence Theorem together with (3.10) and (3.11) we can show that $lim_{t\to\infty}\mathbb{E}[\rho_t|B] = 1$ and $lim_{t\to\infty}\mathbb{E}[\rho_t|S] = 0$

Proposition 3.2. ***(Benabou and Laroque (1992)–Learning Under Full Memory)***[10]

There is complete learning in this game:

$$\lim_{t\to\infty} \mathbb{E}[\rho_t|B] = 1,$$

$$\lim_{t\to\infty} \mathbb{E}[\rho_t|S] = 0.$$

The intuition for the proof is the following. We know from (3.10) that $\mathbb{E}\left[\rho_{t+1}|\rho, B\right] \geq \rho$ with strict inequality if $\rho \in (0,1)$. From the martingale convergence theorem, $lim_{t\to\infty}\mathbb{E}\left[\rho_{t+1}|\rho, B\right] \to x_\infty$ for some random variable $x_\infty \in [0,1]$ with distribution $d\mu_\infty$. Then, for all period t, it must be true that:

$$\mathbb{E}\left[\rho_{t+1}|B\right] = \int_0^1 \mathbb{E}\left[\rho_{t+1}|\rho, B\right] d\mu(\rho).$$

Taking the limit as $t \to \infty$, means that we are taking the expectation for the reputation of the player conditional on being a good type. This expectation depends on the initial reputation, which is given by some known distribution $\mu(\rho)$. As we take the limit for $t \to \infty$ we want to know the expectation of future reputation when the distribution itself

---

[10] For the formal proof see Benabou and Laroque (1992, p953-954).

converged to μ_∞. From (3.10) we know that $\mathbb{E}\left[\rho_{t+1}|\rho, B\right] > \rho$. Thus, for $\rho \in (0,1)$ it must be that the reputation should converge to one of the extremes: zero or one. It should be intuitively clear, though, that there can be no mass at 0. Thus, the reputation must converge to one. The same reasoning is true for the case of a strategic type.

3.4 Learning or "Type Separation"

In the game with full memory, complete learning (or type separation) means that there are some equilibria such that the types are separated statistically. In other words, the uninformed player eventually learns the actual type. The beliefs about the actual type either converges to one (if it is a behavioral type) or to zero (if it is a strategic type).

In a world with bounded memory, the definition of learning is not straight forward, since the player will only hold one of a finite number of beliefs in equilibrium. Thus, convergence of beliefs has a rather different meaning than in the full memory case.

The claim that we make is that under bounded memory, the definition of learning means that, at some point in time, the uninformed player will almost surely have a good impression (if behavioral type) or a bad impression (strategic type) of the informed player.

Definition 3.4. ***(Complete Learning)***

An equilibrium will exhibit ***type separation*** *(or complete learning) if and only if there are two disjoint subsets of $\mathcal{M}$, call them M_S and M_B such that the two conditions below hold:*

$$\lim_{t\to\infty} \Pr\left(s_i \in M_B|B,t\right) = 1, \tag{3.13}$$

$$\lim_{t\to\infty} \Pr\left(s_i \in M_S|S,t\right) = 1.$$

Although the definition is only a claim about the distribution of the types across the memory states, it implies extreme beliefs. This will be shown in section 3.6.1.

3.5 Learning with a Two-State Automaton

In this section we will show that it is possible to separate types even with a two-state automaton. This is not, in any way, to say that this is the automaton that gives the boundedly rational player the highest payoff. It is only to show that if memory is not a conscious process, then screening is possible. Perhaps very expensive, but possible, even with a minimal amount of memory.

Consider the case of a two state automaton where $\varphi_H(1,2) = 1$, $\varphi_T(1,2) = 0$, $\varphi_T(2,1) = 0$ and $\varphi_H(2,1) = 0$. Also, lets take $a_2 = 1$ and the initial state to be s_1. Now consider the case in state s_1 player 1 is indifferent between playing heads or tail. However, in state s_2 she strictly prefers to play tail. Thus, for any belief that she has, she always prefers to play tail. This implies that $q_2 = 0$ and we will loot at the case where $q_1 = 0$ as well.

At state s_1 the payoff of the strategic player is given by $U_1(Tail) = \frac{2a_1-1}{1-\delta}$ and $U_1(Head) = 1 - 2a_1 + \delta U_2$. At state s_2 it is given by: $U_2 = 2a_2 - 1 + \delta U_2$. Thus, we have $U_2 = \frac{1}{1-\delta}$. Now, substituting for U_1, gives us:

$$U_1 = 1 - 2a_1 + \delta\frac{1}{1-\delta},$$

using the fact that $U_1(T) = \frac{2a_1-1}{1-\delta}$ and solving for a_1 :

$$\frac{2a_1 - 1}{1-\delta} = 1 - 2a_1 + \delta\frac{1}{1-\delta}.$$

Thus, we must have that:

$$a_1 = \frac{1}{2-\delta}. \tag{3.14}$$

This strategy profile: $\sigma = (\tau, a, q)$ induces $f^S = (1, 0)$. The stochastic matrix that determines the invariant distribution of the commitment type is given by:

$$T^B = \begin{pmatrix} 1-\bar{q} & \bar{q} \\ 0 & 1 \end{pmatrix}.$$

The long run distribution is $f_1^B = 0$ and $f_2^B = 1$. Since a_1 is given by (4.8), and $q_1 = 0$, the strategic type of player 1 will never leave state s_1. This means that $f_1^S = 1$ and $f_2^S = 0$, and thus the condition of type separation in definition 3.4 is satisfied.

Thus, it is not the restriction of two memory states *per se* that will impair learning. It is the equilibrium condition as we will see on the following section.

3.6 Bound on Learning for n Memory States

We now turn to the main result of the chapter. We show that in any Markovian equilibrium a player with n memory states cannot fully separate types, as defined in 3.4.

We first show the result for irreducible memories. I.e., if a memory is such that the only ergodic set is the entire set of states, then types are never fully separated. In the second part of the section, we show the proof for the case of a reducible memory.

The main intuition for the result is the following. The player has only a finite set of states, thus the number of posteriors after a particular action induced in equilibrium is finite. If the memory state is ever to be reached in equilibrium, by incentive compatibility, the uninformed player must have had an incentive to do so. This implies that the beliefs in all memory states that are reached through the transition rule must not be "too far apart".

3.6.1 Irreducible memories

Lets consider only irreducible memories for now, we will leave the case of reducible memories for the following section. Irreducible memory is a set of states with no transient states (in which you leave with probability one and never return). If there was complete learning in these cases, then the action would not reveal enough extra information to the bounded memory player to move to another state. Thus, the state would be absorbing, contradicting irreducibility.

The proposition below is the main result of the chapter. It shows that under irreducible memories there is an upper bound k, which is a function of the parameters, on the distribution of types over the memory states. This implies that learning is bounded away from one.

Before showing the result, note that under irreducible memories, the transition rule τ will define a unique ergodic distribution over the memory states.

Proposition 3.3. ***(Bound on Learning)***

For $\forall \rho, \delta$ if (τ, a, q) is an equilibrium in the n memory state game, then there exists some bound $k_{\rho,\delta} > 0$ such that:

$$Pr\left(M|S,t\right) \geq 1-k \Rightarrow Pr\left(M^C|B,t\right) \leq 1-k \qquad \forall t, \forall\ M \in \mathcal{M}.$$

Proof. Consider only the case where memory is irreducible. We first note that there exists at least one $m_i \in M$ such that $\rho_i \leq \frac{\rho\varepsilon}{\rho\varepsilon+(1-\rho)(1-\varepsilon)}$. For suppose not. Then, for $\forall i \in M$ it is true that:

$$\rho_i = \frac{\rho f_i^B}{\rho f_i^B + (1-\rho) f_i^S} > \frac{\rho\varepsilon}{\rho\varepsilon + (1-\rho)(1-\varepsilon)},$$

this implies that $(1-\rho)\left[f_i^B(1-\varepsilon) - f_i^S\varepsilon\right] > 0$. Which in turn implies that:

$$f_i^B(1-\varepsilon) > f_i^S\varepsilon. \tag{3.15}$$

Summing the equation (3.15) for all the state in M we have that:

$$\sum_{i\in M} f_i^B(1-\varepsilon) > \sum_{i\in M} f_i^S\varepsilon. \tag{3.16}$$

However, since we know that $\sum_{i\in M} f_i^S \geq 1-\varepsilon$ and $\sum_{i\in M} f_i^B \leq \varepsilon$. We can then write (3.16) as:

$$\varepsilon(1-\varepsilon) \geq \sum_{i\in M} f_i^B(1-\varepsilon) > \sum_{i\in M} f_i^S\varepsilon \geq \varepsilon(1-\varepsilon). \tag{3.17}$$

It is then true that (3.17) does not hold. Thus, we must have that for at least on state $m_i \in M$

$$\rho_i \leq \frac{\rho\varepsilon}{\rho\varepsilon + (1-\rho)(1-\varepsilon)}. \tag{3.18}$$

For this state not to be absorbing, we need that $p_i^H \geq p_j^T$ for at least some state m_j where $j \neq i$. Computing the posterior in state m_i gives us $p_i^H = \frac{\rho_i\bar{q}}{\pi_i}$.

At the same time, we must have that: $p_j^T = \frac{\rho_j(1-\bar{q})}{1-\pi_j}$. Since $p_i^H \geq p_j^T$, we have that: $\frac{\rho_j(1-\bar{q})}{1-\pi_j} \leq \frac{\rho\varepsilon}{\rho\varepsilon+(1-\rho)(1-\varepsilon)}\frac{\bar{q}}{\pi_i}$, which in turn implies that:

$$\rho_j \leq \frac{\rho\varepsilon}{\rho\varepsilon + (1-\rho)(1-\varepsilon)}\frac{(1-\pi_j)\bar{q}}{\pi_i(1-\bar{q})}. \tag{3.19}$$

Equation (3.19) gives us $p_j^H \leq \frac{\rho\varepsilon}{\rho\varepsilon+(1-\rho)(1-\varepsilon)} \frac{\bar{q}^2(1-\pi_j)}{\pi_i(1-\bar{q})\pi_j}$. From state m_j we move to some state $m_{j'}$ and this gives us the following condition: $p_j^H \geq p_{j'}^T$. We then repeat the procedure above and note that the highest upper bound for the posterior p_{n-1}^H is given when the transition occurs without any jump and it is:

$$p_{n-1}^H \leq \frac{\rho\varepsilon}{\rho\varepsilon+(1-\rho)(1-\varepsilon)} \frac{\bar{q}^{n-1}(1-\pi_2)(1-\pi_3)...(1-\pi_{n-1})}{\pi_1(1-\bar{q})^{n-2}\pi_2\pi_3...\pi_{n-1}}.$$

On the other hand, we must also have that $p_{n-1}^H \geq p_n^T \geq \frac{\rho(1-\varepsilon)}{\rho(1-\varepsilon)+\rho\varepsilon} \frac{(1-\bar{q})}{(1-\pi_n)}$. This implies the following inequality:

$$\frac{\rho\varepsilon}{\rho\varepsilon+(1-\rho)(1-\varepsilon)} \frac{\bar{q}^{n-1}(1-\pi_2)(1-\pi_3)\times...\times(1-\pi_{n-1})}{\pi_1(1-\bar{q})^{n-2}\pi_2\pi_3\times...\times\pi_{n-1}} \geq \frac{\rho(1-\varepsilon)}{\rho(1-\varepsilon)+(1-\rho)\varepsilon} \frac{(1-\bar{q})}{(1-\pi_n)}.$$

Thus, we have that:

$$\begin{aligned} \frac{\rho(1-\varepsilon)\varepsilon+(1-\rho)\varepsilon^2}{\rho\varepsilon(1-\varepsilon)+(1-\rho)(1-\varepsilon)^2} &\geq \frac{(1-\bar{q})^{n-1}\pi_1\pi_2\pi_3\times...\times\pi_{n-1}}{\bar{q}^{n-1}(1-\pi_2)(1-\pi_3)\times...\times(1-\pi_n)} \\ &\geq \left(\frac{1-\bar{q}}{\bar{q}}\right)^{n-1}. \end{aligned}$$

Lets denote: $\varphi(\varepsilon) = \frac{1}{1-\varepsilon}$ and construct $\phi(\varepsilon)$, such that:

$$\phi(\varepsilon) = \frac{\rho(1-\varepsilon)\varepsilon+(1-\rho)\varepsilon^2}{\rho\varepsilon+(1-\rho)(1-\varepsilon)}.$$

For $\varepsilon \in (0,1)$ we have that $\varphi(\cdot)$ is continuous and $\varphi'(\cdot) > 0$. Also, for $\rho < 0.5$ it is immediate to show that $\phi'(\varepsilon) > 0$. For the case in which $\rho > 0.5$, note that $\frac{\partial\phi(\cdot)}{\partial\rho} > 0$, for any $\varepsilon < 0.5$. To show that there is a bound on learning, note that:

$$\lim_{\varepsilon \to 0} \frac{\rho (1-\varepsilon) \varepsilon + (1-\rho) \varepsilon^2}{\rho\varepsilon (1-\varepsilon) + (1-\rho)(1-\varepsilon)^2} = 0.$$

This, together with $\left(\frac{1-\bar{q}}{\bar{q}}\right)^{n-1} > 0$, and with the fact that the product $\varphi(\varepsilon)\phi(\varepsilon)$ is continuous and strictly increasing for ε, gives us our bound on learning. □

Thus, if ε is small enough, then there is no possibility of learning through time. For example, for $n = 2$ if $\rho = 0.5$ and $\bar{q} = \frac{2}{3}$ then $\varepsilon \leq \frac{3}{8}$ is the bound.

The bound is for both f_1^B and f_2^S at the same time. I.e., it might happen that $f_1^B < \varepsilon$ but $f_2^S > \varepsilon$. It would be meaningless if the definition was only for "one-sided" learning. Consider for example the case with only one memory state: $f_1^S = 1$ and also $f_1^B = 1$, but there is no learning whatsoever.

Note that this bound is not tight. In other words, the agent will typically learn much less than the bound that we calculated. In any case, as the number of memory states increases, $n \to \infty$, the bound quickly converges to zero, as one would suspect. An interesting open question is to find the equilibrium memory rule that would lead to faster learning given the number of memory states n.

3.6.2 Reducible case

We now turn to the reducible case. Although it is intuitively clear that a reducible memory cannot have more learning than an irreducible one, because there are less states to "dilute" the posteriors, we prove the result in this section.

Consider a memory with k recurrent classes $\mathcal{R}_1, \mathcal{R}_2, ..., \mathcal{R}_k$ and a set of transient states

$\mathcal{T}$.

We want to show that in this case there is also a bound on learning and, in fact, it is smaller than in the irreducible case.

Proposition 3.4. ***(Bound on Learning: the Reducible Case)***

For $\forall\rho,\delta,n$ if (τ,a,q) is an equilibrium in the $n-$memory state game with k recurrent classes and one transient state, then there exists a bound $k_{\rho,\delta}>0$ such that:

$$Pr\,(M|S,t)\geq 1-k\Rightarrow Pr\left(M^C|B,t\right)\leq 1-k \qquad \forall t,\forall\; M\in\mathcal{M}.$$

Proof. Suppose not. There are two sets of states M_S and M_B such that $\forall\varepsilon>0\;\exists\;\bar{t}>0$ such that $\Pr\left(M_S|S,t\right)\geq 1-\varepsilon$ and $\Pr\left(M_B|B,t\right)\geq 1-\varepsilon$ for $\forall t>\bar{t}$. Consider only the case where $M_S\subseteq R_i$ and $M_B\subseteq R_j$ where $j\neq i$. For if $j=i$ then the result of proposition 3.3 applies and we have that the bound on learning is $\left(\frac{1-\bar{q}}{\bar{q}}\right)^{n_i-1}$, where n_i is the number of states in $\mathcal{R}_i$. Also, if M_S or M_B belong to $\mathcal{T}$ then the result clearly does not hold: $\Pr\left(\mathcal{T}|S,t\right)=0$ as $t\to\infty$, similarly for $\Pr\left(\mathcal{T}|B,t\right)$.

Now for every recurrent class $\mathcal{R}_i$ consider the memory state in which p_i^T is lower. Also denote m_h as the memory state in the transient class that has the highest posterior $p_{\mathcal{T}}^H$. It must be the case that $p_{\mathcal{T}}^H\geq p_i^T$. This is true since all the mass of strategic types that enter a recurrent class will stay there forever. Similarly, if $p_{\mathcal{T}}^T$ is the lowest posterior in $\mathcal{T}$, and p_j^H is the highest posterior in some recurrent class, then: $p_j^H\geq p_{\mathcal{T}}^T$.

Thus, the lowest bound is achieved (or highest possibility of learning) when every state in the transient class connects to each other and every state in the recurrent classes also connect to each other. Using the same reasoning as in proposition 3.3 we have that the bound is given by:

$$\left(\frac{1-\bar{q}}{\bar{q}}\right)^{n_i+n_j+n_{\mathcal{T}}-1}.$$

However, this number is smallest when $n_i + n_j + n_{\mathcal{T}} = n$. In other words, the bound is smallest when all the states communicate with each other. □

3.7 Conclusion

Our chapter contributes to the literature on reputation and repeated games with incomplete information. A celebrated recent result in this literature is that the play of the game converges asymptotically to the play of a complete information game. This means that players can profit from a "false" reputation only in the short-run. Constant opportunistic behavior will lead to statistical revelation of the actual type, which means no long-run reputation.

We show that under bounded memory we will not have full learning (or type separation) even in the long-run.

The result of no learning is due to the incentive compatibility constraints, since it does not apply for the case of a non-deterministic finite automaton, even for the minimal one with two states. Therefore, with initial uncertainty about types and bounded memory on the uninformed player long-term reputations can be sustained even in the extreme case where agents have opposite preferences.

The recent results on reputations and long-term relationships are shown to be robust to different underlying games and different monitoring technologies. From what we showed in this chapter though, it is not robust to cognitive constraints on the individuals.

3.8 Appendix

We show here the proof of proposition **??**. We show the result through two separate lemmas.

Lemma 3.2. *For any* $\rho \in (0,1)$, *the sender always weakly prefers to play tail:*

$$U_S\left(Tail|\rho\right) \geq U_S\left(Head|\rho\right).$$

Proof. Suppose there exists a reputation level ρ such that $U_S\left(Head|\rho\right) > U_S\left(Tail|\rho\right)$. Then, it must be that after a tail the receiver correctly updates the reputation of the player to be 1, since in this case a tail could be observed only through a good type. In the case of completely separating the types, the continuation game goes on with no uncertainty about the opponent's type, i.e. $U_S\left(1\right) = \frac{1}{1-\delta}$. Thus:

$$1 - 2a + \delta U_S\left(\rho^H\right) > 2a - 1 + \delta U_S\left(1\right) = 2a - 1 + \frac{\delta}{1-\delta},$$

$$a < \frac{1}{2} + \frac{\delta}{4}\left\{U_S\left(\rho^H\right) - \frac{1}{1-\delta}\right\}.$$

Since $\frac{1}{1-\delta} > U_S\left(\rho^H\right)$ then it must be that $a < \frac{1}{2}$. However, if $U_S\left(H|\rho\right) > U_S\left(T|\rho\right)$ then the strategic type plays head with probability one: $q = 1$ and $\pi = \rho\bar{q} + (1-\rho)\,q = 1 - \rho\left(1-\bar{q}\right) > 0.5$.

However, if $\pi > 0.5$ this implies that the receiver can best reply with $a = 1$. Thus, it is not an equilibrium. $\square$

This is not surprising since we had obtained the same result when the good type plays a pure strategy ($\bar{q} = 1$). The departure from the pure strategy case is shown below.

Lemma 3.3. *For $\delta > \frac{2}{3}$ and any $\rho \in (0,1)$, the sender always weakly prefers to play head:*

$$U_S\left(Head|\rho\right) \geq U_S\left(Tail|\rho\right).$$

Proof. Suppose that there exists a reputation level ρ such that the sender strictly prefers to play tail (like in the pure strategy case – $\bar{q} = 1$): $U_S\left(Tail|\rho\right) > U_S\left(Head|\rho\right)$. This implies that after observing head the receiver can update his posterior to 1 and play a best response from then on.

$$2a - 1 + \delta U_S\left(\rho^T\right) > 1 - 2a + \delta U_S\left(1\right) = 1 - 2a + \frac{\delta}{1-\delta}.$$

Rearranging and solving for $U_S\left(\rho^T\right)$ gives us:

$$U_S\left(\rho^T\right) > \frac{(2-4a)(1-\delta)+\delta}{\delta(1-\delta)}.$$

However, it is also true that

$$\frac{(2-4a)(1-\delta)+\delta}{\delta(1-\delta)} > \frac{(2-4)(1-\delta)+\delta}{\delta(1-\delta)} = \frac{-2(1-\delta)+\delta}{\delta(1-\delta)} = \frac{-2+3\delta}{\delta(1-\delta)}.$$

Then, assuming that $\delta > \frac{2}{3}$ so that $3\delta - 2 > 0$, it must be that:

$$\frac{(2-4a)(1-\delta)+\delta}{\delta(1-\delta)} > \frac{-2+3\delta}{\delta(1-\delta)} > \frac{1}{\delta(1-\delta)} > \frac{1}{(1-\delta)} = U_S\left(1\right).$$

Thus, $U_S\left(\rho^T\right) > \frac{1}{(1-\delta)}$ which cannot be true, since $\frac{1}{(1-\delta)}$ is an upper bound for the sender's utility. □

Chapter 4

Why Contracts? A Theory of Credibility under (No) Commitment

4.1 Introduction

In long-term relationships with uncertainty, learning occurs over time and is often costly. The ability to commit to a contract can make learning faster and improve a player's payoff. Classical examples are in the study of oligopoly models, and the theory of contracts.

In this chapter we study a credibility game in which commitment will not do any better than sequential rationality. In a credibility model, perhaps counter intuitively, the ability to contract will be beneficial only when the adverse selection problem is less severe.

The underlying game is based on Sobel's (1985) credibility model. A policy maker receives messages from an informed advisor. The advisor can be one of two types: a behavioral type, a loyal employee that will always inform the policy maker about the true state of the world, or a strategic type, one with opposite preferences to the policy maker. This game is shown to have a unique equilibrium in which the advisor builds reputation through time, until a period in which all "bad" types will have lied and the policy maker will play a complete information game.

The question that we ask is whether the policy maker could screen faster than the equilibrium outcome. Consider a game in which this policy maker could write a contract. He could then screen the types in the first period so as to play a complete information game from the second period onwards. In fact, the question is more general. Is there any contract in which the receiver could commit to ex-ante that would improve his payoff? We show that commitment is ineffective; the receiver cannot improve his payoff by committing to a contract ex-ante, even in the case where he can credibly commit to a random device, i.e. commit to a particular mixed strategy. In other words, the policy maker will decide to play the strategy that he would otherwise play and thus, the optimal contract will reproduce Perfect Bayesian equilibrium (PBE).

The complete information version of the game studied here will have the property that commitment is ineffective. We show that after adding initial uncertainty, this incomplete information game will extend the results of the complete information game. The intuition for this result is exactly the expert's motives. The fact that the bad type has opposite interests to the policy maker drives the result. Whenever the sender is of a bad type, the game is a zero-sum game and commitment will play no role.

In section 4.2 we outline the model; first, the game with no ability to commit, and second the game modified with the ability to commit ex-ante to a strategy. In section 4.3

		Sender	
		Head	Tail
Receiver	Head	$1,-1$	$-1,1$
	Tail	$-1,1$	$1,-1$

Figure 4.1: Matching Pennies

we show the main result of the chapter for the special case of two periods. We then show an example when commitment does improve the player's payoff. This is done in section 4.4. Section 4.5 extends the result of section 4.3 for any finite repetitions of this game. We conclude in section 4.6.

4.2 A Credibility Model

4.2.1 Standard no-commitment case

The game is a repeated zero-sum game with incomplete information. We will consider the two-player matching pennies case. There are two players: one informed (sender) and one uninformed (receiver). Before the first stage game, nature draws one of two possible types for the sender: $k \in \{B, S\}$; either a behavioral type (B) or a strategic type (S). The receiver is not aware of the sender's true type. The payoffs of the stage game are shown in figure 1. The repeated game payoff is discounted by a rate $\delta < 1$.

The behavioral type of the sender is committed to a pure strategy known by both players. He is analogous to the honest expert that will always report truthfully to the policy maker (see Sobel (1985)). The strategic type maximizes his payoffs, which are shown in figure 4.1. Note that this strategic sender has opposite preferences to the receiver and thus, his goal is to make the receiver worse-off. This game was studied by Monte (2006), although in another context.

We will show that two properties are crucial for the argument of the chapter. The opposite preferences between the players and the fact that the behavioral type is committed to only one action. Interestingly, the receiver's degree of risk aversion plays no role in the argument.

Let h be a history in the game and h_t a history of length t. The set H_t is the set of all possible histories of same length t. The behavioral strategy of the receiver is a map from the set of histories to the set of actions in the stage game. We denote this strategy by

$$a_t : H_t \rightarrow \Delta \{H, T\}.$$

With slight abuse of notation we denote $a(h_t)$ as the probability of the receiver playing head when the history is h_t. The behavioral type of the sender is committed to a pure strategy, so that the probability of playing head is one, regardless of the history. The behavioral strategy for the strategic type of the sender is denoted

$$q_t : H_t \rightarrow \Delta \{H, T\}.$$

We say that $q(h_t)$ is the probability that the strategic type will play head if the actual history of the game is h_t.

The receiver starts the game with some prior ρ on the probability of facing the behavioral type. We denote ρ_t as the receiver's belief at the beginning of time t on the sender being a behavioral type.

We first look at the case where the receiver cannot commit to a strategy ex-ante. The receiver is learning over time the true type of his opponent. At the first time that the sender plays tail, the receiver's belief on the behavioral type drops to zero and the continuation game is a complete information game with zero-sum opponents. This complete information game has a unique equilibrium where both players play both actions with equal probabilities. In the interpretation of the credibility game, it represents the case where the policy maker no longer pays attention to the sender's reports. Whenever heads is played in the stage game, the receiver's belief on the behavioral type is updated according to Bayes' rule, given the strategy profile $\sigma = (a, q)$. Suppose that at the beginning of a stage game after some history h_t the receiver's belief on the behavioral type is ρ_{h_t} and the strategic sender play head with probability q_{h_t}. Then, if heads is played in the stage game, the posterior belief on the behavioral type will be given by:

$$\rho_{(h_t,(H,\cdot))} = \frac{\rho_{h_t}}{\rho_{h_t} + \left(1 - \rho_{h_t}\right) q_{h_t}}.$$

In the proposition below we show that this game has a unique equilibrium regardless of the time horizon. We denote the incomplete information game above as Γ_N, if the game above is repeated N times.

Theorem 4.1. ***[Unique Equilibrium]***

Given the number of stage games t and the prior on the behavioral type ρ, the game Γ_t has a unique equilibrium. This is true for any $t \geq 1$.

Proof. Let $\sigma^* = (a^*, q^*)$ be an equilibrium strategy profile. Then, at the beginning of some stage game t, the history of the game must be one of two possible histories. Either a history such that there was at least one tail played so far by the sender or a history that contains only heads by the sender. In the first case, the unique PBE is the matching pennies equilibrium. In the second case, the belief of the receiver is given by some $\rho_t(h_t)$.

Consider a game of length t. Given a particular history of length $t-1$, say h_{t-1}, the receiver holds a belief at the beginning of stage game $t-1$ given by $\rho(h_{t-1}, \sigma^*)$. The sender will then use his behavioral strategy $q^*(h_{t-1})$. If the sender played heads in stage $t-1$, the receiver's belief at the beginning of stage game t, regardless of his own action, is given by:

$$\rho_{(h_{t-1},(\cdot,H))} = \frac{\rho(h_{t-1},\sigma^*)}{\rho(h_{t-1},\sigma^*) + (1-\rho(h_{t-1},\sigma^*))\, q^*(h_{t-1})}.$$

We know that for every given belief, the last stage game must have a unique outcome. This argument shows that in the next to last stage game, a deviation from the equilibrium action $a^*(h_{t-1})$ to some other action a'_{t-1} will not change the outcome in the last stage game. In other words, if the equilibrium action $a^*(h_{t-1})$ is not the optimal one in the next to last stage game, then the receiver will deviate to his myopic best one, say a'_{t-1}. This deviation will give him a higher payoff in the next to last stage game and the same payoff in the last stage game. Thus, the last stage game is myopic and the next to last stage game must specify a myopic best response from the receiver. We still need to show that the sender is also playing myopically in the next to last stage game.

At the last stage game, $a_t \geq 0.5$, regardless of the history. Thus, we can write

$$U(H|t-1) = 1 - 2a_{t-1} + \delta(2a_t - 1),$$

and

$$U(T|t-1) = 2a_{t-1} - 1.$$

We also know that the receiver plays myopically in stage game $t-1$. We will analyze all the possible cases in stage $t-1$ according to the receiver's belief about the actions being played.

First consider the case where the probability that the receiver's belief about the sender playing head is $\pi_{t-1} > 0.5$. Since the receiver plays myopically, this implies $a_1 = 1$. In this case:

$$U(T|t-1) = 2a_{t-1} - 1 = 1 > -1 + \delta(2a_t - 1) = U(H|t-1).$$

This implies that the sender would find it profitable to play tail in stage game $t-1$, $q_{t-1} = 0$. Thus, $\pi_{t-1} = \rho_{t-1}$. Since the receiver plays myopically, this can only hold if $\rho_{t-1} > \frac{1}{2}$. And in this case, there is a unique equilibrium.

Now consider the case where $\pi_{t-1} = 0.5$. This implies that any a_{t-1} is a myopic best response for the receiver. For this belief be equal to one half it must be true that

$$q_{t-1} = \frac{\frac{1}{2} - \rho_{t-1}}{1 - \rho_{t-1}}.$$

There is a unique q_{t-1} that satisfies this equation and it can only hold if $\rho_{t-1} \leq 0.5$. Moreover, this implies that $\rho_t = 2\rho_{t-1}$ which, if, greater than $\frac{1}{2}$ implies $a_t = 1$ and a_{t-1} just enough to be consistent with mixing for the sender. I.e. $2a_{t-1} - 1 = 1 - 2a_{t-1} + \delta$, solving for a_{t-1} gives us: $a_{t-1} = \frac{2+\delta}{4}$ and again, unique value. If $\rho_2 = 2\rho_1 \leq 0.5$ then in the second stage we have that $a_2 = 0.5$ and the unique value of a_1 that is consistent with this is $a_1 = 0.5$.

The last case to consider is when the strategy profile σ^* and the actual history h_{t-1} are such that the receiver's belief before the next to last stage game is $\pi_{t-1} < 0.5$. We already know that the receiver plays myopically in the next to last stage game. Thus, in the case where $\pi_{t-1} < 0.5$, the receiver's actions is to play tail: $a_{t-1}(h_{t-1}) = 0$. Given this action, the value for the sender to play tail is:

$$U(T|t-1) = 2a_{t-1} - 1 = -1, \tag{4.1}$$

whereas by playing heads the sender can get a value of:

$$U(H|t-1) = 1 - 2a_{t-1} + \delta(2a_t - 1) \geq 1. \tag{4.2}$$

We can see that (4.1) is smaller than (4.2). Thus, telling the truth is profitable and $q(h_{t-1}) = 1$ which implies that the belief about heads being played must be one, and not smaller than 0.5, i.e. $\pi_{t-1} = 1$.

We saw in the above lemma that for a given history h_{t-1}, the PBE in the next to last stage is unique, given a belief ρ_{t-1}. The argument can be easily extended for any stage game that precedes time period $t-1$. □

Remark 4.1. *Note that for constructing this argument we did not need to separate two different conditions: when deviations can be detected and when they cannot. In this game, deviations cannot be detected and randomization is private, but this is not relevant for the argument.*

4.2.2 Commitment case

Now we consider the case in which the receiver can commit to a map of actions. The timing of the game changes. Before the first stage game, the receiver announces and commits to a contract in which he specifies a probability distribution over the set of actions after every possible history. Then, the sender plays an action, gets a payoff and plays the next stage game and so on.

The behavioral strategy for the sender is then given by the map:

$$q_t : C \times H_t \to \Delta \{H, T\},$$

which depends on the history and on the contract specified by the receiver. The strategy of the receiver is to specify a contract C which is a vector of probability distribution for every possible history. I.e. the chosen contract C will be a map:

$$C : H \to \Delta \{H, T\}.$$

If the game is repeated t–times, for example, then a contract C must assign a specific probability distribution over the actions for each different history.

4.3 The two-period case

We first show the result for the two-period case, which we denote Γ_2. This will provide intuition for the general argument.

4.3.1 Two period with no-commitment

In this section we show the two-period case. We will first construct the equilibrium of this game, and argue that this equilibrium is unique. This is a special case of the general theorem 4.1.

	a_1	q_1	ρ_2	a_2	q_2
$\rho > \frac{1}{2}$	1	0	1	1	–
$\frac{1}{2} \geq \rho > \frac{1}{4}$	$\frac{2+\delta}{4}$	$\frac{\frac{1}{2}-\rho_1}{1-\rho_1}$	$2\rho_1 > \frac{1}{2}$	1	0
$\rho \leq \frac{1}{4}$	$\frac{1}{2}$	$\frac{\frac{1}{2}-\rho_1}{1-\rho_1}$	$2\rho_1 \leq \frac{1}{2}$	$\frac{1}{2}$	$\frac{\frac{1}{2}-\rho_2}{1-\rho_2}$

Table 4.1: Equilibrium in two-stage game

At the second stage game there are no reputation effects, so both players will play their myopic best response. First, note that the receiver will always play head with probability greater or equal to one half. If this was not true, then $a_2 < 0.5$, which would imply that the sender would play heads with probability one $q_2 = 1$, and thus the receiver's belief that heads is played is $\pi_2 = 1$. This implies that $a_2 < 0.5$ is not a best response.

The receiver's action will be greater than 0.5 if the belief that the sender's type is behavioral is also greater than 0.5. In this case, the receiver's belief that heads will be played is $\pi_2 = \rho_2 > 0.5$ and $a_2 = 1$ is a best response regardless of q_2. In the case where the belief on the sender's type is smaller than 0.5, the unique outcome in the second stage game must specify $a_2 = 0.5$ and q_2 just enough to keep $\pi_2 = 0.5$.

Knowing this, the receiver will play his myopic best response since he knows that his own action will not change his belief on ρ_2, which in turn determines his payoff in the second stage. This implies that if $\rho > 0.5$, the game has a unique equilibrium where $a_1 = 1$ and $q_1 = 0$ and the second stage game is a complete information game. In this case, we say that there is perfect screening in the first period. When the initial prior on the sender's type is smaller than 0.5, then the unique equilibrium is such that $\pi_1 = 0.5$ and the belief in the second stage is either 0.5 or ρ_2, whichever is higher.

Thus, we can summarize the equilibrium of this two-stage game in the table below.

Table (4.1) shows that if $\rho > \frac{1}{2}$, then there is perfect screening in the first period. However, if $\rho \leq 0.5$, the receiver will play an incomplete information game in the second stage game as well.

The receiver's utility in the unique equilibrium for an initial prior ρ is given by:

$$U_R = \rho\left(2a_1 - 1 + \delta\right) + \left(1-\rho\right)\left(q_1\left(2a_1 - 1 - \delta\right) + \left(1 - q_1\right)\left(1 - 2a_1\right)\right). \tag{4.3}$$

Since we are considering the case where the prior is between $\frac{1}{4}$ and $\frac{1}{2}$, we can substitute a_1 as described in table (4.1). This gives us:

$$U_R = \rho\left(\frac{2+\delta}{2} - 1 + \delta\right) + \left(1-\rho\right)\left(q_1\left(\frac{2+\delta}{2} - 1 - \delta\right) + \left(1-q_1\right)\left(1 - \frac{2+\delta}{2}\right)\right),$$

which, if we note that if $q_1 > 0$ then the sender will weakly prefer to play tail. If he is indifferent, because of the zero-sum nature of the game, then so is the receiver. We can then write the receiver's payoff as if $q_1 = 0$.

$$U_R = \rho\left(\frac{2+\delta}{2} - 1 + \delta\right) + \left(1-\rho\right)\left(\frac{2+\delta}{2} - 1 - \delta\right),$$

which simplifying gives us the following expression:

$$U_R = \left(\frac{2+\delta}{2} - 1\right) + \rho\delta - \delta\left(1-\rho\right). \tag{4.4}$$

Another way of writing (4.4) is the following:

$$U_R = \left(\frac{2+\delta}{2} - 1\right)\left(2\rho - 1\right) + \rho\delta. \tag{4.5}$$

4.3.2 Two period with commitment

In the event where the receiver can write down an enforceable contract, he could screen perfectly in the first period. This contract with perfect screening would allow him to play a complete information game in the second stage, giving him maximum payoff at that stage.

In fact, consider a very small increase in a_1 from the action in table (4.1) such that the receiver screens both types in the first period. Specifically, consider the case where the prior is $\frac{1}{2} \geq \rho > \frac{1}{4}$ and the receiver commits to an action $a_1^C = \frac{2+\delta}{4} + \varepsilon$, for some $\varepsilon > 0$, and $a_2 = 1$. Suppose also that, just as in the no commitment case, if tail is played by the sender in the first stage, then the receiver's action in the second stage is $a = 0.5$.

In this case, the strategic sender finds it profitable to play tail with probability one. I.e. $q_1^C = 0$ and there is perfect screening. The second stage game reproduces the outcome of a complete information game.

To provide an intuition for the result, we consider at first, a specific deviation, which, perhaps would be the most intuitive one. The general result, for any deviation will be shown in the end of the section.

Consider a prior in the range where the receiver screens in the first stage game when playing the no-commitment game. Suppose that in the commitment game the receiver increases the action in stage game 1 by a very small amount, just enough to break the indifference condition for the sender and induce the strategic type to play tail with probability one. Every other action remains the same. With this deviation, the receiver's utility becomes:

$$U_R = \rho \left(2\left(\frac{2+\delta}{4} + \varepsilon\right) - 1 + \delta\right) + (1-\rho)\left(1 - 2\left(\frac{2+\delta}{4} + \varepsilon\right)\right). \tag{4.6}$$

Rearranging the terms we have that:

$$U_R = \left(\frac{2+\delta}{2} - 1\right)(2\rho - 1) + 2\rho\varepsilon + (1-\rho)(-2\varepsilon) + \rho\delta. \tag{4.7}$$

Thus, the receiver's utility under commitment will be greater than his utility under equilibrium, $U_R^C \geq U_R$, if and only if

$$2\rho\varepsilon + (1-\rho)(-2\varepsilon) > 0,$$

which happens iff

$$2\rho\varepsilon - \varepsilon > 0.$$

This condition is never satisfied since it would imply that $2\rho - 1 > 0$. But, by assumption we had that $\rho < \frac{1}{2}$. Thus, for this particular deviation, the receiver's utility of commitment is never greater than the utility from playing equilibrium.

Before we can prove the general result that the equilibrium outcome weakly dominates any contract that the receiver could write, we will first show some properties that the optimal contract must have.

Proposition 4.1. ***[Necessary Conditions]***

If C is the optimal contract chosen and (C, q) is the equilibrium in Γ_2^C, then the contract C must satisfy the following properties:

I. The action played by the receiver in the second period if the sender played tail in *the first period is 0.5:*

$$a_{2(H,T)} = a_{2(T,T)} = 0.5.$$

II. *The sender weakly prefers to play tail in every contingency:*

$$U_S\left(T|h\right) \geq U_S\left(H|h\right), \textit{for any history } h.$$

III. *If the optimal contract is such that the sender strictly prefers to play tail in the first period, then in the second stage game, the action played by the receiver is one if the sender played heads previously:*

$$U\left(T|t=1\right) > U\left(H|t=1\right) \Rightarrow a_{2(H,H)} = a_{2(T,H)} = 1.$$

Proof. First note that the contract must specify five different actions–one for each different possible history. The proof is by contradiction. Let a contract C^* be an optimal one for some prior ρ.

For the claim in (I), suppose first that given this optimal contract C^*, the sender weakly prefers to play tail than to play head. Suppose also that $a_{2(H,T)}$ and $a_{2(T,T)}$ are different then 0.5. Now, consider another contract C' that is exactly the same as the optimal contract C^* except after histories in which the sender played tail in the first stage. At these histories, $a_{2(H,T)}$ and $a_{2(T,T)}$ are 0.5. In this second contract C', the payoff of the receiver is the same when the sender is a behavioral type. The change only affected the histories after which tail was played by the sender, which can happen only when the sender is strategic. If the sender is strategic and weakly prefers to play tail under C^*, he is worse-off under C', since the utility of playing tail has decreased. Given that the strategic sender is playing a zero-sum game with the receiver, if he is worse-off, the receiver must be better off. Thus, in the case where the sender weakly prefers to play tail, the optimal contract must specify an action of 0.5 after the sender plays tail. In the case where the sender strictly prefers to play head in the first stage, the histories after tail are never reached. When we change the contract to make $a_{2(H,T)}$ and $a_{2(T,T)}$ equal to 0.5 we make the incentives for playing tail even worse and this does not change the outcome of the game. In this case, the receiver's payoff under the optimal contract C^* is exactly the same as under the altered contract $C^{'}$. This leads us to conclude that we can concentrate w.l.o.g. on contracts that specify $a_{2(H,T)} = a_{2(T,T)} = 0.5$.

Now suppose that under the optimal contract C^* we have that the sender strictly prefers to play head in the first period. In this case, $q_1 = 1$. We already know that after playing tail, the sender will face a situation in which the receiver is committed to play $a = 0.5$, giving an expected payoff of zero to the players, regardless of the sender's action. If in this contract we also have that $a_1 = 1$ then, the sender's utility of playing tail is 1. His utility for playing heads is given by the utility in the first stage game, which is zero, plus the utility in the second stage game, which is at most 1. Therefore, for a contract with $a_1 = 1$ and a discount factor smaller than 1, i.e. for $\delta < 1$, it must be that the sender strictly prefers to play tail in the first stage game. Consider the case where the sender strictly prefers to play head in the first period and $a_1 < 1$. A second contract C' that specifies a higher a_1, high enough that the sender is indifferent between playing tail or heads, dominates C^*. This is true because when the sender is a behavioral type, the receiver clearly benefits from a higher initial action. When the sender is a strategic type, this increased action will make the strategic sender worse-off, thus benefiting the receiver. This proves claims (II).

For claim (III), suppose that an optimal contract is such that the sender strictly prefers to play tail in the first stage and that either $a_{2(H,H)}$ or $a_{2(T,H)}$ is smaller than 1. Then, another contract that is exactly like C^* except in the event of the sender playing head in the first period, but with a higher $a_{2(H,H)}$ and $a_{2(T,H)}$ until the sender is indifferent in the first stage, or until $a_{2(H,H)} = a_{2(T,H)} = 1$ is profitable for the receiver. If the sender is a behavioral type, the result is straightforward. If the sender is a strategic type, an increase in $a_{2(H,H)}$ or $a_{2(T,H)}$ such that the sender still weakly prefers to play tail, does not change the his payoff, and thus, does not change the receiver's payoff as well. □

We now prove our general result in this section. The result shows that the ability to commit does not improve the receiver's payoff.

Proposition 4.2. ***[Optimal contract is Sequentially Rational]***

The equilibrium in Γ_2 is the same as in Γ_2^C.

Proof. We know that we can consider only contracts that satisfy properties I, II and III in proposition 4.1. Thus, we can have only two different types of contracts. We can have a contract in which the sender strictly prefers to play tail in the first period–and the second period specifies $a_2 = 1$. In this case action in the first period must be:

$$a_1 \geq \frac{2+\delta}{4}. \tag{4.8}$$

In this contract, the receiver's objective function is to solve:

$$\max_{a_1} \rho \left\{2a_1 - 1 + \delta\right\} + (1-\rho)\left\{1 - 2a_1\right\},$$

where $a_1 \in \left[\frac{2+\delta}{4}, 1\right]$. The first order conditions are to set the action $a_1 = 1$ when $\rho > \frac{1}{2}$ and $a_1 = \frac{2+\delta}{4}$ otherwise.

The second possibility of contract must be such that the sender is indifferent between head and tail in the first stage, in this case, the action in the first period must be:

$$a_1 = \frac{2 + \delta\,(2a_2 - 1)}{4}. \tag{4.9}$$

The optimal contract is such that:

$$\max_{a_1, a_2} \rho \left\{2a_1 - 1 + \delta\,(2a_2 - 1)\right\} + (1 - \rho)\left\{1 - 2a_1\right\} \tag{4.10}$$

We can substitute (4.9) in (4.10)and we have that the receiver's problem is to solve the following equation:

$$\max_{a_1} \rho \left\{2a_1 - 1 + \delta \frac{4a_1 - 2}{\delta}\right\} + (1 - \rho)\left\{1 - 2a_1\right\}, \tag{4.11}$$

subject to $a_1 \in \left[0.5, \frac{2+\delta}{4}\right]$. The first order conditions of equation (4.11) is such that when $\rho > \frac{1}{4}$, the optimal action is $a_1 = \frac{2+\delta}{4}$ and $a_2 = 1$. When, in the other hand, the prior is $\rho < \frac{1}{4}$ then the optimal action is $a_1 = 0.5$, and, following equation (4.9), the action in the second stage is also $a_2 = 0.5$.

Therefore, when $\rho < \frac{1}{4}$ the optimal contract is to set $a_1 = a_2 = 0.5$, since this dominates setting $a_1 = \frac{2+\delta}{4}$ which is the optimal contract under the first type of contracts. For a prior on the sender's type on the interval $\frac{1}{2} \geq \rho \geq \frac{1}{4}$ we have that the optimal contract under both possible types is to set $a_1 = \frac{2+\delta}{4}$ and $a_2 = 1$. In this case, there are multiple equilibria, since any $q_1 \in [0, 1]$ and $q_2 = 0$ is an equilibrium. Finally, when $\rho > \frac{1}{2}$ we have that the

		Sender	
		Head	Tail
Receiver	Head	$1,-1$	$0,1$
	Tail	$-1,1$	$1,-1$

Figure 4.2: Modified Matching Pennies

optimal contract is the first type, which specifies $a_1 = 1$. This contract dominates $a_1 = \frac{2+\delta}{4}$, which is the optimal one under contract of type two.

These results, exactly mimic the perfect Bayesian equilibrium shown in table 4.1. □

4.4 Example when commitment is effective

We present a very simple example to show that commitment will indeed be effective when we move from the zero-sum situation. Consider the case where the game has the same types of sender as before, but with the slight alteration in the receiver's payoff as follows:

Consider this modified matching pennies for two periods only. Also, lets focus on a particular case where $\delta = 1$ and $\rho = \frac{1}{4}$. The equilibrium (with no possibility of commitment) in this case is the following. The receiver plays head with probability $a_1 = \frac{2+\delta}{4}$ and $a_2 = 1$ if the sender played head in the first period. If the sender played tail in the first period, then the action of the receiver is 0.5. The sender plays head in the first period with probability $\frac{1}{9}$ and 0 in the second period. Thus, the receiver's payoff is given by:

$$U_R = \rho\left(2a_1 - 1 + \delta\right) + (1-\rho)\left(q_1\left(2a_1 - 1 + \delta 0\right) + (1-q_1)\left(1 - 2a_1 + \delta\frac{1}{3}\right)\right) \quad (4.12)$$

$$U_R = \frac{5}{12}$$

Now consider a contract that specifies $a_1 = 1$ and also $a_2 = 1$ with the actions of both players being the same (as in the equilibrium) in the case where the first period's action of the sender is tail. Then, the receiver's payoff in this case is the following:

$$U_R = \rho (1+\delta) + (1-\rho)\left(0 + \delta \frac{1}{3}\right) \tag{4.13}$$
$$U_R = \frac{3}{4}$$

Therefore, the receiver is better off if he can write a contract.

4.5 Commitment in a finitely repeated game

In this section, we extend the results of section 4.3 for a finitely repeated game. We denote Γ_N the game described in section 4.2.1 and repeated N times. Similarly, we denote Γ_N^C the commitment game described in section 4.2.2 that lasts N periods.

The general version of proposition (4.1) is given below.

Proposition 4.3. ***[Necessary Conditions for Optimal Contracts]***

If C is the optimal contract chosen and (C,q) is the equilibrium in Γ_N^C, then the contract C must satisfy the following properties:

I. The action played by the receiver after the sender has played tail at least once 0.5:

$$a_h = 0.5, \quad \forall h \text{ that contains at least one tail played by the sender.}$$

II. *The sender weakly prefers to play tail in every contingency:*

$$U_S(T|h) \geq U_S(H|h), \textit{for any history } h.$$

III. *If the optimal contract is such that the sender strictly prefers to play tail after a particular history, then in all the subsequent periods of the game, the action played by the receiver is one if the sender played head previously:*

$$U(T|h_t) > U(H|h_t) \Rightarrow a_{h_{t+i}} = 1, \quad \forall i > 1, \forall h_{t+1} \textit{ if the sender has only played head and } \forall h_t$$

Proof. First note that the contract must specify a different action for each different possible history. The proof is very similar to the proof in (4.1). Let a contract C^* be an optimal one for some prior ρ.

Suppose that C^* is the optimal contract. and that claim I does not hold for all histories in which the sender has ever played tail. Now consider a modification in this contract such that after a contingency in which the sender played tail, the receiver now plays 0.5 until the last stage game. This modification will have no effect in the case where at that contingency the sender strictly preferred to play head. In the case where the sender weakly preferred to play tail, by decreasing his utility after playing tail, this may lead to a change on his behavior, but it will not increase his payoff (and thus, not decrease the receiver's payoff) and might decrease it, thus benefiting the receiver.

The same argument is true for claim II. Consider a contingency after which the sender strictly prefers to play head. An increase in the receiver's action until the sender's utility has decreased enough to make him indifferent between head and tail at that contingency will increase the receiver's payoff given a behavioral type and will not increase the sender's payoff.

Finally, when the sender strictly prefers to play tail, the receiver is better off by playing an action of one in the period after that. By increasing the action afterwards until the sender is indifferent or until the action is 1, the receiver increases his payoff given a behavioral type and does not increase the sender's payoff. □

Therefore, we can concentrate on contracts that satisfy the properties described in proposition (4.3). We use this fact to show that among these contracts, the best ones are the ones that mimic the perfect Bayesian equilibrium outcomes. But, before we present the result, we need a simple lemma.

Lemma 4.1. $4+\sum_{t=1}^{k^*} 2^{t+1} = 2^{k^*+2}$

Proof. Lets denote S the following sum:

$$S = 4 + 2^2 + 2^3 + ... + 2^{n-1} + 2^n$$

Then, we have that:

$$\begin{aligned} S &= 2\left(2+2+2^2+...+2^{n-2}+2^{n-1}\right) \\ S &= 2\left(S-2^n\right) \\ S &= 2S-2^{n+1} \\ S &= 2^{n+1} \end{aligned}$$

□

Proposition 4.4. ***[General optimal contract is Sequentially Rational]***

The equilibrium in Γ_N is the same as in Γ_N^C.

Proof. When selecting the optimal contract in Γ_N^C, the receiver is maximizing his ex-ante expected payoff, which is given by:

$$U_R = \rho \left(\sum_{t=1}^{N} \delta^{t-1} (2a_t - 1) \right) + (1-\rho)(-U_S),$$

where U_S is the utility that the sender has at the initial stage game. We know from proposition (4.3) that at all stages the sender weakly prefers to play tail. Thus, we can substitute U_S for the utility of playing tail in the first stage game. The objective function of the receiver is then given by:

$$U_R = \rho \left(\sum_{t=1}^{N} \delta^{t-1} (2a_t - 1) \right) + (1-\rho)(1-2a_1). \tag{4.14}$$

This is the payoff that the principal is maximizing when he chooses the contract. However, the sender is playing a best response against the chosen contract, so the receiver must maximize (4.14) subject to incentive compatibility constraints.

The receiver's problem is the following:

$$\begin{aligned} & \max_{\{a_t\}_{t=1}^{N}} \rho \left(\sum_{t=1}^{N} \delta^{t-1} (2a_t - 1) \right) + (1-\rho)(1-2a_1) \\ subject\ \ to \quad & : \\ a_t \quad & \geq \quad \frac{2+\delta}{4} \Rightarrow a_{t+1} = 1,\ \forall t \\ a_t \quad & < \quad \frac{2+\delta}{4} \Rightarrow a_{t+1} = \frac{4a_t - 2 + \delta}{2\delta},\ \forall t \end{aligned} \tag{4.15}$$

The receiver's problem can be reduced to choosing the optimal a_1 subject to the incentive compatibility constraints above. We can rewrite these constraints as follows:

$$a_t < \frac{2+\delta}{4} \Rightarrow a_{t+1} = \frac{4a_t - 2 + \delta}{2\delta} = \frac{2a_t}{\delta} + k, \ \forall t,$$

where the constant k is defined as:

$$k \equiv \frac{\delta - 2}{2\delta}.$$

Thus, inductively, we can write:

$$a_{t+2} = \frac{2a_{t+1}}{\delta} + k = \frac{4a_t}{\delta^2} + \frac{2}{\delta}k + k.$$

In general, we can write:

$$\tilde{a}_i = \frac{2^{i-1}}{\delta^{i-1}}\tilde{a}_i + k_i, \tag{4.16}$$

where the constant k_i is:

$$k_i = \sum_{s=0}^{i-2} \frac{2^s}{\delta^s} k = \frac{\delta - 2}{2\delta} \sum_{s=0}^{i-2} \frac{2^s}{\delta^s}.$$

It will be convenient to define the functions $V(x)$ and $f(x)$, such that:

$$\begin{aligned} V(x) &\equiv \delta(2f(x) - 1) + \delta V(f(x)), \\ f(x) &\equiv \begin{cases} 1 & \text{if } x \geq \frac{2+\delta}{4} \\ \frac{4x+\delta-2}{2\delta} & \text{if } x < \frac{2+\delta}{4} \end{cases}. \end{aligned}$$

Thus, the problem of the receiver, incorporating the incentive compatibility constraints, can be written as:

$$\max_a \rho\left(2a - 1 + V\left(a\right)\right) + \left(1-\rho\right)\left(1-2a\right). \tag{4.17}$$

The first order condition is:

$$\rho\left(4 + \frac{\partial V\left(a\right)}{\partial a}\right) - 2 = 0.$$

However, $V\left(\cdot\right)$ is not a continuous function. Thus, we write the solution to the problem as:

$$\rho\left(4 + \sum_{t=1}^{k^*} 2^{t+1}\right) - 2 \geq 0, \tag{4.18}$$

where k^* is defined as follows. If $a \in \left[\tilde{a}_{i+1}, \tilde{a}_i\right]$, then $k^* = i$. Note that k^* is uniquely defined for every $a \in [0.5, 1]$.

If $k^*\left(a\right)$ is such that $\rho\left(4 + \sum_{t=1}^{k^*} 2^{t+1}\right) - 2 > 0$, then the receiver can do better by offering a contract with a higher a, which implies a lower k^*.

From lemma (4.1) we have that:

$$4 + \sum_{t=1}^{k^*} 2^{t+1} = 2^{k^*+2}.$$

Thus, condition (4.18) can be written as:

$$\rho 2^{k^*+1} - 1 \geq 0,$$

which implies that when $\frac{1}{2^n} > \rho > \frac{1}{2^{n+1}}$ then the first order condition is such that:

$$2^{k^*} > 2^n \Rightarrow \rho 2^{k^*+1} - 1 \geq 0.$$

We must have that the optimal contract requires that $k^*(a) = n$. The argument goes as follows. If $k^*(a) > n$ then $\rho 2^{k^*+1} - 1 > 0$ and there is a better contract with a higher a (lower k^*). If, however, $k^*(a) < n$ then $\rho 2^{k^*+1} - 1 < 0$ and there is a better contract with a lower a (higher k^*). When, in fact, $k^*(a) = n$, we also have that $\rho 2^{k^*+1} - 1 > 0$ and the optimal contract is to increase a. This action a is increased exactly to the point where $a = \tilde{a}_n$. Thus, given (4.16), we must have that the optimal action when $\rho \in \left[\frac{1}{2^{n+1}}, \frac{1}{2^n}\right]$, is the following:

$$\begin{aligned} a &= \frac{2^{n-1}}{\delta^{n-1}} a_1 + k_n \\ a &= \frac{2^{n-1}}{\delta^{n-1}} \frac{2+\delta}{4} + \frac{\delta - 2}{2\delta} \sum_{s=0}^{n-2} \frac{2^s}{\delta^s} \end{aligned}$$

This shows that the receiver will choose a such that $k^*(a) = n$, which implies that a will be the same as the initial action in the equilibrium with sequential rationality constraints and no ability to commit. □

4.6 Conclusion

In this chapter we considered a dynamic adverse selection problem where the receiver faces two extreme types, with either opposite preferences or a mechanical behavior. We showed that the receiver's ability to commit does not improve his payoff compared to the PBE.

The result suggests that contracting can only be beneficial if preferences are not extreme, as we showed in the example of section 4.4.

Bibliography

D. Abreu, A. Rubinstein, "The Structure of Nash Equilibrium in Repeated Games with Finite Automata," *Econometrica*, 56, 1259-1281, (1988).

L. Anderlini, D. Gerardi and R. Lagunoff "A Super Folk Theorem for Dynastic Repeated Games," Georgetown University, mimeo, (2006).

R. Aumann, S. Hart, and M. Perry, "The Absent-Minded Driver," *Games and Economic Behavior* 20, 102-116, (1997).

R. J. Aumann, M. B. Maschler, and R. E. Stearns. "*Repeated Games with Incomplete Information*," MIT Press, Cambridge, (1995).

H. Bar-Isaac, "Something to Prove: Reputation in teams," doctoral thesis, University of London, (2004) .

R. Benabou, G. Laroque, "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility," *Quarterly Journal of Economics*, 921-58, (1992).

D. Bernheim, R. Thomadsen, "Memory and Anticipation," *The Economic Journal*, 115, 271-304, (2005).

P. Bolton, A. Faure-Grimaud, "Thinking Ahead: The Decision Problem," mimeo, Princeton University, (2005).

Chou, C. F. and Geanakoplos, J. "The Power of Commitment", Cowles Foundation Discussion Paper, 1988.

H. L. Cole, J. Dow, and W.B. English, "Default, Settlement and Signalling: Lending Resumption in a Reputational Model of Sovereign Debt," *International Economic Review*, 36, 365-385, (1995).

Conlisk, J. "Why Bounded Rationality?," *Journal of Economic Literature*, XXXIV June, 669-700, (1996).

M. Cripps, G. Mailath and L. Samuelson, "Imperfect Monitoring and Impermanent Reputation," *Econometrica*, March, 72.2, 407-432, (2004).

J. Dow, "Search Decisions with Limited Memory," *Review of Economic Studies*, 58, 1-14, (1991).

M. Ekmecki, "Sustainable Reputations with Rating Systems," job market paper, Princeton University, (2005).

D. Fudenberg and D. Levine, "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, July, 57.4, 759-778, (1989).

D. Fudenberg and D. Levine, "Maintaining a Reputation When Strategies Are Imperfectly Observed," *Review of Economic Studies*, July, 59.3, 561-579, (1992).

L. Frisell, J. Lagerlof, "Lobbying, Information Transmission, and Unequal Representation," *WZB Discussion Paper*, no SPII, (2005).

R. Fryer, M. Jackson, "Categorical Cognition: A Psychological Model of Categories and Identification in Decision Making," *NBER working paper*, 9579 (2003).

I. Gilboa, "A Comment on the Absent Minded Driver's Paradox," *Games and Economic Behavior*, 20, 25-30, (1997).

I. Gilboa, D. Samet, "Bounded versus Unbounded Rationality: The Tyranny of the Weak," *Games and Economic Behavior,* 1, 213-221, (1989).

J. Greenberg, "Avoiding Tax Avoidance: A (Repeated) Game Theoretic Approach," *Journal of Economic Theory*, 32, 1-13, (1984).

W. Harrington, "Enforcement Leverage when Penalties are Restricted," *Journal of Public Economics*, 37, 29-53, (1988).

S. Hart, "Non-Zero-Sum Two-Person Repeated Games with Incomplete Information," *Mathematics of Operations Research*, 10(1), 117-153, (1985).

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

M. Hellman, T.M. Cover, "On Memory Saved by Randomization," *Annals of Mathematical Statistics*, 42, 1075-1078, (1971).

B. Holmstrom, "Managerial Incentive Problems - A Dynamic Perspective," *Review of Economic Studies*, January, 66.1, 169-182, (1999).

S. Huck, R. Sarin, "Players with Limited Memory," *Contributions to Theoretical Economics*, 4, 1109-1109, (2004).

J. Isbell, "Finitary Games," in Contributions to the Theory of Games III, Princeton, NJ: Princeton Univ. Press, 79-96, (1957)

M. O. Jackson, E. Kalai, "Reputation versus Social Learning," *Journal of Economic Theory*, 88(1), 40-59, (1999).

E. Kalai and E. Lehrer, "Rational Learning Leads to Nash Equilibrium," *Econometrica*, vol. 61, No. 5, 1019-1045, (1993).

E. Kalai, A. Neme, “The Strength of a Little Perfection,” *International Journal of Game Theory*, 20, 335-55, (1992).

E. Kalai, E. Solan, “Randomization and Simplification in Dynamic Decision-Making,” *Journal of Economic Theory*, 111, 251-264, (2003).

E. Kalai, W. Stanford, "Finite Rationality and Interpersonal Complexity in Repeated Games," *Econometrica*, 56, No. 2, 387-410, (1988).

D. Kreps, P. Milgrom, J. Roberts and R. Wilson, “Rational Cooperation In The Finitely Repeated Prisoner’s Dilemma,” *Journal of Economic Theory*, 27.2, 245-252, (1982).

R. Lagunoff, A. Matsui, “Organizations and Overlapping Generations Games: Memory, Communication, and Altruism,” *Review of Economic Design*, 8, 383- 411, (2004).

E. Lehrer, “Repeated Games with Stationary Bounded Recall Strategies,” *Journal of Economic Theory*, 46, 130-144, (1988).

B. Lipman, “Information Processing and Bounded Rationality: A Survey,” *Canadian Journal of Economics*, 27, 42-67, (1995).

B. Lipman, “More Absentmindedness,” *Games and Economic Behavior*, 20, 97-101, (1997).

P. Lo, “Fuzzy Memory and Categorization,” mimeo Yale University, (2005).

G. Mailath, and L. Samuelson, “Who Wants a Good Reputation?,” *Review of Economic Studies*, 68, 415-441, (2001).

D. Monte, “Bounded Memory and Limits on Learning,” mimeo, Yale University, (2006).

S. Morris, “Political Correctness,” *Journal of Political Economy*, 109, 231-265, (2001).

S. Mullainathan, “Thinking through Categories,” working paper, MIT, (2001).

S. Mullainathan, “A Memory Based Model of Bounded Rationality,” *Quarterly Journal of Economics,* 17, 735 -774, (2002).

A. Neyman, “Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoner’s Dilemma,” *Economic Letters*, 19, 227-229, (1985).

C. H. Papadimitriou, “On players with a bounded number of states,” *Games and Economic Behavior*, 122– 131, (1992).

M. Piccione, A. Rubinstein, “On the Interpretation of Decision Problems with Imperfect Recall,” *Games and Economic Behavior* 20, 3-24, (1997).

M. Rabin, "Psychology and Economics," *Journal of Economic Literature*, Vol. XXXVI, 11-46, (1998).

R. Radner, “Team decision problems,” *Annals of Mathematical Statistics*, 33, 857-881, (1962).

R. Reis, "Inattentive Consumers," *Journal of Monetary Economics*, forthcoming, (2007).

A. Rubinstein, "Finite Automata Play the Repeated Prisoners' Dilemma," *Journal of Economic Theory*, 39, 83-96, (1986).

A. Rubinstein, *Modeling Bounded Rationality*, Zeuthen Lecture Book Series, MIT Press, Cambridge, MA (1998).

I. Sarafidis, "What Have You Done for Me Lately? Release of Information and Strategic Manipulation of Memories," *The Economic Journal*, forthcoming, (2007).

C. Sims, "Implications of Rational Inattention," *Journal of Monetary Economics*, 50, 665-690, (2003).

J. Sobel, "A Theory of Credibility," *Review of Economic Studies*, LII, 557-573, (1985).

S. Sorin, "Merging, Reputation, and Repeated Games with Incomplete Information," *Games and Economic Behavior*, 29, 274-308, (1999).

N. Stokey, and R. Lucas, "*Recursive Methods in Economic Dynamics,*" Cambridge, MA: Harvard University Press, (1998).

R. H. Strotz, "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies,* 23, 165-180, (1956).

A. Wilson, "Bounded Memory and Biases in Information Processing," Job market paper, Princeton University, (2003).