# Reputation and Bounded Memory *

Daniel Monte†
Yale University

JOB MARKET PAPER

November 21, 2006

**Abstract**

This paper is a study of bounded memory in a game. 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).

*I am very thankful to Ben Polak, Dirk Bergemann, and Stephen Morris for their constant guidance. I am also very grateful for conversations with Attila Ambrus, Eduardo Faingold, Hanming Fang, Ezra Friedman, Dino Gerardi, Itzhak Gilboa, Ehud Kalai, Pei-yu Lo, George Mailath, Wolfgang Pesendorfer, Maher Said, Lones Smith, Peter Sorensen and Joel Watson. Finally, I thank the seminar participants at the European Meetings of the Econometric Society, International Conference on Game Theory at Stony Brook, North American Meetings of the Econometric Society, Canadian Economic Theory Conference, Midwest Economic Theory, Zeuthen Workshop and at the Yale Theory Lunch.

†Homepage: http://pantheon.yale.edu/∼dm297/. E-mail: daniel.monte@yale.edu

# 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 paper 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 paper 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 paper 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 paper 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 and 3 necessary conditions for equilibria. In particular, we show that the updating rule must be weakly increasing as long as the receiver obtains 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.

Propositions 4 and 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

[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.

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 paper.

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 paper 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 Monte (2006) we consider an extension of the model studied here. In that paper 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 paper. 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 paper, 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 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

[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).

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

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 paper is organized as follows. Section 2 consists of the description of the model and the definitions of memory, strategies, as well as the equilibrium concept. The case of two memory states is shown in section 3. Section 4 gives the main result of the paper: 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 5. In section 6 we present a discussion of the incentive compatibility concept and a comparison with an automaton model. Section 7 concludes the paper. Most of the proofs are in the appendix.

## 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 $\rho$ 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 $\omega_t$ 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*.

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 $\eta$. This variable is capturing an exogenous probability that the relationship will end. We will focus on the case where this probability $\eta$ is very small so that the players expect the game to go on for a very

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

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

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 paper 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 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:

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

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

$$\varphi : \mathcal{M} \times \{T, L\} \to \Delta(\mathcal{M})\,. \tag{2}$$

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 $\varphi$.

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 $(\varphi, a)$ 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 $\sigma$ 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 $\sigma$. For each history $h$ and memory state $s_i$, let the receiver's belief be given by the relative frequency as defined below.

[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.

**Definition 1** ***(Consistency)***

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

$$\mu\left(h|s_i,\sigma\right) = \frac{f\left(h|\sigma\right)}{\sum\limits_{h' \in s_i} f\left(h'|\sigma\right)}. \tag{3}$$

Notice that denominator in expression (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 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 (3) would be $\frac{1}{\eta}$, where recall that $\eta$ 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:

$$\rho_i \equiv \Pr\left(B|s_i,\sigma\right) = \sum_{h \in s_i \cap H_B} \mu\left(h|s_i,\sigma\right). \tag{4}$$

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:

$$\pi_i = \rho_i + \left(1-\rho_i\right) q_i. \tag{5}$$

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 (4) and (5).

$$p_i^B = \frac{\rho_i}{\pi_i}. \tag{6}$$

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 (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 $\sigma$. 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 $\sigma$. Thus, a player might decide to deviate at a particular time, but he cannot trigger a sequence of deviations.

We say that a pair $(\mu, \sigma)$ 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 $(\varphi, a)$. Since the sender is unconstrained and conditions 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 $(\varphi, a)$ 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 $\sigma$ 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 $\sigma$, every memory state will have an associated expected continuation

[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".

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:

$$v_i^B = -\left(1-\pi_i\right)^2 + \left(1-\eta\right)\delta \sum_{j\in\mathcal{M}} \varphi_T\left(i,j\right) v_j^B. \tag{7}$$

The first term on the right of (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 $\varphi$. 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:

$$v_i^S = -q_i\left(1-\pi_i\right)^2 - \left(1-q_i\right)\pi_i^2 + \left(1-\eta\right)\delta\left(q_i \sum_{j\in\mathcal{M}} \varphi_T\left(i,j\right) v_j^S + \left(1-q_i\right)\sum_{j\in\mathcal{M}} \varphi_L\left(i,j\right) v_j^S\right). \tag{8}$$

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 state $i$ the specified action is different then the myopic best one. If the receiver deviates to the best current action and 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** ***(Incentive Compatibility: Transition Rule)***

*If a strategy $\sigma = (\varphi, a, q)$ is* incentive compatible*, then the transition rule $\varphi$ 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.

**Definition 3** ***(Incentive Compatible Equilibrium)***

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

# 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 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.

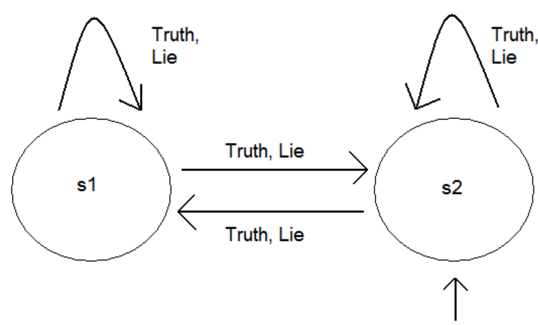


Fig. 1: Updating rule

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.

[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.

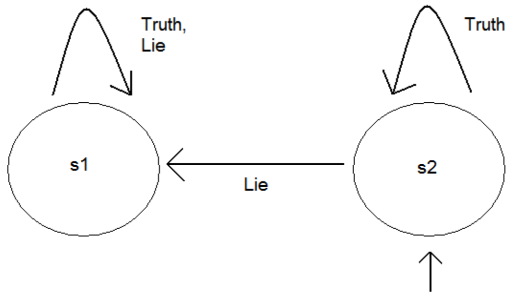


Fig. 2: Rule that Separates after a Lie

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 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 $\eta$

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 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:

$$\lim_{\eta\to 0}\pi_2 = \frac{\rho}{\eta+\rho(1-\eta)} = 1. \tag{9}$$

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.

# 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 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

[10] See Papadimitriou (1992).

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

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 paper 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).

### 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.

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.

[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.

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 $\eta$, and which allows for redundant states.

**Proposition 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 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 $\varphi_T(j, k) > 0 \Rightarrow \varphi_T(j, m) = 0$,*

   *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 $\forall\ m' < m$.*

[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 3 hold without any additional restrictions.

**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:

$$p_i^B\left(v_k^B - v_m^B\right) + p_i^S\left(v_k^S - v_m^S\right) = 0. \tag{10}$$

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

$$p_j^B\left(v_k^B - v_m^B\right) + p_j^S\left(v_k^S - v_m^S\right) = 0. \tag{11}$$

If $p_i^B \neq p_j^B$ then (10) and (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:

$$\begin{aligned} p_i^B\left(v_{k+1}^B - v_k^B\right) + p_i^S\left(v_{k+1}^S - v_k^S\right) &\geq 0, &(12)\\ p_i^B\left(v_k^B - v_{k-1}^B\right) + p_i^S\left(v_k^S - v_{k-1}^S\right) &\leq 0. &(13)\end{aligned}$$

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

$$\begin{aligned} p_j^B\left(v_{k+1}^B - v_k^B\right) + p_j^S\left(v_{k+1}^S - v_k^S\right) &\leq 0, &(14)\\ p_j^B\left(v_k^B - v_{k-1}^B\right) + p_j^S\left(v_k^S - v_{k-1}^S\right) &\geq 0. &(15)\end{aligned}$$

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:

$$p_j^B\left(v_m^B - v_{m\prime}^B\right) + p_j^S\left(v_m^S - v_{m\prime}^S\right) \geq 0. \tag{16}$$

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:

$$p_i^B\left(v_m^B - v_{m\prime}^B\right) + p_i^S\left(v_m^S - v_{m\prime}^S\right) > 0, \tag{17}$$

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 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 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 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 and 3 is depicted in figure 3 below.

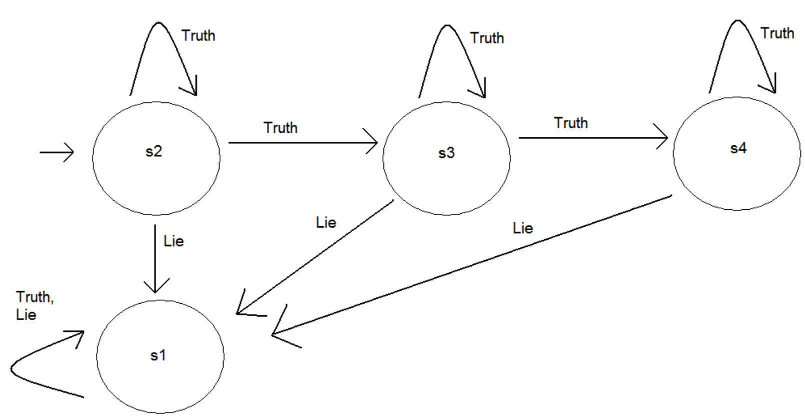


Fig. 3: A Class of Equilibrium Memory Rules

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 and 3.

## 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 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 $\rho_n^*$ 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 $\pi_{n-1}^*$, which depends on the parameters $\delta, n$ and $\eta$. 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 $\delta, n$ and $\eta$.

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

**Proposition 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 $\rho$ is higher than the threshold $\rho_n^*$, 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 6** ***(Deterministic Transition Rule: Uniqueness)***
*Fix the number of memory states $n$ and the initial prior $\rho$. There is at most one equilibrium with deterministic transition rule without redundant states.*

In fact, we show that the threshold is "almost" sufficient for ensuring that the equilibrium involves only deterministic transition rules. The result is that if the transition rule after a truth assigns only positive probability on the current memory state and on the one immediately following, after i.e., $\varphi_T(i,j) = 0$ for $\forall j \in \mathcal{M}$ with $j \neq i, i+1$, then the threshold is sufficient. Thus, we say that it is "almost" sufficient since we could still have "bad" equilibria for the receiver, involving redundant states and jumping.

The result is the following:

**Proposition 7** ***(Deterministic Transition Rule: "Almost" Sufficient Condition)*** *Given the number of memory states n, if the prior on the behavioral type of the sender is greater than the threshold for deterministic rules, $\rho > \bar{\rho}_n$, then there is no equilibrium with randomization in which $\varphi_T(i,j) = 0$ for $\forall j \in \mathcal{M}$ with $j \neq i, i+1$ and all states are non-redundant.*

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 3 would still hold.

# 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. The lowest state is equivalent to a babbling state, 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 $\pi_2$, 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.

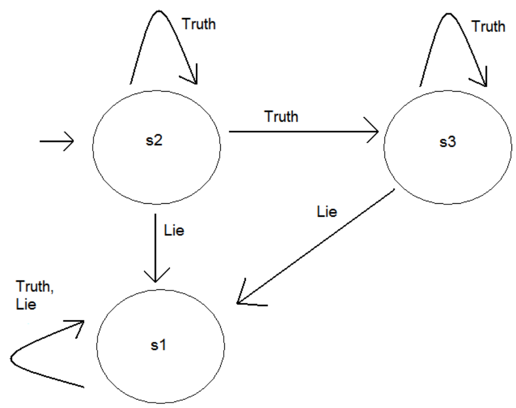

Fig. 4: Three Memory States

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:

$$\pi_2^2 + \delta\frac{1}{4} = (1-\pi_2)^2 + \delta\left(\varphi_T(2,2)\,\pi_2^2 + (1-\varphi_T(2,2))\right). \tag{18}$$

This gives us the following quadratic equation:

$$\delta\varphi_T(2,2)\,\pi_2^2 - 2\pi_2 + 1 + \delta(1-\varphi_T(2,2)) - \delta\frac{1}{4} = 0. \tag{19}$$

Solving for the belief $\pi_2$ gives us:

$$\pi_2 = \frac{1-\sqrt{1-\delta\varphi_T(2,2)\left(1+\delta(1-\varphi_T(2,2))-\delta\frac{1}{4}\right)}}{\delta\varphi_T(2,2)}. \tag{20}$$

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** ***(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 (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 (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:

$$p_2 = \frac{(\pi_2+1)(1-\delta\varphi_T(2,2))}{1-\pi_2+(\pi_2+1)(1-\delta\varphi_T(2,2))}. \tag{21}$$

Now the two conditions missing are that these beliefs $\pi_2$ and $p_2^T$ must be consistent in equilibrium, according to 3. The posterior $p_2^T$ can be written as:

$$p_2^T = \gamma_1\frac{\rho}{\rho+(1-\rho)q_1} + \gamma_2\frac{\rho_2}{\rho_2+(1-\rho_2)q_2} + \gamma_3\frac{\rho_3}{\rho_3+(1-\rho_3)q_3} + ... \tag{22}$$

where the $\gamma's$ indicate the Bayesian updating of time periods. First note that: $\rho_2+(1-\rho_2)q_2 = \frac{\rho+(1-\rho)q_1q_2}{\rho+(1-\rho)q_1}$.

In general, we will have:

$$\rho_t + (1-\rho_t)\, q_t = \frac{\rho + (1-\rho)\, q_1 q_2 ... q_t}{\rho + (1-\rho)\, q_1 ... q_{t-1}}. \tag{23}$$

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:

$$\gamma_t = \frac{(1-\eta)^{t-1} (\varphi_T (2,2))^{t-1} (\rho + (1-\rho)\, q_1 q_2 \times ... \times q_t)}{\Delta}. \tag{24}$$

Before we compute what (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, (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:

$$p_2^T = \frac{\rho}{\Delta} \frac{1}{(1-(1-\eta)\,\sigma)}. \tag{25}$$

And, for the Markovian case, the term $\Delta$ 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:

$$\Delta = \frac{\rho + (1-\rho)\, q - (1-\eta)(\varphi_T (2,2))\, q}{(1-(1-\eta)(\varphi_T (2,2)))(1-(1-\eta)(\varphi_T (2,2))\, q)}. \tag{26}$$

Thus, substituting (26) in (25) gives us the following expression for the receiver's posterior:

$$p_2^T = \frac{\rho\,(1-(1-\eta)(\varphi_T (2,2))\, q)}{\rho + (1-\rho)\, q - (1-\eta)(\varphi_T (2,2))\, q}. \tag{27}$$

Similarly, for the belief $\pi_2$ we have that:

$$\pi_2 = \begin{bmatrix} (1-(1-\eta)(\varphi_T(2,2)))(1-(1-\eta)(\varphi_T(2,2))q)+ \\ +(1-\eta)(\varphi_T(2,2))(\rho+(1-\rho)q-(1-\eta)(\varphi_T(2,2))q) \end{bmatrix}^{-1} \\ [\rho(1-(1-\eta)(\varphi_T(2,2))q)+(1-\rho)q(1-(1-\eta)(\varphi_T(2,2)))] \tag{28}$$

If the beliefs and strategies $\left(p_2^H, \pi_2, \varphi, a, q\right)$ satisfy the system of equations described by (20), (21), (27) and (28) above, then we have an equilibrium in the 3 memory state case.

# 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 paper focuses 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.

[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.

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 $\rho$ 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]

| | Bounded | Memory | | Automata | | |
|---|---|---|---|---|---|---|
| $\rho$ | $\pi_2$ | $\varphi_T(2,2)$ | $U_R$ | $\pi_2^A$ | $\varphi_T(2,2)$ | $U_R^A$ |
| 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 1: Automata: more testing than Bounded Memory

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.

# 7 Conclusion

This paper 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.

[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.

Our view is that, although forgetful, players have the ability to control what to remember and what to forget. A player might think that the 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]

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 paper 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 paper 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 (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 paper confirms this intuition in the context of

[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 memory.

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 paper 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.

# 8 Appendix

## 8.1 $n$ Memory States

This section is divided as follows. First, we show a general version for proposition 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 $\varphi$ as well as on $U_S(s_j)$ for all $j \in \mathcal{M}$.

## 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 8** ***(Increasing Updating Rule: General version of Proposition 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 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 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:

$$\begin{aligned} U_S(j) &= (1-\pi_j)^2 + (1-\eta)\,\delta \sum_{s\in\mathcal{M}} \varphi_T(j,s)\,(1-\pi_s)^2 + ... \\ &\quad + (1-\eta)^t\,\delta^t \pi_k^2 + (1-\eta)^{t+1}\,\delta^{T+1} U_S(l). \end{aligned} \tag{29}$$

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 $(1-\pi_j)^2 \le \pi_l^2$, and also that $\pi_j^2 \le \pi_n^2$. In state $s_n$ we can write the utility for the sender as:

$$U_S(L|n) = \pi_n^2 + (1-\eta)^t\,\delta^t \pi_l^2 + (1-\eta)\,\delta\pi_l^2 + ... + (1-\eta)^{t+1}\,\delta^{t+1} U_S(l), \tag{30}$$

and since we have that:

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

We can substitute in (29) and (30) to get that: $U_S(j) \le U_S(L|n), \forall j$. In particular this holds for $j = n$. ■

**Corollary 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 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:

$$\begin{aligned} v_l^S &= q_l \left\{ -(1-\pi_l)^2 + (1-\eta)\,\delta \Sigma_{j^*} \varphi_T(l, j^*)\, v_{j^*}^S \right\} + \\ &\quad + (1-q_l)\left\{ -\pi_l^2 + (1-\eta)\,\delta v_l^S \right\}. \end{aligned} \tag{31}$$

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 (31) can be written as:

$$v_l^S = -\pi_l^2 + (1-\eta)\,\delta v_l^S. \tag{32}$$

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:

$$v_1^S = -\pi_1^2 + (1-\eta)\,\delta v_l^S. \tag{33}$$

However we can compare the expected payoff on equations (32) and (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 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 $\pi_2$ such that $\pi_2 = \pi_1$ and $v_2^S = v_1^S$, we must have that $\varphi_L(i,1) = 1$.

**Corollary 2** *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 $\pi_i$, 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 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 7** *$\pi_i$ 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:

$$-\pi_i^2 + (1-\eta)\,\delta v_1^S > -\pi_j^2 + (1-\eta)\,\delta v_1^S, \tag{34}$$

but (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 8** *For states reached with positive probability, $\pi$ 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,

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

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

**Lemma 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 $\pi_i$ 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, $\varphi_T\left(i, n\right) = 1$. ■

**Lemma 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 $\varphi_T\left(n, u\right) = 1$ (since $q_n = 0$).

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

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 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 $\pi_i \leq \pi_j$.

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

$$\pi_i^2 + (1-\eta)\,\delta U_S(1) > (1-\pi_i)^2 + (1-\eta)\,\delta U_S(B). \tag{35}$$

But, we also have that:

$$\pi_j^2 + (1-\eta)\,\delta U_S(1) = (1-\pi_j)^2 + (1-\eta)\,\delta \Sigma_{j^*} \varphi_T(i, j^*)\, U_S(j^*). \tag{36}$$

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 (35) and (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 (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 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:

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

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:

$$v_i^S = -\left(1 - \pi_i\right)^2 + \sum_{i^*} \varphi^T(i, i^*)\, v_{i^*}^S. \tag{37}$$

The expected continuation payoff of moving to state $s_i$ after observing the truth in state $s_j$ can be written using (37) as:

$$p_j^B v_i^B + \left(1 - p_j^B\right) v_i^S = -\left(1 - \pi_i\right)^2 + \sum_{i^*} \varphi^T(i, i^*) \left(p_j^B v_{i^*}^B + \left(1 - p_j^B\right) v_{i^*}^S\right). \tag{38}$$

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:

$$p_j^B v_j^B + \left(1 - p_j^B\right) v_j^S = -\left(1 - \pi_j\right)^2 + \sum_{j^*} \varphi_T(j, j^*) \left(p_j^B v_{j^*}^B + \left(1 - p_j^B\right) v_{j^*}^S\right). \tag{39}$$

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

$$p_j^B v_{j^*}^B + \left(1 - p_j^B\right) v_{j^*}^S \geq p_j^B v_{i^*}^B + \left(1 - p_j^B\right) v_{i^*}^S. \tag{40}$$

Using (40) in (38) and (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 13** *The receiver always starts either at the lowest memory state or at the lowest after the babbling state:*

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

**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 (8), we have that $\rho < p_j^B, \forall j > 1$. Thus, if $\varphi_0(i') > 0$,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 8. To relate this proposition with the one presented in the text, we need two additional results:

**Lemma 14** *The beliefs are extreme:*

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

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

$$
p_l^B = \sum_{h\in s_l \cap H_B} \mu\left((h,T)\,|s_l\right) \tag{41}
$$

However, given the results on 8.1 and 8.3 from proposition 8 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 $\eta \to 0$, $\Pr(h_1|s_l) \to 0$ where $h_1$ means that the time period is 1 and therefore $\Pr(B|s_l) \to 0$. By incentive compatibility it will then imply that $\varphi_T(l,l) = 1$ 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 $U_S(L|u) > U_S(T|u), \forall u \in \mathcal{U}$ 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 $\pi_i$ and $\pi_j$ ,$\forall i,j$ 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 $\pi$ 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 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 $(\pi_i)$ 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. ■

### 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 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:

$$\pi_{n-1}=\frac{1}{2}+(1-\eta)\delta\frac{3}{8}. \tag{42}$$

Thus, we need to find the lower bound on prior or, equivalently, the highest $q$ that can support $(42)$. 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:

$$\rho_{n-1}\geq\frac{\pi_{n-1}+\pi_{n-1}^2}{2}. \tag{43}$$

For any $\rho_{n-1}$ that is smaller than the threshold above, we need more $q$ to induce the $\pi$ needed for (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 (43) then we need a lower $q$ and (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 (42) and (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 (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 (42) holds with equality. More on this appears later. We use the equation:

$$\pi_{n-2} = \frac{1}{2} + (1-\eta)\frac{\delta}{2}\left(\pi_{n-1}^2 - \frac{1}{4}\right), \tag{44}$$

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:

$$\rho_{n-2} \geq \left(\frac{\pi_{n-1}+\pi_{n-1}^2}{2}\right)\pi_{n-2}. \tag{45}$$

If $\rho_{n-2}$ is smaller than in equation (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 $i \leq n-2$ :

$$\rho_i \geq \frac{\pi_{n-1}+\pi_{n-1}^2}{2}\prod_{k=i}^{n-2}\pi_k. \tag{46}$$

■

**Corollary 4** *As the number of memory states increase $n \to \infty$, the threshold computed in (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 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.**

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

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

To prove proposition 5 in the text, we show two lemmas.

**Lemma 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.

$$\Pi_{eq} = -\rho_i \left( \sum_{k=i}^{n} (1-\pi_i)^2 \right) - (1-\rho_i) \left\{ q_i \left( (1-\pi_i)^2 + \delta U_S(i+1) \right) + (1-q_i) \left( \pi_i^2 + \delta \frac{1}{4} \frac{1}{1-\delta} \right) \right\}. \tag{47}$$

We want an appropriate way to write (47) so that we can compare with the payoff from a deviation. Note that we can write $\rho_i + (1-\rho_i) q_i q_{i+1} = \pi_{i+1}\pi_i$, $\rho_i + (1-\rho_i) q_i q_{i+1} q_{i+2} = \pi_{i+2}\pi_{i+1}\pi_i$, and so on. However, $(1-\rho_i) q_i (1-q_{i+1}) = (1-\pi_{i+1})\pi_i$; $(1-\rho_i) q_i q_{i+1} (1-q_{i+2}) = (1-\pi_{i+2})\pi_{i+1}\pi_i$ and so on. We can then write (47) as:

$$\begin{aligned}
\Pi_{eq} &= -\pi_i(1-\pi_i) - \delta \left( \pi_i \pi_{i+1} (1-\pi_{i+1}) + (1-\pi_i) \frac{1}{4} \frac{1}{1-\delta} \right) - \\
&\quad -\delta^2 \left( \pi_i \pi_{i+1} \pi_{i+2} (1-\pi_{i+2}) + \pi_i (1-\pi_{i+1}) \frac{1}{4} \frac{1}{1-\delta} \right) + \ldots
\end{aligned} \tag{48}$$

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 (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 $q$s 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 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 (47), and again we can write as in equation (48). We can further change the $q$ and write (48) with $q_{n-s} = 0$,instead. This change in $q_{n-s}$ will not change the value of $\Pi_{eq}$ 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:

$$\begin{aligned} \Pi_{dev} &= -\rho_i \left(\sum_{k=i+s}^{n} (1-\pi_k)^2\right) - (1-\rho_i) \qquad (49) \\ &\left\{q_i^{dev} \left((1-\pi_{i+s})^2 + \delta U_S(i+s+1)\right) + \left(1-q_i^{dev}\right) \left(\pi_{i+s}^2 + \delta \frac{1}{4}\frac{1}{1-\delta}\right)\right\}. \end{aligned}$$

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 (49) can be written as:

$$\begin{aligned} \Pi_{dev} &= -\left[\pi_i (1-\pi_{i+1})^2 + \pi_i \pi_{i+1}^2\right] - \\ &\quad -\delta\left(\pi_i\left[\pi_{i+1}(1-\pi_{i+2})^2 + (1-\pi_{i+1})\pi_{i+2}^2\right] + (1-\pi_i)\frac{1}{4}\frac{1}{1-\delta}\right) - \\ &\quad -\delta^2\left(\pi_i\pi_{i+1}\left[\pi_{i+2}(1-\pi_{i+3})^2 + (1-\pi_{i+2})\pi_{i+3}^2\right] + \pi_i(1-\pi_{i+1})\frac{1}{4}\frac{1}{1-\delta}\right) + ... \end{aligned} \tag{50}$$

We now want to compare the payoffs in (48) but with $q_{n-1} = 0$ and (50) *period by period.* Note that according to lemma (9) we have that the payoff in (48) is greater than the payoff in (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 6.** Let $\pi$ and $\pi'$ 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.o.l.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 $\pi'_{n-1} \geq p_{n-2}'^B$. ■

## References


[1] D. Abreu, A. Rubinstein, "The Structure of Nash Equilibrium in Repeated Games with Finite Automata," *Econometrica*, 56, 1259-1281, (1988).

[2] L. Anderlini, D. Gerardi and R. Lagunoff "A Super Folk Theorem for Dynastic Repeated Games," Georgetown University, mimeo, (2006).

[3] R. Aumann, S. Hart, and M. Perry, "The Absent-Minded Driver," *Games and Economic Behavior* 20, 102-116, (1997).

[4] R. J. Aumann, M. B. Maschler, and R. E. Stearns. "*Repeated Games with Incomplete Information*," MIT Press, Cambridge, (1995).

[5] R. Benabou, G. Laroque, "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility," *Quarterly Journal of Economics*, 921-58, (1992).

[6] D. Bernheim, R. Thomadsen, "Memory and Anticipation," *The Economic Journal*, 115, 271-304, (2005).

[7] P. Bolton, A. Faure-Grimaud, "Thinking Ahead: The Decision Problem," mimeo, Princeton University, (2005).

[8] Conlisk, J. "Why Bounded Rationality?," *Journal of Economic Literature*, XXXIV June, 669-700, (1996).

[9] M. Cripps, G. Mailath and L. Samuelson, "Imperfect Monitoring and Impermanent Reputation," *Econometrica*, March, 72.2, 407-432, (2004).

[10] J. Dow, "Search Decisions with Limited Memory," *Review of Economic Studies*, 58, 1-14, (1991).

[11] L. Frisell, J. Lagerlof, "Lobbying, Information Transmission, and Unequal Representation," *WZB Discussion Paper*, no SPII, (2005).

[12] R. Fryer, M. Jackson, "Categorical Cognition: A Psychological Model of Categories and Identification in Decision Making," *NBER working paper*, 9579 (2003).

[13] I. Gilboa, "A Comment on the Absent Minded Driver's Paradox," *Games and Economic Behavior*, 20, 25-30, (1997).

[14] I. Gilboa, D. Samet, “Bounded versus Unbounded Rationality: The Tyranny of the Weak,” *Games and Economic Behavior,* 1, 213-221, (1989).

[15] J. Greenberg, “Avoiding Tax Avoidance: A (Repeated) Game Theoretic Approach,” *Journal of Economic Theory*, 32, 1-13, (1984).

[16] W. Harrington, “Enforcement Leverage when Penalties are Restricted,” *Journal of Public Economics*, 37, 29-53, (1988).

[17] M. Hellman, T.M. Cover, “Learning with Finite Memory,” *Annals of Mathematical Statistics*, 41, 765-782, (1970).

[18] M. Hellman, T.M. Cover, “On Memory Saved by Randomization,” *Annals of Mathematical Statistics*, 42, 1075-1078, (1971).

[19] S. Huck, R. Sarin, “Players with Limited Memory,” *Contributions to Theoretical Economics*, 4, 1109-1109, (2004).

[20] J. Isbell, “Finitary Games,” in Contributions to the Theory of Games III, Princeton, NJ: Princeton Univ. Press, 79-96, (1957)

[21] M. O. Jackson, E. Kalai, “Reputation versus Social Learning,” *Journal of Economic Theory*, 88(1), 40-59, (1999).

[22] E. Kalai, A. Neme, “The Strength of a Little Perfection,” *International Journal of Game Theory*, 20, 335-55, (1992).

[23] E. Kalai, E. Solan, “Randomization and Simplification in Dynamic Decision-Making,” *Journal of Economic Theory*, 111, 251-264, (2003).

[24] E. Kalai, W. Stanford, "Finite Rationality and Interpersonal Complexity in Repeated Games," *Econometrica*, 56, No. 2, 387-410, (1988).

[25] R. Lagunoff, A. Matsui, “Organizations and Overlapping Generations Games: Memory, Communication, and Altruism,” *Review of Economic Design*, 8, 383- 411, (2004).

[26] E. Lehrer, “Repeated Games with Stationary Bounded Recall Strategies,” *Journal of Economic Theory*, 46, 130-144, (1988).

[27] B. Lipman, “Information Processing and Bounded Rationality: A Survey,” *Canadian Journal of Economics*, 27, 42-67, (1995).

[28] B. Lipman, “More Absentmindedness,” *Games and Economic Behavior*, 20, 97-101, (1997).

[29] P. Lo, “Fuzzy Memory and Categorization,” mimeo Yale University, (2005).

[30] D. Monte, “Bounded Memory and Limits on Learning,” mimeo, Yale University, (2006).

[31] S. Morris, “Political Correctness,” *Journal of Political Economy*, 109, 231-265, (2001).

[32] S. Mullainathan, “Thinking through Categories,” working paper, MIT, (2001).

[33] S. Mullainathan, “A Memory Based Model of Bounded Rationality,” *Quarterly Journal of Economics,* 17, 735 -774, (2002).

[34] A. Neyman, “Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoner’s Dilemma,” *Economic Letters*, 19, 227-229, (1985).

[35] C. H. Papadimitriou, “On players with a bounded number of states,” *Games and Economic Behavior*, 122– 131, (1992).

[36] M. Piccione, A. Rubinstein, “On the Interpretation of Decision Problems with Imperfect Recall,” *Games and Economic Behavior* 20, 3-24, (1997).

[37] M. Rabin, "Psychology and Economics," *Journal of Economic Literature*, Vol. XXXVI, 11-46, (1998).

[38] R. Radner, “Team decision problems,” *Annals of Mathematical Statistics*, 33, 857-881, (1962).

[39] R. Reis, “Inattentive Consumers,” *Journal of Monetary Economics*, forthcoming, (2007).

[40] A. Rubinstein, “Finite Automata Play the Repeated Prisoners’ Dilemma,” *Journal of Economic Theory*, 39, 83-96, (1986).

[41] A. Rubinstein, *Modeling Bounded Rationality*, Zeuthen Lecture Book Series, MIT Press, Cambridge, MA (1998).

[42] I. Sarafidis, “What Have You Done for Me Lately? Release of Information and Strategic Manipulation of Memories,” *The Economic Journal*, forthcoming, (2007).

[43] C. Sims, "Implications of Rational Inattention," *Journal of Monetary Economics*, 50, 665-690, (2003).

[44] J. Sobel, "A Theory of Credibility," *Review of Economic Studies*, LII, 557-573, (1985).

[45] R. H. Strotz, "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies,* 23, 165-180, (1956).

[46] A. Wilson, "Bounded Memory and Biases in Information Processing," Job market paper, Princeton University, (2003).