Reputation and Bounded Memory

Daniel Monte
Yale University
Department of Economics
E-mail: daniel.monte@yale.edu

Preliminary Draft

March 18, 2006

Abstract

This paper is a study of bounded memory in a reputation game. In particular, in a repeated cheap talk game with incomplete information on the sender's type. The receiver is assumed to be constrained by a finite number of memory states and the memory rule is itself part of his strategy. The first result of this paper shows that in this reputation game the updating rule will be rather simple: monotonic and increasing. The second main result of this paper shows that when memory constraints are severe the updating rule will involve randomization before reaching the extreme states. The key intuition is that in a two-player game with incomplete information randomization is used as a memory saving device and also as a strategic element: to test the opponent and give incentives for types to be revealed early in the game. The results in this paper extend to general reputation games where the normal type is a zero-sum player and the commitment type is playing a pure strategy.

JEL Classification Codes: C72, D82, D83.
KEYWORDS: bounded memory, reputation, repeated games, cheap-talk.

1 Introduction

The behavioral constraints of an individual might prevent him from realizing the history of a game and from updating according to Bayes' rule. Memory constraint is an example of such cognitive restrictions and this paper studies the role of bounded memory in a strategic context.

The study of the implications of an imperfect memory has taken two different modelling strategies in the literature. One approach is to make explicit assumptions on the way that memory works while assuming that the agent is not aware of these limitations. The goal is then to look for the implications of this particular memory process in different contexts. This memory process could be, for example, bounded recall, where the agent is only able to record the information of the last k periods, or it could be based on memory decay, such as described by Mullainathan (2002) and Sarafidis (2001).

The other approach in modeling memory restrictions is to assume constraints on the agent's memory, but such that the agent is fully aware of these limitations and can decide on the optimal strategy given this constraint. The memory rule itself becomes part of the player's strategy. The absentminded driver model of Piccione and Rubinstein (1997) is a representative paper on this second view. Particularly important papers also include Dow (1991) and Wilson (2004).

I take the latter approach in this paper. Memory is assumed to be a finite set of states and the player has to decide on a transition rule and on a map from states to actions. At every point in time all the player's memory about the history of the game must be contained in one of the states. The information in the current period is known by the player, but he is forgetful between periods. Memory can be interpreted as a collection of agents where at each stage game a different agent plays the game, and their payoff takes into account the payoffs of all future agents as well. The equilibrium in the bounded memory model is the equilibrium between this collection of agents and the opponent. This way, memory is a metaphor for an organization where the workers don't communicate perfectly, but through a limited number of messages. Given this equilibrium concept, the bounded memory model differs substantially from an automata. In the latter, the ex-ante optimal solution might not be an interim equilibrium as is required in the former. I discuss this issue further in section 5.

Bounded memory as modeled in this way has been studied in single player games only. The results are well understood and briefly discussed below, but the effects of this particular memory restriction in multi-player games is still an open problem[1]. In this paper I tackle this question in the context of a repeated cheap talk game with incomplete information.

Wilson (2004) was the first paper in economics to model the player's memory as a finite set of states[2]. She had an extensive game in which a decision maker receives signals about the true state of the world and has to make a decision after receiving 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 Wilson's paper for a binary signal case is that the transition rule is random only at the extreme states, which implies that the decision maker will ignore information only when he is as convinced as his finite memory allows. A similar problem had been studied before by Hellman and Cover (1970). Their model describes the best way (ex-ante optimal) to keep track of the signals received for the two-hypothesis testing problem using a finite automaton. Again, the transition rules are random at the extreme states and the interpretation is that, perhaps counter intuitively, randomization is used as a memory saving device.

In this paper the analysis is of a two-player game, where one of the players (the receiver) exhibits the same type of memory constraint as Wilson's decision maker, while the other player (the sender) is unconstrained.

The unbounded sender will condition his strategy on the entire history of the game, but the bounded memory receiver will choose an action rule given his memory constraint and how to update the beliefs after verifying the true state of the world. This model aims to capture the heuristics on how agents keep track of information in this context. I describe the neces-

[1]There are papers on multi-player games with bounded recall (Lehrer (1988) and Huck and Sarin (2004), for example), which is different from bounded memory. While the updating rule is the crucial aspect of the present study, in the bounded recall literature the memory rule is exogenously given.

[2]Mullainathan (2001) and Fryer and Jackson (2003) also study models where agents are restricted to hold a finite set of posteriors. However, in their models the updating rule (categorization) is given exogenously, and is not part of the player's strategy, it falls in the "first approach" in modeling memory that I discussed before.

sary conditions for this updating rule to be an equilibrium and show some properties of the equilibrium strategies. In particular I present necessary and sufficient conditions on the parameters for the bounded memory player to play a pure strategy and thus mimic a Bayesian updating rule, with no loss of information.

In other words, if the prior on the opponent's type is sufficiently high, then memory will not be binding and the bounded memory player uses a pure strategy: his updating rule mimics Bayes rule.

There is yet another threshold on the prior such that receiver will still use deterministic transition rules whenever the prior falls short on the "Bayesian threshold", but is above this second cutoff. In this case, there will be information loss only in the extreme state. The receiver is fulled by his memory in the exact moment that his belief is the highest one.

When the prior is low relatively to the size of the player's memory (prior lower than the second threshold discussed above), such that the player will require a long sequence of signals before being convinced about the opponent's type, then the receiver will use random transition rules in the initial states. In this last case, it will involve testing before updating. Knowing that it is very likely that the sender is of a bad type, and that his memory will not allow him to store all the information before being sufficiently convinced, the receiver will choose to ignore good signals for part of the time, while waiting to catch the sender whenever he lies.

This result suggests that 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), and Wilson (2004)[3]. However, unlike these single player models, here there is also a strategic reason for random transition rules: test the opponent and give incentives for types to be revealed early in the game.

The repeated cheap talk game studied in this paper is taken from Sobel (1985). At every stage game of Sobel's model the sender observes the true state of the world in that period and sends a message to the receiver. The receiver then takes an action and the payoffs are realized. Payoffs depend only on the true state of the world in the current period and on the action taken by the receiver. At this point 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 belief (Bayesian updating in Sobel's paper) on the sender's type.

[3]Kalai and Solan (2003) also emphasize the role of randomization when agents are not fully rational (bounded complexity).

The results of this game extend to any finite game where the behavioral type is committed to playing a pure strategy and the normal type (or the strategic type) is playing a zero-sum game with the bounded memory player. I chose this reputation game to study the role of memory because there are several papers studying the issues of bounded memory in a single player context, but none in a game setting. This would better be done in a game in which memory plays a central role, and a reputation game critically depends on the memory of the uninformed player (receiver). Thus, by constraining the receiver's memory I can give a sharp characterization of the implications on beliefs and updating rules.

The reason why the repeated cheap talk model was chosen among the class of reputation games with zero-sum normal types is the fact that there are several possible applications to this particular model. One interpretation is in the context of a policy maker that listens to a different informed advisor in every different area. Each advisor can be one of two types. He could be an honest type, a loyal employee that will always inform the policy maker about the true state of the world (a behavioral type), or he could be an advisor with a secret interest in becoming the next election's running up candidate (a strategic type). This strategic type of advisor gains whenever the policy maker is worse off. Since the policy maker has to deal with a huge amount of information for different matters, it becomes hard to keep track of all the signals that the advisors have send. I describe a rule of thumb on how this policy maker will keep track of the information received given his memory imperfection[4].

The study of bounded memory in a repeated game with incomplete information is such that the information structure of the restricted player has two dimensions: a belief about the opponent's type and a belief about the current time period. In the model studied in this paper the decision of which of these two aspects of information is to be stored is very clear on the player's decision. I show that for a minimal memory, the time period dimension is completely ignored and the memory is used only to keep track of the opponent's type. It is as if the receiver started the game with long run beliefs. This result remains as we increase the memory size of the player: the long run beliefs of the receiver in the Bayesian version of this game will always be the same as the extreme beliefs of the bounded memory player. However,

[4]Several interesting applications of repeated cheap-talk games can be found in Sobel (1985), Benabou and Laroque (1992), Frisell and Lagerlof (2005), and Morris (2001).

as memory increases there will be extra storage space to partially keep track of time, and updating becomes more frequent. The information structure in this paper is asymmetric and one of the signals (a lie) will always be remembered, even for a minimal memory, but for small memories there is no storage capacity to keep all the signals, thus the other signal (truth) will often be ignored.

The paper is organized as follows. The basic structure of the model and the definitions of memory and strategies in this game are in section 2 where I also present a discussion about the equilibrium concept and the assumptions of the model. The case with two memory states is shown in section 3, where I compare my results with the equilibrium in the Bayesian world. I start with two periods and then extend to an infinite horizon. Section 4 shows the condition for the receiver to have deterministic transition rules in an N state memory and the properties of the memory rule when this condition is not met. In section 5 there is a discussion of the incentive compatibility concept and a comparison with the issues in a single player game. Also there is a comparison of this paper with an automata model in a reputation context. Section 6 concludes the paper. Most of the proofs are in the appendix.

2 Model

This is a repeated cheap talk game with incomplete information on the sender's type. The model is based on Sobel (1985). The sender can be one of two possible types. With probability ρ the sender is a behavioral type that always tells the truth. This behavioral type will be denoted H as in "honest type". Nothing will be said about the preferences or strategy of this behavioral type of sender, it will just be assumed that at every period of the game the message send by this type is equal to the true state of the world. With probability $(1-\rho)$ the sender is a "strategic type" S, with utility opposite to the receiver's, as will be described below.[5]

The timing of every stage game is the following. Nature draws a state of the world in the period, $\omega_t \in \Omega = \{0, 1\}$ each happening with probability $\frac{1}{2}$. The sender observes ω_t and sends a message $m_t \in \{0, 1\}$ to the receiver. This message has no direct influence on the player's payoffs, thus a cheap talk game. The receiver then observes this message and takes an action. After he takes the action, the payoffs are realized and the states are verified. At this

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

point, the receiver can tell whether the sender has lied to him or not about the state of the world. Based on this information, the receiver updates his belief on the sender's type.

For every stage game the payoff of the receiver is a quadratic loss function: $U_R = -(a_t - \omega_t)^2$ while the utility of the strategic sender is $U_S = (a_t - \omega_t)^2$. With this particular functional form of utility the receiver has a unique best response for every message send, the action rule will be uniquely determined given the message. The use of a different functional form for the utility would require randomization in the equilibrium actions. This would take the focus out of our main concern, which is to characterize the updating rule, rather than the action rule.

2.1 Memory and Strategies

I define the memory of the receiver to be a finite set where each element corresponds to a different memory state. Let M be the memory of the receiver. Thus, $M = \{1, 2, ..., N\}$ means that the receiver is constrained to N memory states.

The transition rule is a map from the memory and the signal (true or lie) received, to a subset of the memory. In other words, the transition rule $\sigma : M \times \{T, L\} \to \Delta(M)$ is the receiver's updating rule. Let s_i represent an element from the set M.

It is also part of the receiver's strategy to decide at time $t = 0$, before the first period starts, at which memory set he will start the game. Let g_0 be this initial distribution over the memory states: $g_0 \in \Delta(M)$.

Finally, the receiver's strategy also includes an action rule, a map from memory and messages to actions: $a : M \times \{0, 1\} \to \mathbb{R}$. I call the tuple (σ, a, g_0) the memory rule of the receiver.

To define the sender's strategy I need two additional assumptions. First, assume that at every period of the game the sender knows in which memory state the receiver is. This assumption will simplify the analysis. To drop this assumption, I would have to condition the sender's strategy on a subset of the receiver's memory according to the equilibrium transition rule rather than on the exact memory state that the receiver is in. Note, however, that because of the action taken by the receiver, the sender would be able to infer the memory state that he was in the previous period. His beliefs would then be restricted to the states that have positive probability of being reached by the equilibrium strategy. For the purpose of this paper, assume that there is a mechanism through which the sender would be able to find out the exact memory state that he is in. Although it seems a strong assumption, it will be innocuous in the deterministic cases. In the cases of non-deterministic transition rules, the transition rule would be different under this assumption, but the shape of the memory rule, which is one of the insights of this paper will be the same regardless of this assumption. This assumption is also innocuous in the two memory state case described next section.

The second assumption is that the sender's strategy will be symmetric with respect to the state of the world, i.e. after any history in the game we have that $\Pr(m_i = 1|\omega_i = 1) = \Pr(m_i = 0|\omega_i = 0)$. Together these two assumptions imply that the strategy of the sender is $q : H \times M \to [0, 1]$, where H is the set of all possible histories. Under this symmetry assumption, the payoffs are also symmetric with respect to the state of the world: $U_R(m_i = 1|\omega_i = 1) = U_R(m_i = 0|\omega_i = 0)$.

2.2 Incentive Compatibility and Games with Imperfect Recall

To find the equilibrium we use the notion of incentive compatibility as described by Piccione and Rubinstein (1997)[6] and Wilson (2004). A strategy is incentive compatible if an agent cannot gain by one shot deviation 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.

As usual in games with imperfect recall[7], the equilibrium strategy might not be time consistent in the usual sense: in the middle of the game the player might find it in his best interest to revise his entire strategy. I take the view that in games with imperfect recall the player is not allowed to change his entire strategy, since he cannot control his future selves' actions. The crucial assumption here is that the player will not recall his deviation, thus he cannot revise his entire strategy, but only make local deviations. When deciding on

[6]Piccione and Rubinstein (1997) refer to this condition as "modified multiself consistency".

[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.

an action to take, the interim player assumes that his future selves will play the equilibrium strategy regardless of his current decision. For a further discussion of imperfect recall, time consistency and incentive compatibility, see Aumann et al (1997), Gilboa (1997) and Piccione and Rubinstein (1997).

The equilibrium in this model is defined as a strategy that is a best response for the unbounded player given the equilibrium strategy of the bounded memory player and accordingly this strategy of the bounded memory player must be incentive compatible given the opponent's actions. In other words the strategies of both players must form an equilibrium between the unbounded player and all the interim selves of the boundedly rational player.

The result of Piccione and Rubinstein that in single player games with no discounting ex-ante optimality of a strategy implies incentive compatibility (Piccione and Rubinstein (1997), Proposition 3) will be of no use in the context of a two player game. Their result shows that in single player games if memory is modeled as we suggest in this paper, then the optimal solution ex-ante will be the same as the interim equilibrium. This is the reason why the automata described in Hellman and Cover (1970) is the same as Wilson's (2004) decision maker. In this paper, however, assuming that an automata is playing the game and disconsidering incentive compatibility constraints is the same as assuming that the receiver can credibly commit to actions, which could change the sender's behavior. I take the view that the equilibrium must be incentive compatible, and a strategy profile that is a Nash equilibrium ex-ante but not for the interim players is not an equilibrium in this game, because it would involve credible commitment on the part of the bounded memory player. I present some results and a further discussion on this topic in section 5.

In games with imperfect recall it must be made clear the assumptions of the modeler. There might be multiple equilibria even in one person games, and whereas the ex-ante decision maker will coordinate his actions in the most profitable equilibrium is something that should be made explicit by the modeler. In a two player game these issues are also present and typically there are multiple equilibria. I will take the view that in this paper there are compelling reasons to assume that the receiver can coordinate on the most profitable equilibrium, since the memory rule will describe the agents' heuristics on updating beliefs, and coordinating on a second best would rule out any evolutionary argument. Thus, one way to think about this problem is of a mechanism design, where the principal is the ex-ante player and the agents are the unbounded opponent and all the interim selves (each self correspond to a decision node) 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

The goal of this section is to show through the simplest possible case how beliefs are updated in this world with finite memory. The intuition obtained in this section will carry on to more general memories.

3.1 Two Periods

In this section I study the case where the game described above is played for two periods. I want to compare the results of a game in which the receiver is constrained to two memory states only with the results of a game where the receiver performs Bayesian updating. For this two period case, consider no discounting on the periods.

The proposition below shows the equilibrium when the receiver is updating his beliefs on the sender's type using Bayes rule. For the proof of this proposition see Sobel (1985).[8]

Proposition 1 *For each prior ρ there is a unique equilibrium in the two period game with a Bayesian receiver. Moreover, for any $\rho > \frac{1}{4}$ the equilibrium is such that in the initial period the receiver believes the sender with probability greater than $\frac{1}{2}$.*

The equilibrium is the following. If the probability of a behavioral type is very high, $\rho \geq \frac{7}{8}$, the receiver will have a very high belief that the truth will be told in the first period and therefore it will be very costly for the strategic sender to build a reputation. This implies that the best response for the sender is to lie on the first period, $q_1 = 0$.

If the prior on the sender's type is instead very low, $\rho \leq \frac{1}{4}$, then the receiver believes with high probability that he is dealing with a strategic type. Therefore the only possible equilibrium is babbling in both periods.

[8] In fact, Sobel shows that this result holds more generally and not only for the two period case.

The equilibrium with reputational concerns occurs for intermediate priors, in particular for $\rho \in \left[\frac{1}{4}, \frac{7}{8}\right]$. In this case, the proposition above shows that there will also be a unique equilibrium. On the first period the sender will be mixing with some probability $q > 0$. This implies that the probability that the receiver thinks that the message is true is $\pi_1 \equiv \Pr(T|t=1) = \rho + (1-\rho)q$, from now on let T denote true and L denote lie. On the second period after having observed a true signal, the sender will lie with probability 1, $q_{2T} = 0$, since it will be the last period and there will be no reputational concern in the last period. The belief in the second period following a true signal in the first period is the proportion of behavior type in the last period, i.e. $\pi_{2T} = \frac{\rho}{\rho+(1-\rho)q}$. Finally, on the second period after observing a lie, babbling is the only possible outcome ($q_{2L} = \frac{1}{2}$, and $\pi_{2L} = \frac{1}{2}$).

Note that if I had a two period game and a receiver with bounded memory but restricted to three memory sets, I would be able to exactly reproduce the equilibrium described above. To see this, denote the initial state by state 2, and attach the belief $\rho + (1-\rho)q$ to that state. To the state labeled 3, attach the belief $\frac{\rho}{\rho+(1-\rho)q}$ and finally state number 1 is associated to babbling. Formally, if the memory of the receiver is $M = \{1,2,3\}$ then consider the following transition function: $\sigma^T(1,1) = 1, \sigma^T(2,3) = 1$, and $\sigma^T(3,3) = 1$ together with $\sigma^L(i,1) = 1, \forall i \in M$. This memory, with this transition rule together with $g_0(2) = 1$ (i.e. the probability of starting the game at memory state 2 is equal to 1) will reproduce the equilibrium above.

In fact, the result above holds for any finite or infinite horizon game: a bounded memory player is always capable to reproduce the Bayesian equilibrium, as long as he is given enough memory states.

Lemma 1 *For any $\langle \Gamma, (H,S), (\rho, 1-\rho), (U_S, U_R) \rangle$ $\exists$ M , $|M| < \infty$, such that the equilibrium is identical to the one with a Bayesian receiver (no memory constraints).*

Proof. If Γ_T is finite, with T periods, then let $|M| = N+1$, with deterministic transition rules. The argument then goes exactly as described for the two state case. For an infinite horizon game Γ_∞, there always exist T^* such that $\pi_{T^*} = 1$ (Sobel (1985)). At this point, both types are revealed and the game is over. If $|M| = T^* + 1$, with deterministic transition rules we have the same equilibrium as in the Bayesian case. ■

Thus, the interesting cases arrive when the player is not only bounded in his memory, but also has a "short" memory, $|M| < T^* + 1$. This will be discussed further in section 4.

In this section I restrict attention to the two memory states and two periods case, since this is the simplest setting where you depart from the Bayesian updating. Denote this game with only two periods and allowing the receiver to use only two memory states by Γ_2.

The updating rule in this context is the probability of switching out of the initial state after receiving a true signal and another probability after receiving a false signal, as depicted in figure 1.

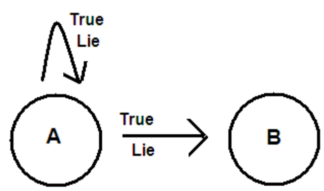

Fig. 1: Updating Rule

There are two questions in this setting: which are the possible equilibria in this game? Also, among these equilibria, which one gives the receiver a higher ex-ante expected payoff?

The first issue in this game before I can answer the questions above is how to compute the beliefs on the memory states. Given the equilibrium strategies, I can compute the beliefs in each state and the posterior on the sender's type. Let $M = \{A, B\}$ and denote π_i as the probability of truth given that the receiver is in memory state i. In this case we have that $\pi_A \equiv \Pr(T|A)$ and $\pi_B \equiv \Pr(T|B)$.

For the posterior on the sender's type, denote p_i^H as the probability that the receiver believes that the sender is a behavioral type after verifying if the sender lied or told the truth in state i $\left(p_A^H \equiv \Pr(H|T, A)\right)$. Note that since there is no noise on the sender's information structure, whenever the receiver observes a lie, he can be sure that the sender is a strategic type.

To compute the beliefs in each memory set, the agent uses Bayesian updating *given* his beliefs on the information set, which is the same approach as in Piccione and Rubinstein (1997). The intuition here is to think of the probabilities of time periods as long run frequencies.

Consider as an example the case where the transition rule is: $\sigma^T(A, A) = 1$ and $\sigma^L(A, B) = 1$. This transition rule is depicted in figure 2. If this is the transition rule, the best response for the sender is to lie right away $(q_A = 0)$ since he will not be able to build reputation by telling the truth.

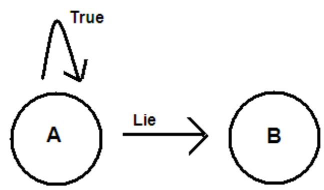

Fig. 2: Separate the Liars

Since the transition rule is completely separating the liars in the second period, whenever the receiver reaches memory set B he can be sure that he is dealing with a strategic type of sender, thus the only possible belief in that state is the one associated to babbling: $\pi_B = \frac{1}{2}$.

To compute the belief in memory state A we have to compute the probabilities of the time periods.

$$\pi_A = \Pr(t = 1|A)\Pr(T|t = 1, A) + \Pr(t = 2|A)\Pr(T|t = 2, A) \qquad (1)$$

We can think of these probabilities as long run frequencies. First note that $\Pr(t = 2|A) = \rho \Pr(t = 1|A)$ but we also have that the probabilities of the time periods should sum to 1. Solving for these two equations we get that: $\Pr(t = 1|A) = \frac{1}{1+\rho}$ and $\Pr(t = 1|B) = \frac{\rho}{1+\rho}$. Moreover, the probability of truth in the initial period is the proportion of behavioral types $\Pr(T|t = 1, A) = \rho$, while in the second period, we have $\Pr(T|t = 2, A) = 1$, which means that if the receiver reaches the second period and is still in state 1, then it must mean that he is dealing with the behavioral type of sender, therefore he may expect to receive a true signal with probability one. Thus, equation (1) gives us $\pi_A = \frac{2\rho}{1+\rho}$.

The equilibrium in this game is a strategy for the sender that is a best response for him, together with an incentive compatible memory rule (σ, a, g_0). Where incentive compatibility means that the interim receivers find it in their best interest to follow the memory rule given the sender's strategy and given that his future selves are also following it (as discussed in section 2.2).

There are three classes of equilibria in this two state game with deterministic initial state. Below we show the definition of incentive compatibility for the transition rule σ in this game Γ_2. Define $p_i^H \equiv \Pr(H|T, i, t = 1)$ and similarly $p_i^S \equiv \Pr(S|T, i, t = 1)$, then we have the following definition for an incentive compatible strategy.

Definition 1 *(Incentive Compatibility in two-period games)*
A transition rule σ in Γ_2 is said to be incentive compatible if:

$$\begin{aligned}\sigma^T(i,j) &> 0 \implies -p_i^H(1-\pi_j)^2 - p_i^S\pi_j^2 \geq -p_i^H(1-\pi_{j'})^2 - p_i^S\pi_{j'}^2 \\ \sigma^L(i,j) &> 0 \implies -\pi_j^2 \geq -\pi_{j'}^2\end{aligned}$$

For $\forall i, j, j' \in M = \{A, B\}$.

If, in equilibrium $\pi_A = \pi_B$ then a deterministic transition rule $\sigma^T(A,B) = \sigma^L(A,B) = 1$ can support this equilibrium. This result extends to a more general case with an $N-$ state memory: there will always be a trivial equilibrium in which the states have equal beliefs. In this trivial equilibrium, the receiver is in fact wasting memory sets. The more interesting cases come when we look at memory rules where the receiver is not wasting memory states. One rationale for focusing in memories without redundant states is that I believe that memories with useless memory states would likely have been ruled out by evolution and competition. Memory as studied here is finite, so the bounded memory player will choose a memory rule that does not waste more of his already scarce resources. Next section I show a result that allows us to concentrate on memories without identical states.

Proposition 2 *The equilibrium rules in Γ_2 when $\pi_A \neq \pi_B$ involve only deterministic transition rules.*

For $\rho \leq \frac{1}{3}$,babbling in both states is the only possible equilibrium. Where babbling is characterized by a belief of $\frac{1}{2}$ in both states and with the strategic sender telling the truth with probability $\frac{1}{2}$ in both periods.

In equilibrium with two memory states and two periods there are no random transition rules. In fact this will be true for the infinite horizon case as well. The interesting property of the equilibrium in which the receiver starts at the highest belief (which corresponds to the transition rule depicted in figure 2) is that the receiver keeps track of the liars. Worth noting that 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 A, so he doesn't know if he has already separated all the liars or not. This inflates the belief in state A and gives the sender a high payoff in the initial period.

3.2 Infinite Horizon

To proceed in the analysis and study the infinite horizon game we make a slight modification on the game. I assume that at every period there is an exogenous probability η that the game will end. This assumption will be helpful since as we depart from the finite horizon case, we need well defined priors over the time periods. This death rate will give us this distribution. The analysis here is to the case where $\eta \rightarrow 0$, so that the players still expect the game to continue for a very long period before it is over. Assume also a discount factor δ on the time periods.

The first result in the paper is that the receiver will hold the "extreme" beliefs, that completely separate the types.

Proposition 3 *For the two memory state game and infinite horizon, the unique non trivial equilibrium is such that:* $\pi_1 = \frac{1}{2}$ *and* $\lim\limits_{\eta \rightarrow 0} \pi_2 = 1$.

Proof. $\pi_2 = \Pr(t = 1|s_2)\rho + \Pr(t = 2|s_2) + \Pr(t = 3|s_2)\ldots$ Where the probabilities of time periods are given by:

$$\Pr(t = 1|s_2) = \frac{1}{1 + (1-\eta)\rho + (1-\eta)^2\rho + \ldots} = \frac{\eta}{\eta + (1-\eta)\rho}$$

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

$$\lim_{\eta \rightarrow 0} \pi_2 = \frac{\rho}{\eta + \rho(1-\eta)} = 1 \qquad (2)$$

■

The interpretation is that having a very short memory, the receiver will start the game with "long run beliefs". The result above follows from the fact that in an infinite horizon game, the only way that the message received in state 2 can be a false one is if it comes from a strategic sender, and this can only be true if the period is the first one, since the babbling state is absorbing. However, the probability of being at period 1 goes to zero as the exogenous death rate $\eta \rightarrow 0$.

The equilibrium is such that the receiver starts at state 2 if his prior is high enough and at state 1 otherwise and the transition rules are deterministic: $\sigma^L(i,1) = 1$, $\sigma^T(1,1) = 1$, and $\sigma^T(2,1) = 0$.

From now on the results in the paper refer only to the infinite horizon case. The assumption on the exogenous death rate creates a stationary environment for the players, such that the expected continuation payoffs at every memory state, given the sender's type, are the same regardless of the time period.

4 N Memory States

To analyze the general setting with N memory sates and infinite horizon I need extra notation to proceed in the analysis. The expected continuation payoff for the receiver at memory state i given that the sender is a behavioral type is denoted by v_i^H, where v_i^H is equal to:

$$v_i^H = -\left(1-\pi_i\right)^2 + \left(1-\eta\right)\Sigma_{j^*}\sigma^T\left(i,j^*\right)v_{j^*}^H$$

The expected continuation payoff for the receiver given a strategic sender is denoted by v_i^S. Since we don't know the behavior of the strategic sender, the only thing that we can tell is that $v_i^S = -U_S\left(i\right)$. Where $U_S\left(i\right)$ is the expected continuation payoff for the strategic sender and $U_S\left(i\right) = \max\left\{U_S\left(T|i\right), U_S\left(L|i\right)\right\}$.

In this section I define incentive compatibility for a general memory with N memory states. It is a generalization of the definition presented on the previous section.

Definition 2 *Incentive Compatibility*

The memory rule is incentive compatible if for $\forall i, j$ *we have that:*

$$\begin{aligned}\sigma^T\left(i,j\right) &> 0 \implies p_i^H v_j^H + p_i^S v_j^S \geq p_i^H v_{j'}^H + p_i^S v_{j'}^S, \forall j' \\ \sigma^L\left(i,j\right) &> 0 \implies v_j^S \geq v_{j'}^S, \forall j'\end{aligned}$$

The action rule of the receiver is:

$$\max_a -P\left(\omega=1|m\right)\left(a-1\right)^2 - P\left(\omega=0|m\right)a^2$$

$$a_i\left(m\right) = P\left(\omega=1|m\right)$$

Equilibrium in this game is defined as an incentive compatible memory rule for the receiver together with a strategy for the sender that is a best response for him.

Definition 3 *Equilibrium*

An equilibrium is a tuple $(\sigma, a, g_0; q)$ *that satisfies the following conditions: For every state* i *and history* h, q_{ih} *is such that*[9]*:*

$$\begin{aligned} q_{ih} &< 1 \Longrightarrow U_S(L|i) \geq U_S(T|i) \\ q_{ih} &> 0 \Longrightarrow U_S(T|i) \geq U_S(L|i) \end{aligned} \qquad \text{(E1)}$$

The memory rule (σ, a, g_0) *is incentive compatible, given* q. *(E2)*

4.1 Deterministic Transition Rules

When the receiver's memory is not binding, we say that the receiver is updating his belief's using Bayes rule, with no bias whatsoever. In this case, the receiver is using deterministic transition rules. There are cases, though, in which the transition rule is deterministic, but the updating differs from Bayesian because of the last state. The memory of the receiver will confuse him in this extreme state and there will be biases in information processing in that point. In this section I show the conditions on the parameters under which the receiver will play pure strategy, i.e. will use deterministic transition rules (the algorithm uses the same reasoning whether you want to compute the threshold for Bayesian updating or for deterministic rules only).

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

Lemma 2 *Fix* N *and* ρ. *There is at most one equilibrium in pure strategies for the receiver without redundant states.*

However, we need to investigate whether such an equilibrium will exist at all, and also under what conditions. Thus, I present two results. 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 the threshold the receiver will not play pure strategy. This is a necessary condition for the receiver to 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, saying that if the sender is using a best response and the transition rules are not random,

[9] We omit the subscript h from the sender's utility since for all i, $U_S(\cdot|i, h)$ is the same, regardless of the history h.

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 an infinite number of memory states.

Lemma 3 *Given a memory size N, there exists a threshold ρ_N^* such that if the prior is smaller than this threshold: $\rho < \rho_N^*$ then there is no equilibrium in pure strategy.*

The proof of this lemma is by induction. 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 N-1 is at least as high as some threshold π_{N-1}^*, which depends on the parameters δ, N and η. If the belief is lower than this threshold, the sender will prefer to tell the truth and be updated with probability one to the highest state. Moreover, by incentive compatibility there is a lower bound on the posterior state $N-1$, i.e. 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 for an additional period. Together, this implies that there is a lower bound on the prior on the sender's type at that stage game. However, the prior on state $N-1$ is the posterior of state $N-2$. 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 I show how to compute this lower bound given the parameters δ, N and η.

Next lemma shows a sufficient condition for deterministic transition rules.

Lemma 4 *Let the transition rules be deterministic: $\sigma^T(i, i+1) = 1$, the posterior on the sender's type be computed as: $p_i^H = \frac{\rho_{i-1}}{\pi_{i-1}}$ and the strategy for the sender be a best response for him. Then it will be incentive compatible for the receiver to move only one state after a true signal:*

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

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

4.2 Non Deterministic Transition Rules

When the condition on the threshold described on the previous section is not met, there are no equilibria with deterministic transition rules anymore (besides the trivial one) and randomization is needed.

Before I present the main result of the section, the lemma below shows the implications on beliefs of a game that is arbitrarily long. I will be interested in the cases where the exogenous death rate $\eta \to 0$[10], and I first show the following results:

Lemma 5 *As $\eta \to 0$, we have that:*

1) $\pi_1 = \frac{1}{2}$, moreover state 1 is absorbing.

2) $\pi_N = 1$.

The intuition for the results above is straightforward. At state N there are no reputation incentives and the bad type of sender will lie right away, thus, the only chance of having a strategic sender in this state is if the receiver reaches the state for the first time. However, if the sender is an honest type, the state is absorbing. Thus, the probability of being there for the first time goes to zero as the death rate goes to zero. Also, the same happens at state 1. If it is not the initial state, then only the strategic type of sender reaches that state, in which case the result is obvious. If it is the initial state, the probability of having a strategic sender at that state goes to one as the death rate goes to zero[11].

Proposition 5 below is the main result of this paper. As I have pointed out, there are multiple equilibria in this game, in particular, there are many rules in which the receiver has redundant states. In the appendix I wrote a variation of the proposition below to allow for these bad equilibria, but the most intuitive way to understand the proposition is to have in mind a rule without the redundant states, i.e. with N different memory states (holding different beliefs in equilibrium). As we will see later in this section, there are compelling reasons to focus only on the equilibria that gives the receiver the highest payoff, and in this case lemma 6 below shows that we can ignore the redundant states without loss of generality. The result on redundant states tells us that any equilibrium in which the receiver is using

[10] Again, the proposition in the appendix holds for any η, and not only for when $\eta \to 0$.

[11] Since this state is absorbing, in equilibrium it will not be the initial state.

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 6 *If a receiver has memory M with N states and (σ, a, g_0, q) gives the receiver a payoff U_R^* and is such that $\pi_i = \pi_j$, then there $\exists$ $(\sigma, a, g_0, q)'$ for memory $M'$with $N-1$ states and that gives the receiver utility U_R^*.*

From now on, label the states in non-decreasing order. I.e. if, in equilibrium, the beliefs of the receiver on the sender's messages are $\pi_j > \pi_i$ then label the states $j > i$.

The main result is stated below, ignoring the redundant states and assuming that $\eta \to 0$.

Proposition 4 *If (σ, a, g_0, q) is an equilibrium, then:*

1) After Lie: $\sigma^L(j, 1) = 1$ (there is always a dumping state)

2) If $U_S(L|i) > U_S(T|i) \implies \sigma^T(i, N) = 1$

3) After True: $\pi_j > \pi_i \implies \sigma^T(j, i) = 0$ (don't go back after a True signal)

4) $g_0(2) = 1$

The proposition above shows that any memory rule, in equilibrium, has to be such that the receiver separates the liars after a lie is observed. This state is absorbing, which means that the receiver will not forget if one has ever lied to him.

Another result is that while the receiver might ignore true signals, by not updating after receiving them, he will never update to a worse belief after a true. 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.

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.

However, we still want to understand how this bounded memory receiver updates after true signals. the lemma below tells us part of the story. All the results in this direction depend on a condition that the posteriors on the sender's type are different on the states. To weaken this restriction, in the appendix we prove the following lemma: $\pi_j > \pi_i \implies p_j^H \geq p_i^H$. Thus, we only have to worry about the cases where $\pi_j > \pi_i$ but $p_j^H = p_i^H$.[12]

Lemma 7 *Consider only memory states where* $p_i^H \neq p_j^H$

1) Single crossing

$\sigma^T(i,k) > 0$,$\sigma^T(i,m) > 0$ *and* $\sigma^T(j,k) > 0 \implies \sigma^T(j,m) = 0.$
$\forall k, m$ *such that* $\pi_k \neq \pi_m$ *and* $\forall i, j$ *such that* $p_i^H \neq p_j^H$

2) No Jumps

$\sigma^T(i,k-1) > 0$,$\sigma^T(i,k+1) > 0 \implies \sigma^T(j,k) = 0$ $\forall i, j$ *such that* $p_i^H \neq p_j^H$

3) Monotonicity

If $\sigma^T(i,m) > 0 \implies \sigma^T(j,m') = 0$ $(\forall\ m' < m), \forall i, j$ *such that* $p_j^H > p_i^H$

5 Related Models

5.1 Incentive Compatibility and the Absent-minded Driver Revisited

We are back to the well known model of Piccione and Rubinstein. The story is the following. A driver leaves the bar late at night, after a few drinks and little consciousness. He hits the road in the hope of getting back home, which is in the second exit of the road. However, in the first exit there is a bad outcome and there is no coming back. If he never exits, he reaches a safe hotel, which is better than the first exit, but worse than home. Whenever he reaches a point where he can either exit or continue, he cannot recall whether he has faced the exact same decision minutes ago or not. What shall he do?

In this problem the driver has a very severe memory constraint: he will not remember if he is in the first or second point in the road. In other words,

[12] It can be easily shown that for $N \leq 4$ $\pi_j > \pi_i \Rightarrow p_j^H > p_i^H$, in which case the properties in lemma 6 hold without any additional restrictions.

he has only one memory state. His choice is confined to figure out an action rule in his unique memory state.

Assume that the driver is incapable of designing a strategy, but only of checking whether it is a best response or not. There is a friendly bartender that will choose the best strategy for the driver and will tell him what to do before he leaves the bar[13]. If the bartender chooses optimally, it will turn out that the driver will not have any incentive to deviate from it in either of the nodes. This result is formally proved in Piccione and Rubinstein (1997).

Now lets add a twist. Consider another situation where this drunk driver faces a very long road. In this road there are two exits: a good one in one side of the road and a bad one on the other side. After every mile the drunk driver receives a signal about the true side of the good exit, the signal might be a sign posted by the city patrol, a tree that he can vaguely recall being in the true side, or any other sign. The driver can't recall all the signals (that would be impossible after the drinks!), but again, the benevolent bartender has taught him a rule of thumb on how to keep track of these signals.

In the model described above the driver has several memory states and decides on an updating rule as well as on an action rule. He will recall the bartender's instructions at every point in the road[14], and will always find it in his best interest to follow the suggested rule of thumb, both in the bar and at the road.

Lets change the environment once again. The drunk driver will leave the bar with what he thinks is his wife. The woman next to him is not drunk at all and knows the way home perfectly well. The way home is very long, and includes infinitely many turns. Assume that every correct turn increases the driver's payoff, whereas a wrong turn decreases it. Assume also that the initial turns have a higher value for both the driver and the woman. At every mile the woman will tell him where to turn.

Before leaving the bar the, now worried, bartender warns the drunk men that the woman might be his wife or his evil mistress. The mistress wants to force the driver out of his way home. The bartender will now give the driver a rule of thumb on how to keep track of the woman's suggestions (given that keeping track of all of them is not an option after the drinks). The driver is drunk, but like in the other cases above, he can check the optimality of the

[13]One can eliminate the bartender and assume instead that the driver decided himself on a rule before he started drinking (knowing that in the way back he would be drunk).

[14]Think of this as a rule of thumb on updating beliefs that he has adopted in many other similar situations.

rule of thumb at every point. If at some point he believes that the strategy is not 'optimal anymore' he will deviate.

The bartender knows this and will give a suggestion that the driver will always want to obey. If the woman tells him the wrong path, he will dump her and keep driving alone, with no clue about the way back. The bartender's suggestion if the driver is very drunk (memory constraints are severe) will be to ignore that the woman has told the truth for part of the time. This will give the evil mistress additional incentives to lie faster, and get some benefit from it in the beginning of the road. At some point, he will start thinking of her as his true wife, but if his memory is very restricted, chances are that what he thinks is his true wife is in fact his evil mistress.

This story of the absentminded driver illustrates the differences between the models and the results on incentive compatibility for different settings. There are two key points here. The first difference is in the memory of the driver in the three games. In the first model, with the two exits, the driver is constrained to one memory set only, thus, updating rule didn't make sense in that context. The driver did not decide on the transition rule, but only on the action rule. On the other two models, memory is finite but with more than one state. Thus, the memory rule involved not only an action rule, but also a transition between states.

The second key difference between the three examples refers to the ex-ante choice of the memory rule. In both the first and second model, the memory rule is chosen optimally by the ex-ante player (the bartender), but with no conflict between what the bartender chooses and what the driver will decide to do at every decision node given the bartender's choice.

In the third model, which is the one studied in this paper, the bartender's suggestion will be constrained to a subset of all possible 'rules of thumb'. There will be rules in which the drunk driver would rather deviate in the middle. The bartender takes that into account and chooses the best rule for the driver, given this additional constraint.

To build a bridge between my model and the automata literature, consider a forth example of this story. We take the third example and add a further constraint on the driver: that he is so drunk that he cannot check for best responses, thus he will simply follow the bartender's suggested 'rule of thumb'. In this case, we wouldn't have to worry about incentive compatibility issues, instead we would have an automata model as will be described below. Worth mentioning that with this additional constraint the driver can in fact perform better than the bounded memory player. Bounded rationality is playing the role of credible threat as in Gilboa (1988).

5.2 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. It has been used in economics mostly to capture bounded rationality in implementing a strategy.

It was studied in the context of finite memory by Hellman and Cover (1970). 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 inexistent: Piccione and Rubinstein (1997) show that the ex-ante optimal strategy will also be incentive compatible.

In this paper there are two reasons of why an automaton could differ from a bounded memory player. The first one is the same as in a single player game with discounting. An automaton would allow an individual to commit to actions and avoid 'temptations' of his future selves to deviate. The second reason is the ability to commit against an opponent, in much the same way as a Cournot oligopolist would benefit from committing to a level of output. Thus, modeling the player's memory as an automaton would require a further assumption. Namely that the player can credibly commit to his strategy[15]. This could be the case of an institution where different workers control the memory states and are obliged to comply with a set of rules, without calculating the best response at that point in time.

I take the view that both approaches have their own interest, but this paper is only about the case where incentive compatibility is indeed an issue. I show that in some situations the automaton can do better than the bounded memory player, while in others they do just as well (obviously, the 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, I show some results for the three state case, where the automaton does better than the bounded memory player.

[15]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. Further research is needed to understand the role of commitment in this particular reputation game.

To find the results for a game with a player restricted to 3 memory states only I use proposition (5) where we have that: the belief in the lowest state is equivalent to the babbling state $\pi_1 = \frac{1}{2}$, moreover this lowest state is absorbing. Also, the belief in the highest state is one ($\pi_3 = 1$) and finally, the receiver will start at the intermediate state $g(2) = 1$. Remains to calculate π_2 and the transition rules, $\sigma^T(2,2)$ and $\sigma^T(2,3)$, as well as the strategy for the sender. In order to compute this equilibrium, lets focus on Markov equilibrium only, i.e. when the strategy of the sender depends only on the current memory state $q : M \to [0,1]$. The picture below shows the equilibrium rule, and table 1 shows the results for different parameter values.

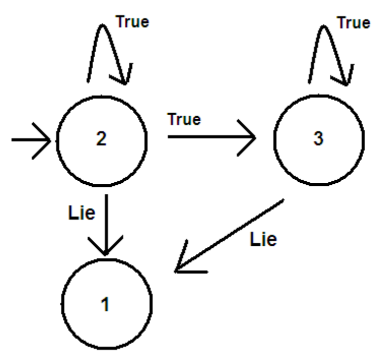

Fig. 3: Equilibrium for 3 States

One thing to note on table (I) below is that the lower the prior on the sender's type, the higher $\sigma^T(2,2)$ which means that the receiver will test more the sender. These are the equilibria for which the receiver is mixing on his updating rule. If ρ is very high (in this case the threshold is 0.72), there will be no randomization. All these equilibria were computed for $\eta = 10^{-60}$ and $\delta = 0.8$.

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

	Bounded	Memory		Automata			
ρ	π_2	$\sigma^T(2,2)$	U_R	π_2^A	$\sigma^T(2,2)$	U_R^A	
0.1	0.5	1	-1.25	0.5313	0.93588	-1.2415	
0.2	0.6135	0.7474	-1.1755	0.5919	0.80077	-1.1730	
0.3	0.6534	0.6394	-1.0727	0.6385	0.68141	-1.0716	
0.4	0.6892	0.5275	-0.95186	0.6791	0.56073	-0.9514	(I)
0.5	0.7232	0.4023	-0.81802	0.7168	0.42747	-0.8178	
0.6	0.7571	0.2512	-0.67359	0.7540	0.26638	-0.6735	
0.7	0.7925	0.0502	-0.51982	0.7920	0.05320	-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	

6 Conclusion

This paper is a study of bounded memory in a reputation game. It differs from the existing literature on imperfect memory by considering the effects of bounded memory in a game in which the memory rule is chosen by the player. The equilibrium is such that the memory process must be a best response for all the interim selves of the bounded memory player. This incentive compatibility constraint was not yet studied in multi-player games.

I showed that the updating rule is rather simple and will always be monotonic and increasing. In particular, for any (finite) memory there is a range of priors such that the bounded memory player will do as well as if using Bayes' rule (memory is not binding). There is yet another range for which the bounded memory player will keep using deterministic transition rules, but will suffer loss (as compared to a Bayesian player) on the extreme state, when he gets confused about the time period.

The second important contribution of this paper is to show the updating rule when memory constraints are severe. In these cases the receiver will not use pure strategy anymore and will, instead, use random transition rules in the initial states. Despite the multiplicity of equilibrium that games with bounded memory have, there are necessary conditions on the updating rule that suggests 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 is used for two different reasons. First 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.

In a broader sense, this paper is part of an emerging literature in restricted capacity to deal with information. Agents fail to use Bayes rule due to some constraint on their technology. This departure from Bayes rule could be due to a cost on updating new information (Reis (2005)), on a restriction to acquire new information (Sims (2003)), on a cost to think through the implications of a particular action (Bolton and Faure-Grimald (2005)), or on memory constraints. In a repeated interaction, this ability to sort information is very important due to 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 leads to inertia and inattention. Due to a restricted capacity in dealing with information, agents cannot execute Bayes rule and will choose the information to memorize, and to acquire. In other words, the agents will sort the information received and ignore part of it. This paper confirms this intuition in the context of a two player game, showing that the agents will ignore information and update only sporadically when their memory is constrained.

It is still not clear what are the implications of bounded memory in sustaining cooperation in repeated interactions. In the model presented, the strategic sender and the receiver had opposite preferences, and the zero-sum nature of this relation didn't leave any room for cooperation when the bad type of sender was caught. 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.

An important lesson of this model is that the incentive compatibility issues are very important in two player games. If they are not considered, the bounded memory player becomes an automaton, which is equivalent to a receiver with commitment power. It is a natural and interesting extension to understand how these models work on the presence of credible commitment, and I show part of this story in this paper: automata will in general do better than the bounded memory player.

The application considered in this paper was one of an uninformed player receiving signals from an informed expert. 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.

Finally, this paper has focused on a problem of bounded memory in a game where the opponent is unbounded. The next step is to study models of bounded memory when both players are constrained. This way the memory of one player depends on the payoffs and on the types, but also on the opponent's memory. One player's memory rule will depend on the other player's memory rule. In other words, the way agents form beliefs in this world is endogenous to the model. Beliefs are determined by payoffs in a way that is not captured by fully rational models.

7 Appendix

7.1 Two Memory States and Two periods

In this section we prove propositions 2 in the paper. It shows that the only possible equilibria in the two state two period game where the receiver does not waste memory states are the ones that involve no mixing in the transition rule.

Proposition 2 *If $\pi_A \neq \pi_B$, then the only rules that are incentive compatible are the ones that involve no mixing.*

Proof. To prove this lemma, first note that since there are only two periods, and both states are such that the belief is greater or equal than 0.5, the strategic sender will lie when he reaches period 2 (or will be indifferent between lying and telling the truth, if $\pi = \frac{1}{2}$).

If in equilibrium $\pi_i > \pi_j \implies \sigma_{iFj} = 1$. This comes directly from the definition of incentive compatibility. Assume that the receiver starts deterministically in state A and consider the two cases separately, first when $\pi_A > \pi_B$ and then when $\pi_A < \pi_B$.

First note that if in equilibrium $\pi_A > \pi_B$ then it must be that $q_1 = 0$. Since the utilities of the sender given their actions are given by:

$$\begin{aligned} U_S(LL) &= \pi_A^2 + \sigma^L(A,A)\,\pi_A^2 + \sigma^L(A,B)\,\pi_B^2 \\ U_S(TL) &= (1-\pi_A)^2 + \sigma^T(A,A)\,\pi_A^2 + \sigma^T(A,B)\,\pi_B^2 \end{aligned}$$

But, $\pi_A^2 \geq \sigma^T(A,A)\,\pi_A^2 + \sigma^T(A,B)\,\pi_B^2$ and $\sigma^L(A,A)\,\pi_A^2 + \sigma^L(A,B)\,\pi_B^2 \geq (1-\pi_A)^2$. Thus, $U_S(LL) \geq U_S(TL)$ and therefore, $q_1 = 0$. Thus, $p_A^H = 1$ which, in turn, implies by incentive compatibility that $\sigma^T(A,A) = 1$.

Finally, we want to show that if in equilibrium $\pi_B > \pi_A \implies \sigma^T(A,B) = 1$.

$p_A^H = \frac{\rho}{\rho+(1-\rho)q_1}$ and $\pi_B = p_A^H$, since only the behavioral type will tell the truth in the second period when the state is informative (has a belief higher than $\frac{1}{2}$). Thus:

Suppose the memory rule is incentive compatible meaning that:

$$\sigma^T(A,A) > 0 \implies -p_A^H(1-\pi_A)^2 - p_A^S\pi_A^2 \geq -p_A^H(1-\pi_B)^2 - p_A^S\pi_B^2$$

Using the fact that $\pi_B = p_A^H$, we have that:

$$-\pi_B(1-\pi_A)^2 - (1-\pi_B)\pi_A^2 \geq -\pi_B(1-\pi_B)^2 - (1-\pi_B)\pi_B^2 = -\pi_B(1-\pi_B)$$

Which in turn implies that $-\pi_B + 2\pi_B\pi_A - \pi_B\pi_A^2 - \pi_A^2 + \pi_B\pi_A^2 \geq -\pi_B + \pi_B^2$. Finally, we have that:

$$2\pi_B\pi_A - \pi_A^2 \geq \pi_B^2 \implies \pi_B^2 + \pi_A^2 - 2\pi_B\pi_A \leq 0 \implies (\pi_B - \pi_A)^2 \leq 0 \quad (3)$$

Which is a contradiction.

Therefore, if the memory rule is incentive compatible, then it must be that $\sigma^T(A,A) = 0$. ■

7.2 N Memory States

This section is divided as follows. First, I show a general version for proposition 4 in the text. This theorem is true regardless if the transition rule is deterministic (in which case it is trivially true) or not. Then I proceed to show in which cases the receiver will use deterministic transition rules. The reason that I have switched the order that was presented in the text is that I will use the results of proposition 4 to show the case of deterministic rules.

In proposition 4 the equilibrium set is restricted to a set of memory rules, and the deterministic rules will have to satisfy the properties shown in the proposition.

7.3 Random Transition Rules

In this section I prove a general version of proposition 4[16]. Define l as the state with highest expected continuation payoff if the receiver is facing a strategic sender. Formally: $D \equiv \{l \in M | v_l^S \geq v_i^S, \forall i \in M\}$, similarly define: $U \equiv \{u \in M | v_u^H \geq v_i^H, \forall i \in M\}$.

[16] In this proposition I make no use of lemma 5, that was used in proposition 4.

Proposition 5 *If* (σ, a, g_0, q) *is an equilibrium, then:*

1) *After Lie:* $\sigma^L(j, l') = 0$ *where* $l' \notin \{l | \pi_l = \min_i \pi_i\}$ *(there is always a dumping state)*

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

3) *After True:* $\pi_j > \pi_i \Longrightarrow \sigma^T(j, i) = 0$ *(don't go back after a True signal)*

4) $g_0(i) = 0, \forall \pi_i > \pi_{(2)}$

The first result comes from incentive compatibility. If $\Pr(H|i, F) = 0$, $\forall i$,we must have that after a lie, the receiver moves to a state with 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

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

Thus, the payoff of the sender after lying is:

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

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

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

Lemma 8 $j \notin D \Longrightarrow \sigma^L(i, j) = 0, \forall i \in M.$

Proof. By incentive compatibility, $\sigma^L(i, j) > 0 \Longrightarrow v_j^S \geq v_{j'}^S, \forall j'$ ■

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 to the sender than telling the truth in a state with belief higher than $\frac{1}{2}$.

Lemma 9 *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.

$$\begin{aligned} U_S(L|i) &= \pi_N^2 + (1-\eta)\,\delta U_S(l) \\ U_S(T|i) &= (1-\pi_N)^2 + (1-\eta)\,\delta \sum_{j^*} \sigma^T(i, j^*)\, U_S(j^*) \end{aligned}$$

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

$$\begin{aligned} U_S(j) &= (1-\pi_j)^2 + (1-\eta)\,\delta \sum_{j^*} \sigma^T(j, j^*)\,(1-\pi_{j^*})^2 + ... \\ &\quad + (1-\eta)^t\,\delta^t \pi_k^2 + (1-\eta)^{t+1}\,\delta^{T+1} U_S(l) \end{aligned}$$

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.

$$\begin{aligned} (1-\pi_j)^2 &\leq \pi_l^2, \forall j \\ \pi_j^2 &\leq \pi_N^2, \forall j \end{aligned}$$

$$U_S(L|N) = \pi_N^2 + (1-\eta)^t\,\delta^t \pi_l^2 + (1-\eta)\,\delta \pi_l^2 + ... + (1-\eta)^{t+1}\,\delta^{t+1} U_S(l)$$

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

Thus, we have that:
$U_S(j) \leq U_S(L|N), \forall j \Longrightarrow$In particular this holds for $j = N$. ■

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

$$\pi_i = 1 \Longrightarrow U_S(L|i) > U_S(T|i)$$

Lemma 10 *Sender weakly prefers to lie in all the states:*

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

Proof. Suppose $U_S(T|i) > U_S(L|i) \Longrightarrow q_i = 1 \Longrightarrow \pi_i = 1$.

By the corollary above, we have a contradiction. ■

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

From now on, we write q_i instead of q_{it}. We do this w.l.o.g. because the argument has to hold for any time period.

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

$$\begin{aligned} v_l^S &= q_l \left\{-(1-\pi_l)^2 + (1-\eta)\,\delta\Sigma_{j^*}\sigma^T(l, j^*)\,v_{j^*}^S\right\} + \\ &\quad + (1-q_l)\left\{-\pi_l^2 + (1-\eta)\,\delta v_l^S\right\} \end{aligned} \tag{4}$$

However, in this state i the strategic sender weakly prefers lying than telling the truth. For if is this not the case, $q_i = 1 \Longrightarrow \pi_i = 1$, which implies that lying is actually better for the sender. So we have to consider only the case where $(1-\pi_i)^2 + (1-\eta)\,\delta\Sigma_{j^*}\sigma^T(i, j^*)\,U_S(j^*) \leq \pi_i^2 + (1-\eta)\,\delta U_S(i)$

Thus equation (4) can be written as:

$$v_l^S = -\pi_l^2 + (1-\eta)\,\delta v_l^S \tag{5}$$

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:

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

Again, we have only to consider the case where:

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

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:

$$v_1^S = -\pi_1^2 + (1-\eta)\,\delta v_l^S \tag{6}$$

However we can compare the expected payoff on equations (5) and (6) 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. $\sigma^L(i,1) = 1, \forall i$.

Lemma 11 *Memory state 1 has highest expected payoff given a strategic sender:*$1 \in D$.

Proof.

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

However,$(1-\pi_l)^2 + (1-\eta)\,\delta\Sigma_{j^*}\sigma_{lTj^*}U_S(j^*) \leq \pi_l^2 + (1-\eta)\,\delta U_S(l)$, 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), which 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 on state 1. Again, we need only to consider the case where

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

Thus, we can write v_1^S as:

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

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

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

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

The corollary below shows an immediate consequence of this lemma is that unless there is a state π_2 such that $\pi_2 = \pi_1$ and $v_2^S = v_1^S$, we must have that $\sigma_{iF1} = 1$.

Corollary 2 *All the states with lowest expected continuation payoff for the sender must have the same belief:*

$$i \in D \Longrightarrow \pi_i = \pi_1$$

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

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

If $\pi_l > \pi_1 \Longrightarrow v_l^S < v_1^S$. Which is a contradiction. ■

Corollary 3 *For any state j such that $\pi_j > \pi_1$ then by incentive compatibility it must be true that $\sigma^L(i,j) = 0$.*

Proof. Since $\Pr(H|i,L) = 0, \forall i \Longrightarrow$ Then by incentive compatibility: $p_i^S v_1^S > p_i^S v_j^S \Longrightarrow \sigma^L(i,j) = 0$. ■

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 to a state where his expected payoff is v_1^S. Again, this lemma relies on the first result of this section, that says that lying is always weakly preferred by the sender.

Lemma 12 *π_i and v_i^S have the exact opposite ordering.*

Proof. $v_i^S = -\pi_i^2 + (1-\eta)\,\delta v_1^S, \forall i$

If $\pi_j > \pi_i\,(<)$ then

$$-\pi_i^2 + (1-\eta)\,\delta v_1^S > -\pi_j^2 + (1-\eta)\,\delta v_1^S \Longrightarrow v_i^S > v_j^S$$

■

This lemma leads us to the following corollary: the order of states will be the same as the order by v_i^H. 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, than there must not exist another state that has higher expected continuation payoff for the receiver for both types of sender (i.e. higher v_i^S and v_i^H). Since a state with lower belief has higher v_i^S it must be that this state with lower belief has lower v_i^H, otherwise for whatever posterior the receiver holds, it is always strictly better to move to this lower belief state than to the original state.

Corollary 4 *For the states reached with positive probability, π_i and v_i^H have the exact same ordering.*

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

$$p_{i^*}^H v_j^H + p_{i^*}^S v_j^S \geq p_{i^*}^H v_{j'}^H + p_{i^*}^S v_{j'}^S, \qquad \forall j'$$

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

$$p_{i'}^H v_j^H + p_{i'}^S v_j^S \geq p_{i'}^H v_k^H + p_{i'}^S v_k^S, \quad \forall i' \in N$$

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

Lemma 13 *If the receiver knows with probability one that the sender is behavioral type, she will update to the state with highest expected continuation payoff given a behavioral type of sender:*

$$U_S(L|i) > U_S(T|i) (\Longrightarrow q_{it} = 0, \forall t) \Longrightarrow \sigma^T(i,h) = 1$$

Proof. $q_i = 0 \Longrightarrow \Pr(H|i,T) = 1$. Since we know that $v_h^H \geq v_{i'}^H, \forall i'$ and also that v_i^H and π_i have the same ordering, we must have that:

$$N = \arg\max_{i'} p_i^H v_{i'}^H + p_i^S v_{i'}^S = \arg\max_{i'} v_{i'}^H$$

Thus, $\sigma^T(i,N) = 1$ ■

Lemma 14 *$N \in U$ and $\pi_h = \pi_N$,$\forall h \in U$.*

Proof. First we show that $v_N^H = v_h^H, h \in U$.
$q_N = 0$. Suppose $v_h^H > v_N^H \Longrightarrow \sigma^T(N,h) = 1$ (since $q_N = 0$).

$$\begin{aligned} v_h^H &= -(1-\pi_h)^2 + (1-\eta)\,\delta \sum_{h^*} \sigma^T(h,h^*)\, v_{h^*}^H \\ &\leq -(1-\pi_N)^2 + (1-\eta)\,\delta \sum_{h^*} \sigma^T(h,h^*)\, v_{h^*}^H \\ &\leq -(1-\pi_N)^2 + (1-\eta)\,\delta v_h^H = v_N^H \end{aligned}$$

Thus, $v_h^H > v_N^H$ cannot happen. The proof that $\pi_h = \pi_N$ 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 15 *If the sender strictly prefers to lie on state i and is indifferent in state j, then $\pi_i > \pi_j$:*

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

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

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

$$\pi_i^2 + (1-\eta)\,\delta U_S(1) > (1-\pi_i)^2 + (1-\eta)\,\delta U_S(h) \tag{7}$$

But, we also have that:

$$\pi_j^2 + (1-\eta)\,\delta U_S(1) = (1-\pi_j)^2 + (1-\eta)\,\delta \Sigma_{j^*}\sigma^T(i,j^*)\,U_S(j^*) \tag{8}$$

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,

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

Finally, from (7) and (8) we have that:

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

Which is a contradiction with equation (8). ■

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. The current payoff is higher and also the future payoff.

Lemma 16 *The Receiver will only go up chain after a true signal:*

$$\pi_j > \pi_i \Longrightarrow \sigma^T(j,i) = 0.$$

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

$$p_j^H v_i^H + p_j^S v_i^S \geq p_j^H v_j^H + p_j^S v_j^S$$

However, it can also be written as:

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

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 $U_S(L|i) > U_S(T|i) \Longrightarrow \sigma^T(i,N) = 1$, implying that $\pi_i > \pi_j$ (see lemma (20) that implies that if $q_i = 0$ and $q_j > 0 \Longrightarrow \pi_i > \pi_j$). Thus, we conclude that $U_S(L|i) = U_S(T|i)$.

Therefore, $v_i^S = -U_S(T|i) = -\left((1-\pi_i)^2 + \sum_{i^*} \sigma^T(i,i^*)\, U_S(i^*)\right)$

$$\begin{aligned} v_i^S &= -(1-\pi_i)^2 + \sum_{i^*} \sigma^T(i,i^*)\, v_{i^*}^S \\ p_j^H v_i^H + p_j^S v_i^S &= -(1-\pi_i)^2 + \sum_{i^*} \sigma^T(i,i^*) \left(p_j^H v_{i^*}^H + p_j^S v_{i^*}^S\right) \end{aligned}$$

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

$$p_j^H v_j^H + p_j^S v_j^S = -(1-\pi_j)^2 + \sum_{j^*} \sigma^T(j,j^*) \left(p_j^H v_{j^*}^H + p_j^S v_{j^*}^S\right)$$

By incentive compatibility: $p_j^H v_{j^*}^H + p_j^S v_{j^*}^S \geq p_j^H v_{i^*}^H + p_j^S v_{i^*}^S$. Thus:

$$p_j^H v_j^H + p_j^S v_j^S \geq p_j^H v_i^H + p_j^S v_i^S$$

■

Lemma 17 *The receiver always starts either at the lowest memory state or at the lowest after the babbling state):*

$$g_0(i) = 0, \forall \pi_i > \pi_{(2)}$$

Proof. The ex-ante receiver chooses in which memory state to start the game by finding the solution to $\max_i \rho v_i^H + (1-\rho) v_i^S$. Given points 1 and 3 in proposition 4, we have that $\rho < p_j^H, \forall j > 1$. Thus, if $g_0(i') > 0$,for some $\pi_{i'} > \pi_{(2)}$, then state i' is not reached with positive probability in the game, except for time $t = 0$. ■

This concludes the proof of proposition 5. To relate this proposition with the one presented in the text, we need two additional results:

Lemma 5: *As* $\eta \to 0$, *we have that:*

1) $\pi_1 = \frac{1}{2}$, *moreover state 1 is absorbing.*

2) $\pi_N = 1$.

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

$$p_l^H = \sum_{k=1}^{\infty} \Pr(t = k|l) \Pr(H|T, l, t = k) \tag{9}$$

However, given (1) and (3) from proposition 5 together with the fact that the strategic senders will either remain on one of the states in D for ever 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 $\eta \to 0$, $\Pr(t = 1|l) \to 0$ and therefore $\Pr(H|l) \to 0$. By incentive compatibility it will then imply that $\sigma^T(l, l) = 1$ and consequently $\pi_l = 0.5$.

The argument for point 2) is essentially the same. As $\eta \to 0$ we have that the strategic senders will be locked in the lowest state and also that $U_S(L|u) > U_S(T|u), \forall u \in U$ since in the highest states there are no reputation incentives. Thus, as $\eta \to 0$, $\Pr(H|u) \to 1$. ■

Below I prove the lemma that said that any memory rule with redundant states can be reduced to a rule with less memory states but non identical.

Lemma 6 *If a receiver has memory* M *with* N *states and* (σ, a, g_0, q) *gives the receiver a payoff* U_R^* *and is such that* $\pi_i = \pi_j$, *then there* $\exists$ $(\sigma, a, g_0, q)'$ *for memory* M'*with* $N - 1$ *states and that gives the receiver utility* U_R^*.

Proof. Let $\pi_i = \pi_j$. This implies immediately that $v_i^S = v_j^S$. Thus, if both states are reached in equilibrium it must be that $v_i^H = v_j^H$. The receiver is always completely indifferent between both states i and j after a truth or lie.

If $p_i^H = p_j^H$, then the states are identical and we can consider them as being a single state (just rewrite the transition rules). If $p_i^H > p_j^H$, then they must have the same transition rules, or else $v_i^H = v_j^H$ would not hold, but if they have the same transition rules then again they are identical and we can group them as one. ■

The following results show how the movement up chain should be. We show the three points in lemma 7.

Lemma 18 *Single Crossing (for states where $p_i^H \neq p_j^H$)*
$\sigma^T(i,k) > 0$,$\sigma^T(i,l) > 0$ *and* $\sigma^T(j,k) > 0 \Longrightarrow \sigma^T(j,l) = 0$.

Proof. Proof of Single Crossing: $\sigma^T(i,k) > 0$ *and* $\sigma^T(i,m) > 0 \Longrightarrow$

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

Suppose $\sigma^T(j,k) > 0$ and $\sigma^T(j,m) > 0$

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

If $p_i^H \neq p_j^H$ then (10) and (11) cannot hold at the same time. ■

The lemma below shows a "no jump" result for states where p_i^H and p_j^H are different.

Lemma 19 *No jumps (for states where $p_i^H \neq p_j^H$):*

$$\sigma^T(i,k-1) > 0, \sigma^T(i,k+1) > 0 \Longrightarrow \sigma^T(j,k) = 0 \qquad \forall i,j;\ p_i^H \neq p_j^H$$

Proof. $\sigma^T(i,k+1) > 0$ *and* $\sigma^T(i,k-1) > 0 \Longrightarrow$

$$p_i^H\left(v_{k+1}^H - v_k^H\right) + p_i^S\left(v_{k+1}^S - v_k^S\right) \geq 0 \tag{12}$$
$$p_i^H\left(v_k^H - v_{k-1}^H\right) + p_i^S\left(v_k^S - v_{k-1}^S\right) \leq 0 \tag{13}$$

If $\sigma^T(j,k) > 0 \Longrightarrow$

$$p_j^H\left(v_{k+1}^H - v_k^H\right) + p_j^S\left(v_{k+1}^S - v_k^S\right) \leq 0 \tag{14}$$
$$p_j^H\left(v_k^H - v_{k-1}^H\right) + p_j^S\left(v_k^S - v_{k-1}^S\right) \geq 0 \tag{15}$$

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

We now show results on the order of the posteriors, the goal of this section is to show that the order of beliefs is the same as the order of posteriors.

Consider two sates π_i and π_j ,$\forall i, j$ such that: $\pi_j > \pi_i$ but also such that in equilibrium the posteriors have different order: $p_j^H < p_i^H$. We want to show that this is a contradiction. First, I need a lemma on the transition rule of two states with different posteriors.

Lemma 20 *Monotonicity (Let $p_i^H > p_j^H$.)*

If $\sigma^T(j, m) > 0 \Longrightarrow \sigma^T(i, m-1) = 0$ (holds for any $m' < m$, not only for $m-1$)

Proof. First, note that by incentive compatibility:

$$\sigma^T(j, m) > 0 \Longrightarrow p_j^H v_m^H + p_j^S v_m^S \geq p_j^H v_{m-1}^H + p_j^S v_{m-1}^S$$

This means that:

$$p_j^H \left(v_m^H - v_{m-1}^H\right) + p_j^S \left(v_m^S - v_{m-1}^S\right) \geq 0 \tag{16}$$

Note that $\left(v_m^H - v_{m-1}^H\right) \geq 0$ and $\left(v_m^S - v_{m-1}^S\right) \leq 0$.
Thus, since $p_i^H > p_j^H$ (and consequently $\left(p_j^S > p_i^S\right)$), we have that:

$$p_i^H \left(v_m^H - v_{m-1}^H\right) + p_i^S \left(v_m^S - v_{m-1}^S\right) > 0 \tag{17}$$

■

What this monotonicity result is telling us is that for any two states with different posteriors we must have that 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.

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

Lemma 21 *The beliefs of the states are weakly ordered according to the posteriors:*

$$\pi_j > \pi_i \Longrightarrow p_j^H \geq p_i^H$$

Proof. Consider any two states i and j such that: $\pi_j > \pi_i$ and $p_j^H < p_i^H$. 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

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

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

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

At the same time we have that:

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

$U_S(T|i) > U_S(T|j)$ and $U_S(L|i) < U_S(L|j)$
Thus, $U_S(T|i) = U_S(L|i) \Longrightarrow U_S(L|j) > U_S(T|j)$
However,

$$U_S(L|j) > U_S(T|j) \Longrightarrow q_j = 0 \Longrightarrow p_j^H = 1$$

Which is a contradiction.

It could alternatively be that:

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

and again we have a contradiction. ■

7.4 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.

Lemma 22 *Given a memory size N, $\exists\ \rho_N^*$ such that $\rho < \rho_N^*$ there is no equilibrium in pure strategy.*

Proof. This lemma is to show what is 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 $\sigma_{N-1TN} = 1$.

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 $\pi_{N-1} = \rho_{N-1}$. Whereas 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:

$$\pi_{N-1} = \frac{1}{2} + (1-\eta)\,\delta\frac{3}{8} \tag{18}$$

Thus, we need to find the lower bound on prior or, equivalently, the highest q that can support (18). 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 IC constraint as well.

To compute the IC of the receiver, note that: $v_N^H = 0$; $v_{N-1}^H = -(1-\pi_{N-1})^2$; $v_N^S = -1 - \frac{(1-\eta)\delta}{1-(1-\eta)\delta}\frac{1}{4}$; and $v_{N-1}^S = -\pi_{N-1}^2 - \frac{(1-\eta)\delta}{1-(1-\eta)\delta}\frac{1}{4}$.

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

$$p_{N-1}^H\left(v_N^H - v_{N-1}^H\right) + p_{N-1}^S\left(v_N^S - v_{N-1}^S\right) \geq 0$$

In this context, it translates to: (rearranging terms and substituting ps and vs) :

$$\frac{\rho_{N-1}}{\rho_{N-1} + \left(1-\rho_{N-1}\right) q_{N-1}}\left(v_N^H - v_{N-1}^H\right) + \left(1 - \frac{\rho_{N-1}}{\rho_{N-1} + \left(1-\rho_{N-1}\right) q_{N-1}}\right)\left(v_N^S - v_{N-1}^S\right) \geq 0$$

Which happens if and only if:

$\rho_{N-1}(1-\pi_{N-1})^2 + \left(\pi_{N-1} - \rho_{N-1}\right)\left(\pi_{N-1}^2 - 1\right) \geq 0 \iff$
$(1-\pi_{N-1})\left[\rho_{N-1}(1-\pi_{N-1}) - \left(\pi_{N-1} - \rho_{N-1}\right)(\pi_{N-1}+1)\right] \geq 0 \iff$
$\rho_{N-1}(1-\pi_{N-1}) - \left(\pi_{N-1} - \rho_{N-1}\right)(\pi_{N-1}+1) \geq 0 \iff$

$$\rho_{N-1} \geq \frac{\pi_{N-1} + \pi_{N-1}^2}{2} \tag{19}$$

For any ρ_{N-1} that is smaller than the threshold above, we need more q to induce the π needed for (18) 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 (19) then we need a lower q and (18) is maintained.

We showed that $\sigma_{N-1TN-1} = 0$ by incentive compatibility, moving forward is better for the receiver.

The conclusion of this result is that if we arrive at state $N-1$ 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 $\sigma_{N-1TN-1} > 0$. If we arrive at state $N-1$ 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 lets look at state $N-2$ and generalize the argument for states $i = N-2, N-3, ...1$. The necessary conditions for $\sigma_{N-2TN-1} = 1$ are the following.

Suppose (18) and (19) so that the last two states the receiver plays pure strategy. We want to find conditions for $\sigma_{N-2TN-1} = 1$.

If (18) does not hold with equality, i.e. if it is better for the sender to lie in state $N-1$, then the lower bound is higher , thus we focus on the case where (18) holds with equality. More on this later.

$$\pi_{N-2} = \frac{1}{2} + (1-\eta)\frac{\delta}{2}\left(\pi_{N-1}^2 - \frac{1}{4}\right) \tag{20}$$

Together with $\rho_{N-1} \geq \frac{\pi_{N-1}+\pi_{N-1}^2}{2}$ which is the same as $\frac{\rho_{N-2}}{\pi_{N-2}} \geq \frac{\pi_{N-1}+\pi_{N-1}^2}{2}$. We can write this condition as:

$$\rho_{N-2} \geq \left(\frac{\pi_{N-1}+\pi_{N-1}^2}{2}\right)\pi_{N-2} \tag{21}$$

If ρ_{N-2} is smaller than in equation (21) then when we get to state $N-1$ the receiver will rather stay put than go forward.

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

$$\rho_i \geq \frac{\pi_{N-1}+\pi_{N-1}^2}{2}\prod_{k=i}^{N-2}\pi_k \tag{22}$$

■

Corollary 5 *As $N \to \infty$, $\rho_N^* \to 0$.*

The lemma above guarantees that $\sigma_{N-1TN} = 1$. But what guarantees that $\sigma_{i-1Ti} = 1, \forall i$, or any other deviation from the specified deterministic transition rule? Next lemma answers this question. I show that if the receiver is playing pure strategy $\sigma^T(i, i^*) = 1$, 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 23 $-\pi_j(1-\pi_j) > -\pi_j(1-\pi_{j'})^2 - (1-\pi_j)\pi_{j'}^2$, $\forall j, j'$

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 $\pi_j, \pi_{j'}$. ■

Lemma 24 *Let* $\sigma^T(i, i+1) = 1$, $p_i^H = \frac{\rho_{i-1}}{\pi_{i-1}}$ *and the strategy for the sender is a best response for him. Then:* $p_{i-1}^H v_i^H + \left(1 - p_{i-1}^H\right) v_i^S \geq p_{i-1}^H v_{i-s}^H + \left(1 - p_{i-1}^H\right) v_{i-s}^S$, $\forall s > 0$.

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

$$\Pi_{eq} = -\rho_i\left(\sum_{k=i}^{N}(1-\pi_i)^2\right) - (1-\rho_i)\left\{q_i\left((1-\pi_i)^2 + \beta U_S(i+1)\right) + (1-q_i)\left(\pi_i^2 + \beta\frac{1}{4}\frac{1}{1-\beta}\right)\right\} \tag{23}$$

We want an appropriate way to write (23) 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 (23) as:

$$\begin{aligned}\Pi_{eq} &= -\pi_i(1-\pi_i) - \beta\left(\pi_i\pi_{i+1}(1-\pi_{i+1}) + (1-\pi_i)\frac{1}{4}\frac{1}{1-\beta}\right) - \quad (24)\\ &\quad -\beta^2\left(\pi_i\pi_{i+1}\pi_{i+2}(1-\pi_{i+2}) + \pi_i(1-\pi_{i+1})\frac{1}{4}\frac{1}{1-\beta}\right) + \ldots\end{aligned}$$

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 $(1-\pi_{i-1})^2 + \beta U_S(i) = \pi_{i-1}^2 + \beta\frac{1}{4}\frac{1}{1-\beta}$ and therefore, any $q_i^{dev} \in [0,1]$ will not change equation (24). In particular, consider $\tilde{q}_i = q_i^{eq}$ in fact, consider the same modification for the entire strategy for the sender, i.e. $\tilde{q}_j = q_j^{eq}, \forall j \geq i$.

Lets rewrite the deviation payoff replacing the qs in the way suggested above. We want to compare the payoffs period by period. At all periods before reaching state $N-s$ lemma (23) tells us that the equilibrium payoff is higher. Remains to show what happens at state $N-s$. 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: $-\prod_{k=i}^{N-1}\pi_k\left[\pi_N^*(1-\pi_{N-s})^2 + (1-\pi_N^*)\pi_{N-s}^2\right]$.

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

$$\begin{aligned}1-\pi_N^* &< \pi_N^* - 2\pi_N^*\pi_{N-s} + \pi_N^*\pi_{N-s}^2 + \pi_{N-s}^2 - \pi_N^*\pi_{N-s}^2 \iff\\ 1-\pi_N^* &< \pi_N^* - 2\pi_N^*\pi_{N-s} + \pi_{N-s}^2 \iff\\ 0 &< 2\pi_N^* - 2\pi_N^*\pi_{N-s} + \pi_{N-s}^2 - 1 \iff\\ 0 &< 2\pi_N^*(1-\pi_{N-s}) - \left(1-\pi_{N-s}^2\right) \iff\\ 0 &< (1-\pi_{N-s})\left\{2\pi_N^* - (1+\pi_{N-s})\right\} \iff 2\pi_N^* > 1+\pi_{N-s}\end{aligned}$$

However, a necessary condition for equilibrium in pure strategy was that the it should be incentive compatible for the receiver to update in state $N-1$ as well and this condition is that $\rho_{N-1} \geq \frac{\pi_{N-1}+\pi_{N-1}^2}{2}$,knowing that we have that $\pi_N^* = p_{N-1}^H = \frac{\rho_{N-1}}{\pi_{N-1}}$, 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 25 *Under deterministic transition rules we must have that:*

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

Proof. The equilibrium payoff is again given by (23), and again we can write as in equation (24). We can further change the q and write (24) with $q_{N-s} = 0$,instead. This change in q_{N-s} will not change the value of Π_{eq} since $U_S(L|N-s) = U_S(T|N-s)$ or, $(1-\pi_{N-s})^2 + \beta U_S(N-s+1) = \pi_{N-s}^2 + \beta$.The deviation payoff is:

$$\begin{aligned} \Pi_{dev} &= -\rho_i \left(\sum_{k=i+s}^{N} (1-\pi_k)^2 \right) - (1-\rho_i) \qquad (25) \\ & \left\{ q_i^{dev} \left((1-\pi_{i+s})^2 + \beta U_S(i+s+1) \right) + \left(1 - q_i^{dev}\right) \left(\pi_{i+s}^2 + \beta \frac{1}{4}\frac{1}{1-\beta} \right) \right\} \end{aligned}$$

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 for the sender, i.e. $\tilde{q}_{j+1} = q_j^{eq}$.Once we use $\tilde{q}$ as the deviation probabilities for the sender, then (25) can be written as:

$$\begin{aligned} \Pi_{dev} &= -\left[\pi_i (1-\pi_{i+1})^2 + \pi_i \pi_{i+1}^2\right] - \qquad (26) \\ & -\beta \left(\pi_i \left[\pi_{i+1} (1-\pi_{i+2})^2 + (1-\pi_{i+1})\, \pi_{i+2}^2\right] + (1-\pi_i)\, \frac{1}{4}\frac{1}{1-\beta} \right) - \\ & -\beta^2 \left(\pi_i \pi_{i+1} \left[\pi_{i+2} (1-\pi_{i+3})^2 + (1-\pi_{i+2})\, \pi_{i+3}^2\right] + \pi_i (1-\pi_{i+1})\, \frac{1}{4}\frac{1}{1-\beta} \right) + ... \end{aligned}$$

We now want to compare the payoffs in (24) but with $q_{N-1} = 0$ and (26) *period by period*. Note that according to lemma (9) we have that the payoff in (24) is greater than the payoff in (26) 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:

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

Which can be written as

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

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. ■

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

Lemma 26 *Fix N and ρ. There is at most one equilibrium in pure strategies for the receiver without redounding states.*

Proof. Let π and π' be the vectors of beliefs associated to two different equilibria in pure strategies (if the beliefs are identical, then we must have that the equilibria is in fact unique). Assume w.o.l.g. that $\pi_i > \pi'_i$ for some $i \in M \implies \pi_{i+1} > \pi'_{i+1}, \forall i < N-1$. This result is true by 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 ι in one of the two equilibrium. This would imply that the sender is a deterministic transition rule in state i in one of the two equilibria. Given this result, now lets examine two possibilities:

$\pi_{N-1} = \pi'_{N-1} \implies \pi_{N-2} = \pi'_{N-2}; \pi_{N-3} = \pi'_{N-3}$; and so on. This is a contradiction.

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

$\blacksquare$

References

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

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

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

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

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

[6] Conlisk, J. (1996) "Why Bounded Rationality?", *Journal of Economic Literature*, XXXIV June (2005), 669-700.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[28] R. Reis, "Inattentive Consumers", mimeo Harvard University, (2004).

[29] A. Rubinstein, "Finite Automata Play the Repeated Prisoners' Dilemma," *Journal of Economic Theory*, 39 (1986), 83-96.

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

[31] I. Sarafidis, "What Have You Done for Me Lately? Release of Information and Strategic Manipulation of Memories", Yale working paper (2001).

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

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

[34] A. Wilson, "Bounded Memory and Biases in Information Processing", Job market paper, Princeton (2004).