Learning with Bounded Memory in Games*

Daniel Monte†
Sao Paulo School of Economics - FGV

March 25, 2014

Abstract

We study learning with bounded memory in zero-sum repeated games with one-sided incomplete information. The uninformed player has only a fixed number of memory states available. His strategy is to choose a transition rule from state to state, and an action rule, which is a map from each memory state to the set of actions. We show that the equilibrium transition rule involves randomization only in the intermediate memory states. This is in contrast with the earlier literature on optimal finite memory, as in Hellman and Cover (1970) and Wilson (2003), where randomization occurred only at the extreme states. Such randomization, or less frequent updating, is interpreted as a way of testing the opponent, which generates inertia in the player's behavior and is the main short-run bias in information processing exhibited by the bounded memory player.

JEL classification : C72, D82, D83
Keywords: bounded memory, incomplete information games, repeated games.

1 Introduction

In this paper we assume that people categorize the world in a coarse way. This is consistent, for example, with the fact that consumer reports often come in the form of finite ratings. It is also consistent with the view that it may not be possible to distinguish beliefs about two agents or products that differ only by a negligible fraction (in a particular dimension). In fact this is also how some authors in psychology model working memory.[1] To illustrate, assume that an agent thinks of his opponent as being someone "trustworthy", "not trustworthy" or "unclear," instead of having a precise posterior distribution about the opponent's true underlying type.

In our model, bounded memory captures this coarse categorization. Apart from bounded memory, the agent is rational, and updating from one category to another is part of her strategy. A

*This paper is a substantially revised version of my Ph.D. dissertation at Yale University. I am very thankful to Ben Polak, Dirk Bergemann, and Stephen Morris for their constant guidance. Many people have helped me throughout the process and I am very thankful to them. I am particularly thankful for conversations with Attila Ambrus, Faruk Gul, Pei-Yu Lo, George Mailath, Tymofiy Mylovanov, and Arthur Robson.

†Sao Paulo School of Economics - FGV. Email: daniel.monte@fgv.br. Phone: (55-11)-3799-3727.

[1] See for example Miller (1956), and Cowan (2001).

bounded memory player has only a fixed number of memory states available. All she knows about the history of the game is her current memory state. The player is aware of her memory constraints, and her strategy is to choose a transition rule from state to state and an action rule, which is a map from each memory state to the set of actions. We show how this agent updates her beliefs and the implications of this chosen updating rule in a strategic setting. In particular, we show that the equilibrium updating rule in a reputation game involves randomization at the intermediate memory states (as opposed to what has been found in decision problems), and we argue that this inertia is a short-run bias in information processing in the behavior of the bounded memory player.

Given that the player is only constrained in her memory, we view memory as a "conscious" process, i.e. the player is subject to "incentive compatibility constraints" (sequential rationality constraints extended to games with bounded memory). The action that the player chooses at each memory state and the transition from state to state must be optimal given her beliefs at that state, and taking as given her *own* strategy - both action and transition rules - at all other states. In contrast to sequential rationality in models with perfect recall, incentive compatibility constraints here do not imply an optimal continuation strategy. The reason a bounded-memory player must take her own strategy at all periods (including future periods) as given when deciding on an action or on which state to move is that if she deviates today, she will not remember it tomorrow.

Conscious memory distinguishes our model from other models of bounded rationality in the literature, in particular, from the standard finite automata. Like ours, a standard automaton has a fixed set of states, a transition rule, and an action rule. However, a standard automaton is assumed to fully (and credibly) commit to a strategy at an ex-ante stage, and, hence, it does not face incentive compatibility constraints.[2] The idea of incentive compatibility constraints, or sequential rationality in bounded memory, was introduced by Piccione and Rubinstein (1997) and Wilson (2003), but these authors studied single-person decision problems for which commitment ex-ante has been shown to be ineffective. Here, we study games, in which the inability to commit matters.

To be able to isolate the effects of bounded memory from complexity considerations, we focus on zero-sum games. An infinitely repeated zero-sum game with complete information requires minimal complexity. We look at the incomplete information case in which the probability of a commitment type is sufficiently low, in the spirit of Kreps et al. (1982).

The precise setting of this study is an infinitely repeated two player game with incomplete

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

information. One player is uncertain about his opponent's true type, which we assume to be either a commitment type that is restricted to only one action or a normal type that has opposite preferences to those of the uninformed player. Our main propositions apply to a general class of zero-sum games, but memory restrictions will only bind in games in which learning the true type of the opponent is profitable for the uninformed player. Thus, we focus on games in which the equilibrium payoff of the uninformed player in the complete information game against the commitment type is strictly higher than her equilibrium payoff in the incomplete information game (we want to rule out uninteresting cases, such as when all payoffs are the same). The matching pennies is a canonical example for our stage game. We assume that only the uninformed player has finite memory, whereas the informed player is an unbounded player.

Every period the two agents play a simultaneous finite game. The payoffs are realized and the actions are observed with certainty. Then, the uninformed player updates her beliefs (according to her chosen memory rule) on the opponent's true type.

Equilibrium will always exist and, typically, not be unique. In propositions 1 and 2, we show conditions that must be satisfied in any equilibrium of the game. In particular, we show that the updating rule must be weakly increasing after observing actions that are consistent with the action a commitment type would play. This implies that the uninformed player's belief on the commitment type is stochastically higher after observing the commitment type's action. In addition, because a different action outside this set leaves no uncertainty in the mind of the uninformed player, she moves to her "lowest" memory state after such an action. We show that, in this "lowest" state, her belief on the informed player being the commitment type is close to zero, and, in her "highest" state, the belief is close to one. This result holds even for the minimal case of only two memory states, or one-bit memory.

In the cases in which the prior belief on the commitment type is sufficiently small, the equilibrium transition rule will require randomization in any equilibrium in which memory states are used in an "efficient" way. Informally, we interpret this result as indicating that when the uninformed player does not have enough memory to keep track of all the actions played by the opponent, she will use randomization to overcome the memory problem and to "test" the opponent before updating.

The role of random transition rules in an optimal finite memory has been studied in single person decision problems. Hellman and Cover (1970) studied the two-hypothesis testing problem with a finite automaton (with ex-ante commitment to the strategy). A decision maker has to make a decision after a very long sequence of signals. However, the decision maker cannot recall all the sequence and, instead, has to choose the best way to store information given his finite set of memory states. A key result of that paper is that, for the discrete signal case, the transition rule is random in the extreme states. The authors concluded that, perhaps counter intuitively, the decision maker uses randomization as a memory-saving device.[3]

Wilson (2003) studied a problem similar to Hellman and Cover (1970). In her model the decision maker was subject to sequential rationality constraints. The optimal memory rule obtained was similar to Hellman and Cover's rule and it included randomization in the extreme states. She showed that this randomization at the extreme states can explain several biases in information processing, such as confirmatory bias and overconfidence/underconfidence bias.

In contrast to the results of Hellman and Cover (1970) and Wilson (2003), we show that in our setting the randomization occurs at the intermediate states and not at the extreme states. Thus, the main behavioral bias exhibited by the bounded memory player in this class of games is infrequent updating in the short-run, which is a result of randomization in the transition rule of memory states associated with intermediate beliefs. We interpret this infrequent updating as excessive inertia and occasional overreaction (when updating does occur).

The results of this paper depend on the perfect monitoring assumption, or similarly that the commitment type is playing a pure strategy. The absorbing state depends on this assumption, and, in turn, it allows us to characterize the necessary conditions for equilibria, stated in proposition 1. In a separate paper (Monte (2013)) we study a similar problem, but in a game in which the commitment type is playing a mixed strategy. In that paper, we focus on the implications of bounded memory in long-run reputations. In particular, we show that the important recent result in repeated games with incomplete information in which the play of the game converges to the play of a complete information game (e.g., Cripps et al. (2004)) may not hold if the uninformed player has bounded memory.

This paper is organized as follows. Section 2 consists of the description of the model and the definitions of memory, strategies, and the equilibrium concept. In section 3, we briefly discuss the full memory case. Section 4 we show the main result of the paper, i.e., the characterization of the memory rule in any equilibrium of the game. Section 5 shows that for low prior beliefs on the commitment type, the transition rule will involve randomization at the intermediate states. In Section 6 we present a result on the value of memory. Section 7 presents the conclusions of the

[3] "It is somewhat surprising that randomization is needed at all, since randomization usually decreases information." (Hellman and Cover (1970), p 781).

paper. The proofs not in the main text are left in the Appendix.

2 Model

The setting of our study is an infinitely repeated game in which one player has incomplete information on the opponent's type. The informed player can be one of two possible types: with probability ρ he is a commitment type, committed to a known pure strategy. This commitment type will be denoted c. With probability $(1-\rho)$ the informed player is a rational player which we call the normal type and denote r, with utility opposite to the uninformed player's. We denote the uninformed player by b.

The stage game is a finite game Γ where the set of actions for the uninformed player is A_b and for the normal type it is A_r. The commitment type is playing a fixed action $\bar{a} \in A_r$. At every period both players choose actions from the sets A_b and A_r, then they both observe the outcome and the uninformed player updates his beliefs on the informed player's type. Denote $u_b(a, a')$ to be the utility of the uninformed player in the stage game if the action that he chose was $a \in A_b$ and the action chosen by the informed player was $a' \in A_r$. Similarly, denote $u_r(a, a')$ to be the normal type's stage game payoff. We also denote u_i, for $i = b, r$ to be the mixed extension of u_i. Thus, $u_i(\tau_b, \tau_r) = Eu_i(a_b, a_r)$, for $a_b \in A_b$ and $a_r \in A_r$, where the expectation is with respect to the measure over $A_b \times A_r$ defined by the mixed action profile $(\tau_b, \tau_r) \in \Delta(A_b) \times \Delta(A_r)$.

The game between the uninformed player and the normal type is a zero-sum game. We will assume that the game is such that there are gains form learning the opponent's true type; and that learning is not trivial. Memory restrictions will bind in games in which the normal type has a strict incentive to play some action other than $\bar{a}$, whenever the uninformed player is playing a best response to $\bar{a}$. For simplicity, we make the following assumption on the stage game Γ: each player has two actions only, and there are no (strictly or weakly) dominated actions.

This implies that the complete information game against the normal type has a unique Nash equilibrium in the stage game, denoted by $\tau = (\tau_b^*, \tau_r^*)$. It can be shown that in this equilibrium the uninformed player plays the pure strategy best response to $\bar{a}$, denoted by $b(\bar{a})$, with positive probability and the normal type plays $\bar{a}$ with positive probability, $\tau_r^*(\bar{a}) > 0$. Moreover, the uninformed player's stage game payoff when he is choosing $b(\bar{a}) \in A_b$ and the informed player is choosing $\bar{a}$ is higher than the uninformed player's stage game equilibrium payoff in the complete information game against the normal type:

$$u_b\left(b\left(\bar{a}\right),\bar{a}\right) > u_b\left(\tau_b^*,\tau_r^*\right).$$

The reason we make these assumptions is to rule out games in which there are no gains from learning, and thus, no need for memory (e.g. when all outcomes yield the same payoff to both players). An example that satisfies our restrictions is given by the traditional matching pennies game, where the commitment type is playing one of the two actions with probability 1.[4]

We assume that the game is repeated, but after every period there is a positive exogenous stopping probability $\eta < 1$. We will focus on the case where this probability η is very small so that the players expect the game to go on for a very long horizon. The players discount their repeated game payoff using the discount factor $\delta = (1-\eta) < 1$.[5]

In the model with no memory constraints, the trade-off for the normal type of the informed player is between building reputation or revealing himself. He might want to mimic the commitment type and build reputation for the following stage game. Or, he might want to choose an action $a \neq \bar{a}$ and reveal himself. Once he reveals himself, he plays a zero-sum game with the uninformed player, and there is a unique equilibrium payoff for this subgame. We will later see that this trade-off is still present in the game with a bounded memory uninformed player.

2.1 Memory and Strategies

A history in this game is defined as nature's choice of the actual type of the informed player, the sequences of action profiles, and the memory states of the uninformed player. A history of length $t \geq 1$ is denoted $h_t = \left(k, \left\{\left(a_b^l, a^l\right)\right\}_{l=1}^t, \left\{\left(s_i^l\right)_{l=1}^t\right\}\right)$, where $k \in \{c, r\}$, $a_b \in A_b$, $a \in A_r$ and $s_i \in \mathcal{M}$. The initial history is denoted t_0. The set of histories of length t is denoted $\mathcal{H}_t$, and let $\mathcal{H} = \cup_{t=0}^\infty \mathcal{H}_t$. We will assume throughout the paper that beliefs are public, so that memory states are observed by the opponent. Although this is an assumption that is hard to justify, it is not crucial for our main result, namely that randomization occurs only at intermediate memory states.[6] The case of public beliefs will give tractability to this model and we will be able to provide a complete characterization

[4] A more applied game with bounded memory can be found in Monte (2007). There, we study the same underlying bounded memory problem, but the stage game is a cheap talk game, very much based on the credibility model of Sobel (1985).

[5] The assumption of the stopping probability η is done for technical reasons, which will be discussed in section (2.3).

[6] This assumption can be motivated in two different ways. First, a slightly modified model in which the uninformed player takes an observable continuous action before the agent acts together with a single peaked utility function would give us the same results as in this paper– without the assumption of observable memory states. Second, one can think of this as literally being public beliefs, such as institutions that publish credit ratings.

of the equilibrium updating rule. [7]

The informed player, who is unconstrained, will condition his strategy on the observed history of the game.

The normal type's behavioral strategy is:

$$\tau_r : \mathcal{H} \to \Delta(A_r).$$

Let Σ_r be the set of all behavioral strategies τ_r. The commitment type's repeated game strategy is τ_c and it is defined to be the constant play of $\bar{a}$, regardless of the history.

The memory of the uninformed player is defined as a finite set of states $\mathcal{M} = \{s_1, s_2, ..., s_n\}$. A typical element of $\mathcal{M}$ is denoted by s_i or s_j.

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

$$\tau_b : \mathcal{M} \to \Delta(A_b). \tag{1}$$

At the end of each period, the uninformed player must decide to which memory state to move to based on his current memory state and on the opponent's action.[8] Allowing for the possibility of randomization, we write the transition rule as a map

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

We denote $\varphi_a(i,j)$ as the probability of moving from state s_i to state s_j given that the informed player chose $a \in A_r$. This transition rule will determine how the uninformed player updates beliefs. One way to think of this is that the bounded memory player's knowledge about the history of the game is summarized by an n–valued statistic s_i, which is updated according to the map φ.

Finally, it is also part of the uninformed player's strategy to decide, before the first stage game, his initial distribution over the memory states $g \in \Delta(\mathcal{M})$. The strategy for the uninformed player is the triple (g, φ, τ_b) and we denote the strategy profile by $\sigma = ((g, \varphi, \tau_b), \tau_r)$.

[7] In the case where the informed player does not observe the current memory state, he must keep track of the history to infer the current memory state of the uninformed player. This private belief of the informed player will impose an additional inference problem on the bounded memory player: he is learning about the opponent's type but also about the opponent's private belief of his current memory state. With private beliefs there is no recursive form, which greatly complicates the analysis. This is in the same way that much of the 'tractability' is lost when we move from a repeated game with public monitoring to a repeated game with private monitoring.

[8] We could allow the uninformed player to condition his transition rule in his own current action as well. However, since he is forgetful, he will choose to condition his transition only on relevant new information (the informed player's action). We discuss this in section 2.4. In the automata literature, this is the distinction between full and exact automata, see Kalai and Neme (1992).

Given the assumption on public beliefs, we focus on Markovian equilibria and the memory state will serve as the state variable, so with slight abuse of notation, we will refer to $\tau_{r,i}(a)$ as the probability of choosing a given a memory state s_i.

2.2 Payoffs

For every strategy profile σ, every memory state will have an associated expected continuation payoff for the bounded memory player conditional on the actual type of the informed player. The expected continuation payoff for the bounded memory player at the beginning of a stage game at memory state s_i, given that the informed player is a commitment type, is denoted by v_i^c. This expected continuation payoff is the sum of the stage game payoff and the continuation payoff induced by the strategy profile. Formally, this expected continuation payoff v_i^c can be written as:

$$v_i^c = u_b\left(\tau_{b,i}, \bar{a}\right) + \delta \sum_{j=1}^{n} \varphi_{\bar{a}}\left(i, j\right) v_j^c, \tag{3}$$

where $\tau_{b,i}$ is the probability distribution over the set of actions A_b in memory state s_i, according to τ_b. The first term on the right of (3) is the payoff of the bounded memory player in the stage game. This payoff is given by the equilibrium behavioral strategy τ_b and the strategy of the commitment type of the informed player, which is to choose $\bar{a}$ with probability 1. The second term is the expected continuation payoff of the continuation game. This depends on the transition rule and on the associated continuation payoffs of all states reached with positive probability given the transition rule φ. The expected continuation payoff for the uninformed player given a normal type of the informed player is denoted by v_i^r, and we can write this expected payoff as:

$$v_i^r = u_b\left(\tau_{b,i}, \tau_{r,i}\right) + \delta \left(\sum_{a_l \in A_r} \tau_{r,i}\left(a_l\right) \sum_{j=1}^{n} \varphi_{a_l}\left(i, j\right) v_j^r \right), \tag{4}$$

where $\tau_{r,i}$ is the probability distribution over the set of actions A_r used by the normal type in memory state s_i and $\tau_{r,i}(a_l)$ is the probability of choosing a_l at memory state s_i.

When deciding on an action to take, and on which state to move, the bounded memory player makes his decisions based on the expected continuation payoffs associated with his decisions.

Given a strategy profile σ, we denote the informed player's expected continuation payoff in state s_i as $U_r^{\sigma}(s_i)$, and when there is no confusion, we omit the superscript and write simply $U_r(s_i)$. This utility is given by a stage-game payoff and a discounted expected continuation payoff that depends on the transition rule φ as well as on $U_r(s_j)$ for all $s_j \in \mathcal{M}$ reached through φ. If the strategy of the informed player depends only on the current memory state of the uninformed player, the informed player's expected utility $U_r(s_i)$ can be written as:

$$U_r(s_i) = u_r(\tau_{b,i}, \tau_{r,i}) + \delta \left\{ \sum_{a_l \in A_r} \tau_{r,i}(a_l) \sum_{j=1}^{n} \varphi_{a_l}(i,j) U_r(s_j) \right\}. \tag{5}$$

Note that if the true type of the informed player is the normal type, then the informed player and the uninformed player are playing a zero-sum game. This implies that their expected continuation payoffs are exactly the opposite, i.e. $v_i^r = -U_r(s_i)$.

2.3 Beliefs

As described, we view memory as a conscious process. Players know that they are forgetful. At every memory state they will hold a distribution of beliefs over the set of histories. Given a strategy profile $\sigma = ((g, \varphi, \tau_b), \tau_r)$, let $\mu(h|s_i, \sigma)$ denote the belief of the uninformed player in state s_i that the correct history is h, given the strategy profile σ. As usual, at any information set the beliefs about all histories must sum up to one

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

Following Piccione and Rubinstein (1997), we assume that the beliefs correspond to "relative frequencies" as follows.[9] Let $f(h|\sigma)$ be the probability that the play of the game passes through the history h given the strategy profile σ. Here, note that it is possible that $f(h|\sigma) + f(h'|\sigma) > 1$ for two histories on the same information set. Indeed, this is the case if h is reached with probability greater than 0.5 under strategy profile σ and from h the game moves to h' with probability 1 (and the bounded memory player does not distinguish h from h'). In other words, while the beliefs in any particular information set must sum to one, the relative frequencies $f(\cdot)$ need not (see Remark 2).

With abuse of notation, we will denote $h \in s_i$, for each history h and memory state s_i, for a history with its last component being memory state s_i. That is, at each time t, the history $h_t \in s_i^t$ if $h_t = \left(k, \left\{\left(a_b^l, a^l\right)\right\}_{l=1}^{t}, \left\{\left(s_i^l\right)_{l=1}^{t}\right\}\right)$.

We will say that the strategy profile σ is consistent with beliefs μ if these beliefs are given by the relative frequencies as defined below.

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

Definition 1 ***(Consistency)***

A strategy profile σ is consistent with the beliefs μ if, for every memory state s_i and for every history $h \in s_i$, we have that the beliefs are computed as follows:[10]

$$\mu\left(h|s_i,\sigma\right)=\frac{f\left(h|\sigma\right)}{\sum_{h'\in s_i} f\left(h'|\sigma\right)}. \tag{6}$$

Remark 1 *The exogenous probability of the game ending, η, is important in infinite horizon games with imperfect recall. When computing beliefs in an infinite horizon game, the bounded memory player must have priors about the time periods. Given that the game is infinite, a uniform prior is not feasible. The exogenous stopping probability generates a well defined probability distribution over the different time periods.*

Remark 2 *Notice that the denominator in expression (6) can be greater than one. The underlying reason for this is that the uninformed player only keeps track of the time (the period of the game) insofar as his transition rule allows. Thus, for example, depending on the transition rule, a t-period history and its parent $(t-1)$-period sub-history could place the uninformed player in the same memory state. This contrasts with what would be the uninformed player's information sets in the standard game without bounded memory. In the extreme case of one memory state, all histories must be in the same state and the denominator in (6) would be $\frac{1}{\eta}$, where recall that η is the exogenous stopping probability. Even in this case, however, the exogenous stopping probability ensures that the denominator in (6) is bounded and, thus, that the beliefs are well defined (the bounded memory player will have well defined priors over the time periods).*

Let $\mathcal{H}_c$ be the set of histories for which the actual type is c. Similarly, $\mathcal{H}_r$ is the set of histories for which the actual type is r; hence, $\mathcal{H}_c \cup \mathcal{H}_r = \mathcal{H}$. At the beginning of a stage game, given some memory state s_i, the *prior* belief that the opponent is a commitment type is denoted by:[11]

$$\rho_i \equiv \Pr\left(c|s_i,\sigma\right)=\sum_{h\in s_i\cap\mathcal{H}_c} \mu\left(h|s_i,\sigma\right). \tag{7}$$

At the beginning of every stage game, we denote $\pi_i \equiv \Pr\left(\bar{a}|s_i,\sigma\right)$ as the probability that the informed player will choose the action $\bar{a}$ in that stage game, given the current memory state s_i. If

[10] In our analysis we will only consider memory states that are reached on equilibrium path, so the ratio is well-defined.

[11] Recall that we say that a history $h \in s_i$ if the last memory state of h is memory state s_i Thus, with abuse of notation, a memory state $h \in \cap\mathcal{H}_c$ if $h \in \mathcal{H}_c$ and the last memory state of history h is s_i.

the informed player is using a Markovian strategy, we can write the probability of $\bar{a}$ being chosen at memory state s_i as:

$$\pi_i = \rho_i + (1 - \rho_i)\,\tau_{r,i}(\bar{a})\,. \tag{8}$$

After observing whether the action played, the bounded memory player updates his belief concerning the probability that the informed player is a commitment type. We denote this *posterior* after observing $\bar{a}$ as $p_i \equiv \Pr(c|\bar{a}, s_i)$. These beliefs are computed using (7) and (8).

$$p_i = \frac{\rho_i}{\pi_i}. \tag{9}$$

Whenever action $a \neq \bar{a}$ is played, the uninformed player's posterior belief that the informed player is a commitment type is zero, since the commitment type always chooses $\bar{a}$ (we assume that this is true even following histories off-equilibrium path).

In a game with full memory, the player's posterior in the end of a stage game is also his prior in the next stage game. This is not true in general for games with bounded memory players. In any stage game, the player does not necessarily know which was the previous stage game; or the belief he held in the last period. Upon reaching a memory state s_i, the uninformed player will hold a belief about his opponent given by (7), regardless of the actual history. Since all his knowledge about the history of the game is given by the statistic s_i, the belief he holds in s_i must depend only on this information.

2.4 Equilibrium

In our concept of equilibrium, we use the notion of incentive compatibility as described by Piccione and Rubinstein (1997) and Wilson (2003).[12] The assumption that we make is that at every information set the player holds beliefs induced by the strategy profile σ. If there is a deviation in the play of the game, the agent will not remember it, and his future beliefs will still be the ones induced by the strategy σ. Thus, a player might decide to deviate at a particular time, but he cannot trigger a sequence of deviations.

We say that a pair (μ, σ) is *incentive compatible* when it satisfies two specific conditions. First, the strategy of the normal type of the informed player is a best response for him at every history, given the strategy of the bounded memory player $(\varphi_0, \varphi, \tau_b)$. Second, the strategy of the bounded memory player is a best response for him at every point in time, given his beliefs and taking as

[12] Piccione and Rubinstein (1997) refer to this solution concept as "modified multiself consistency". The term incentive compatibility is taken from Wilson (2003).

given the strategy of the informed player and his *own* strategy at all memory states. Again, the reason for taking his own strategy as given when deciding on which action to take or what state to move is that a deviation is not remembered in future periods and the beliefs in the following periods will be the ones obtained by assuming that the strategy profile is σ.

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

We define equilibrium in this game using the notion of incentive compatibility. An equilibrium in this game is a pair (μ, σ) such that the strategies and beliefs are consistent, and the strategies are incentive compatible.

Definition 2 ***(Incentive Compatible Equilibrium)***
The strategy profile $\sigma = ((g, \varphi, \tau_b), \tau_r)$ *is an* incentive compatible equilibrium *if there exists a belief* μ *such that the pair* (μ, σ) *is consistent and* σ *satisfies the conditions below:*

1. $U_r^{((\varphi_0,\varphi,\tau_b),\tau_r)}(s_i) \geq U_r^{((\varphi_0,\varphi,\tau_b),\tau_r')}(s_i)$, $\forall \tau_r' \in \Sigma_r, \forall s_i \in \mathcal{M}$.

2. *For all states* $s_i \in \mathcal{M}$*, we must have:*

$$\begin{aligned} \rho_i v_i^c + (1-\rho_i) v_i^r \quad \geq \quad & \rho_i \left[u_b(a, \bar{a}) + \delta \sum_{j=1}^{n} \alpha_j(\bar{a}) v_j^c \right] + \\ & + (1-\rho_i) \left(u_b(a, \tau_{r,i}) + \delta \left(\sum_{a_l \in A_r} \tau_{r,i}(a_l) \sum_{j=1}^{n} \alpha_j(a_l) v_j^r \right) \right), \end{aligned}$$

for any $a \in A_b$*, and* $\forall \alpha : A_r \to [0,1]$ *such that for* $\forall a_l \in A_r$, $\sum_{j=1}^{n} \alpha(a_l) = 1$*, and* $\alpha_j(a_l) \geq 0$*, for* $\forall j \in \{1, ..., n\}$.

In the context of this game, and given that the player is not forgetful within the period, but only across periods, the incentive compatibility constraint of the bounded memory player (condition 2 in the definition above) can be written as two separate conditions: one condition for the transition rule and another one for the action rule.

The condition for incentive compatibility on the action rule of the uninformed player requires that he takes the myopic best action at all stage games. For suppose not: at some memory state s_i the specified action is different from the myopic best one. If the uninformed player deviates to the best current action he will not remember it in the following period. Since histories are private, the informed player will only punish the uninformed player for this deviation if this punishment is profitable even in the case of no deviations. This implies that it must not be profitable, and thus, the uninformed player should deviate and play the myopic best one. For the same reason, the uninformed player will not consider his own current action when deciding on which state to move next.

The incentive compatibility condition for the transition rule requires that the uninformed player moves to the memory state that gives him the highest expected payoff given his beliefs. Thus, if his transition rule assigns positive probability to move from state s_i to state s_j after the action played by the informed player $a \in A_r$, then given his beliefs at state s_i, it must be optimal for him to do so. We state the definition of incentive compatibility in the uninformed player's transition rule.

Definition 3 ***(Incentive Compatibility: Transition Rule)***
If a strategy $\sigma = ((g, \varphi, \tau_b), \tau_r)$ is incentive compatible*, then the transition rule φ satisfies the following condition. For any states s_i, s_j, and $s_{j'} \in \mathcal{M}$:*

$$\begin{aligned}
\varphi_{\bar{a}}(i,j) &> 0 \Rightarrow p_i v_j^c + (1-p_i)\, v_j^r \geq p_i v_{j'}^c + (1-p_i)\, v_{j'}^r, \\
\varphi_{a_l}(i,j) &> 0 \Rightarrow v_j^r \geq v_{j'}^r, \forall a_l \in A_r \backslash \{\bar{a}\}
\end{aligned}$$

In games with imperfect recall there are typically multiple equilibria (even in one person games).[13] Some authors take the view that in decision problems, the ex-ante agent will coordinate his actions in the most profitable equilibrium, as discussed by Aumann et al (1997). In a multi-player game this assumption is less appealing, since instead of ex-ante optimality we look for equilibria. Therefore, in this paper, we characterize all equilibria and point out the ones in which memory states are not used in a redundant way.

3 Full Memory

We first study the game under the assumption that both players have full memory. A version of this game has been studied by Sobel (1985). To construct the perfect Bayesian equilibrium we first

[13] 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.

note that whenever the informed player chooses an action different from $\bar{a}$, the unique equilibrium in the following subgame is the zero-sum equilibrium, since the normal type of the informed player has revealed its type. When the informed player chooses $\bar{a}$, the uninformed player updates his beliefs using Bayes rule. The trade-off for the normal type at every period is between getting a high stage game payoff (by choosing $a' \neq \bar{a}$) or getting a worse stage game payoff but building reputation (by playing $\bar{a}$).

A second property of the equilibrium is that the uninformed player plays a myopic best response to his beliefs and the normal type's strategy at every stage game, regardless of the discount factor δ. For suppose not. Then there exists a history h with corresponding belief on the commitment type ρ_h, in which the uninformed player plays an action $a \in A_r$ with a probability different from the one specified by the myopic best response. The expected continuation payoff for the uninformed player following the play of this stage game is: $p_h U_b(\sigma_b, \tau_c) + (1 - p_h) U_b(\sigma_b, \tau_r)$. If the uninformed player deviates to the myopic best response, his stage game payoff is at least as high and his expected continuation payoff if he sticks to his continuation strategy from the following period onwards is: $p_h U_b(\sigma_b, \tau_c) + (1 - p_h) U_b(\sigma_b, \tau_r')$. Suppose that $U_b(\sigma_b, \tau_r) > U_b(\sigma_b, \tau_r')$. Given that this is a zero-sum game, it immediately implies that $U_r(\sigma_b, \tau_r') > U_r(\sigma_b, \tau_r)$. Therefore, τ_r is not consistent with perfect Bayesian equilibrium, which is a contradiction.

The perfect Bayesian equilibrium is characterized by the informed player choosing $\bar{a}$ with positive probability in all histories but one: when the uninformed player's belief (on the commitment type) is so high that the uninformed player plays a best response against the commitment type's action. When this is the case, the informed player strictly prefers to play some other $a' \neq \bar{a}$.

We have that if $p > \tau_r^*(\bar{a})$, then the uninformed player plays a best response to $\bar{a}$, which we have denoted $b(\bar{a})$. The normal type plays a stage game best response to $b(\bar{a})$, which, given the restrictions that we have made in Γ, is different from $\bar{a}$. The types are then revealed and from the following stage game on, the game is of complete information.

As long as the informed player plays $\bar{a}$, the uninformed player's belief increases monotonically until it becomes greater than $\tau_r^*(\bar{a})$. When this happens, the normal type of the informed player plays a best response to $a \neq \bar{a}$ with probability one and the following continuation game is a complete information game.

Let the equilibrium strategy of the informed player be such that at the i^{th} stage game he plays $\bar{a}$ with probability $\tau_{r,i}(\bar{a})$, after histories in which the actions played by the informed player are only $\bar{a}$ in the last $i-1$ stage games. Moreover, players play (τ_b^*, τ_r^*) at every stage game following a single play of $a \neq \bar{a}$ by the informed player. Formally, for a history in which the informed player has only played $\bar{a}$, the uninformed player updates his beliefs as follows:

$$\rho_{i+1} = \frac{\rho_i}{\rho_i + (1 - \rho_i)\, \tau_{r,i}(\bar{a})},$$

if $\bar{a}$ has been played and

$$\rho_{i+1} = 0,$$

if $a \neq \bar{a}$ was played by the informed player.

Reputation increases monotonically until the normal type reveals itself. When this happens, if it does, the reputation immediately drops to zero where it stays forever. We contrast this game and this equilibrium with the case where the uninformed player has a finite number of memory states, so that he can only hold a finite number of beliefs in equilibrium.

4 Equilibrium Memory Rule

As we have discussed previously, there are typically multiple equilibria in games with imperfect recall. In our game, there are many equilibria in which the memory rule of the uninformed player has redundant states. In particular, there is always an equilibrium in which all memory states induce the same belief. In the appendix we show necessary conditions for *any* equilibria, but in this section we focus on equilibria with non-redundant states. First, we construct a non-trivial equilibrium in the two-state case. Then, we proceed to the general case of n memory states, for which we characterize the equilibrium transition rule of the bounded memory player. We show that it satisfies a weak monotonicity condition. Hence, the resulting updating rule resembles Bayesian updating whenever possible.

From what follows, we label the states in increasing order of continuation payoffs given a commitment type. Thus, if $i > j$ then $v_i^c > v_j^c$. This is w.l.o.g. because all the equilibria are the same up to relabeling.

One of our main results is to show that in the two extreme states, s_1 and s_n, the beliefs of the bounded memory player on the opponent's type are zero, in the lowest state, and very close to one – in some precise sense – in the highest state. This is true even for the simple case of two memory states only. The outcome in the intermediate states is shown to depend critically on the size of the memory and the initial prior on the commitment type.

Throughout the paper we consider only strategies in which all states are reached with positive probability in equilibrium. Suppose, for now, that all states have different v_i^c and, consequently, different v_i^r (memory states with the same expected continuation payoff $v_i^c = v_j^c$ will be shown to be redundant and are considered in the appendix).

4.1 2 Memory States

In this section we restrict attention to the two-memory state case. This is a very special case, since the memory is minimal: one bit only. We can interpret a two-memory state person as someone who thinks only through two categories; he either thinks of his opponent as a "high type" or as a "low type." The resulting equilibrium in this two-state world will help us understand the outcome on the extreme states of more general memories ($n > 2$), which we will examine in detail in the next section.

An updating rule for the two-memory state case is a probability of switching from state 1 to state 2 and vice-versa, after observing each action. A general updating rule is depicted in the figure below.

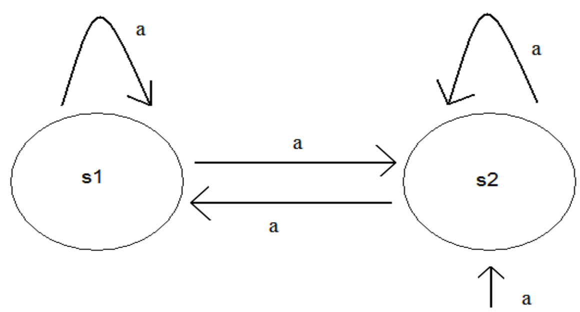

Fig. 1: General Updating Rule

There are multiple equilibria in this two-state case: one equilibria is when both states induce the same belief (the prior) and thus any transition from one state to the other is interim optimal. In this case, it is as if the bounded memory player is playing an infinite sequence of one-shot games. As we have mentioned, in this section we focus on the case where the states are not identical and induce different actions from the bounded memory player. This excludes the case where the prior on the commitment type is very small.[14]

We now construct this equilibrium in what follows. Given our convention, we fix $v_2^c > v_1^c$. Thus, the expected continuation payoff given a normal type is higher in state 1: $v_1^r > v_2^r$ (and, because of the zero-sum game, $U_r(s_2) > U_r(s_1)$). To construct this equilibrium, we will proceed in 4 steps.

[14] For a small prior on the commitment type, the constant play of the unique stage game Nash equilibrium is the only possible incentive compatible equilibrium in the infinite horizon game. We prove this result formally in section 5, where we also show how to find this threshold prior.

First, note that after observing the action $a \neq \bar{a}$, the bounded memory player becomes convinced that the opponent is a normal type, that is, her posterior on the normal type is one. Thus, by incentive compatibility, $\varphi_a(i,1) = 1$, for $i = 1,2$ for $a \neq \bar{a}$.

Thus, the utility of the informed player at state s_2 if she plays $a \neq \bar{a}$ can be written as:

$$U_r(a|s_2) = u_r(\tau_{b,2}, a) + \delta U_r(s_1). \tag{10}$$

Second, in equilibrium, it must be that

$$u_r(\tau_{b,2}, a) \geq u_r(\tau_b^*, \tau_r^*) \geq u_r(\tau_{b,2}, \bar{a}), \tag{11}$$

To prove this, we use the fact that $U_r(s_1) \leq \sum_{j=1}^{2} \varphi_{\bar{a}}(2,j) U_r(s_j)$ and the restrictions on the stage game. Suppose, by contradiction that $u_r(\tau_{b,2}, \bar{a}) > u_r(\tau_{b,2}, a)$, then:

$$\begin{aligned}
U_r(a|s_2) &= u_r(\tau_{b,2}, a) + \delta U_r(s_1) \\
&< u_r(\tau_{b,2}, \bar{a}) + \delta\left(\varphi_{\bar{a}}(2,1) U_r(s_1) + \varphi_{\bar{a}}(2,2) U_r(s_2)\right) \\
&= U_r(\bar{a}|s_2).
\end{aligned}$$

This would imply that the informed player should play $\bar{a}$ with probability one at every history in memory state 2. The myopic best response of the uninformed player should be to play $b(\bar{a})$ and $u_r(b(\bar{a}), a) > u_r(b(\bar{a}), \bar{a})$ by assumption. This leads to a contradiction. Therefore, in equilibrium (11) must hold.

We now proceed to argue that $U_r(a|s_2) > U_r(\bar{a}|s_2)$, but before, note that $U_r(s_i) \geq \frac{u_r(\tau_b^*, \tau_r^*)}{1-\delta}$, for any $s_i \in \mathcal{M}$. Suppose, by contradiction, that $U_r(\bar{a}|s_2) \geq U_r(a|s_2)$. The informed player's payoff at memory state s_2 is $U_r(s_2) = U_r(\bar{a}|s_2)$ and

$$U_r(s_2) \leq u_r(\tau_{b,2}, \bar{a}) + \delta U_r(s_2),$$

Therefore:

$$\begin{aligned}
U_r(s_2) &\leq \frac{u_r(\tau_{b,2}, \bar{a})}{1-\delta} \\
&\leq \frac{u_r(\tau_{b,2}, a)}{1-\delta} \leq U_r(s_1)
\end{aligned}$$

However, note that we have assumed that the memory states do not induce the same payoffs in equilibrium, and, in particular, we have focused on the case in which $U_r(s_2) > U_r(s_1)$, which implies a contradiction. Therefore, $U_r(a|s_2) > U_r(\bar{a}|s_2)$ and the normal type plays $a \neq \bar{a}$ in every history at state s_2. The intuition for this result is that the trade-off between stage-game payoff and reputation incentives does not exist in this highest state: given that the state is the highest one, in the subsequent period the normal type will at most be as well-off as she is in the current period.

The transition rule in memory state s_2 must be such that after observing $\bar{a}$, the bounded memory player chooses to stay in that state, that is: $\varphi_{\bar{a}}(2,2) = 1$. The transition rule must be such that either the bounded memory player starts at memory state s_2, or the probability of moving away from that state after action $\bar{a}$ must be positive. Following the argument, the long-run probability of having a commitment type in state s_1, is zero. This is true since, given a commitment type, the bounded memory player either starts at state s_2, or eventually reaches that state and stays there forever. On the other hand, given a normal type, the bounded memory player will visit state s_1 infinitely often, since whenever she reaches state s_2, she immediately goes back to state s_1. As the stopping probability η goes to zero, or equivalently, as $\delta \to 1$, the expected length of the game increases. Therefore in state s_1 the bounded memory player's belief on a normal type goes to zero as $\delta \to 1$. We prove this result formally in the appendix for a memory with n states.

The equilibrium in the case of a not very small prior and non-redundant memory states is depicted below.

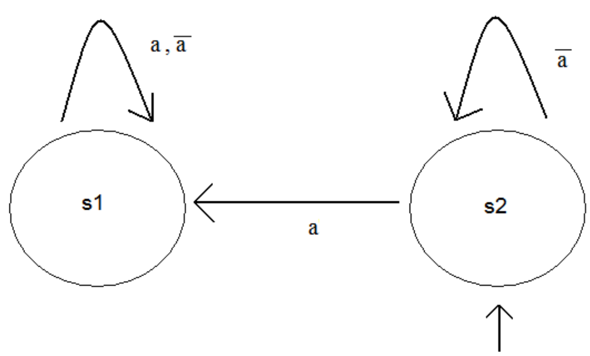

Fig. 2: Equilibrium Rule with Separation

This shows that the bounded memory player, having a very small memory, will start the game with "long run beliefs". Another interesting property of the equilibrium is that the bounded memory player keeps track of the normal types. The normal type will still benefit from the fact that his opponent is boundedly rational, but not because the bounded memory player forgets her actions, but because the bounded memory player doesn't know the period that she is in when she starts the game. In other words, the bounded memory player is confused about the time period when she is in state s_2, so she doesn't know if she has already separated the types.

In the initial state, the bounded memory player not knowing the actual period will think that it is much more likely that the types have already been separated (as $\delta \to 1$ the probability that she attaches to the time period being the first one–the only period in which the normal type is also possible at that state– is very close to zero). This "inflates" the belief in state s_2 and gives the informed player a high payoff in the initial period. For the same reason, conditional on the commitment type, the expected payoff is indeed higher for the bounded memory player than it is for her counter factual Bayesian version.

4.2 n Memory States

Consider now the general case in which the bounded memory player is restricted to n memory states, where $n \geq 2$.

Our main result is shown in proposition 1 below. We show that any equilibrium memory rule will satisfy a weakly increasing property. The equilibrium updating rule is such that the uninformed player separates the types. Since only the normal type can play action $a \neq \bar{a}$, this actions is completely revealing. Thus, the uninformed player's posterior belief after such action is zero. He will then move to his lowest state, and therefore $\varphi_a(i, 1) = 1$, $a \neq \bar{a}$ for any memory state s_i. The same intuition holds for histories in which the normal type strictly prefers to play $a \neq \bar{a}$. In this case, action $\bar{a}$ is completely revealing. The uninformed player then moves to his highest state with probability one.

While the uninformed player might ignore information by not updating after $\bar{a}$, he will never update to a worse belief after observing $\bar{a}$. The uninformed player will get a better payoff from staying in the same state rather than moving to a lower state. One interpretation of this result is that the uninformed player might ignore new information, but he will not forget the information that he already holds.

Finally, the extreme states s_1 and s_n must have beliefs about the opponent's type that are zero and one, or close to one, respectively. The intuition is that at state s_n there are no reputation incentives, thus the normal type of the informed player will reveal himself right away. If the uninformed player is at this memory state, the only chance that the informed player is the normal type is if this is the first stage game being played at this memory state. In other words, the normal type of the informed player will stay in this state for at most one period. On the other hand, if the informed player is a commitment type, the state is absorbing and this type will be in state s_n forever. The probability of being at state s_n for the first time goes to zero as the stopping probability gets smaller. The same argument holds for what happens at state s_1. If this is not the initial state, as we will show for the case of η close to zero, then only the normal type of informed player can reach this state. In this case, the result is obvious. If this was the initial state, the probability of having a normal type at that state would go to one as the stopping probability goes to zero, therefore this state would be absorbing and in equilibrium, it would not be the initial state.

We should remark here that while an equilibrium always exists in our game, we were unable to establish under what conditions an equilibrium in which all states are used in a non-redundant way exists. For that reason, the results of sections 4 and 5 are necessary conditions for equilibria.

We state the result below for the case where the stopping probability is very small, $\delta \to 1$. In the appendix we prove a more general version of the proposition, which holds for any stopping probability η, and which allows for redundant states.

Proposition 1 ***(Updating Rule)***

If the strategy profile $\sigma = ((g, \varphi, \tau_b), \tau_r)$ is an equilibrium with non redundant states, then, we have that:

1. *After observing the action $a \neq \bar{a}$ move to an absorbing state:* $\varphi_a(j, 1) = 1$.
2. *Never go back after observing $\bar{a}$:* $\rho_j > \rho_i \Rightarrow \varphi_{\bar{a}}(j, i) = 0$.
3. *The initial state is the lowest state after the absorbing one:*

$$\lim_{\delta \to 1} g(s_1) = 0 \text{ and } \lim_{\delta \to 1} g(s_2) > 0.$$

4. *The lowest belief approaches zero:* $\lim_{\delta \to 1} \rho_1 = 0$.
5. *The highest belief approaches one:* $\lim_{\delta \to 1} \rho_n = 1$.

In our game, in any equilibrium with non-redundant states, the normal type is revealed either in finite time or asymptotically. The result that types are eventually revealed is derived from results 4 and 5 of the proposition above. They imply that the informed player reveals her type immediately in the case of 2 states – as we have mentioned in the end of section 4.1– and that she will reveal her type at the first period in which the memory state s_n is reached. Recall that in the Bayesian version of this game, that is, with both players fully rational, the informed player reveals her type in finite time.[15] That is, given the stage game and the prior on the commitment

[15] This game was analyzed by Sobel (1985) under a somewhat different stage game –with cheap talk and conflicting preferences, but with the same properties of my game.

type, one can calculate the maximum number of periods before which the normal type reveals her type with probability one. Under bounded memory, depending on the memory size and the initial prior on the commitment type, randomization in the transition rule might be a necessary condition for equilibrium (see section 5), implying that there are arbitrarily long histories on-equilibrium path in which the normal type does not reveal her type. However, these histories are reached with probability that approach zero exponentially. This comes from the fact that at every period the normal type plays $a \neq \bar{a}$ with a positive probability.[16] Thus, similarly to the Bayesian version, types are also revealed in the bounded memory world, either in finite time or asymptotically. We state this result below.

Corollary 1 *For any given prior ρ and a number of memory states n, and any $\varepsilon > 0$, there exists a period $\bar{t} \geq 1$, after which the normal type is revealed with probability at least $1 - \varepsilon$. That is, conditional on the normal type, $f(h_t|\sigma) > 1 - \varepsilon$, for every $h_t \in s_1 \cap \mathcal{H}_r$, $\forall t > \bar{t}$.*

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 types. The proposition below tells how the updating happens after action $\bar{a}$. All the results depend on a condition that the posteriors about the informed player's type are different on the states. To weaken this restriction, in the appendix we prove the following result: $\rho_j > \rho_i \Rightarrow p_j \geq p_i$.

Proposition 2 ***(Monotonicity)***
For any two states $s_i, s_j \in \mathcal{M}$ such that $p_i \neq p_j$, i.e. states with different posteriors, then we have the following.

1. *(Single crossing)* $\varphi_{\bar{a}}(i,k) > 0$, $\varphi_{\bar{a}}(i,m) > 0$ *and* $\varphi_{\bar{a}}(j,k) > 0 \Rightarrow \varphi_{\bar{a}}(j,m) = 0$,

   for $\forall k, m$ *such that* $\rho_k \neq \rho_m$.

2. *(Weak Monotonicity)* $\varphi_{\bar{a}}(i,k') > 0$,$\varphi_{\bar{a}}(i,k'') > 0 \Rightarrow \varphi_{\bar{a}}(j,k) = 0$,

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

3. *(Increasing Rule)* *If* $\varphi_{\bar{a}}(j,m) > 0 \Rightarrow \varphi_{\bar{a}}(i,m') = 0$,

   for $\forall\ m' < m$, *and* $p_i > p_j$.

[16] This is proved in the appendix, Lemma 5 and corollary 1.

Proof. For the single crossing property, suppose that $\varphi_{\bar{a}}(i,k) > 0$ and also that $\varphi_{\bar{a}}(i,m) > 0$. This implies that:

$$p_i(v_k^c - v_m^c) + (1 - p_i)(v_k^r - v_m^r) = 0. \tag{12}$$

Suppose now that $\varphi_{\bar{a}}(j,k) > 0$ and $\varphi_{\bar{a}}(j,m) > 0$, then

$$p_j(v_k^c - v_m^c) + (1 - p_j)(v_k^r - v_m^r) = 0. \tag{13}$$

If $p_i \neq p_j$ then (12) and (13) cannot hold at the same time. Thus, two states must have at most one state in common in their transition rules.

The next step is to show a "weak monotonicity" result for states where p_i and p_j are different. Suppose that $\varphi_{\bar{a}}(i,k+1) > 0$ and $\varphi_{\bar{a}}(i,k-1) > 0$. This implies that:

$$p_i\left(v_{k+1}^c - v_k^c\right) + (1 - p_i)\left(v_{k+1}^r - v_k^r\right) \geq 0, \tag{14}$$

$$p_i\left(v_k^c - v_{k-1}^c\right) + (1 - p_i)\left(v_k^r - v_{k-1}^r\right) \leq 0. \tag{15}$$

If in addition we also have that $\varphi_{\bar{a}}(j,k) > 0$. Then it must be true that :

$$p_j\left(v_{k+1}^c - v_k^c\right) + (1 - p_j)\left(v_{k+1}^r - v_k^r\right) \leq 0, \tag{16}$$

$$p_j\left(v_k^c - v_{k-1}^c\right) + (1 - p_j)\left(v_k^r - v_{k-1}^r\right) \geq 0. \tag{17}$$

The equations above cannot hold for $\rho_{k+1} > \rho_k > \rho_{k-1}$ and $p_i \neq p_j$.

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

$$\varphi_{\bar{a}}(j,m) > 0 \Rightarrow p_j v_m^c + (1 - p_j) v_m^r \geq p_j v_{m\prime}^c + (1 - p_j) v_{m\prime}^r,$$

which means that:

$$p_j(v_m^c - v_{m\prime}^c) + (1 - p_j)(v_m^r - v_{m\prime}^r) \geq 0. \tag{18}$$

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

$$p_i(v_m^c - v_{m\prime}^c) + (1 - p_j)(v_m^r - v_{m\prime}^r) > 0, \tag{19}$$

which proves our last condition. ■

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

As we mentioned before, there are always equilibria with redundant sates. Lemma 1 below shows that we can ignore the redundant states without loss of generality. This result tells us that any equilibrium in which the uninformed player is using a redundant state can be reproduced with a strategy without redundant states. Therefore, we can focus only on rules where all states have different beliefs.

Lemma 1 ***(Redundant States)***
Consider an uninformed player with memory $\mathcal{M}$ that has only n states. The strategy $\sigma = ((g, \varphi, \tau_b), \tau_r)$ gives the uninformed player a payoff of U_b^*. Now suppose that $\rho_i = \rho_j$, for some $s_i, s_j \in \mathcal{M}$. Then, there $\exists\ \sigma'$ for some other memory $\mathcal{M}'$ with $n-1$ states and that also gives the uninformed player a payoff U_b^*.

Proof. Let $\rho_i = \rho_j$. From the proof of proposition 1 this implies that $v_i^r = v_j^r$. Thus, if both states are reached in equilibrium it must be that $v_i^c = v_j^c$. The uninformed player is always completely indifferent between the two states i and j after any action played by the opponent. If $p_i = p_j$, then the states are identical and we can consider them as being a single state (just rewrite the transition rules). If $p_i > p_j$, then they must have the same transition rules, or else $v_i^c = v_j^c$ would not hold. But, if they have the same transition rules then again they are identical and we can group them as one. ■

A class of memory rules that satisfies propositions 1 and 2 is depicted in figure 3 below.

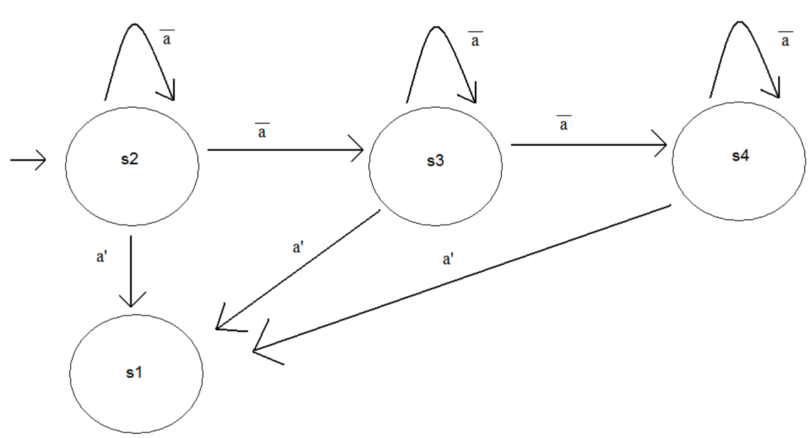

Fig. 3: A Class of Equilibrium Memory Rules

In the following section, we will show the conditions under which the only equilibria with non-redundant memory states involve random updating rule, $\varphi_{\bar{a}}(i, i+1) \in (0, 1)$ for some states s_i, with $n > i > 1$.

5 Random Updating Rule and Inertia

In the previous section, we characterized the equilibrium transition rule. In this section, we show that for sufficiently small prior on the commitment type the equilibrium transition rule will involve randomization, but not at the extreme states. A bounded memory player updates less frequently than a Bayesian player. We interpret this result as indicating that the bounded memory player ignores relevant information as a way of overcoming his limited memory. This infrequent updating in the transient states (the intermediate states) is our interpretation of inertia as a bias in the short run for the uninformed player. It is as if he is being 'skeptical' and testing the informed player before updating him to a higher reputation.

Given a memory of size n, there is a threshold in the prior space such that if the prior is smaller than this threshold, the uninformed player will not use deterministic transition rules.

Proposition 3 ***(Random Transition Rule)***
Given any number of memory states $n > 2$, there exists a threshold on the prior on the commitment type, ρ_n^*, such that if the actual prior is smaller than this threshold, $\rho < \rho_n^*$, then an equilibrium with non-redundant states must involve randomization in the transition rule at the intermediate memory states.

Proof. The proof is by induction. Consider first the two last states, s_{n-1} and s_n. We will compute a threshold on the prior of memory state s_{n-1} such that the uninformed player will choose $\varphi_{\bar{a}}(n-1, n) = 1$.

The informed player will weakly prefer to play $a \neq \bar{a}$ in state s_{n-1} if:

$$u_r(\tau_{b,n-1}, a) + \delta u_r(\tau_b^*, \tau_r^*) \geq u_r(\tau_{b,n-1}, \bar{a}) + \delta u_r(b(\bar{a}), a). \tag{20}$$

We need to find the highest $\tau_{r,n-1}(\bar{a})$ that can support (20). The intuition is that if $\tau_{r,n-1}(\bar{a})$ is too high, the posterior on state s_{n-1} will be too low and the uninformed player will not want to move to state s_n. Thus, we need to consider the uninformed player's incentive compatibility constraint as well. For the uninformed player's incentive compatibility to hold, we need that:

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

Given that at state s_n, the normal type plays $a \neq \bar{a}$, for non-trivial equilibria (in which induced beliefs are different), we must have that that $\rho_n \geq \tau_r^*(\bar{a})$ (otherwise, $\tau_{b,n} = \tau_b^*$ and all states would be identical). If the equilibria is in pure strategies, $p_{n-1} = \rho_n$. Therefore, the prior on memory state s_{n-1} must be such that:

$$\frac{\rho_{n-1}}{\rho_{n-1} + \left(1 - \rho_{n-1}\right) \tau_{r,n-1}(\bar{a})} \geq \tau_r^*(\bar{a}). \tag{21}$$

At memory state s_i, in which $i < n$, it must be that $\rho_i + (1 - \rho_i)\,\tau_{r,i}(\bar{a}) = \tau_r^*(\bar{a})$, since if this equality does not hold, then the bounded memory player would not mix in her actions in that stage game and, as result, the informed player would not either. If $\rho_i + (1 - \rho_i)\,\tau_{r,i}(\bar{a}) > \tau_r^*(\bar{a})$, then $\tau_{b,i}(b(\bar{a})) = 1$ and, the normal type would play $a \neq \bar{a}$, which means that $\tau_{r,i}(\bar{a}) = 0$, which can only be consistent with the inequality if $\rho_i \geq \tau_r^*(\bar{a})$. On the other hand, if $\rho_i + (1 - \rho_i)\,\tau_{r,i}(\bar{a}) < \tau_r^*(\bar{a})$, then $\tau_{b,i}(b(\bar{a})) = 0$, and the normal type should play $\bar{a}$ with probability one, which is a contradiction. Therefore a lower bound on ρ_i is obtained by setting $\rho_i + (1 - \rho_i)\,\tau_{r,i}(\bar{a}) = \tau_r^*(\bar{a})$. Using (21) we get that:

$$\begin{aligned} \rho_{n-1} &\geq \tau_r^*(\bar{a}) \left(\rho_{n-1} + \left(1 - \rho_{n-1}\right) \tau_{r,n-1}(\bar{a})\right) \\ \rho_{n-1} &\geq \left(\tau_r^*(\bar{a})\right)^2 \end{aligned}$$

By induction, we have that:

$$\rho \geq \left(\tau_r^*(\bar{a})\right)^n$$

■

6 Payoffs: The Value of Memory

In this section we compute a bound on the value of memory. We show that the difference between the payoff obtained by an unconstrained uninformed player and the payoff obtained by a bounded memory uninformed player is bounded by the maximum difference of payoffs in one stage game. In equilibrium, a bounded memory player benefits from the type separation, but pays a high screening cost.

Consider the class of two-player finite stage games specified in Section 2 and their incomplete information version with a commonly known prior ρ that the informed player is a commitment type. In the benchmark case where both players are unconstrained in their memory capacities (and otherwise), denote by U_{Bayes} the expected payoff of the uninformed player in the perfect Bayesian equilibrium of the game.

Now consider the same game above but in which the uninformed player is restricted to limited memory. We write $U_b((g, \varphi, \tau_b), \tau_r)$ to be the expected discounted repeated game payoff obtained by the bounded memory player when the incentive compatible equilibrium strategy profile is $((g, \varphi, \tau_b), \tau_r)$.

Before we proceed, let us introduce some new notation. Let $M = \max_{(a_b,a_r)} u_b(a_b, a_r)$ and $m = \min_{(a_b,a_r)} u_b(a_b, a_r)$, $\forall a_b \in A_b, \forall a_r \in A_r$. Define $\mathcal{S}_N$ to be the set of incentive compatible equilibria in the repeated game in which the bounded memory player is restricted to N memory states. Denote by U_N the upper bound on the payoff that the uninformed player restricted to N memory states can achieve in any incentive compatible equilibrium of the repeated game with incomplete information. Formally, define $U_N \equiv \sup\{U_b((g, \varphi, \tau_b), \tau_r) \,|\, ((g, \varphi, \tau_b), \tau_r) \in \mathcal{S}_N\}$.[17] We now prove the following result: $U_N \geq U_2$, $\forall N > 2$. The intuition here is straightforward: for any number of memory states N, there always exists an incentive compatible equilibrium that is equivalent to the most profitable equilibrium of a player restricted to $m \leq N-1$ memory states. The extra state(s) could be off-equilibrium path, or, in case of equilibria in which all states are reached on-equilibrium path, some states could be redundant, as we show below.

Lemma 2 ***(Payoffs: Non-decreasing in the Number of States)***
The upper bound on the payoffs of the equilibria with non-redundant states are non-decreasing in the uninformed player's number of memory states:

$$U_N \geq U_2, \forall N \geq 2.$$

Proof. Consider a memory $\mathcal{M}$ that has only N states. Let $\sigma = ((g, \varphi, \tau_b), \tau_r)$ be an incentive compatible equilibrium (without redundant states, see lemma 1) in such a world. In this equilibrium, denote the expected repeated game payoff of the bounded memory player by U_b^*. Now consider a different game in which the uninformed player is restricted to $N+\Delta$ memory states, with $\Delta \geq 1$. Then, there exists an incentive compatible equilibrium $\sigma' = ((g', \varphi', \tau_b'), \tau_r')$ in which the bounded memory player gets payoff U_b^*. To see this, let the strategy of the bounded memory player, (g', φ', τ_b'), be such that: (1) $g' = g$; (2) $\tau'_{b,i} = \tau_{b,i}$ $\forall i = 1, ..., N$ and $\tau'_{b,j} = \tau_{b,1}$, $\forall j = N+1, ..., N+\Delta$; and (3) $\varphi'_{\bar{a}}(i, j) = \varphi_{\bar{a}}(i, j)$, $\forall i, j = 2, ..., N$, moreover, $\varphi'_a(i, 1) = \varphi_a(i, 1) = 1$, $a \neq \bar{a}, \forall i = 1, ..., N$, but $\varphi'_{\hat{a}}(j, j+1) = \varphi'_{\hat{a}}(1, N+1) = \varphi'_{\hat{a}}(N+\Delta, 1) = 1$, $\forall \hat{a} \in A_r$, $\forall j = N+1, .., N+\Delta-1$. Let the strategy of the normal type be $\tau'_{r,i} = \tau_{r,i}$, $\forall i = 1, ..., N$ and $\tau'_{r,i} = \tau_{r,1}$, $\forall i = N+1, ..., N+\Delta$. It is straightforward to notice that if $\sigma = ((g, \varphi, \tau_b), \tau_r)$ is an incentive compatible equilibrium, then σ' must also be an equilibrium. ■

We are now ready to prove our main result in this section:

[17] Recall that the set $\{U_b((g, \varphi, \tau_b), \tau_r) \,|\, ((g, \varphi, \tau_b), \tau_r) \in \mathcal{S}_N\}$ is bounded, since $\frac{M}{1-\delta} \geq U_N \geq \frac{m}{1-\delta}$, and is non-empty.

Proposition 4 ***(Payoff Difference)***

The difference in the uninformed player's equilibrium payoff in the case in which he is a fully Bayesian player and the case in which he has bounded memory is bounded by the largest difference in the payoffs of the stage game. That is, $\forall N > 1$, $\exists\ \bar{\delta} < 1$, *such that* $\forall \delta > \bar{\delta}$*:*

$$U_{Bayes} - U_N \leq M - m.$$

Proof. To prove this result, we will compare the payoff of a Bayesian uninformed player with the payoff achieved by a bounded memory player restricted to two memory states and we use the result of lemma 2 above. We can write U_{Bayes} as follows:

$$U_{Bayes} = \rho \left\{ u_1^c + \delta u_2^c + ... + \delta^{T-1} u_T^c + \delta^T \frac{u_{T+1}^c}{1-\delta} \right\} + (1-\rho) \left\{ u_1^r + \delta \frac{u_b\left(\tau_b^*, \tau_r^*\right)}{1-\delta} \right\},$$

where $u_t^c \equiv u_b\left(\tau_b^t, \bar{a}\right)$ and τ_b^t is the probability distribution over A_b induced by the equilibrium strategy σ_b following the history $h_{t-1} = \left\{\left(a_b^l, \bar{a}\right)\right\}_{l=1}^{t-1}$, $\forall a_b^l \in A_b$. Similarly, $u_1^r \equiv u_b\left(\tau_b^1, a\right)$, where $a \neq \bar{a}$. We use here two direct implications of the full memory perfect Bayesian equilibrium discussed in Section 3 (see Sobel (1985) for a more detailed construction of such equilibria). In the first term, we note that there exists $T > 0$ such that $u_{t'}^c = u_{T+1}^c$ for any $t' > T$. That is, for a fixed game and prior ρ, there exists a time period after which the subsequent subgame is a complete information game. The second term in the expression above follows from the fact that in the perfect Bayesian equilibrium, the normal type weakly prefers to play action $a \neq \bar{a}$ in the initial period.

Now consider the case in which the uninformed player is constrained to two memory states only. Denote his payoff in the incentive compatible equilibrium in which there are no redundant states by U_2. We can write this term as:

$$U_2 = \rho v_2^c + (1-\rho) v_2^r = \rho \frac{u_b\left(\tau_{b,2}, \bar{a}\right)}{1-\delta} + (1-\rho) \left\{ u_b\left(\tau_{b,2}, a\right) + \delta \frac{u_b\left(\tau_b^*, \tau_r^*\right)}{1-\delta} \right\}.$$

We can write:

$$\begin{aligned} U_{Bayes} - U_2 \quad = \quad & \rho \left\{ \left(u_1^c - u_b\left(\tau_{b,2}, \bar{a}\right)\right) + ... + \delta^{T-1}\left(u_T^c - u_b\left(\tau_{b,2}, \bar{a}\right)\right) + \delta^T \frac{\left(u_{T+1}^c - u_b\left(\tau_{b,2}, \bar{a}\right)\right)}{1-\delta} \right\} \\ & + (1-\rho)\left\{u_b\left(a_b^1, a\right) - u_b\left(\tau_{b,2}, a\right)\right\} \end{aligned}$$

Recall that in equilibrium the uninformed player (Bayesian or not) always chooses a probability over the stage-game-actions that is a myopic best response to the strategy of the normal type given his beliefs at that point in the game (see discussion in Section 2.4 and Section 3). Moreover, recall from our equilibrium construction of the two-memory state case in Section 4 that $\lim_{\delta \to 1} \rho_1 = 0$

and $\lim_{\delta\to 1}\rho_2 = 1$ (see Proposition 5 in the appendix for the detailed proof). These two results together immediately imply that for any fixed finite stage game Γ satisfying the conditions that we have described in Section 2, it must be that $\exists\ \bar{\delta} > 0$ such that for any $\delta > \bar{\delta}$, the equilibrium strategy of the bounded memory player, (g, φ, τ_b), specifies that $\tau_{b,2}$ is the degenerate play of $b(\bar{a})$. Therefore, we have that for any given $\delta > \bar{\delta}$:

$$\begin{aligned} v_2^c &= \frac{u_b(b(\bar{a}), \bar{a})}{1-\delta}, \\ v_2^r &= u_b(b(\bar{a}), a) + \delta \frac{u_b(\tau_b^*, \tau_r^*)}{1-\delta}. \end{aligned}$$

Note that by definition of $b(\bar{a})$ it must be that $u_t^c = u_b\left(\tau_b^t, \bar{a}\right) \leq u_b(b(\bar{a}), \bar{a})$, $\forall t$. Therefore, we have the following result: $\forall \delta > \bar{\delta}$ it must be that:

$$\begin{aligned} U_{Bayes} - U_2 &\leq (1-\rho)\left\{u_b\left(\tau_b^1, a\right) - u_b\left(\tau_{b,2}, a\right)\right\}, \\ &\leq (1-\rho)(M-m). \end{aligned}$$

From lemma (2), we know that $U_N \geq U_2$, which completes the proof. ■

This result reminds us that the value of (even a minimal) memory is quite large: if a player has no memory at all (1 memory state), the unique incentive compatible equilibrium is the constant play of the static Bayesian Nash equilibrium. If the prior on the commitment type is relatively high, the normal type will benefit from the limited resources of his opponent indefinitely. On the other hand, a player with minimal memory (2 states) is able to perfectly screen the opponent in the initial period and thus will pay a cost of at most one-period. Additional memory states (i.e. more than 2 states) may diminish the one-period screening cost, but observing the bound obtained in the previous proposition. In particular, the benefit from additional states is not larger than the difference $(M-m)$.

This stark result stems from the properties of the underlying game and the monitoring assumption. In a game with imperfect monitoring, this result no longer holds. In Monte (2013) we show that the normal type will benefit indefinitely from the bounded memory of his opponent creating cycles of permanent reputations.

7 Conclusion

This paper is a study of bounded memory in a reputation game. It differs from the existing literature on imperfect memory by considering a game in which the memory rule is chosen by the player and satisfies incentive compatibility constraints. Most models of bounded memory assume that, during the play of a game, people have no control whatsoever over what to remember or what to forget. Our view is that, although forgetful, players have some ability over what information to retain. A player that is aware of his limitations will sort the information received and, in particular, may choose to ignore relevant information.

We characterized the updating rule and showed that in this game it is rather simple: monotonic and weakly increasing. We show that for a sufficiently low probability of a commitment type, the updating rule will involve random transition rules in the intermediate states. Our interpretation is that this randomization in the transition rule is being used for two different reasons. First, it is used to overcome the memory problem by not storing all the signals. This intuition was also present in single player games. Second, in a multi-player game, randomization may be used as a strategic element: to test the opponents before updating. We interpreted this infrequent updating as indicating that, in the short-run, a bounded memory player may exhibit inertia in her behavior, updating less frequently than what we would expect from a fully Bayesian player.

8 Appendix

The appendix is divided as follows. First, we show a general version for proposition 1 in the text. The result holds for transition rules that are random or deterministic (in which case it is trivially true). Second, we show under which conditions the equilibria with non redundant memory states must involve random transition rules.

Fix a strategy profile $\sigma = ((g, \varphi, \tau_b), \tau_r)$. This induces an expected continuation payoff for every memory state. Define s_l to be a state with highest expected continuation payoff if the uninformed player is facing a normal type, given σ. Formally: $\mathcal{D}_\sigma \equiv \{s_l \in \mathcal{M} | v_l^r \geq v_i^r, \forall s_i \in \mathcal{M}\}$, similarly define: $\mathcal{U}_\sigma \equiv \{s_u \in \mathcal{M} | v_u^c \geq v_i^c, \forall s_i \in \mathcal{M}\}$.

We label the states in an increasing order of induced reputation. If two states are such that $i > j$ then it must be that $\rho_i \geq \rho_j$. Moreover, if a state is reached with positive probability in equilibrium, then there must not exist another state that has higher expected continuation payoff for the bounded memory player regardless of the type of the informed player (i.e. higher v_i^c and v_i^r). If there exists a state s_j with v_j^c higher than v_i^c it must be that v_j^r is lower than v_i^r, otherwise for whatever posterior the bounded memory player holds, he will find it better to move to state s_j and s_i would never be reached in equilibrium.

Finally, we also have that if $\rho_i \geq \rho_j$ then $v_i^c > v_j^c$. The intuition for this result is that in equilibrium it must be that

$$\rho_i \left(v_i^c - v_j^c\right) + (1 - \rho_i) \left(v_i^r - v_j^r\right) \geq 0,$$

otherwise, the bounded memory player can behave 'as if' he was in state s_j. I.e., he can choose $\tau_{b,j}$ and $\varphi_a(j, j')$, instead of $\tau_{b,i}$ and $\varphi_a(i, j')$, $\forall a \in A_r, \forall s_{j'} \in \mathcal{M}$. Now suppose that $\rho_i > \rho_j$ and $v_i^c < v_j^c$, then

$$\rho_j \left(v_j^c - v_i^c\right) + \left(1 - \rho_j\right) \left(v_j^r - v_i^r\right) \geq 0 \Rightarrow \rho_i \left(v_j^c - v_i^c\right) + (1 - \rho_i) \left(v_j^r - v_i^r\right) \geq 0,$$

which means that state s_i would be identical to state s_j. Thus, $\rho_i > \rho_j$ implies that $v_i^c \geq v_j^c$.

With these results in mind, we can state and prove our main result of the paper. We will denote $U_r(a|s_i)$ as the expected repeated game payoff of the normal type r if he plays $a \in A_r$ and the current memory state of the bounded memory player is s_i.

Proposition 5 ***(Increasing Updating Rule: General version of Proposition 1)***
If the strategy profile $\sigma = ((g, \varphi, \tau_b), \tau_r)$ is an equilibrium, then:

1. *After an action $a \neq \bar{a}$, always move to the lowest state: $\varphi_a(j, l') = 0$, for any state $s_{l'}$ such that $v_{l'}^r > \min_i v_i^r$ and $a \neq \bar{a}$.*

2. *If there is some state $s_i \in \mathcal{M}$, in which the informed player strictly prefers to play $a \neq \bar{a}$, then the transition rule of the bounded memory player must be such that he would move to the highest state after observing $\bar{a}$: $U_r(a \neq \bar{a}|s_i) > U_r(\bar{a}|s_i) \Rightarrow \varphi_{\bar{a}}(i, h\prime) = 0$, for any state $s_{h'}$ such that $v_{h'}^c < \max_i v_i^c$.*

3. *After $\bar{a}$, never move to a lower state: $v_j^c > v_i^c \Rightarrow \varphi_{\bar{a}}(j, i) = 0$.*

4. *Always start at the state with lower reputation or the second lowest: $g(s_1) > 0$ or $g(s_2) > 0$. Moreover: $\lim_{\eta \to 0} g(s_1) = 0$ and $\lim_{\eta \to 0} g(s_2) > 0$.*

5. *The prior belief in the lowest state is close to zero: $\lim_{\eta \to 0} \rho_l = 0$, $\forall s_l \in \mathcal{D}_\sigma$.*

6. *The prior belief in the highest state is close to one:$\lim_{\eta \to 0} \rho_u = 1$, $\forall s_u \in \mathcal{U}_\sigma$.*

We show the proof of this proposition through lemmas 3 to 10. These are conditions that must hold in any equilibrium.

Our first result comes from incentive compatibility. After $a \neq \bar{a}$, the uninformed player knows the actual type of the informed player and moves to a state with the highest expected continuation payoff given a normal type. As defined above, this is a state $s_l \in \mathcal{D}_\sigma$.

Lemma 3 *The uninformed player always moves to the state with lowest continuation payoff after an action other than $\bar{a}$:*

$$s_j \notin \mathcal{D}_\sigma \Rightarrow \varphi_a(i,j) = 0,\ \ a \neq \bar{a} \text{ and } \forall s_i \in \mathcal{M}.$$

Proof. After $a \neq \bar{a}$, the uninformed player knows with probability one that the informed player is the normal type. Thus, by incentive compatibility, whenever $\varphi_a(i,j) > 0$, for $a \neq \bar{a}$ it must be true that $v_j^r \geq v_{j'}^r, \forall s_{j'} \in \mathcal{M}$. ■

Therefore, we can write the payoff of the informed player for not playing $\bar{a}$ at state s_i as:

$$U_r(a|s_i) = u_r(\tau_{b,i}, a) + \delta U_r(s_l). \tag{22}$$

Lemma 4 *In equilibrium, it must be that*

$$u_r(\tau_{b,i}, a) \geq u_r(\tau_b^*, \tau_r^*) \geq u_r(\tau_{b,i}, \bar{a}),$$

for all $s_i \in \mathcal{M}$.

Proof. Suppose that there exists a state s_i such that $u_r(\tau_{b,i}, \bar{a}) > u_r(\tau_{b,i}, a)$, for $a \neq \bar{a}$. We know from lemma 3 that the utility of the informed player at state s_i if he plays $a \neq \bar{a}$ can be written as (22). His utility for playing $\bar{a}$ is:

$$U_r(\bar{a}|s_i) = u_r(\tau_{b,i}, \bar{a}) + \delta \sum_{j=1}^{n} \varphi_{\bar{a}}(i,j) U_r(s_j).$$

Moreover, since $v_l^r = -U_r(s_l)$ it is also true that $U_r(s_l) \leq \sum_{j=1}^{n} \varphi_{\bar{a}}(i,j) U_r(s_j)$. This, in turn implies that $U_r(\bar{a}|s_i) > U_r(a|s_i)$, which implies that the informed player plays $\bar{a}$ with probability one at every history in that memory state. The myopic best response of the uninformed player should be to play $b(\bar{a})$ and $u_r(b(\bar{a}), a) > u_r(b(\bar{a}), \bar{a})$ by assumption. This is a contradiction, and therefore, in equilibrium it must be that $u_r(\tau_{b,i}, a) \geq u_r(\tau_{b,i}, \bar{a})$. If it holds with equality, then $u_r(\tau_{b,i}, a) = u_r(\tau_b^*, \tau_r^*) = u_r(\tau_{b,i}, \bar{a})$, otherwise, given the restrictions in the stage game, it must be that $u_r(\tau_{b,i}, a) > u_r(\tau_b^*, \tau_r^*) > u_r(\tau_{b,i}, \bar{a})$. ■

Lemma 5 *If at a state the normal type has a lower expected stage game payoff for revealing its type than the payoff at state s_n, then it must be that at that state her payoff for revealing his type is at least as high as for mimicking the commitment type. Formally:*

$$u_r(\tau_{b,n}, a) > u_r(\tau_{b,j}, a) \Rightarrow U_r(a|s_j) \geq U_r(\bar{a}|s_j)$$

Proof. Suppose that for some state s_j, $u_r(\tau_{b,j}, a) \geq u_r(\tau_{b,n}, a)$ and the normal type strictly prefers to play $\bar{a}$: $U_r(\bar{a}|s_j) > U_r(a|s_j)$. This implies that the informed player plays $\bar{a}$ with probability one, and thus the reputation in that state is one, $\rho_j = 1$ and the uninformed player plays $b(\bar{a})$. Her expected payoff at that state is $U_r(a|s_j) = u_r(b(\bar{a}), a) + \delta U_r(s_l) \geq U_r(\bar{a}|s_j)$, where the last inequality comes from lemmas 2 and 3 above (the normal type will get a payoff less than $u_r(\tau_b^*, \tau_r^*)$ as long as she plays $\bar{a}$ and will move to state s_l the time that he plays $a \neq \bar{a}$). ■

Given this result, we have that for any state s_j such that $u_r(\tau_{b,n}, a) > u_r(\tau_{b,j}, a)$, the informed player's utility can be written as:

$$U_r(s_j) = U_r(a|s_j) = u_r(\tau_{b,j}, a) + \delta U_r(s_l) < U_r(a|s_n) \tag{23}$$

The next result shows that whenever the informed player reaches the highest state, she will strictly prefer to reveal his type. The intuition for this result is that, given lemma 3, the informed player can get an expected stage game payoff greater than $u_r(\tau_b^*, \tau_r^*)$ at most once in the game. Thus, in the state with highest stage game payoff, her best response is to play $a \neq \bar{a}$.

Lemma 6 *In the highest state the utility of the normal type is strictly higher for playing $a \neq \bar{a}$ (except for the trivial equilibrium where all the states are the same):*

$$U_r(a|s_n) > U_r(\bar{a}|s_n).$$

Proof. We can write the utility of the normal type as:

$$\begin{aligned} U_r(a|s_n) &= u_r(\tau_{b,n}, a) + \delta U_r(s_l), \\ U_r(\bar{a}|s_n) &= u_r(\tau_{b,n}, \bar{a}) + \delta \sum_{j=1}^{n} \varphi_{\bar{a}}(n, j) U_r(s_j). \end{aligned}$$

Suppose, by contradiction, that $U_r(\bar{a}|s_n) \geq U_r(a|s_n)$. Then, by definition, $U_r(s_n) = U_r(\bar{a}|s_n)$. Consider two different cases: i) $U_r(s_n) \geq U_r(s_j)$ for all states s_j such that $\varphi_{\bar{a}}(n, j) > 0$; ii) $U_r\left(s_{j'}\right) > U_r(s_n)$ for some state $s_{j'}$ such that $\varphi_{\bar{a}}(n, j') > 0$.

For the first case, we have that:

$$\begin{aligned} U_r(s_n) &\leq u_r(\tau_{b,n}, \bar{a}) + \delta U_r(s_n), \\ U_r(s_n) &\leq \frac{u_r(\tau_{b,n}, \bar{a})}{1-\delta} \leq u_r(\tau_{b,n}, a) + \delta U_r(s_l). \end{aligned}$$

Where the last step follows directly from lemma 4 (remember that $U_r(s_l) \geq \frac{u_r\left(\tau_b^*, \tau_r^*\right)}{1-\delta}$).

For the second case let $s_{j'}$ be such that $U_r\left(s_{j'}\right) \geq U_r\left(s_i\right)$, $\forall s_i \in \mathcal{M}$ with $\varphi_{\bar{a}}(n,i) > 0$. Together with (23), this gives us the following inequality:

$$\begin{aligned}
U_r\left(\bar{a}|s_n\right) &\leq u_r\left(\tau_{b,n},\bar{a}\right) + \delta U_r\left(s_{j'}\right) \\
&\leq u_r\left(\tau_{b,n},\bar{a}\right) + \delta\left(u_r\left(\tau_{b,j'},a\right) + \delta U_r\left(s_l\right)\right) \\
&\leq u_r\left(\tau_{b,n},a\right) + \delta\left(u_r\left(\tau_{b,l},a\right) + \delta U_r\left(s_l\right)\right) \\
&= U_r\left(a|s_n\right).
\end{aligned}$$

This proves our result. Thus, the normal type reveals himself with probability one in the highest state. ■

Corollary 2 *If the uninformed player plays $b(\bar{a})$ with probability 1 then the normal type strictly prefers to play $a \neq \bar{a}$.*

The reputation in state all states $s_l \in \mathcal{D}_\sigma$ must be the lowest ones. The uninformed player always moves to some state $s_l \in \mathcal{D}_\sigma$ after $a \neq \bar{a}$, which means that

$$\begin{aligned}
v_l^r &= u_b\left(\tau_{b,l},a\right) + \delta v_l^r, \\
v_l^r &= \frac{u_b\left(\tau_{b,l},a\right)}{1-\delta}.
\end{aligned}$$

It must also be true that:

$$v_1^r \geq \frac{u_b\left(\tau_{b,1},a\right)}{1-\delta}.$$

Since, by definition, $v_l^r \geq v_1^r$ it must be that $u_b\left(\tau_{b,l},a\right) \geq u_b\left(\tau_{b,1},a\right)$. Thus, in fact, $u_r\left(\tau_{b,1},a\right) \geq u_r\left(\tau_{b,l},a\right)$.

Lemma 7 *If the uninformed player knows with probability one that the informed player is a commitment type, he will update to the state with highest expected continuation payoff given a commitment type of informed player. In particular:*

$$U_r\left(a|s_i\right) > U_r\left(\bar{a}|s_i\right) \Rightarrow \varphi_{\bar{a}}(i,j) = 0 \textit{ for all states } s_j \textit{ such that } \rho_j < \rho_n.$$

Proof. If the normal type strictly prefers to play $a \neq \bar{a}$ at some state s_i, the uninformed player's posterior on the commitment type after $\bar{a}$ is one: $p_i = 1$. The uninformed player then updates to the state with highest expected payoff given a commitment type, i.e. to the state with highest v_i^c. (This result is in fact stronger: if $\tau_{r,i}(\bar{a}) = 0$ then $\varphi_{\bar{a}}(i,j) = 0$ for all s_j such that $\rho_j < \rho_n$. ■

The lemma below shows that the uninformed player will not move to a lower state after observing $\bar{a}$.

Lemma 8 *Following the action $\bar{a}$, the transition rule is weakly increasing:*

$$\rho_j > \rho_i \Rightarrow \varphi_{\bar{a}}(j,i) = 0.$$

Proof. First note that,

$$\rho_j \left(v_j^c - v_i^c\right) + \left(1-\rho_j\right)\left(v_j^r - v_i^r\right) \geq 0.$$

A posterior following $\bar{a}$ is always such that: $p_k \geq \rho_k$, $\forall s_k$. Therefore:

$$p_i \left(v_j^c - v_i^c\right) + \left(1-p_i\right)\left(v_j^r - v_i^r\right) > 0.$$

■

Lemma 9 *For any prior on the commitment type ρ, $\exists\ \bar{\eta} > 0$, $\bar{\delta} < 1$ such that in any equilibrium with non-redundant states of a game in which $\eta < \bar{\eta}$, and $\delta > \bar{\delta}$, it must be true that state s_1 is absorbing. The reputation in the lowest state approaches zero:*

$$\lim_{\eta \to 0} \rho_1 = 0.$$

Proof. We need to consider only two possibilities of equilibria. First, suppose that $\exists\ \eta^* > 0$ and $k_{\eta^*} < 1$ such that for $\forall \eta < \eta^*$ the equilibrium is such that $\varphi_{\bar{a}}(1,1) < k_{\eta^*}$. We know from lemmas (3) and (8) that as $\eta \to 0$, if the uninformed player starts at state s_1, then given a commitment type of the informed player, the uninformed player eventually leaves that state and never returns. On the other hand, given a normal type of the informed player the uninformed player visits state s_1 infinitely often. Since the probability of remaining in the same state after $\bar{a}$ is bounded by k_{η^*}, as the stopping probability goes to zero the uninformed player will believe that it is unlikely that the informed player is a commitment type if he finds himself in state s_1. Formally, remember that $\rho_i = \sum_{h \in s_i \cap H_c} \mu(h|s_i,\sigma)$, where:

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

For ease of notation, denote $\varphi_{\bar{a}}(1,1)$ simply by φ_1. We can write the prior ρ_i as:

$$\begin{aligned} \rho_i &= \frac{\rho}{\sum\limits_{h' \in s_i} f(h'|\sigma)}\{1 + \delta\varphi_1 + \delta^2\varphi_1^2 + \ldots\} &(24)\\ &= \frac{\rho}{\sum\limits_{h' \in s_i} f(h'|\sigma)}\frac{1}{1-\delta\varphi_1}. \end{aligned}$$

We know that the informed player plays $a \neq \bar{a}$ at memory state s_i with probability $1 - \tau_{r,i}(\bar{a})$. We also know that this probability is bounded above by the stage game Nash Equilibrium probability $\tau_r^*(\bar{a})$. Denote $\tilde{\tau}_r = \max\{\tau_{r,1}(\bar{a}), \tau_{r,2}(\bar{a}), \tau_{r,3}(\bar{a}), ..., \tau_{r,n}(\bar{a})\}$. Then, at all periods the probability that the normal type of the informed player is at memory state s_1 is at least $(1 - \tilde{\tau}_r)$. Therefore, we have a bound on the denominator of (24):

$$\sum_{h' \in s_1} f\left(h'|\sigma\right) \geq (1 - \tilde{\tau}_r)\left(1 + \delta + \delta^2 + ...\right) = \frac{(1 - \tilde{\tau}_r)}{1 - \delta}$$

This, in turn, gives us an upper bound on the belief that the informed player is a commitment type, at the beginning of a stage game when the current memory state is s_1.

$$\rho_i = \frac{\rho}{\sum\limits_{h' \in s_i} f\left(h'|\sigma\right)} \frac{1}{1 - \delta\varphi_1} \leq \frac{\rho}{1 - \delta\varphi_1} \frac{\eta}{(1 - \tilde{\tau}_r)}$$

Finally, we have that as the stopping probability goes to zero:

$$\lim_{\eta \to 0} \rho_1 \leq \lim_{\eta \to 0} \frac{\rho}{1 - \delta\varphi_1} \frac{\eta}{(1 - \tilde{\tau}_r)} = 0$$

Now lets consider the second possibility. Suppose that for any $k < 1$ and any η^* we have that $1 > \varphi_{\bar{a}}(1,1) > k$ for some $\eta < \eta^*$. The uninformed player must be indifferent between staying at state s_1 and moving to some state s_j after $\bar{a}$. Otherwise, either $\varphi_{\bar{a}}(1,1) = 1$ in which case we would have an equilibrium with only one memory state used in equilibrium, or, we could have $\varphi_{\bar{a}}(1,1) = 0$ for any $\eta > \eta^*$, but in this case we are back to the first argument of this proof and $\lim_{\eta \to 0} \rho_1 = 0$. Before the start of the first stage game, the uninformed player decides where in her memory she should start the game. She chooses based on her prior belief ρ. We consider here only the decision between state s_1 and state s_n. We will show that if the equilibrium transition rule $\varphi_{\bar{a}}(1,1) \to 1$ as $\eta \to 0$, then $g(s_1) = 0$, i.e., the uninformed player will not start the game in state s_1. In this case, we have that $\rho_1 = 0$ immediately from lemmas (3) and (8). If the uninformed player starts at state s_1, it must be that her payoff from starting at state s_1 is greater or equal than starting at any other state, in particular at state s_n. Thus:

$$\rho\left(v_n^c - v_1^c\right) + (1 - \rho)\left(v_n^r - v_1^r\right) \leq 0$$

First, note that given the result in lemma (8), we have that once the uninformed player reaches memory state s_n, if the informed player is a commitment type, she will be in that state forever. Thus, we can write:

$$\sum_{h \in s_n \cap H_c} f(h|\sigma) = \frac{\sum_{\tilde{h} \in H} f\left(\tilde{h}|\sigma\right)}{1 - \delta},$$

where $\tilde{h}$ is a history where the uninformed player reaches memory state s_n for the first time, given a commitment type.

The normal type of the informed player plays $\bar{a}$ less often than the commitment type does. Given lemma (3), the probability that the uninformed player reaches state s_n from state s_1 is lower given a normal type than given a commitment type. Moreover, given a normal type, the uninformed player remains in state s_n for only one period and then moves back to state s_1. Thus, we have:

$$\sum_{h \in s_n \cap H_r} f(h|\sigma) \leq \frac{\sum_{\tilde{h} \in H} f\left(\tilde{h}|\sigma\right)}{1 - \delta \sum_{\tilde{h} \in H} f\left(\tilde{h}|\sigma\right)}.$$

Therefore:

$$\begin{aligned} \lim_{\eta \to 0} \frac{\sum_{h \in s_n \cap H_r} f(h|\sigma)}{\sum_{h \in s_n \cap H_c} f(h|\sigma)} &\leq \lim_{\eta \to 0} \frac{\sum_{\tilde{h} \in H} f\left(\tilde{h}|\sigma\right)}{1 - \delta \sum_{\tilde{h} \in H} f\left(\tilde{h}|\sigma\right)} \frac{\eta}{\sum_{\tilde{h} \in H} f\left(\tilde{h}|\sigma\right)} \\ &= \lim_{\eta \to 0} \frac{\eta}{1 - \delta \sum_{\tilde{h} \in H} f\left(\tilde{h}|\sigma\right)} = 0. \end{aligned}$$ [18]

Thus, if $\varphi_1 \to 1$ as $\eta \to 0$, then the belief in state s_n approaches one:

$$\lim_{\eta \to 0} \rho_n = \lim_{\eta \to 0} \frac{\sum_{h \in s_n \cap H_c} f(h|\sigma)}{\sum_{h \in s_n \cap H_c} f(h|\sigma) + \sum_{h \in s_n \cap H_r} f(h|\sigma)} = 1. \tag{25}$$

For equilibria with more than one state, we must have that the informed player is mixing between both actions in state s_1 as we saw in lemma (5).

From lemma (5):

$$v_i^r = u_b(\tau_{b,i}, a) + \delta v_1^r.$$

Therefore,

$$\begin{aligned} v_n^r - v_1^r &= u_b(\tau_{b,n}, a) - u_b(\tau_{b,1}, a) \\ &\leq u_b(b(\bar{a}), a) - u_b(\tau_b^*, \tau_r^*), \end{aligned}$$

where the last inequality comes from lemma (4). The r.h.s. of the inequality above depends on the particular stage game payoffs, not on the equilibrium strategies. In particular, it is not a function of the transition rule φ and also not a function of the stopping probability η.

Moreover, for the commitment type, we have that

$$\begin{aligned} v_1^c &\leq u_b(\tau_{b,1}, \bar{a}) + \delta\{\varphi_1 v_1^c + (1 - \varphi_1) v_n^c\} \\ v_1^c &\leq \frac{u_b(\tau_{b,1}, \bar{a})}{1 - \delta\varphi_1} + \frac{\delta(1 - \varphi_1)}{1 - \delta\varphi_1} v_n^c, \end{aligned}$$

and v_n^c can be written as

$$\begin{aligned}
v_n^c &= u_b(\tau_{b,n}, \bar{a}) + \delta v_n^c \\
v_n^c &= u_b(\tau_{b,n}, \bar{a}) + \delta \{\varphi_1 v_n^c + (1-\varphi_1) v_n^c\} \\
v_n^c &= \frac{u_b(\tau_{b,n}, \bar{a})}{1-\delta\varphi_1} + \frac{\delta(1-\varphi_1)}{1-\delta\varphi_1} v_n^c
\end{aligned}$$

This implies that

$$\begin{aligned}
v_n^c - v_1^c &\geq \frac{u_b(\tau_b(s_n), \bar{a})}{1-\delta\varphi_1} - \frac{u_b(\tau_b(s_1), \bar{a})}{1-\delta\varphi_1} \\
&\geq \frac{u_b(\tau_b(s_n), \bar{a}) - u_b(\tau_b^*, \tau_r^*)}{1-\delta\varphi_1}
\end{aligned}$$

Given (25) we have that

$$v_n^c - v_1^c \geq \frac{u_b(b(\bar{a}), \bar{a}) - u_b(\tau_b^*, \tau_r^*)}{1-\delta\varphi_1}.$$

Finally, note that the l.h.s. of the inequality above is decreasing in η and increasing in φ_1. For a discount factor not too small (as a function of ρ and the stage game payoffs), it follows directly that

$$\rho(v_n^c - v_1^c) + (1-\rho)(v_n^r - v_1^r) \geq 0,$$

for η small enough and φ_1 high enough. Therefore, the bounded memory player would prefer to start at memory state s_n, implying that $\rho_1 = 0$. ■

Lemma 10 *The uninformed player starts either at the lowest memory state or at the second lowest state:*

$$\begin{aligned}
\lim_{\eta\to 0} g(s_1) &= 0, \\
\lim_{\eta\to 0} g(s_2) &> 0.
\end{aligned}$$

Proof. The ex-ante uninformed player chooses in which memory state to start the game. She will start the game at state s_{i_0} that maximizes her ex-ante payoff. Thus, she will choose s_{i_0} such that:

$$\rho v_{i_0}^c + (1-\rho) v_{i_0}^r \geq \rho v_i^c + (1-\rho) v_i^r \quad \text{for } \forall s_i \in \mathcal{M}.$$

Given the results 1 and 3 in proposition (5), we have that $\rho < p_j^T, \forall j > 1$. Thus, if $g(s_{i'}) > 0$,for some $i' > 2$, then state s_2 is not reached with positive probability in the game, except for time $t = 0$. Thus, if $g(s_{i'}) > 0$ and $g(s_2) = 0$ then s_2 would never be reached. On the other hand, if the uninformed player starts at memory state s_1, she will locked in that state forever. Thus, in an equilibrium where she uses all her states, she will start the game in state s_2. ■

References

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

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

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

[4] Cowan, N., "The Magical Number 4 in Short-term Memory: A Reconsideration of Mental Storage Capacity," *Behavioral and Brain Sciences*, 24, 87-185, 2001.

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

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

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

[8] M. Hellman, T.M. Cover, "On Memory Saved by Randomization," *Annals of Mathematical Statistics*, 42, 1075-1078, 1971.

[9] E. Kalai, A. Neme, "The Strength of a Little Perfection," *International Journal of Game Theory*, 20, 335-55, 1992.

[10] E. Kalai, E. Solan, "Randomization and Simplification in Dynamic Decision-Making," *Journal of Economic Theory*, 111, 251-264, 2003.

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

[12] D. Kreps, P. Milgrom, J. Roberts, R. Wilson, "Rational cooperation in the finitely repeated prisoners' dilemma," *Journal of Economic Theory*, vol. 27(2), pages 245-252, 1982.

[13] E. Lehrer, "Repeated Games with Stationary Bounded Recall Strategies," *Journal of Economic Theory*, 46, 130-144, 1988.

[14] Miller, G. A., “The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity For Processing Information,” *Psychological Review*, 63, 81-97, 1956.

[15] D. Monte, “Reputation and Bounded Memory in Repeated Games with Incomplete Information,” Yale University PhD Dissertation, 2007.

[16] D. Monte, “Bounded Memory and Permanent Reputations,” *Journal of Mathematical Economics*, v. 49, p. 345-354, 2013.

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

[18] M. Rabin, “Psychology and Economics,” *Journal of Economic Literature*, Vol. XXXVI, 11-46, 1998.

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

[20] J. Sobel, “A Theory of Credibility,” *Review of Economic Studies*, LII, 557-573, 1985.

[21] A. Wilson, “Bounded Memory and Biases in Information Processing,” Job market paper, Princeton University, 2003.