



Hu, T.-W. (in press). Forgetful updating and stubborn decision-makers. *Economic Theory*. <https://doi.org/10.1007/s00199-022-01423-7>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1007/s00199-022-01423-7](https://doi.org/10.1007/s00199-022-01423-7)

[Link to publication record on the Bristol Research Portal](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Springer at <https://doi.org/10.1007/s00199-022-01423-7> .Please refer to any applicable terms of use of the publisher.

University of Bristol – Bristol Research Portal

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/brp-terms/>



Forgetful updating and stubborn decision-makers

Tai-Wei Hu¹

Received: 23 July 2021 / Accepted: 5 March 2022
© The Author(s) 2022

Abstract

A decision-maker receives an informative signal each period and is randomly required to make a terminal action based on the signals received so far. The decision-maker is restricted to use a (stochastic) finite automaton no larger than a given size to process information. In contrast to the existing literature that focuses on very low probability of termination, I consider information structures with a (nearly) revealing signal, in which analytical solutions are available for all probability values of termination. Results from that model reveal two robust predictions regarding constrained optimal behaviour. First, it is optimal to ignore small (in terms of informativeness) signals. Second, when deterministic schemes are optimal, big signals with similar strengths should be treated similarly; otherwise, randomization takes a lexicographic order according to the strengths of the signals. I also identify a new behavioural bias, information stubbornness, according to which the decision maker does not respond to further informative signals after seeing a nearly revealing signal. As a result, the decision-maker can persistently choose the wrong action even after an unbounded number of informative signals.

Keywords Imperfect recall · Bounded rationality · Bounded memory · Absent-minded · Behavioural biases

JEL Classification D8 · D9

1 Introduction

Recently there has been a surge of interests in behavioural biases in economic agents' decision-making processes. One way to understand such biases is to introduce frictions/capacity constraints in the agent's ability to process information, and derive

I am grateful to Kalyan Chatterjee for discussions that benefit the paper in early stages.

✉ Tai-Wei Hu
taiwei.hu@bristol.ac.uk

¹ University of Bristol, Bristol, UK

behavioural biases as the endogenous responses to those constraints. In particular, Hellman and Cover (1970) and its recent incarnation, Wilson (2014), derive constrained optimal decision rules under limited memory modelled by finite automata constrained by a given number of memory states.¹ Wilson (2014) shows that a constrained decision-maker (DM) only responds to extreme signals but ignores any others, and a positive extreme signal moves one memory state up and a negative one moves one down (although the movements can be randomized), regardless of their relative strengths. However, these results are obtained at the limit case where the expected time horizon for a decision is infinitely far away relative to the memory capacity, and it is not obvious whether or which of these results are robust when the expected time horizon is bounded away from that limit.

In this paper, I revisit these results for a class of information structures that allow for tractable analysis. I adopt the same framework as (Wilson 2014), which features two possible states of nature, H and L . In each period the DM receives an informative signal, and there is a chance that she has to make a terminal decision. Without memory constraints, the optimal decision rule is to take the optimal action with respect to the posterior from the signals received so far according to the Bayes' rule. Instead, a memory constraint is imposed so that only decision rules that are implementable by a finite automaton of a given size are allowed. This consists of finitely many memory states and a transition rule that governs its evolution by specifying the next state to go to (which can be randomized) conditional upon the signal received. Each memory state is associated with an action, which would be taken if a terminal action is called for when the DM happens to land on that memory state.

Wilson (2014) focuses on the limit case where the probability of terminal action, η , vanishes. Assuming that the distribution of signals has a full support, she proves two salient features of the optimal finite automaton of a given size for η sufficiently small. First, small signals are ignored in the sense that they do not trigger a transition to a different memory state. Surprisingly, however, all signals but the most informative signal in either direction are *small* signals, and hence the optimal automaton only reacts to extreme signals. Second, the memory states can be ranked in the sense that the extreme signal (the *high* signal) that increases the posterior on H would cause a transition to a higher memory state, and the extreme signal that decreases it (the *low* signal) would cause a downward transition. Here comes the second surprise: any transition can only occur between adjacent memory states (although sometimes with different probabilities), even though the high signal and the low signal can have very different relative strengths in terms of Bayesian updating.

In contrast to Wilson (2014), I consider arbitrary termination probability for a class of information structures that allow for tractable analysis. First I show that the decision problem in Wilson (2014) with memory constraints implies intrinsic biases: no finite automaton can implement the unconstrained optimal decision rule under the full-support condition. This makes the model intractable and only asymptotic results are

¹ There is a long history of using finite automata to model economic agents. Earlier works include (Rubinstein 1986) on repeated games. Kalai and Solan (2003) show that randomization is necessary in general Markov decision problems when confining implementation to finite automata. More recently, Monte and Said (2014) extends the analysis of Hellman and Cover (1970) to changing worlds where state of nature may switch over time.

available as η vanishes. In contrast, if one of the signals fully reveals one of the states of nature, then the unconstrained optimum is implementable with a finite automaton, whose size depends on the prior belief. A two-signal information structure with one of them being fully revealing is called a “model of breakthroughs,” with the revealing signal being the “breakthrough” signal.²

In the model of breakthroughs I obtain analytical solutions for all η 's. In particular, in any optimal finite automaton, the revealing signal, say the low signal, will cause a transition to the lowest memory state which is absorbing. I then show that this feature remains for the constrained optimal finite automaton when the low signal is sufficiently informative, but it does not need to be fully revealing. This new behavioural bias, called *information stubbornness*, is not included in Wilson (2014)'s results, although I also maintain the full-support condition. This feature is proved using results from the model of breakthroughs and the modified multi-self consistency, a concept invented by Piccione and Rubinstein (1997) and then extended by Wilson (2014) to the framework here. Multi-self consistency implies that in the optimal finite automaton, each memory state is associated with a posterior (potentially biased) belief, and both the optimal action and optimal transitions are determined by a version of “sequential rationality” according to those beliefs. This necessary condition for optimality then traces biases in behaviour back to the biases in the associated beliefs. A key observation here is that these beliefs are continuous in the underlying parameters governing the information structure. Since it is strictly optimal to transit to the lowest memory state when the low signal is fully revealing, the main part of the proof shows that it remains so when the low signal is sufficiently informative.

Information stubbornness highlights a new behavioural bias implied by limited memory. Without memory constraints, the DM would revert to the action corresponding to state of nature H when she receives sufficiently many high signals. In contrast, the constrained optimal rule dictates the action corresponding to L once a low signal is received, without responding to any signals thereafter, despite the fact that they are informative. Furthermore, the result also differs significantly from those in Wilson (2014), who shows that any transition can only occur between adjacent memory states at the limit where η is arbitrarily close to zero. In contrast, information stubbornness implies that the transition “jumps” from any memory state to the lowest one when a low signal is received, and it stays there thereafter.

Besides information stubbornness, I obtain two other analytical features of the constrained optimal rule by adding a third signal to the model of breakthroughs. First, I consider the situation where the third signal is sufficiently uninformative, and show that it is optimal to completely ignore it in the constrained optimal rule. This result is consistent with Wilson (2014)'s result for η close to zero but here it is generalized to all η in my environment. Thus, besides information stubbornness, the bias “sticky information” identified in Wilson (2014) still holds here. Second, I consider the case where the two unrevealing signals have similar strengths in the model of breakthroughs. In contrast to Wilson (2014), the DM treats signals of similar strengths similarly, albeit in a lexicographical order. When randomization is optimal, the stronger signal is more

² This terminology is borrowed from Che and Mierendorff (2019).

fluid—it moves the memory state to the next until the weaker one is completely sticky. However, both signals can be fluid even for η arbitrarily small.

2 The model

The model is essentially (Wilson 2014). There are two states of nature, $\theta \in \{H, L\}$, and the prior probability over $\theta = H$ is $\mathbf{P}_0(H) = p_0$. The model has an unbounded number of periods, and in each period, with probability η the DM has to choose between two actions, $a \in A = \{a^H, a^L\}$, with utility function $u(a, \theta)$ given by

$$u(a^H, H) = u^H > 0, u(a^L, L) = u^L > 0, u(a^H, L) = 0 = u(a^L, H).$$

Once the action is chosen the game is over. With probability $1 - \eta$ the game continues. Note that the chance to choose the terminal action is exogenously given.

The state of nature, however, is not observable to the DM. Instead, the DM can observe a sequence of signals, which is i.i.d. conditional on the state of nature. Here I focus on information structures with two signal, $\mathcal{S} = \{h, \ell\}$. A decision rule is a function $D : \mathcal{S}^* \rightarrow A$, where $\mathcal{S}^* = \bigcup_{t=0}^{\infty} \mathcal{S}^t$ is the set of all partial histories of signal realizations, including the empty one. The decision rule maps each possible partial history of signal realizations to the terminal action if the DM is called upon to make one. The unconstrained optimal rule can be fully characterize by the posterior, p : the optimal decision is to take a^H whenever $p > p^* \equiv u^L / (u^H + u^L)$ and to take a^L whenever $p < p^*$. The posterior is computed according to the Bayes rule, for which it is convenient to work with likelihood ratios. Define

$$\rho_0 = \frac{p_0}{1 - p_0}, \rho^* = \frac{p^*}{1 - p^*} = \frac{u^L}{u^H}, \text{ and } \xi(s) = \frac{\mu_s^H}{\mu_s^L} \text{ for both } s = h, \ell. \quad (1)$$

I normalize the labels so that $\xi(h) > 1 > \xi(\ell)$. I use ρ to denote the likelihood ratio $p / (1 - p)$ for a generic posterior p on H , and from ρ and signal s , the likelihood ratio for the new posterior is $\rho' = \rho \xi(s)$. Hence, h increases the posterior on H while ℓ increases the posterior on L .

Finite automata and implementation

Given the set of signals, \mathcal{S} , a stochastic finite-state automaton (SFSA) consists of a list $M = \langle Q, \tau, d, g \rangle$, where Q is the set of *memory states*, $\tau : Q \times \mathcal{S} \rightarrow \Delta(Q)$ is the *transition rule*, $d : Q \rightarrow \Delta(A)$ is the *action rule*, and $g \in \Delta(Q)$ is the distribution over initial states. I use $\tau(q, s; q')$ to denote the transition probability from q to q' when receiving s , and $d(a; q)$ to denote the probability of taking action a at memory state q when called upon for terminal action. When the finite automaton is deterministic (abbreviated as DFSA), I use $\tau(q, s) = q'$ to denote the transition rule and $d(q) = a$ to denote the decision rule. The following result fully characterizes the signal structures for which a finite automaton can implement the unconstrained optimum.

Proposition 2.1 *Suppose that $\xi(h) < \infty$ and that $\eta \in (0, 1)$. Then, the unconstrained optimum is implementable with a finite automaton if and only if $\xi(\ell) = 0$.*

Proposition 2.1 is proved by the use of the Myhill–Nerode Theorem (Nerode 1958), which gives a full characterization of what finite automata of a given size can do in terms of partitions of the partial histories.³ For self-containment I give the relevant version of the theorem in the Appendix. Proposition 2.1 then shows that except for the case where one of the two signals fully reveals the state of nature, no finite automaton can implement the unconstrained optimum, and hence the constrained optimal behaviour must exhibit some biases relative to the full rationality benchmark.

The impossibility result can be easily extended to more than two signals, as more signals can only make the problem more complicated. As a benchmark, in the next section I study a model where the ℓ -signal does fully reveal the state of nature, and in Section 4 I use the results from there to characterize the constrained optimal behaviour when the ℓ -signal is not fully revealing but only nearly so.

3 A model of breakthroughs

Here I consider the case where $\mu_h^H = 1$, that is, $\xi(\ell) = 0$.⁴ Later I will extend the analysis to allow for $\mu_h^H < 1$. To simplify notation, denote $\mu_\ell^L \equiv \mu$ with $\mu \in (0, 1)$. A low signal, ℓ , fully reveals the state of nature L (hence a *breakthrough*), while a high signal, h , only gradually increases the posterior on H . Without memory constraint, the optimal rule would dictate that action a^L to be taken whenever an ℓ -signal appears and thereafter, while a^H is taken only after sufficiently many h -signals without seeing any ℓ -signal. As shown later, this can be implemented by a finite automaton with sufficiently many memory states. I will also characterize the constrained optimal rule when memory is constrained.

First, suppose that $\rho_0 \geq \rho^*$, where the likelihood ratios are defined in (1). In this case, the unconstrained optimum can be implemented by a two-state DFSA with $Q = \{q_H, q_L\}$, $q^o = q_H$, $\tau(q_H, h) = q_H$, $\tau(q_H, \ell) = q_L$, $\tau(q_L, h) = q_L = \tau(q_L, \ell)$, and $d(q_H) = a^H$ and $d(q_L) = a^L$, as depicted in Fig. 1. Thus, the memory state q_L is self-absorbing for both signals, and the action is a^L ; q_H is absorbing only for signal- h , and the action is a^H . This DFSA, labelled M_2^b , with q_H as the initial state, implements the unconstrained optimum: since $\rho_0 \geq \rho^*$, it is optimal to take a^H when receiving high signals, and it is optimal to switch to a^L and continue using that whenever a low signal is received. Note that this implementation does not depend on η .

Now suppose that $\rho_0 < \rho^*$. In this case, the unconstrained optimum can still be implemented by a DFSA, but its size depends on the distance between $\ln \rho_0$ and $\ln \rho^*$ relative to $\ln \xi(h)$. In particular, let

³ I am grateful to Pathikrit Basu for pointing out that the Myhill–Nerode Theorem can be useful for this result.

⁴ By assuming that $\mu_s^\theta > 0$ for all s and for both $\theta = H, L$, Wilson (2014)’s analysis excludes this case. However, as we will see below, this is not a trivial case and it helps us solve some of the cases under the full support condition, but gives a rather different implications than what Proposition 4 in Wilson (2014) suggests.

Fig. 1 The DFSA, M_2^b , that implements unconstrained optimum when $\rho_0 \geq p^*$

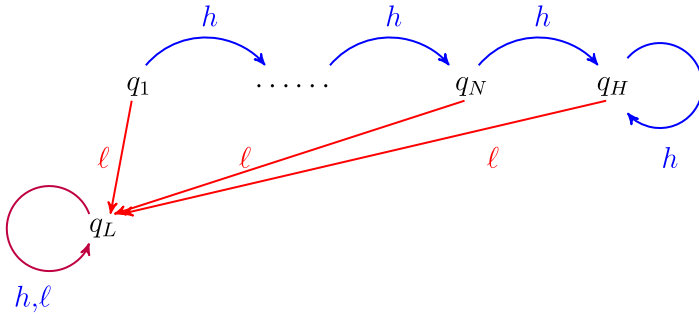
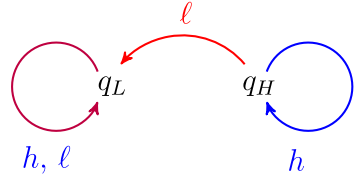


Fig. 2 The DFSA, M_{N+2}^b , that implements unconstrained optimum when $\rho_0 < p^*$

$$N = N(\rho_0) \equiv \left\lceil \frac{\ln \rho^* - \ln \rho_0}{\ln \xi(h)} \right\rceil. \tag{2}$$

Then, the optimal DFSA requires $N + 2$ states, given by $Q = \{q_H, q_L, q_1, \dots, q_N\}$, $\tau(q_L, \ell) = q_L = \tau(q_L, h)$, $\tau(q_i, h) = q_{i+1}$ and $\tau(q_i, \ell) = q_L$ for $i = 1, \dots, N - 1$, $\tau(q_N, h) = q_H$ and $\tau(q_N, \ell) = q_L$, and $\tau(q_H, h) = q_H$ and $\tau(q_H, \ell) = q_L$, as depicted in Fig. 2, with q_1 as the initial memory state. The action rule is given by $d(q_L) = d(q_i) = a^L$ for all $i = 1, \dots, N$, and $d(q_H) = a^H$. This DFSA, labelled M_{N+2}^b , implements the unconstrained optimum for all $\eta \in (0, 1)$. When $N = 0$, this is equivalent to M_2^b , and hence we can set $N(\rho_0) = 0$ in for $\rho_0 \geq \rho^*$.

These results then show that whether the memory constraint, $|Q| \leq K$, binds or not depends only on ρ_0 but not on η , and it binds if and only if $K < N(\rho_0) + 2$. Conversely, for a given K , let ρ_0^K be the smallest ρ_0 such that $N(\rho_0) + 2 \leq K$. Then, the memory constraint is binding if and only if $\rho_0 < \rho_0^K$. The following proposition shows that strict randomization is optimal if and only if the memory constraint is binding and the optimal SFSA is not trivial. In fact, it shows that the following SFSA is optimal, which deviate from M_K^b by having $\tau(q_i, h; q_i) = \alpha = 1 - \tau(q_i, h; q_{i+1})$ for $i = 1, \dots, K - 2$, where $q_{K-1} = q_H$, while all the other transitions follow the same rule as M_K^b . This SFSA is denoted by $M_K^b(\alpha)$ and $M_K^b = M_K^b(0)$, and Fig. 3 depicts $M_K^b(\alpha)$ for $K = 5$. Finally, define M_L as the finite automaton with one memory state that takes action a^L all the time.

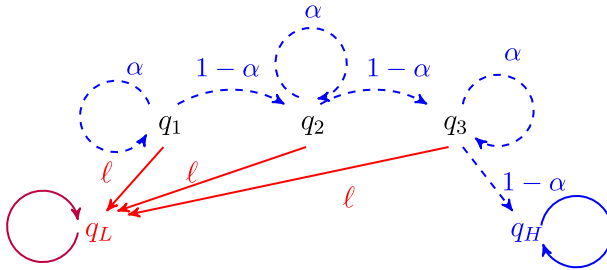


Fig. 3 $M_K^b(\alpha)$ with $K = 5$

Proposition 3.1 *Suppose that $\xi(\ell) = 0$ and $\xi(h) \in (0, 1)$. For each $K \geq 3$ and $\eta \in (0, 1)$, there exists $\underline{\rho}_0 > 0$ such that the smallest optimal SFSA is*

$$\begin{cases} M_{N(\rho_0)+2}^b & \text{if } \rho_0 \geq \rho_0^K, \\ M_K^b(\alpha) \text{ for some } \alpha \in (0, 1) & \text{if } \rho_0 \in [\underline{\rho}_0, \rho_0^K), \\ M_L & \text{if } \rho_0 \leq \underline{\rho}_0. \end{cases} \tag{3}$$

Moreover, $\underline{\rho}_0$ converges to zero as η goes to zero.

Proposition 3.1 gives a full characterization of optimal SFSA in the model of breakthroughs. The optimal SFSA identified uniquely determines the constrained optimal decision rule in each case (except at $\rho_0 = \underline{\rho}_0$), and is the smallest SFSA to implement the corresponding rule. It shows that strict randomization is optimal if and only if the memory constraint binds, that is, $\rho \in (\underline{\rho}_0, \rho_0^K)$. For lower priors, the optimal rule is to take a^L all the time, which reflects the fact that randomization cannot fully substitute for efficient learning. The lower bound for priors under which randomization is optimal, $\underline{\rho}_0$, does depend on η , although its upper bound, ρ_0^K , does not. At the threshold $\underline{\rho}_0$, $M_K^b(\alpha^*)$ under optimal α^* and M_L give exactly the same ex ante payoffs. However, as η converges to zero, the payoff from $M_K^b(\alpha)$ for any $\alpha < 1$ converges to $p_0u^H + (1 - p_0)u^L$, the highest payoff possible, and this implies that $\underline{\rho}_0$ converges to zero. To my knowledge this is the first complete characterization of constrained optimal rules for a class of information structure for all η and for all priors.

4 Information stubbornness

In this section I show information stubbornness when ℓ is sufficiently informative, but not fully revealing, according to which it is optimal for the DM to ignore any further signals once an ℓ -signal is received—hence the stubbornness. To this end, I first need some structural results from Wilson (2014) to characterize optimal rule under memory constraint. Given a state of nature θ and a memory state $q \in Q$, the expected payoff

accumulated from q when called to act conditional on θ is then

$$\begin{aligned} & \eta g(q) \left[\sum_{a \in A} d(a; q) u(a, \theta) \right] + \eta(1 - \eta) \sum_{q_1, s_1 \in \mathcal{S}} g(q_1) \tau(q_1, s_1; q) \mu_{s_1}^\theta \\ & \left[\sum_a d(a; q) u(a, \theta) \right] + \eta(1 - \eta)^2 \sum_{q_1, q_2 \in \mathcal{Q}, s_1, s_2 \in \mathcal{S}} g(q_1) \tau(q_1, s_1; q_2) \mu_{s_1}^\theta \\ & \tau(q_2, s_2; q) \mu_{s_2}^\theta \left[\sum_a d(a; q) u(a, \theta) \right] + \dots \\ & = f(q|\theta) \left[\sum_a d(a; q) u(a, \theta) \right], \end{aligned} \tag{4}$$

where

$$f(q|\theta) = \sum_{T=1}^{\infty} \eta(1 - \eta)^{T-1} \left[\sum_{(q_1, \dots, q_{T-1}), (s_1, \dots, s_{T-1}), q_T=q} g(q_1) \prod_{t=1}^{T-1} \mu_{s_t}^\theta \tau(q_t, s_t; q_{t+1}) \right]. \tag{5}$$

As noted in Wilson (2014), $f(q|\theta)$ is the stationary distribution under the transition probability from q' to q given by

$$T^\theta(q'; q) = \sum_{s \in \mathcal{S}} [\eta g(q) + (1 - \eta) \mu_s^\theta \tau(q', s; q)]. \tag{6}$$

Wilson (2014), extending (Piccione and Rubinstein 1997), then defines the “belief” associated with memory state $q \in \mathcal{Q}$ as

$$\begin{aligned} p(q) &= \frac{p_0 f(q|H)}{p_0 f(q|H) + (1 - p_0) f(q|L)} \text{ and } p(q, s) \\ &= \frac{p_0 f(q|H) \mu_s^H}{p_0 f(q|H) \mu_s^H + (1 - p_0) f(q|L) \mu_s^L}, \end{aligned} \tag{7}$$

with the associated likelihood ratios given by

$$\rho(q) = \rho_0 f(q|H) / f(q|L) \text{ and } \rho(q, s) = \rho(q) \xi(s).$$

To characterize optimal SFSA, I use $V_q(\theta)$ to denote the continuation value when the current memory state is at q conditional state of nature θ , which is characterized by the following recursive equations:

$$V_q(\theta) = \eta \sum_{a \in A} d(a; q) u(a, \theta) + (1 - \eta) \sum_{s \in \mathcal{S}, q' \in \mathcal{Q}} \mu_s^\theta \tau(q, s; q') V_{q'}(\theta). \tag{8}$$

Following (Wilson 2014), two memory states q and q' are called *equivalent* if $V_q(\theta) = V_{q'}(\theta)$ for both $\theta = H, L$.⁵

Proposition 4.1 *Suppose that M is an optimal SFSA with memory constraint K and without equivalent states. Then, we rank the memory states in M according to*

$$\rho(q_1) \leq \rho(q_2) \leq \dots \leq \rho(q_K),$$

with the convention that if $\rho(q_i) = \rho(q_{i+1})$ then $V_{q_i}(H) \leq V_{q_{i+1}}(H)$. Let $\Delta V_{i,j}^\theta = V_{q_i}^\theta - V_{q_j}^\theta$.

1. $\Delta V_{i,j}^H < 0$ and $\Delta V_{i,j}^L > 0$ for all $i < j$, and $\Delta V_{j,i}^H / \Delta V_{i,j}^L \geq \Delta V_{k,j}^H / \Delta V_{j,k}^L$ for all $i < j < k$.
2. Define $\bar{\rho}_i = \Delta V_{i,i+1}^L / \Delta V_{i+1,i}^H$, $i = 1, \dots, K - 1$. Then, in M ,
 - (a) for each q_i ,

$$\rho(q_i) \in [\bar{\rho}_{i-1}, \bar{\rho}_i]; \tag{9}$$

- (b) $\tau(q, s; q_i) > 0$ only if

$$\rho(q, s) \in [\bar{\rho}_{i-1}, \bar{\rho}_i], \tag{10}$$

where $\bar{\rho}_0 = 0$ and $\bar{\rho}_K = \infty$;

- (c) $d(q_i; a^h) > 0$ only if $\rho(q_i) \geq \rho^*$ and $d(q_i, a^\ell) > 0$ only if $\rho(q_i) \leq \rho^*$.

Proposition 4.1 follows directly from Corollary 1 and Lemma 1 in Wilson (2014). For self-containment, however, a full proof of Proposition 4.1 can be found in the Online Appendix. As in Wilson (2014), the proof is mainly based on an extended version of the multiself consistency proposed by Piccione and Rubinstein (1997), and I prove that version in the Online Appendix as well. Proposition 4.1 states that in the optimal SFSA, the DM essentially uses the beliefs $\rho(q)$ to decide the optimal transition and the optimal actions if called upon. Note that Proposition 4.1 applies to *all* information structures, including the model of breakthroughs and information structures that feature full supports.

Now I turn to information stubbornness and consider the case where $\mu_h^H < 1$ but close, and hence the signal ℓ is no longer a breakthrough but still a strong signal for state of nature L . Recall the threshold $\underline{\rho}_0$ given by Proposition 3.1.

Proposition 4.2 *For each $K \geq 3$, $\eta \in (0, 1)$, and $\rho_0 > \underline{\rho}_0$, there exists $\epsilon > 0$ such that for all $\mu_h^H \in (1 - \epsilon, 1]$, any optimal SFSA transits to the lowest memory state after ℓ from any memory state, which is absorbing.*

According to Proposition 4.2, for any given $\eta \in (0, 1)$ and $\rho > \underline{\rho}_0$, for a range of information structures that satisfy full-support, when receiving signal ℓ it is optimal to

⁵ As argued in Wilson (2014), one can obtain the same ex ante expected payoffs by merging equivalent states into a single memory state.

transit all the way back to q_L for all memory states in any optimal SFSA. This is in great contrast to the conclusion from Wilson (2014), where transition only occurs between adjacent memory states for small η 's. The difference lies in the order according to which the limit is taken. In Wilson's case it is to fix $\mu_h^H < 1$ and to take η to zero, in my case it is to fix η (no matter how small) and to consider μ_h^H close to one.

Proposition 4.2 also implies that the DM's behaviour exhibits behavioural biases relative to the full-rationality benchmark. In particular, consider a situation where the DM receives an ℓ -signal at an early stage and hence is stuck in the memory state q_L . Her belief then stays the same, regardless of the signals she receives afterwards. In contrast, a Bayesian DM would shift her beliefs upwards if more h -signals come. In fact, if the underlying state of nature is H , then the constrained DM would almost surely get stuck with a biased belief fully convinced of L , while the Bayesian DM's belief would almost surely converge to H . Note that the result holds for any given K . This finding also highlights a difference between modelling limited memory by finite automata and by bounded recall. In models of bounded recall, only most recent experiences count.⁶ In contrast, here the ℓ -signal can occur some time ago while the DM remains stubborn.

The result is proved by showing that the optimal SFSA takes the form given by (3) in the relevant range of priors. Since Proposition 3.1 also ensures uniqueness, by continuity one only needs to check against small deviations to the transition probabilities in the optimal SFSA. This is guaranteed essentially by the continuity of the beliefs identified in Proposition 4.1 applied to the model of breakthroughs. The proof verifies that when $\mu_h^H = 1$, it is strictly optimal to transit to q_L for all memory states, and this optimality is maintained when μ_h^H is close to one by continuity.

Three signals

Here I extend the results to information structures with three signals where $\mathcal{S} = \{h, h', \ell\}$. For expositional simplicity I only consider the case where $\mu_h^H + \mu_{h'}^H = 1$, i.e., ℓ -signal fully reveals state of nature L as in the model of breakthroughs but signals h and h' both increase the posterior on H . However, all the results here can be extended to the case where μ_h^H is close to one, as in Proposition 4.2. I use two benchmark cases to illustrate how different strengths of the signals h and h' affect optimal randomization. I normalize the parameters so that $\xi(h) > \xi(h')$, that is, h -signal is a stronger signal than h' . When η is close to zero, Wilson (2014) shows that transition can only happen when receiving h but not h' . Here I show that this is true for all η when $\xi(h')$ is sufficiently close to one, that is, when h' -signal is very uninformative. In contrast to her result, however, for $\xi(h')$ close to $\xi(h)$, I show that it is optimal not to ignore h' , but randomization may occur. I also characterize how the randomization is shared between h and h' .

To state these results, I fix $\xi(h)$ but vary $\xi(h')$ to be between 1 and $\xi(h)$, and I do so by parametrizing the conditional probabilities as follows. Fix $\mu_h^H = 1 - \nu = 1 - \mu_{h'}^H \in (0, 1)$ and fix $1 - \mu_h^L = 1 - \mu = \mu_\ell^L + \mu_{h'}^L \in (0, 1)$ with $\mu > \nu$. For

⁶ For a recent example of models of bounded recall, see Barlo et al. (2009) in the context of repeated games.

a given pair $\mu > \nu$, I vary $\mu_{h'}^L$ within the interval $[(1 - \mu)\nu/(1 - \nu), \nu]$. Note that $\mu > \nu$ implies that $\xi(h) = (1 - \nu)/(1 - \mu) > 1$. For all $\mu_{h'}^L \in ((1 - \mu)\nu/(1 - \nu), \nu)$, $\xi(h') < \xi(h)$. When $\mu_{h'}^L = \nu$, $\xi(h') = 1$, that is, h' -signal is uninformative but h -signal is informative. When $\mu_{h'}^L = (1 - \mu)\nu/(1 - \nu)$, $\xi(h') = \xi(h) > 1$ and both signals are equally informative.

In the two extreme cases the optimal SFSA is characterized almost in exactly the same way as in Proposition 3.1. When $\xi(h') = 1$, Proposition 3.1 directly applies by amending the SFSA so that $\tau(q, h') = q$ for all q , that is, the h' -signal is completely ignored. In the other extreme where $\xi(h') = \xi(h)$, the result still holds but with a slightly more subtle modification: randomization at q_i under the optimal SFSA, $M_K^b(\alpha)$, can be split arbitrarily between the two signals as long as

$$\begin{aligned} \mu_h^H \tau(q_i, h; q_i) + \mu_{h'}^H \tau(q_i, h'; q_i) &= \alpha, \\ \text{and } \tau(q_i, h; q_i) + \tau(q_i, h; q_{i+1}) &= 1 = \tau(q_i, h'; q_i) + \tau(q_i, h'; q_{i+1}), \end{aligned} \tag{11}$$

for all $i = 1, \dots, K - 2$, with $q_{K-1} = q_H$. That is, the *average* probability of staying at q_i is equal to α across the signals h and h' . Note that when $\alpha = 0$ is optimal this implies $\tau(q_i, h; q_i) = 0 = \tau(q_i, h'; q_i)$. Let $M_K^b(\beta, \beta')$ denote the SFSA such that $\tau(q_i, h; q_i) = \beta$ and $\tau(q_i, h'; q_i) = \beta'$ for all $i = 1, \dots, K - 2$, with $q_{K-1} = q_H$. Note that $M_K^b(\beta, 1)$ implies that h' -signal is ignored everywhere. Note also that, since $\xi(h)$ is the same in both extreme cases, the unconstrained optimal decision rule is implementable if and only if $\rho_0 \geq \rho_0^K$, where ρ_0^K is determined in exactly the same way as in Section 3 (that is, ρ_0^K is the lowest ρ_0 such that $N(\rho_0) \leq K - 2$, with $N(\rho_0)$ given by (2)). Interestingly, however, for $\mu_{h'}^L$ slightly below ν , the unconstrained optimal rule is not implementable even for $\rho_0 \in (\rho_0^K, \rho^*)$, since for $\xi(h')$ very close to one, it will require more than K h' -signals to cross ρ^* from ρ_0 . For $\rho_0 < \rho_0^K$, the constraint K will be binding for both extreme cases. Note that we still have the threshold $\underline{\rho}_0$ below which taking a^L all the time is the constrained optimal rule, but that threshold would depend on $\mu_{h'}^L$. We have the following proposition.

Proposition 4.3 *Suppose that $\mu_h^H = 1 - \nu = 1 - \mu_{h'}^H \in (0, 1)$ and that $1 - \mu_{h'}^L = 1 - \mu = \mu_{h'}^L + \mu_{h'}^L \in (0, 1)$ with $\mu > \nu$.*

1. *For each $\rho_0 \in (\underline{\rho}_0, \rho^*)$, there exists $\epsilon > 0$ such that for all $\mu_{h'}^L \in [\nu - \epsilon, \nu]$, the optimal SFSA takes the form $M_K^b(\alpha, 1)$, i.e., h' is ignored.*
2. *For each $\rho_0 \in (\underline{\rho}_0, \rho_0^K)$, there exists $\epsilon > 0$ such that for all $\mu_{h'}^L \in (\frac{\nu(1-\mu)}{1-\nu}, \frac{\nu(1-\mu)}{1-\nu} + \epsilon]$, the optimal SFSA takes the form $M_K^b(\beta, \beta')$ with optimal $\beta' = \tau(q_i, h'; q_i) > 0$ for all i . Moreover, in the optimal $M_K^b(\beta, \beta')$, $\beta > 0$ only if $\beta' = 1$.*

Proposition 4.3 only considers priors in the range $(\underline{\rho}_0, \rho^*)$. For $\rho_0 \geq \rho^*$, M_2^b is still optimal. For $\rho_0 < \underline{\rho}_0$, it is still optimal to use M_L . For priors in between, Proposition 4.3 (1) shows that it is optimal to ignore signal- h' when it is sufficiently uninformative, i.e., when $\xi(h')$ is close to the unity. Note that this implies a different kind of biases when the memory constraint is binding. Indeed, when $\rho_0 > \rho_0^K$ but close and when

$\xi(h') > 1$ but close, a rational DM will take action a^H after seeing sufficiently many h' 's but not the constrained DM, who would stay at q_1 and take action a^L . This result extends Wilson's (2014) result that small signals should be ignored, but here it is shown for all η 's in information structures that are close to the breakthroughs.

Proposition 4.3 (2) deals with the case where $\xi(h')$ is close to $\xi(h)$ but slightly lower. Note that for $\rho_0 \geq \rho_0^K$, the unconstrained optimal rule is still implementable when $\xi(h')$ is sufficiently close to $\xi(h)$. According to Proposition 4.3 (2), for lower ρ_0 , randomization is optimal and it would first occur when receiving h' before it occurs when receiving h ; in other words, randomization between staying in the current memory state and the next when receiving the more informative signal h can happen only when it is optimal to ignore the less informative signal, h' . Note that this last point essentially follows from the characterization result, Proposition 4.1, according to which $\beta = \tau(q_i, h; q_i) \in (0, 1)$ implies that $\rho(q_i, h) = \rho(q_i)\xi(h) = \bar{q}_i$ for all $i = 1, \dots, K - 2$, which in turn implies that $\rho(q_i, h') = \rho(q_i)\xi(h') < \bar{q}_i$.

Proposition 4.3 only considers the two extreme cases. For the case where $K = 3$, the model becomes tractable and we have the following proposition.

Proposition 4.4 *Suppose that $\mu_h^H = 1 - v = 1 - \mu_{h'}^H \in (0, 1)$ and that $1 - \mu_h^L = 1 - \mu = \mu_\ell^L + \mu_{h'}^L \in (0, 1)$ with $\mu > v$, and that $K = 3$. For each $\mu_{h'}^L \in (\frac{v(1-\mu)}{1-v}, v)$, there exists $\underline{\rho}_0$ and $\tilde{\rho}_0 \in [\underline{\rho}_0, \frac{\rho^*}{\xi(h)})$ such that the optimal SFSA is $M_b^K(\beta, \beta')$ with*

$$\begin{cases} \beta = 0 = \beta' & \text{if } \rho_0 \in [\frac{\rho^*}{\xi(h')}, \xi^*), \\ \beta = 0 \text{ and } \beta' > 0 & \text{if } \rho_0 \in (\tilde{\rho}_0, \frac{\rho^*}{\xi(h')}), \\ \beta > 0 \text{ and } \beta' = 1 & \text{if } \rho_0 \in (\underline{\rho}_0, \tilde{\rho}_0). \end{cases} \quad (12)$$

According to Proposition 4.4, the optimal SFSA is $M_K^b(\beta, \beta')$ in the three-signal environment under the model of breakthroughs whenever the memory constraint is binding (for priors not low; otherwise M_L is optimal). Note that when $\rho_0 \geq \rho^*/\xi(h')$ the unconstrained optimal rule is implementable. For lower priors, the second line in (12) shows that at q_1 both h and h' are fluid signals, and neither is ignored. This is in contrast to Wilson's (2014) result where only extreme signals are fluid, and note that for this result I do not really need $\xi(\ell) = 0$ but only need ℓ -signal to be sufficiently informative as in Proposition 4.2.

5 Conclusion

An important point in Wilson (2014) is to establish that potential behavioural biases are in fact rational responses to decision-makers' imperfections in information processing abilities. Here I demonstrated that these implications do depend not only on the constraints imposed by such imperfections, but also the underlying environment, such as the probability of terminal actions. My results suggest that two behaviour implications are robust: first, it is almost universally true that constrained optimal rule ignores small signals; second, the DM should treat big signals according to their informativeness, and, when it comes to optimal randomization, use a lexicographical order.

Moreover, a novel behavioural prediction is identified here, information stubbornness, and I have shown that this is a robust feature of the constrained optimal rule when there is a sufficiently informative signal, regardless of the expected horizon of the decision problem. This bias shows that a single experience can lock in a memory-constrained decision-maker’s belief without responding to future informative signals.

Finally, results in the paper are proved mainly based on the extended principle of modified multiself consistency proposed by Piccione and Rubinstein (1997). Hence, I suspect that most of them would survive in environments where that principle also holds true. For example, the result that small signals are ignored only depend on the fact that the belief associated with a memory state according to the modified multi-self consistency lies strictly within the boundaries given by Proposition 4.1, which is a direct consequence of multi-self consistency under binary states of nature.

Appendix: Proofs

Before the proof of Proposition 2.1, I first give a version of the Myhill–Nerode theorem adopted to the framework in the paper. To do so, I need to introduce some concepts first. A relation over the set of partial histories of signal realizations, $R \subset \mathcal{S}^* \times \mathcal{S}^*$, is called *right-invariant* if

$$\mathbf{x}R\mathbf{y} \iff (\mathbf{x} \circ \mathbf{z})R(\mathbf{y} \circ \mathbf{z}) \text{ for all } \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{S}^*, \tag{13}$$

where \circ denotes concatenation. Given a decision rule D , \mathcal{L}_θ^D denotes the set of partial histories under which action a^θ is taken, $\theta = H, L$.

Theorem 5.1 (Myhill–Nerode Theorem) *The rule D can be implementable by a DFSA iff there is a right-invariant equivalence relation R with finitely many equivalence classes such that \mathcal{L}_θ^D is a union of some of those equivalence classes for both $\theta = H, L$.*

The equivalence classes correspond to the memory states in the corresponding DFSA, and the equivalence classes that make up \mathcal{L}_θ^D consist of those action states where action a^θ is taken, for both $\theta = H, L$. Thus, the DFSA gives a finite partition of partial histories that captures the finiteness of the DM’s memory capacity. The right-invariance condition captures the fact that if the DFSA enters the same memory state after two different partial histories, then it will end up in the same memory state (although not necessarily the same as the original one) after any consecutive partial history.

Proof of Proposition 2.1

Note that the “if” part is proved in the analysis of the model of breakthroughs. For the “only if” part, under the assumption of two signals with $\xi(h) \in (1, \infty)$ and $\xi(\ell) \in (0, 1)$, I show that under the unconstrained optimal rule D , any right-invariant equivalence relation R such that \mathcal{L}_θ^D is a union of some of those equivalence classes for both $\theta = H, L$, it must have infinitely many equivalence classes. The “only if” part then follows from the Myhill–Nerode theorem.

Now, let R be such an equivalence relation. Without loss of generality I assume that $\rho_0 \leq \rho^*$. Now, I construct an infinite sequence of partial histories, $\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^n, \dots$ and show that each has to be of a different equivalence class. The construction is simple: each \mathbf{x}^n consists of m_n h -signals, with the sequence $\{m_n\}_{n=1}^\infty$ determined as follows. Let m_1 be the smallest integer $m \geq 1$ such that

$$\ln(\rho_0) + m \ln[\xi(h)] > \ln(\rho^*).$$

Given m_n, m_{n+1} is the smallest integer m such that

$$(m - m_n) \ln[\xi(h)] > -\ln[\xi(\ell)]. \quad (14)$$

This implies that m_n strictly increases with n . Now we show that, for each $n \geq 2$, \mathbf{x}^n must belong to a different equivalence class other than the one \mathbf{x}^i belongs to for all $i < n$. To see this, let k be the smallest integer such that

$$\ln(\rho_0) + m_{n-1} \ln[\xi(h)] + k \ln[\xi(\ell)] < \ln(\rho^*).$$

This implies that $D(\mathbf{x}^i \circ \mathbf{y}) = a^L$ for all $i < n$, where \mathbf{y} consists of k ℓ -signals. However, by (14),

$$\begin{aligned} & \ln(\rho_0) + m_n \ln[\xi(h)] + k \ln[\xi(\ell)] \\ & > \ln(\rho_0) + m_{n-1} \ln[\xi(h)] + (k - 1) \ln[\xi(\ell)] \geq \ln(\rho^*), \end{aligned}$$

and hence $D(\mathbf{x}^n \circ \mathbf{y}) = a^H$. Thus, by right-invariance, \mathbf{x}^n must belong to a different cell than \mathbf{x}^i for all $i < n$.

Proof of Proposition 3.1

We first show that any optimal SFSA takes the form of $M_K^b(\alpha_1, \dots, \alpha_{K-2})$, in which $\sigma(q_i, h)(q_i) = \alpha_i = 1 - \sigma(q_i, h)(q_{i+1})$ for $i = 1, \dots, K - 2$ with $q_{K-1} = q_H$, with other transitions the same as M_K^b . Note that $M_K^b(\alpha)$ is a special case where $\alpha_1 = \dots, \alpha_{K-2} = \alpha$. We may assume that there are no redundant states. By Proposition 4.1, we can rank the memory states. For notational consistency, I use q_L to denote the one with the lowest associated belief and q_H the highest, and rank the others by $\rho(q_1) \leq \rho(q_2) \leq \dots \leq \rho(q_{K-2})$. Now, since an ℓ -signal sends the posterior to zero, it follows that $\rho(q_L) = 0$, and for q_L it is optimal to stay at q_L after any signal. For other memory state, since receiving h can only trigger a transition to go up or to stay, $V_{q_i}(\theta)$ only depends on $V_{q_j}(\theta)$ for $j > i$, while $\rho(q_i)$ only depends on transitions from q_j with $j < i$. Thus, we can take $\rho(q_i)$ as the prior and look for the optimal SFSA with $K - i$ memory states, and use the necessary conditions in Proposition 4.1 to characterize the optimal SFSA. Note that at q_H , the highest memory state, it is optimal to stay there and take action a^H after seeing h -signal.

This allows for an induction argument starting from $i = K - 2$ and then working backwards. We first show by induction that the optimal SFSA the transitions from

q_i, \dots, q_K take the form $M^{b,K-i+1}(\alpha_i, \dots, \alpha_{K-2})$ (part (a)), and then we show that the optimal SFSA has $\alpha_1 = \alpha_2 = \dots = \alpha_{K-2} = \alpha \in (0, 1)$ (part (b)).

(a) For $i = K - 2$ this is immediate, as the only possible transition is either staying in q_{K-2} or going to q_H after seeing an h -signal. Suppose that it holds for $i + 1 \leq K - 2$, and in the optimal SFSA the transitions from q_{i+1}, \dots, q_{K-2} take the form $M^{b,K-i}(\alpha_{i+1}, \dots, \alpha_{K-2})$. Below in (b) we prove that it is optimal to have $\alpha_{i+1} = \dots = \alpha_{K-2}$. This, by a simple computation from the corresponding value functions, implies that for all $j = i + 1, \dots, K - 3, \bar{\rho}_{j+1}/\bar{\rho}_j$ is constant and

$$1 < \bar{\rho}_{j+1}/\bar{\rho}_j \leq \xi(h), \text{ with equality iff } \alpha_{i+1} = 0 = \dots = \alpha_{K-2}, \tag{15}$$

that is, $\bar{\rho}_{j+1}/\bar{\rho}_j = \xi(h)$ iff it is the DFSA $M^{b,K-i}$. Moreover, we have $\rho(q_{i+1}) < \bar{\rho}_{i+1}$: otherwise, $\rho(q_{i+1}) = \bar{\rho}_{i+1}$ and hence $\rho(q_{i+1}, h) = \bar{\rho}_{i+1}\xi(h) > \bar{\rho}_{i+1}$, that is, $\alpha_{i+1} = 0$ and, by (15), this implies that the transitions follow the DFSA $M^{b,K-i}$ and, furthermore, $\rho(q_{i+1}, h) = \bar{\rho}_{i+1}\xi(h) = \bar{\rho}_{i+2}$. This last equality also implies that from q_{i+1} following signal h there is indifference between moving to q_{i+2} and q_{i+3} , implying that there is a redundant updating state. This concludes that $\rho(q_{i+1}) < \bar{\rho}_{i+1}$.

Now, $\rho(q_{i+1}) < \bar{\rho}_{i+1}$ implies, by (15), that $\rho(q_{i+1})\xi(h) < \bar{\rho}_{i+2}$. Since $\rho(q_i) \leq \rho(q_{i+1})$ and $\rho(q_i, h) = \rho(q_i)\xi(h) \leq \rho(q_{i+1})\xi(h) < \bar{\rho}_{i+2}$, by Proposition 4.1, from q_i and an h -signal, randomization can only occur between q_i and q_{i+1} , or between q_{i+1} and q_{i+2} , but not both. Now we show that if the latter happens, then we can eliminate q_{i+1} without affecting the ex ante expected payoff and there is a redundant state; as a result, only the former matters and hence the optimal SFSA takes the form $M^{b,K-i+1}(\alpha_i, \dots, \alpha_{K-2})$ from q_i on. To see this, suppose, by contradiction, that the optimal SFSA randomizes between q_{i+1} and q_{i+2} from q_i after h . This implies that $\rho(q_i) < \rho(q_{i+1})$, and that, by Proposition 4.1, $\rho(q_i, h) = \bar{\rho}_{i+1}$, which in turn implies that $\rho(q_{i+1}, h) > \rho(q_i, h) = \bar{\rho}_{i+1}$. By Proposition 4.1 the last inequality implies that $\alpha_{i+1} = 0$, and by the symmetry noted above, in $M^{b,K-i}(\alpha_{i+1}, \dots, \alpha_{K-2})$ we have $\alpha_{i+1} = \dots = \alpha_{K-2} = 0$. That is, we have a deterministic scheme from q_{i+1} on. Moreover, this also implies that at prior $\rho(q_i)$, $M^{b,K-i}$ and $M^{b,K-i-1}$ give exactly the same payoff, and hence q_{i+1} is a redundant state.

(b) Here I show that it is optimal to set $\alpha_1 = \dots = \alpha_{K-2} \in (0, 1)$. First we compute the value functions under $M_K^b(\alpha_1, \dots, \alpha_{K-2})$:

$$V_{q_i}(H) = \left[\prod_{j=i}^{K-2} \frac{1 - \alpha_j}{1 - (1 - \eta)\alpha_j} \right] (1 - \eta)^{K-i-1} u^H; \tag{16}$$

$$V_{q_H}(L) = \frac{(1 - \eta)\mu u^L}{1 - (1 - \eta)(1 - \mu)}; \tag{17}$$

$$V_{q_i}(L) = \left\{ 1 - \left[\prod_{j=i}^{K-2} \frac{1 - \alpha_j}{1 - (1 - \mu)(1 - \eta)\alpha_j} \right] \frac{\eta(1 - \eta)^{K-i-1}(1 - \mu)^{K-i-1}}{1 - (1 - \eta)(1 - \mu)} \right\} u^L. \tag{18}$$

Note that this implies that the ex ante payoff, $p_0 V_{q_1}(H) + (1 - p_0) V_{q_1}(L)$, is symmetric in $(\alpha_1, \dots, \alpha_{K-2})$ and supermodular, and hence it is optimal to set $\alpha_i = \alpha$ for all i . By doing so, the ex ante payoff is

$$F(\alpha) = p_0 \left[\frac{1 - \alpha}{1 - (1 - \eta)\alpha} \right]^{K-2} (1 - \eta)^{K-2} u^H + (1 - p_0) \times \left\{ 1 - \left[\frac{1 - \alpha}{1 - (1 - \mu)(1 - \eta)\alpha} \right]^{K-2} \frac{\eta(1 - \eta)^{K-2}(1 - \mu)^{K-2}}{1 - (1 - \eta)(1 - \mu)} \right\} u^L. \quad (19)$$

Thus,

$$F'(\alpha) = p_0 \left[\frac{1 - \alpha}{1 - (1 - \eta)\alpha} \right]^{K-3} \frac{-(K - 2)\eta(1 - \eta)^{K-2} u^H}{[1 - (1 - \eta)\alpha]^2} + (1 - p_0) \left[\frac{1 - \alpha}{1 - (1 - \mu)(1 - \eta)\alpha} \right]^{K-3} \frac{(K - 2)\eta(1 - \eta)^{K-2}(1 - \mu)^{K-2} u^L}{[1 - (1 - \eta)(1 - \mu)\alpha]^2}. \quad (20)$$

Thus, $F'(\alpha) > 0$ at $\alpha = 0$ if and only if

$$\frac{p_0}{1 - p_0} < \frac{u^L}{u^H} (1 - \mu)^{K-2},$$

that is, if and only if the memory constraint is binding. Since the best DFSA is $M_K^b = M_K^b(0)$ (other than taking a^L all the time), this shows that strict randomization is better. Note that this also shows that $M_K^b(\alpha)$ is the optimal SFSA for a range of priors below ρ_0^K . Finally, I show the existence of $\underline{\rho}_0$ above which taking a^L all the time is not optimal. For each $\rho_0 \leq \rho_0^K$, let $W(\rho_0)$ be the payoff from $M_K^b(\alpha)$ with the optimal α . By the Theorem of Maximum, $W(\rho_0)$ is continuous in ρ_0 . Now, at $\rho_0 = \rho_0^K$, since M_K^b implements the unconstrained optimum, we have $W(\rho_0^K) > (1 - p_0)u^L$, the latter being the payoff from taking a^L always. Now, $W(\rho_0) \geq (1 - p_0)u^L$ if and only if $F(\alpha^*) \geq (1 - p_0)u^L$, and by (19), this is equivalent to

$$\rho_0 \geq \left[\frac{1 - (1 - \eta)\alpha^*}{1 - (1 - \mu)(1 - \eta)\mu^*\alpha^*} \right]^{K-2} \frac{\eta(1 - \mu)^{K-2}}{1 - (1 - \eta)(1 - \mu)} \rho^*.$$

But we also know that α^* satisfies the FOC with $F'(\alpha^*) = 0$, and by (20), this is then equivalent to

$$1 \geq \frac{1 - (1 - \mu)(1 - \eta)\alpha^*}{1 - (1 - \eta)\alpha^*} \frac{\eta}{1 - (1 - \eta)(1 - \mu)}.$$

Since α^* strictly decreases with ρ_0 , there exists a unique $\underline{\rho}_0 < \rho_0^K$ such that $W(\rho_0) > (1 - \rho_0)u^L$ for all $\rho_0 \in (\underline{\rho}_0, \rho_0^K]$. Finally, from (19) it is straightforward to verify that

$$\lim_{\eta \rightarrow 0} F(\alpha) = p_0u^H + (1 - p_0)u^L$$

for any $\alpha \in [0, 1)$ and hence $\underline{\rho}_0$ converges to zero as η approaches zero.

Proof of Proposition 4.2

Let $\eta \in (0, 1)$ be given. When $\mu_h^H = 1$, from Proposition 3.1 for any $\rho > \underline{\rho}_0$ the optimal SFSA is $M_K^b(\alpha)$ with if $\rho_0 < \rho_0^K$, and the optimal SFSA is $M_{N(\rho_0)+2}^b$ if $\rho_0 \geq \rho_0^K$. Note that in the latter case, when $K > N(\rho_0) + 2$, there can be multiplicity of optimal SFSA up to redundant memory states (i.e., one can split q_i to $q_{i,1}$ and $q_{i,2}$ with identical transition and action rule as q_i , and hence with $V_{q_{i,1}}(\theta) = V_{q_{i,2}}(\theta)$ for both $\theta = H, L$). I prove the results in two parts, first for $\rho_0 < \rho_0^K$ and the second for $\rho_0 \geq \rho_0^K$.

Part (i) I show that if $\rho_0 < \rho_0^K$, then for a range of μ_h^H below one the optimal SFSA still takes the form of $M_K^b(\alpha)$. To prove this, I first show that small deviations from $M_K^b(\alpha)$ are not optimal. More precisely, consider each SFSA (with K memory states) as a vector of all transition probabilities and action probabilities, and consider the distance between them as the maximum difference across those probabilities, and I show that there exists ϵ_0 such that optimal SFSA takes the form of $M_K^b(\alpha)$ among SFSA that deviate from those within distance ϵ_0 .

To simplify notation, denote μ_h^H by $1 - \epsilon$ and μ_ℓ^L by μ . Since the beliefs $\rho(q_i)$ and the value functions $V_{q_i}(\theta)$ are all continuous in μ_h^H and since when $\mu_h^H = 1$, in $M_K^b(\alpha)$ we have $\rho(q_L) = 0$, it follows that for small deviation of the transition probabilities, we still have $\rho(q_i, \ell) < \bar{q}_0$ for μ_h^H slightly below one. Thus, by Proposition 4.1, optimal SFSA has to take the form of $M_K^b(\alpha_1, \dots, \alpha_{K-2})$. We can then compute the continuation values for $M_K^b(\alpha_1, \dots, \alpha_{K-2})$: $V_{q_L}(H) = 0$, $V_{q_L}(L) = u^L$, and

$$V_{q_i}(H) = \left[\prod_{j=i}^{K-2} \frac{1 - \alpha_j}{1 - (1 - \eta)(1 - \epsilon)\alpha_j} \right] [(1 - \eta)(1 - \epsilon)]^{K-i-1} u^H; \tag{22}$$

$$V_{q_H}(L) = \frac{(1 - \eta)\mu u^L}{1 - (1 - \eta)(1 - \mu)}, \quad V_{q_H}(H) = \frac{\eta u^H}{1 - (1 - \eta)(1 - \epsilon)}; \tag{23}$$

$$V_{q_i}(L) = \left\{ 1 - \left[\prod_{j=i}^{K-2} \frac{1 - \alpha_j}{1 - (1 - \mu)(1 - \eta)\alpha_j} \right] \frac{\eta(1 - \eta)^{K-i-1}(1 - \mu)^{K-i-1}}{1 - (1 - \eta)(1 - \mu)} \right\} u^L. \tag{24}$$

By the same arguments as before, it is optimal to set $\alpha_i = \alpha$ for all i , and hence

$$F(\alpha) = p_0 \left[\frac{1 - \alpha}{1 - (1 - \eta)(1 - \varepsilon)\alpha} \right]^{K-2} [(1 - \eta)(1 - \varepsilon)]^{K-2} u^H \tag{25}$$

$$+ (1 - p_0) \left\{ 1 - \left[\frac{1 - \alpha}{1 - (1 - \mu)(1 - \eta)\alpha} \right]^{K-2} \frac{\eta(1 - \eta)^{K-2}(1 - \mu)^{K-2}}{1 - (1 - \eta)(1 - \mu)} \right\} u^L. \tag{26}$$

The function $F(\alpha)$ is quasi-concave in $\alpha \in [0, 1]$, and hence its local optimum is continuous in ε . To see this, note that, disregarding some constants, $F'(\alpha)$ is proportional to

$$\begin{aligned} F'(\alpha) &\simeq - \frac{(1 - \alpha)^{N-1}}{[1 - (1 - \eta)(1 - \varepsilon)\alpha]^{N+1}} \left[\frac{\rho_0}{\rho^*} \xi(h)^N \right] + \frac{(1 - \alpha)^{N-1}}{[1 - (1 - \eta)(1 - \mu)\alpha]^{N+1}} \\ &= \left[\frac{(1 - \alpha)^{N-1}}{[1 - (1 - \eta)(1 - \varepsilon)\alpha]^{N+1}} \right] \left\{ \frac{[1 - (1 - \eta)(1 - \varepsilon)\alpha]^{N+1}}{[1 - (1 - \eta)(1 - \mu)\alpha]^{N+1}} - \left[\frac{\rho_0}{\rho^*} \xi(h)^N \right] \right\}, \end{aligned}$$

and it can be verified that the second term in the last equation strictly decreases with α whenever $\varepsilon < \mu$, and hence can have at most one solution to $F'(\alpha) = 0$ besides $\alpha = 1$. Note that $\alpha = 1$ is excluded by the fact that at $\varepsilon = 0$, the optimal SFSA is $M_K^b(\alpha^*)$ with $\alpha^* \in [0, 1)$, and hence it is cannot be optimal for ε small.

Now, let ε be small so that we have local optimality. To ensure global optimality, consider the set of SFSA M that does not take the form $M_K^b(\alpha_1, \dots, \alpha_{K-1})$ and is at least ε_0 -distance from any of those, denoted by D . This set is compact, and, when $\mu_h^H = 1$,

$$W_{M_K^b(\alpha^*)} > \max_{M \in D} W_M. \tag{27}$$

By continuity again, there exists $\varepsilon_1 \in (0, \varepsilon]$ such that for any $\mu_h^H \in [1 - \varepsilon, 1]$, (27) still holds.

Part (ii) Here I show that if $\rho_0 > \rho_0^K$, then optimal SFSA takes the form M_{N+2}^b with $N = N(\rho_0)$ for a range of μ_h^H below one. The proof follows the same outline as in (i), but now I need to handle replicate memory states as $K > N(\rho_0) + 2$. To do that, notice that in any replica of M_{N+2}^b the beliefs for any replica of q_L and q_i remains the same for $i = 1, \dots, N$ as their correspondents, but the replica for q_H can be higher. However, whatever that belief may be, it is finite and $p(q_H, \ell) = 0$ when $\mu_h^H = 1$ and hence it is still optimal to transit to q_L for a range of μ_h^H below one.

Proof of Proposition 4.3

Proof of (1). Here I consider $\mu_{h'}^L = v - \varepsilon$ for ε small. Let $\rho_0 \in (\underline{\rho}_0, \rho^*)$ and K be given. When $\varepsilon = 0$, the optimal SFSA is $M_{\min\{N(\rho_0)+2, K\}}^b(\beta, 1)$ for appropriate β . One can then compute the value functions and associated beliefs, which are very similar to those computed in the Proof of Proposition 3.1, and in the optimal SFSA, $\rho(q_i) < \bar{\rho}_i$

for all i . Using the same arguments as in the Proof of Proposition 4.2 and continuity of the beliefs and value functions, we can restrict attention to local deviations in the same sense as in the Proof of Proposition 4.2. But this then implies that $\rho(q_i, h') < \bar{\rho}_i$ under any local deviations and for ε small, as a result, by Proposition 4.1, the optimal SFSA has the form of $M_K^b(\beta, \beta')$ with $\beta' = 1$. The fact that it is optimal to have symmetric transition probabilities across memory states follow the same arguments as in Proposition 3.1.

Proof of (2). Here I consider $\mu_{h'}^L = \frac{v(1-\mu)}{1-v} + \varepsilon$ for ε small. Let $\rho_0 \in (\rho_0^L, \rho_0^K)$ be given. When $\varepsilon = 0$, the optimal SFSA takes the form $M_K^b(\beta, \beta')$. Note that when $\varepsilon = 0$, the optimal β and β' are not determined individually but only the joint transition probabilities as given by (11) are uniquely determined. Using the same arguments as in the Proof of Proposition 4.2 and continuity of the beliefs and value functions, for ε small, by Proposition 4.1 we can restrict attention to SFSA in the form of $M_K^b(\beta, \beta')$. The fact that it is optimal to have symmetric transition probabilities across memory states follow the same arguments as in Proposition 3.1.

Now I study the optimal β and β' for ε small, which, as we will see below, are uniquely determined for $\varepsilon > 0$. Let

$$\begin{aligned} \alpha^H &= \mu_h^H \beta + \mu_{h'}^H \beta' = (1 - v)\beta + \beta v, \text{ and } \alpha^L \\ &= \frac{\mu_h^L \beta + \mu_{h'}^L \beta'}{\mu_h^L + \mu_{h'}^L} = \frac{(1 - \mu)\beta + \left(\frac{v(1-\mu)}{1-v} + \varepsilon\right) \beta'}{\frac{(1-\mu)}{1-v} + \varepsilon}. \end{aligned}$$

Now, the continuation values can be computed in the same way. In particular, $V_{q_i}(H)$ still takes the form of (16), but with α replaced by α^H , and $V_{q_1}(L)$ still takes the form of (18), but with α replaced by α^L . Thus, the expected payoff from $M_K^b(\beta, \beta')$ is now

$$\begin{aligned} F(\beta, \beta') &= p_0 \left[\frac{(1 - \alpha^H)(1 - \eta)}{1 - (1 - \eta)\alpha^H} \right]^{K-2} u^H \\ &+ (1 - p_0) \left\{ 1 - \left[\frac{(1 - \eta)\left(\frac{1-\mu}{1-v} + \varepsilon\right)(1 - \alpha^L)}{1 - \left(\frac{1-\mu}{1-v} + \varepsilon\right)(1 - \eta)\alpha^L} \right]^{K-2} \frac{\eta}{1 - (1 - \eta)\left(\frac{1-\mu}{1-v} + \varepsilon\right)} \right\} u^L, \end{aligned}$$

Taking derivatives, we have

$$\begin{aligned} \frac{\partial F}{\partial \beta} &= p_0 \left[\frac{(1 - \eta)(1 - \alpha^H)}{1 - (1 - \eta)\alpha^H} \right]^{K-3} \frac{(-K + 2)\eta(1 - \eta)u^H}{[1 - (1 - \eta)\alpha^H]^2} (1 - v) \quad (28) \\ &+ (1 - p_0) \left[\frac{(1 - \eta)\left(\frac{1-\mu}{1-v} + \varepsilon\right)(1 - \alpha^L)}{1 - \left(\frac{1-\mu}{1-v} + \varepsilon\right)(1 - \eta)\alpha^L} \right]^{K-3} \\ &\frac{(K - 2)\eta(1 - \eta)u^L}{[1 - (1 - \eta)\left(\frac{1-\mu}{1-v} + \varepsilon\right)\alpha^L]^2} (1 - \mu), \end{aligned}$$

and

$$\begin{aligned} \frac{\partial F}{\partial \beta'} = p_0 & \left[\frac{(1-\eta)(1-\alpha^H)}{1-(1-\eta)\alpha^H} \right]^{K-3} \frac{(-K+2)\eta(1-\eta)u^H}{[1-(1-\eta)\alpha^H]^2} v \\ & + (1-p_0) \left[\frac{(1-\eta)\left(\frac{1-\mu}{1-\nu} + \varepsilon\right)(1-\alpha^L)}{1-\left(\frac{1-\mu}{1-\nu} + \varepsilon\right)(1-\eta)\alpha^L} \right]^{K-3} \\ & \frac{(K-2)\eta(1-\eta)u^L}{[1-(1-\eta)\left(\frac{1-\mu}{1-\nu} + \varepsilon\right)\alpha^L]^2} \left[\frac{v(1-\mu)}{1-\nu} + \varepsilon \right]. \end{aligned} \quad (29)$$

Now, for any $\varepsilon > 0$, since the first terms in both (28) and (29) are negative but the second terms are positive and since

$$\frac{1-\nu}{1-\mu} > \frac{v}{\left[\frac{v(1-\mu)}{1-\nu} + \varepsilon \right]} = \frac{1-\nu}{1-\mu + (1-\nu)\varepsilon},$$

$\frac{\partial F}{\partial \beta'} = 0$ implies that $\frac{\partial F}{\partial \beta} < 0$, that is, if optimal $\beta' < 1$, optimal $\beta = 0$. Moreover, this shows that whenever $\rho_0 \leq \rho_K$ and hence at $\alpha^H = \alpha^L = 0$, $\frac{\partial F}{\partial \beta'} \geq 0$ when $\varepsilon = 0$. Thus, for any $\varepsilon > 0$, $\frac{\partial F}{\partial \beta'} > 0$ under $\beta = \beta' = 0$ and hence optimal $\beta' > 0$.

Proof of Proposition 4.4

Without loss of generality, rank the memory states so that $\rho(q_L) < \rho(q_1) < \rho(q_H)$, and assume that from some memory state after seeing ℓ -signal it transits to q_L and $d(q_L) = a^L$. This implies that $\rho(q_L) = 0$, and by Proposition 4.1 it follows that it is optimal to transit to q_L from any memory state after an ℓ -signal. Also, in q_1 , after seeing signal- h or h' the transition cannot go down. Moreover, $d(q_H) = a^H$ for otherwise there is no variation in actions. This also implies that $q^o = q_1$. If $d(q_1) = a^H$, then q_1 and q_H become equivalent memory states. So $d(q_1) = a^L$. It then follows that the optimal SFSA takes the form $M_3^b(\beta, \beta')$.

For $\rho_0 \in \left[\frac{\rho^*}{\xi(h')}, \rho^* \right)$, one h' -signal is sufficient to convince the DM to take action a^H and hence the unconstrained optimal rule can be implemented by $M_3^b(0, 0)$. So consider $\rho_0 < \frac{\rho^*}{\xi(h')}$. We can then compute the ex ante expected payoffs from $M_3^b(\beta, \beta')$ for any $\mu_{h'}^L \in \left(\frac{v(1-\mu)}{1-\nu}, v \right)$, which is given by

$$\begin{aligned} F(\beta, \beta') = p_0 & \left[\frac{(1-\alpha^H)(1-\eta)}{1-(1-\eta)\alpha^H} \right] u^H \\ & + (1-p_0) \left\{ 1 - \left[\frac{(1-\eta)(1-\mu + \mu_{h'}^L)(1-\alpha^L)}{1-(1-\mu + \mu_{h'}^L)(1-\eta)\alpha^L} \right] \frac{\eta}{1-(1-\eta)(1-\mu + \mu_{h'}^L)} \right\} u^L, \end{aligned}$$

where $\alpha^H = (1 - v)\beta + v\beta'$ and $\alpha^L = [(1 - \mu)\beta + \mu_{h'}^L\beta'] / (1 - \mu + \mu_{h'}^L)$. Thus,

$$\frac{\partial F}{\partial \beta} = p_0 \frac{-\eta(1 - \eta)(1 - v)u^H}{[1 - (1 - \eta)\alpha^H]^2} + (1 - p_0) \frac{(K - 2)\eta(1 - \eta)(1 - \mu)u^L}{[1 - (1 - \eta)(1 - \mu + \mu_{h'}^L)\alpha^L]^2}, \tag{30}$$

and

$$\frac{\partial F}{\partial \beta'} = p_0 \frac{-\eta(1 - \eta)v u^H}{[1 - (1 - \eta)\alpha^H]^2} + (1 - p_0) \frac{(K - 2)\eta(1 - \eta)\mu_{h'}^L u^L}{[1 - (1 - \eta)(1 - \mu + \mu_{h'}^L)\alpha^L]^2}. \tag{31}$$

Now, let

$$\tilde{\rho}_0 = \left[\frac{1 - (1 - \eta)v}{1 - (1 - \eta)\mu_{h'}^L} \right]^2 \frac{\rho^*}{\xi(h)} < \rho_0^3 = \frac{\rho^*}{\xi(h)}, \tag{32}$$

where the inequality follows from the fact that $\mu_{h'}^L < v$. Now, it is straightforward to verify, using (30) and (31), that for all $\rho_0 > \tilde{\rho}_0$, the only solutions to the FOC's for β and β' is to have $\beta = 0$ and $\beta' > 0$, and that optimal β' strictly increases as ρ_0 decreases from ρ_0^3 and hits one before ρ_0 reaches $\tilde{\rho}_0$. Similarly, for $\rho_0 > \tilde{\rho}_0$, optimal $\beta > 0$ and $\beta' = 1$. Finally, $\underline{\rho}_0$ is determined in a similar way as in the proof of Proposition 3.1, and, in case $\underline{\rho}_0 < \tilde{\rho}_0$ given by (32), take $\tilde{\rho}_0 = \underline{\rho}_0$.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00199-022-01423-7>.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Barlo, M., Carmona, G., Sabourian, H.: Repeated games with one-memory. *J. Econ. Theory* **144**, 312–336 (2009)
- Che, Y.K., Mierendorff, K.: Optimal dynamic allocation of attention. *Am. Econ. Rev.* **109**, 2993–3029 (2019)
- Hellman, M.E., Cover, T.M.: Learning with finite memory. *Ann. Math. Stat.* **41**, 765–782 (1970)
- Kalai, E., Solan, E.: Randomization and simplification in dynamic decision-making. *J. Econ. Theory* **111**, 251–264 (2003)
- Monte, D., Said, M.: The value of (Bounded) memory in a changing world. *Econ. Theor.* **56**, 59–82 (2014)
- Nerode, A.: Linear automaton transformations, *Proc AMS*, **9** (1958)

- Piccione, M., Rubinstein, A.: On the interpretation of decision problems with imperfect recall. *Games Econom. Behav.* **20**, 3–24 (1997)
- Rubinstein, A.: Finite automata play the repeated Prisoner's dilemma. *J. Econ. Theory* **39**, 83–96 (1986)
- Wilson, A.: Bounded memory and biases in information processing. *Econometrica* **82**, 2257–2294 (2014)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.