SELECTIVE SAMPLING WITH INFORMATION-STOREG CONSTRAINTS*

Philippe Jehiel and Jakub Steiner

A memoryless agent can acquire arbitrarily many signals. After each signal observation, she either terminates and chooses an action, or she discards her observation and draws a new signal. By conditioning the probability of termination on the information collected, she controls the correlation between the payoff state and her terminal action. We provide an optimality condition for the emerging stochastic choice. The condition highlights the benefits of selective memory applied to the extracted signals. Implications—obtained in simple examples—include (i) confirmation bias, (ii) speed-accuracy complementarity, (iii) overweighting of rare events, and (iv) salience effect.

Economic agents often acquire information about the state of the economy before making their decisions. The information is typically modelled as a signal that helps the agent refine the distribution of the state and improve the decision-making. Often, signals come over time and agents can absorb only a small number of them. We capture this information-processing friction by assuming that agents receive as many signals as they wish but can remember only a finite number of them when making their choices. In the simplest setting we analyse, the agent can only remember one signal. A key strategic variable that we consider is to allow the agent to ignore some signals with positive probability and restart the signal extraction process. We allow agents to employ an arbitrary stationary decision process that specifies for each possible signal realisation a probability with which the agent restarts the process as well as the chosen action in case of termination. We do not impose time constraints and costs in the basic formulation so that the friction comes solely from the limited information-storing capacity of the agent.

We ask ourselves: Should the agent optimally make her choice as soon as she receives the first signal whatever the realisation of it is, or could she be better off by rerunning the very same information-acquisition process? Can hesitation—selective repetition of a fixed stochastic decision procedure—be welfare-enhancing?

A general insight is that selective rerunning of the primitive decision procedure is typically optimal. To document this most generally, we provide a simple necessary condition satisfied by the optimal rerunning strategy. The result is an interim indifference condition imposed on the agent who has concluded her decision-making with a plan to choose a particular action. Given the recommended action, the agent’s posterior expected payoff from implementing this

* Corresponding author: Jakub Steiner, CERGE-EI, Politických Vězňů 7, Praha 1, 110 00, Czechia, Email: jakub.steiner@cerge-ei.cz

This paper was received on 4 February 2019 and accepted on 21 November 2019. The Editor was Gilat Levy.

This paper has been previously circulated under the title: ‘On Second Thoughts, Selective Memory, and Resulting Behavioral Biases.’ We thank Mark Dean, Andrew Ellis, Alessandro Pavan, Philip Reny, Larry Samuelson, Colin Stewart, Balazs Szentes, colleagues at the University of Edinburgh, the audiences at Bocconi, Queen Mary, Columbia, Ecole Polytechnique, Zurich and St Andrews universities, workshops and conferences in Erice, Alghero, Faro, Gerzenssee, New York, Cambridge, Vancouver, and Barcelona, the editor (Gilat Levy), and the two referees for helpful comments. Ludmila Matysková and Jan Šedek provided excellent research assistance. Deborah Nováková and Laura Straková have helped with English. Jakub Steiner has received financial support from the Czech Science Foundation grant 16-00703S and from the ERC grant 770652. Philippe Jehiel thanks the ERC grant n° 742816 for funding.

[ 1753 ]
action must be the same as the posterior expected payoff from rerunning the whole decision-making—the whole selective repetitions of the primitive signal extraction—and implementing whichever action the second run of the decision-making will recommend. We refer to this as to the second-thought-free condition.

For illustration, consider a binary decision of whether to make an investment of a fixed size. The agent receives payoff 1 if she invests in the good state of the economy, payoff $-1$ if she invests in the bad state, and receives 0 when she does not invest whatever the state. One of the two states is a priori more likely; for sake of concreteness, let the prior probability of the good state be $2/3$. Both states give rise to a population of good and bad signals, with the share of the good signals at 90% in the good state and 10% in the bad state. The agent draws possibly several signal realisations in sequence but remembers only the last one when making her investment choice. As follows from simple optimisation considerations, assume she invests if and only if the last remembered signal was good. Observe that the decision rule generated by the immediate termination upon the first signal that comes in does not generate a second-thought-free choice rule: An agent whose first observed signal was bad prefers to rerun the decision process, since the new run will either lead to not investing again or will give rise to the signal realisation that conflicts with the first observation and will lead to investing. Since, conditional on two conflicting signals, the a priori more common state is more likely, investing is preferred in this contingency. The agent benefits from having second thoughts when the first observed signal is surprising.

We interpret the probability of terminating the decision process after receiving a particular signal as a search intensity targeting this particular signal. A higher probability of termination at a given information set inflates the likelihood that the agent makes the terminal choice at the set. We show that the failure of the second-thought-free condition with uniform search intensity in the above investment decision example indicates that relative to the uniform search, the agent benefits from decreasing the search intensity for the bad attribute. More generally, the second-thought-free condition follows from the first-order condition imposed on the optimal search intensities.

The model provides microfoundations to a range of behavioural stylised facts. The unifying principle of our behavioural insights is the intuition that the agent targets her search towards the type of evidence that would provide her with more valuable posteriors under the uniform search. This principle generates confirmation bias in the context of the above example, since evidence that confirms the agent’s prior leads to more informed posterior than does evidence that contradicts the prior. An optimally targeted information search also generates speed-accuracy complementarity in the same setting; that is, accuracy of choice declines with the response time. The effect is generated by the confirmation bias: The agent encountering evidence contradicting her prior is likely to disregard the evidence and to have a second thought. Hence, long response times indicate a surprising state of the world, and the constrained-optimal choice rule commits errors in the surprising state relatively often. Overweighting of rare events occurs in a related setting in which the agent’s task is to form a probability belief about an event that is known to be rare, such as a flight accident, by observing a random flight outcome. Since observing a flight accident is far more informative about the probability of future accidents than observing an uneventful flight, the agent optimally biases her search towards eventful flights. In the last behavioural application, we show in a setting with multiple states that distinct states of the world are salient in the sense that they attract the agent’s attention (i.e., trigger higher termination rates in our framework). The effect arises because an indistinct perception stimulus that can be generated by several similar states is less informative than a distinct perception stimulus that is
most likely generated by a specific distinct state. Hence, the optimal information search targets stimuli indicating distinct states.

Our leading interpretation of the model is in terms of a single-person with information storage limitations but perfect ability to adjust optimally her termination strategies as a function of what she remembers. These adjustments can be viewed as a result of a successful trial and error process or as a result of evolutionary pressures in which case the adjustments may not be fine-tuned to each specific problem. Alternatively, one may think of the termination strategies and the final decisions as being chosen by different persons and only the one in charge of the final decision would be subject to information storage limitations, thereby allowing the termination strategies to be optimally determined.

There is a wide range of studies that propose different modelling of optimisation over information structures. Relative to rational inattention models (Sims, 2003), we provide a procedural micro-foundation of our set of the feasible information structures (which must be obtained through variations of the termination strategies and as such can straightforwardly be related to the time dimension of the decision process). Perhaps surprisingly, our memoryless agent shares some of the flexibility in her choice of information structures with the sequential-sampling model of Wald (1945) in which there is perfect information aggregation. But, our approach allows for a simple derivation of the speed-accuracy complementarity, which is less immediate to obtain with sequential-sampling models (see however Fudenberg et al., 2018).

Relative to studies based on finite automata (Hellman and Cover, 1970; Wilson, 2014; Basu and Chatterjee, 2015), our approach yields a simple necessary condition for the optimal choice rule, the second-thought-free condition. This condition arises because our agent chooses the probability of termination at each of her information sets. Such a termination optimisation is absent from the related models with finite automata in which the termination is exogenous (or the objective involves asymptotic performance as time diverges). The second-thought-free condition allows us to characterise the optimal choice rule in the binary settings.

This article belongs to a growing economic literature that explains behavioural stylised facts as the constrained-optimal behaviour of decision-makers facing information processing frictions. For instance, Robson (2001), Rayo and Becker (2007), Netzer (2009), and Khaw et al. (2017) provide microfoundations for risk attitudes; Gabaix and Laibson (2017) endogenise discounting; and Wilson (2014), Compte and Postlewaite (2012), and Leung (2020) establish constrained-optimal ignorance of weakly informative news.2

1. Model

An agent faces a decision under uncertainty. She chooses an action $a \in A$ and receives a payoff $u(a, \theta)$ in the fixed payoff state $\theta \in \Theta$ drawn from an interior prior $\pi \in \Delta(\Theta)$. The sets $\Theta$ and $A$ are finite. The agent chooses a Blackwell experiment $p$, where $p$ is a family of conditional signal distributions $p(x | \theta)$ that depend on $\theta \in \Theta$. The experiment generates a signal realisation $x$ from a finite space $X$. The conditional signal distributions are fully mixed: $p(x | \theta) > 0$ for all $x, \theta$. We allow the agent to choose among possibly several such experiments and we let $\mathcal{P}$ denote the

---

1 In the latter case, while the second-thought free conditions need not be satisfied for each problem in isolation, we would still derive that some degree of selective hesitation is optimal.

2 Somewhat less related is a literature that explores how exogenous analogy-based and extrapolation-driven errors in learning lead to behavioural biases; see coarse learning in Jehiel (2005) and its application to overoptimism in Jehiel (2018). By contrast, in our approach, the agent optimises the error distribution given the constraints.

© 2019 Royal Economic Society.
For each \((p, \beta, \sigma)\), the decision process is a Markov chain evolving on the agent’s states of mind, with transition probabilities that depend on the payoff state \(\theta\). The chain begins in the state of mind \(O\) and transits to states \(x \in X = \{x_1, x_2\}\) with probabilities \(p(x | \theta)\). The process returns to \(O\) with probability \(1 - \beta_x\), or terminates with choice of \(a = \sigma(x)\) with probability \(\beta_x\).

Fig. 1. For each \((p, \beta, \sigma)\), the decision process is a Markov chain evolving on the agent’s states of mind, with transition probabilities that depend on the payoff state \(\theta\). The chain begins in the state of mind \(O\) and transits to states \(x \in X = \{x_1, x_2\}\) with probabilities \(p(x | \theta)\). The process returns to \(O\) with probability \(1 - \beta_x\), or terminates with choice of \(a = \sigma(x)\) with probability \(\beta_x\).

exogenous set of experiments from which she chooses. We impose no restrictions on \(\mathcal{P}\) (other than the full-support of each \(p\)).

The agent can repeat the selected experiment arbitrarily many times, but she is unable to aggregate the information across the repetitions. Each run of the experiment is a cognition that exhausts the agent’s capacity dedicated to the problem being solved. Once the agent hits the constraint at the end of the experiment, she can continue only after she unclogs her capacity by amnesia.

We model this as follows. The agent can condition the repetition of the experiment on the last observed signal realisation. She chooses a vector \(\beta = (\beta_x)_{x \in X} \in B = [0, 1]^{|X|} \setminus \{(0, \ldots, 0)\}\) of termination probabilities \(\beta_x\) for each signal realisation \(x\); we call \(\beta\) a termination strategy. The agent runs the experiment \(p\) for the first time, receives signal realisation \(x^{(1)}\) with probability \(p(x^{(1)} | \theta)\) and terminates the reasoning with probability \(\beta_x^{(1)}\). She restarts with the complementary probability \(1 - \beta_x^{(1)}\), and receives a signal realisation \(x^{(2)}\) from a new run of the process \(p\) with probability \(p(x^{(2)} | \theta)\), terminates with probability \(\beta_x^{(2)}\) or restarts with probability \(1 - \beta_x^{(2)}\), and continues until she terminates after a random number of repetitions of \(p\); see Figure 1. When the agent chooses having distinct \(\beta_x\) for different \(x\), then she implements the familiar idea of selective memory; some facts and observations are easily forgotten whereas others are remembered and they trigger choice. After the agent terminates the reasoning with a terminal signal realisation \(x\), she selects an action \(a = \sigma(x)\) according to an action strategy \(\sigma : X \rightarrow A\).

By excluding the termination strategy \((0, \ldots, 0)\) we force the agent to make a decision \(a \in A\). Since \(\beta \neq (0, \ldots, 0)\) and each feasible experiment \(p\) generates all signal values with a positive probability in each state, the decision process almost surely eventually terminates.

The outcomes of distinct runs of \(p\) are conditionally independent. Thus, the probability that the agent terminates after \(t\) repetitions of the experiment \(p\) resulting in the signal history \(x^t = (x^{(1)}, \ldots, x^{(t)})\) is

\[
\rho \left( x^t | \theta; p, \beta \right) = \beta_{x^{(t)}} p(x^{(t)} | \theta) \prod_{l=1}^{t-1} (1 - \beta_{x^{(l)}}) p(x^{(l)} | \theta). \tag{1}
\]

\footnote{We do not allow for mixed action strategies since the optimum can always be achieved with a pure action strategy.}
We let

\[ r(a \mid \theta; p, \beta, \sigma) = \sum_{t=1}^{\infty} \sum_{x_t' = a} \rho(x_t' \mid \theta; p, \beta) \]

(2)
denote the probability that the agent who employs the experiment \( p \), the termination strategy \( \beta \), and action strategy \( \sigma \) terminates with action \( a \) in state \( \theta \). We call \( r(p, \beta, \sigma) \) the choice rule. The set of feasible choice rules is \( \mathcal{R}(P) = \{ r(p, \beta, \sigma) : p \in \mathcal{P}, \beta \in B, \sigma \in S \} \). Sometimes we abuse notation, omit \( p, \beta, \sigma \) and write \( r(a \mid \theta) \) for the probability of \( a \) in state \( \theta \) under the rule constructed by some \( p, \beta, \sigma \).

The repeated-cognition problem is to select, for a given prior \( \pi \), utility function \( u \) and set \( \mathcal{P} \), a feasible choice rule \( r \) that maximises the expected payoff:

\[ \max_{r \in \mathcal{R}(\mathcal{P})} \sum_{\theta \in \Theta, a \in A} \pi_{\theta} r(a \mid \theta) u(a, \theta). \]

(3)

The optimisation in (3) can be an outcome of selective pressures that favour successful decision procedures via cultural or biological evolution, or via competition of firms differing in their internal procedures. There are no costs to delaying the decision in our model but incorporating such costs would not affect our qualitative insights when these are not too big. We address agents with less severe memory constraints and with exponential time discounting preferences in Section 6.

2. Optimal Cognition Biases

We now derive a necessary optimality condition that the choice rule solving the repeated-cognition problem must satisfy. Generically, the condition requires the agent to engage in selective information processing—that is, to ignore some signals more often than others.

2.1. Second-Thought-Free Choice Rules

We start with a definition of second-thought-free choice rules. If the agent’s decision process generates such a rule, then she has no incentive to rerun the process regardless of the action recommendation with which the process terminates. Our main result below states that the optimal rule is second-thought-free.

Let \( r \) be a generic stochastic choice rule that specifies conditional probabilities \( r(a \mid \theta) \) of each action \( a \in A \) in each state \( \theta \in \Theta \).

DEFINITION 1. The choice rule \( r \) is second-thought-free with respect to the utility \( u \) and prior \( \pi \) if the agent prefers each action recommended by the rule to a new run of the rule \( r \). That is, for each action \( a \) chosen with positive probability,

\[ E_{a}[u(a_1, \theta) \mid a_1 = a] \geq E_{a}[u(a_2, \theta) \mid a_1 = a], \]

(4)

where the expectations are with respect to the random variables \( \theta \) and \( a_2 \), and \( \alpha(\theta, a_1, a_2) = \pi_{\theta} r(a_1 \mid \theta) r(a_2 \mid \theta) \) is the joint distribution of the state and two actions consecutively generated by \( r \).

The definition requires the agent who terminates with an action plan \( a \) to weakly prefer \( a \) to forgetting \( a \) and choosing whichever action a new run of the decision process will recommend.
Although the definition allows for the strict preference against having a second thought, the next lemma shows that if a choice rule is second-thought-free, then the agent is indifferent between terminating and the second thought.

**Lemma 1.** If $r$ is second-thought-free, then (4) is met with equality for each action a chosen with positive probability:

$$E_\alpha[u(a_1, \theta) \mid a_1 = a] = E_\alpha[u(a_2, \theta) \mid a_1 = a].$$

We refer to (5) as the second-thought-free condition.

**Proof.** If (4) holds with strict inequality for some $a$ chosen with positive probability, then

$$E_\alpha[u(a_1, \theta)] = E_\alpha[E_\alpha[u(a_1, \theta) \mid a_1]] > E_\alpha[E_\alpha[u(a_2, \theta) \mid a_1]] = E_\alpha[u(a_2, \theta)],$$

which contradicts that $a_1$ and $a_2$ are conditionally iid. \hfill \Box

### 2.2. Optimality Condition

We provide here a general necessary optimality condition imposed on the stochastic choice rule.

**Proposition 1.** If a choice rule solves the repeated-cognition problem (3), then it is second-thought-free and satisfies (5).

To understand the statement, consider the optimal choice rule $r^*$ generated by a process that consists of a random number of repetitions of a primitive cognition $p$. Once these repetitions of $p$ terminate with a signal realisation $x$ and the agent is about to take an action $a = \sigma(x)$, then, according to the proposition, she must be indifferent between $a$, and running the process associated with $r^*$ from scratch, where the new run of $r^*$ would involve new repetitions of $p$.

To prove Proposition 1, we introduce an effective experiment $s(p, \beta)$. While the primitive experiment $p(x \mid \theta)$ specifies the probability that its one run results in signal $x$, we define $s(x \mid \theta; p, \beta)$ to be the probability that selective repetitions of $p$ according to the termination strategy $\beta$ terminate with $x$. Relative to $p(x \mid \theta)$, the effective probabilities $s(x \mid \theta; p, \beta)$ are inflated for those $x$ at which the agent terminates with a high probability $\beta_x$.

**Lemma 2.** An agent who employs a primitive experiment $p$ and a termination strategy $\beta$ terminates with $x$ in state $\theta$ with probability:

$$s(x \mid \theta; p, \beta) = \frac{\beta_x p(x \mid \theta)}{\sum_{x' \in X} \beta_{x'} p(x' \mid \theta)}.$$  

**Proof.** Experiment $s(p, \beta)$ satisfies the recursive formula

$$s(x \mid \theta; p, \beta) = \beta_x p(x \mid \theta) + \sum_{x' \in X} (1 - \beta_{x'}) p(x' \mid \theta) s(x \mid \theta; p, \beta).$$

The first summand is the probability that the agent terminates with signal $x$ after the first run of the experiment $p$. The second summand is the probability that the agent continues after the first run and terminates with $x$ later. Solving for $s(x \mid \theta; p, \beta)$ gives (6). \hfill \Box

The lemma implies that $s(p, \beta)$ and hence also $r(p, \beta, \sigma)$ are homogeneous of degree zero with respect to $\beta$. Thus, since we abstract from the delay costs, for any optimal termination strategy $\beta^*$, $\alpha \beta^*$ for $\alpha \in (0, 1)$ is optimal too, and it generates the same optimal choice rule $r^*$ as
The action and state sets are binary: $A = \Theta = \{0, 1\}$. To avoid a trivial case, we assume that neither action is dominant. Then, without loss of generality, $u(a, \theta) = u_0 > 0$ if $a = \theta$ and $u(a, \theta) = 0$ otherwise. State $\theta$ is drawn from an interior prior $\pi \in \Delta(\Theta)$. The exogenous set $\mathcal{P}$ of the feasible statistical experiments is finite, and each $p \in \mathcal{P}$ delivers a signal $x$ from a finite signal space $X$ with probability $p(x \mid \theta)$. The agent chooses $p \in \mathcal{P}$, the termination strategy $\beta = (\beta_x)_{x \in X}$ and action strategy $\sigma : X \rightarrow A$ to maximise the expected payoff.

The first result states that there exists a solution in which the agent ignores all but two signal realisations of the chosen experiment $p$. That is, she always repeats the experiment upon
encountering all but two signals. Roughly, the result follows because it is advantageous to consider only the most informative signal realisations.\textsuperscript{4}

**Lemma 3.** There exists a solution in which the termination probability $\beta_x$ is positive for at most two signal values $x \in X$.

See Appendix for the proofs omitted in the main text.

Based on the lemma, we can, without loss of generality, restrict the signal space $X$ to be binary, and identify it with the action and state spaces, $X = A = \Theta$. Again without loss of generality, we choose signal labels in each experiment in such a way that each experiment $p \in \mathcal{P}$ satisfies the monotone likelihood ratio property: $p(1 \mid \theta)/p(0 \mid \theta)$ increases in $\theta$. We continue to assume that $p(x \mid \theta) > 0$ for all $x$ and $\theta$.

Define $\sigma_I$ to be the identity function, and let the agent employ the binary experiment $p$ and the action strategy $\sigma_I$. The next lemma characterises the set $\mathcal{R}_{p,\sigma_I} = \{ r(p, \beta, \sigma_I) : \beta \in B \}$ of the feasible choice rules that such an agent has access to. To characterise this set, we introduce a parameter that we dub *perceptual distance* between states 0 and 1 under the experiment $p$:

$$d_p = \frac{p(1 \mid 1)p(0 \mid 0)}{p(0 \mid 1)p(1 \mid 0)}.$$ 

The perceptual distance is a summary statistic of the experiment $p$. The larger it is, the more $p$ reliably discriminates between the two states. The monotone likelihood property of each $p$ implies that $d_p > 1$. The next lemma states that the perceptual distance is preserved under any termination strategy $\beta$.

**Lemma 4.** $\mathcal{R}_{p,\sigma_I} = \{ r : r(1 \mid 1)r(0 \mid 0) = d_pr(1 \mid 0)r(0 \mid 1) \}.$

That is, a rule $r$ can be constructed from $p$ if and only if it preserves the perceptual distance: $r(1 \mid 1)r(0 \mid 0) = d_p$ (or if it always selects a same action). By controlling the termination strategy $\beta$, the agent trades off the likelihoods $r(0 \mid 0; p, \beta, \sigma_I)$ and $r(1 \mid 1; p, \beta, \sigma_I)$ of the correct choice in the states 0 and 1, respectively. See Figure 2. The set $\mathcal{R}_{p,\sigma_I}$ of rules accessible from $p$ is compact.

Thanks to the chosen labelling of the signals, the agent can equate her choice to the observed signal without a loss:

**Lemma 5.** For any rule $r(p, \beta, \sigma)$ there exists $\beta'$ such that the rule $r(p, \beta', \sigma_I)$ achieves at least as high expected payoff as $r(p, \beta, \sigma)$ where $\sigma_I$ is the identity function.

The solution to the repeated-cognition problem in the binary setting exists since the objective is continuous in the choice rule and the agent optimises on the compact set $\bigcup_{p \in \mathcal{P}} \mathcal{R}_{p,\sigma_I}$ of the rules.

Let $\overline{p}$ be the experiment with the maximal perceptual distance: $\overline{p} \in \arg \max_{p \in \mathcal{P}} d_p$, and let $\overline{d} = \max_{p \in \mathcal{P}} d_p$. In line with the intuition that the agent should go for the most informative experiment, we establish:

**Lemma 6.** There exists a solution to the repeated-cognition problem in which the agent employs the experiment $\overline{p}$ with the maximal perceptual distance.

\textsuperscript{4} This insight exploits the assumption of perfect patience, since impatient agents would trade off informativeness against delay costs. We conjecture that when exponential discounting is considered, then the result that the agent ignores all but two signal realisations continues to hold for sufficiently patient agents and generic signal structures.
Fig. 2. Each point in $[0, 1]^2$ on this graph corresponds to a choice rule. The depicted curves are the sets $R_{p,\sigma_I}$ of the choice rules constructible from experiments $p$ and action strategy $\sigma_I$. The thick curve corresponds to the experiment $\overline{p}$ with the maximal perceptual distance $\overline{d}$. Since the objective is linear in the choice rule, the indifference curves are downward sloping lines. The dashed line is the indifference curve tangential to $R_{\overline{p},\sigma_I}$. The dot depicts the solution of the repeated-cognition problem.

Fig. 3. Confirmation bias with discounting. Action 1 is a priori more attractive: $\pi_1u_1 = 5 \times \pi_0u_0$. The primitive experiment is symmetric: $p(1 \mid 1) = p(0 \mid 0) = 0.9$. The agent terminates immediately when she observes signal value 1, $\beta^*_1 = 1$. When $\delta > 0.71$, then the agent is biased towards state 1: when she encounters signal value 0, then she terminates the decision-process with a probability only $\beta^*_0(\delta) < 1$ (the full curve). The dotted line is $\beta^*_0 / \beta^*_1$ from the baseline model without discounting.
The last lemma implies that all details of the set $P$ relevant for the solution are summarised in the one-dimensional statistic $\overline{d}$ that is independent of the payoff function $u$.

We are now ready to solve the binary setting. The optimal effective choice rule $r^*(a \mid \theta) = r(a \mid \theta; \overline{p}, \beta^*, \sigma_f)$ consists of four unknown probabilities and it is determined by four conditions: the second-thought-free condition (5), the feasibility condition from Lemma 4, and two normalisation conditions. Let parameter $R = \frac{\sigma_1 u}{\sigma_0 u}$ measure the relative a priori attractiveness of action 1.

**Proposition 2.**

1. When $R \geq \overline{d}$, then the agent always chooses action 1;
2. when $R \leq 1/\overline{d}$, then the agent always chooses action 0;
3. when $R \in (1/\overline{d}, \overline{d})$, then the agent chooses both actions with positive probabilities and
   
   \[ r^*(1 \mid 1) = \overline{d} - \sqrt{\overline{d}R - R \overline{p}(0 \mid 1)} / \overline{d} - 1, \]

   \[ r^*(0 \mid 0) = \overline{d} - \sqrt{\overline{d}R - R \overline{p}(0 \mid 1)}, \]

4. When the ex ante attractiveness of one of the actions is too strong relative to the perceptual distance of the two states, then the agent always chooses the ex ante attractive action. The decision process is non-trivial for intermediate incentives: the agent engages in repeated cognition and she chooses both actions with positive probabilities.

4. **Behavioural Applications**

This section presents three behavioural effects illustrated in the binary setting from Section 3: confirmation bias, speed-accuracy complementarity, and overweighting of rare events.

4.1. **Confirmation Bias**

Psychologists and economists distinguish at least three mechanisms leading to confirmation bias: (i) People search for evidence selectively, targeting the evidence type in accord with their priors, e.g., Nickerson (1998), (ii) they selectively memorise and recall the data supporting their priors, e.g., Oswald and Grosjean (2004), and (iii) they selectively interpret ambiguous evidence, e.g., Rabin and Schrag (1999) and Fryer et al. (2018). We focus on the first two mechanisms and interpret them in light of our optimal repeated-cognition result.

**Corollary 1.** When action 1 is a priori more attractive, $R \in (1, \overline{d})$, and the unique primitive binary experiment is symmetric, $p(1 \mid 1) = p(0 \mid 0) > 1/2$, then the agent searches relatively more intensively for signal value 1: $\beta^*_1 > \beta^*_0$.

---

5 Such summary statistic of $P$ continues to exist when $2 < |X| < \infty$. For any pair of signal realisations $(x, x')$ and an experiment $p$, let $d_{x,x',p} = \frac{p(x \mid 0)p(x' \mid 1)}{p(x \mid 1)p(x' \mid 0)}$. Then, $\overline{d}$ is the maximum of $d_{x,x',p}$ over all ordered pairs $(x, x')$ and experiments $p$. © 2019 Royal Economic Society.
PROOF. Since $\beta_1^*/\beta_0^*$ increases in $R$, it suffices to show that $\beta_1^*/\beta_0^* = 1$ when $R = 1$ and the primitive experiment $p$ is symmetric. Indeed, when $R = 1$, then by (9),
\[
\frac{\beta_1^*}{\beta_0^*} = \sqrt{d} \frac{p(0 | 1)}{p(1 | 1)} = \sqrt{\frac{p(0 | 0)p(0 | 1)}{p(1 | 1)p(1 | 0)}} = 1,
\]
where $d = \frac{p(0 | 0)p(1 | 1)}{p(1 | 0)p(0 | 1)}$ and the last equality follows from the symmetry of $p$. □

To see the connection to confirmation bias, consider, like in our introductory example, an agent whose task is to announce the realised state of the world: she receives reward $u_1 = u_0 = 1$ if she makes the correct announcement and 0 otherwise. The agent finds the state $\theta = 1$ a priori more likely than the state 0, $\pi_1 > \pi_0$. Consider the decision process that terminates immediately after the first run of the experiment and chooses the action equal to the observed signal value: $\beta_0 = \beta_1 = 1$, $\sigma = \sigma_I$. To establish that such an unbiased process is suboptimal it suffices to show that it does not satisfy the second-thought-free condition. To see this, we examine the agent who has received the a priori unlikely signal 0 and argue that she benefits from the second thought. Such a surprised agent is better off by restarting instead of terminating with action 0, since if the new run of the process concludes with signal 0 again, then the second thought will have been inconsequential. If, however, the new run of the process concludes with signal and action 1, then the induced switch from action 0 to 1 is beneficial. This is because when the experiment $p$ is symmetric, then, conditional on two conflicting signals, the a priori more common state 1 is relatively more likely. The agent benefits from second thought whenever she receives the surprising recommendation, and thus will deviate from the uniform search in favour of the a priori likely signal value 1.

The optimal strategy resembles the natural process in which the selective memory gives rise to confirmation bias. Consider the fastest optimal strategy, letting $\beta_1^* = 1$. When the agent observes signal 1 that confirms her prior belief, then she terminates and immediately announces the state 1. But if she is surprised, observing signal 0 that contradicts her prior, then she discards the signal with positive probability $\beta_0^*$ and repeats the experiment. Although finding the exact optimal value $\beta_0^*$ may be difficult, the fact that double-checking one’s own reasoning when one arrives at a surprising conclusion is a common practice suggests that people are able to deviate from the unbiased information-acquisition process in the payoff improving direction.

4.1.1. Comments

(1) The above insight can receive an alternative political economy interpretation. The two states $\theta$ represent left vs right wing policy. Consider a right-wing newspaper that targets the right-wing readers viewed as having a prior belief in favour of the right-wing policy. Readers can only absorb one piece of information (the analog of our information storage constraint) and the newspaper has to decide which piece of information $x$ as generated by $p(x | \theta)$ to choose as its headline. Our model explains why such journals would target their search toward evidence favouring the right-wing policy.6

(2) Meyer (1991) studies optimal biases in a sequential-learning problem of an agent who receives a sequence of signals and, unlike our agent, can aggregate the sequence. Meyer’s main insight is that some asymmetries in the signal structure are optimal. Although optimal

6 This is to be contrasted with the reputation-based explanation of Gentzkow and Shapiro (2006). See also Calvert (1985), Suen (2004), and Che and Mierendorff (2019) for constrained-optimal media-bias models.

© 2019 Royal Economic Society.
asymmetries arise both in her and our frameworks, the two papers study distinct optimisations. While our agent controls termination probabilities in a stationary decision process, Meyer’s agent controls the choice of a Blackwell experiment in each round of a non-stationary process.

4.2. Speed-Accuracy Complementarity

Our model generates the speed-accuracy complementarity effect—a stylised fact stating that delayed choices tend to be less accurate than speedy choices; see the psychology studies of Swensson (1972) and Luce (1986). We establish this effect in the setting from the previous subsection.

Let $\varphi(\theta, a, t)$ be the joint probability distribution of the state $\theta$, chosen action $a$, and the reaction time $t$ generated by the solution $(p, \beta^*, \sigma_i)$ of the repeated-cognition problem.

**COROLLARY 2.** When action 1 is a priori more attractive, $R \in (1, d)$, and the unique primitive binary experiment is symmetric, $p(1 \mid 1) = p(0 \mid 0) > 1/2$, then the probability $Pr_{\varphi}(a = \theta \mid t)$ of the correct choice decreases with response time $t$.

Due to the stationarity of the decision process, the probability of the correct choice conditional on the payoff state is independent of the reaction time: $Pr_{\varphi}(a = \theta \mid t, \theta) = Pr_{\varphi}(a = \theta)$. At optimum, this conditional probability of the correct choice is larger in the a priori more attractive state 1 than in the state 0, reflecting the relative weights of the two states in the objective. Overall, unconditionally on the payoff state, the probability $Pr_{\varphi}(a = \theta \mid t)$ of the correct choice depends on the response time because $t$ correlates with $\theta$. A long response time indicates that the agent has repeatedly encountered the signal value 0 and has hesitated to terminate. Hence, conditional on large $t$, the likelihood of the unattractive state 0 becomes high. The longer the agent has hesitated, the more likely it is that she is facing the unattractive state in which she is making more mistakes.

**PROOF.** $\beta_1^* > \beta_0^*$ by Corollary 1. We let $f_0 = \beta^*_1 p(1 \mid \theta) + \beta^*_0 p(0 \mid \theta)$ denote the probability of termination per each round in state $\theta$. The response time $t$ in the state $\theta$ is geometrically distributed with the decision rate $f_0$: $Pr_{\varphi}(t \mid \theta) = f_0 (1 - f_0)^t$ for $t = 0, 1, \ldots$. Since $p(1 \mid 1) = p(0 \mid 0) > p(1 \mid 0) = p(0 \mid 1)$ and $\beta^*_1 > \beta^*_0$, the decision rate is higher in state 1 than in state 0: $f_1 > f_0$. Thus, the likelihood ratio $Pr_{\varphi}(t \mid \theta = 1)/Pr_{\varphi}(t \mid \theta = 0)$ decreases with $t$, and hence $Pr_{\varphi}(\theta = 1 \mid t)$ decreases in $t$. The fact that $\beta^*_1 > \beta^*_0$, and the symmetry of $p$ implies that the probability of the correct choice is larger in state 1 than in state 0:

$$r \left(1 \mid 1; p, \beta^*, \sigma_i\right) = \frac{\beta^*_1 p(1 \mid 1)}{\beta^*_0 p(0 \mid 1) + \beta^*_1 p(1 \mid 1)} > \frac{\beta^*_0 p(0 \mid 0)}{\beta^*_0 p(0 \mid 0) + \beta^*_1 p(1 \mid 0)} = r \left(0 \mid 0; p, \beta^*, \sigma_i\right).$$

Since $Pr_{\varphi}(a = \theta \mid t) = Pr_{\varphi}(\theta = 1 \mid t) r(1 \mid 1; p, \beta^*, \sigma_i) + Pr_{\varphi}(\theta = 0 \mid t) r(0 \mid 0; p, \beta^*, \sigma_i)$, the result obtains. \(\square\)

4.2.1. Comment

The predictions of our model are in line with the evidence on state recognition problems reported in Ratcliff and McKoon (2008) according to which: (i) the posterior probability of correct recognition is higher when announcing the a priori more likely state, and (ii) late announcements are relatively less precise. This is in contrast to the prediction of the traditional Wald model (see © 1990 Royal Economic Society.
4.3. Overweighting of Rare Events

We consider a state-recognition task in which the two actions are a priori equally attractive, \( R = \pi_0 u_0 / \pi_1 u_1 = 1 \), and \( \pi_0 = \pi_1 = 1/2 \). In contrast to previous applications, the distribution of the signal values \( x = 0, 1 \) is asymmetric across states. Specifically, the probability of \( x = 1 \) in state \( \theta \) is \( \rho_\theta \in (0, 1) \) and the probability of \( x = 0 \) is \( 1 - \rho_\theta \). We assume, essentially without loss of generality, that \( \rho_0 < \rho_1 < 1 - \rho_0 \). The a priori probability of event \( x = 1 \) is \( (\rho_0 + \rho_1)/2 < 1/2 \), and thus the event \( x = 1 \) is relatively rarer than \( x = 0 \). The next result states that at the optimum, the agent is relatively more likely to discard the more common event \( x = 0 \) in agreement with Kahneman and Tversky (1979), who observe that agents tend to overweight rare events.

**Corollary 3.** When the two actions are a priori equally attractive, \( R = 1 \), then the agent is biased in favour of the event \( x = 1 \): \( \beta_1^* > \beta_0^* > 0 \) (and her guess of the state equals the observed signal realisation, i.e., \( \sigma = \sigma_I \)).

**Proof.** This task is a special case of our binary setting with the primitive experiment \( p(x \mid \theta) = \rho_{\theta} \) if \( x = 1 \), \( p(x \mid \theta) = 1 - \rho_{\theta} \) if \( x = 0 \) and with equally a priori attractive actions, \( R = 1 \). Since \( \rho_0 < \rho_1 \), the labelling of the signals satisfies the monotone likelihood property. Since \( R = 1 \in (1/d, d) \), Proposition 2 implies that the agent’s behaviour is stochastic, both \( \beta_0^* \) and \( \beta_1^* \) are positive, and the ratio of the search intensities \( \beta_1^* / \beta_0^* \) satisfies (9). Since \( R = 1 \), (9) simplifies to

\[
\frac{\beta_1^*}{\beta_0^*} = \sqrt{\frac{p(0 \mid 1)}{p(1 \mid 1)}} = \sqrt{\frac{p(0 \mid 1)p(0 \mid 0)}{p(1 \mid 1)p(1 \mid 0)}} = \sqrt{\frac{(1 - \rho_1)(1 - \rho_0)}{\rho_1 \rho_0}}.
\]

The inequality \( \beta_1^* > \beta_0^* \) follows from \( \rho_1 < 1 - \rho_0 \).

For illustration, consider a formation of belief over the probability of a flight accident. The accident probability per flight in the safe state of the world is \( 10^{-6} \), whereas it is \( 10^{-5} \) in the dangerous state of the world, and both states are a priori equally likely. The agent can sequentially observe arbitrarily many past flight outcomes, but cannot aggregate the information, and recalls only the last observed flight. She guesses that the state of the world is dangerous if and only if the last observed flight is eventful.

Consider first an agent who always terminates right after the observation of the first data-point. Such an agent benefits from a ‘second thought’ whenever she observes an uneventful flight: either the second observed flight will be uneventful, in which case the second thought will have been inconsequential, or the redrawn flight will be eventful and the agent will switch her assessment from the safe to the dangerous state. Such a switch is beneficial since conditional on two contradicting data-points the dangerous state is relatively more likely. Thus, relative to the immediate termination strategy, the agent will benefit from discarding the uneventful flight observations with positive probability.

---

Fudenberg *et al.* (2018) for an elaboration of Wald model in which the stakes attached to the correct recognition are a priori unknown.)
5. A Tractable Setting with Multiple States

We now present a class of settings with multiple payoff states that admits a simple analytical solution in the form of a system of linear equations for the optimal termination probabilities. Subsection 5.1 applies this solution in a stylised example that explains salience of perceptually distinct states as a second-best adaptation.

The agent faces a perceptual task that requires her to announce a realisation of the state $\theta \in \Theta$ drawn from a fully mixed prior $\pi \in \Delta(\Theta)$, where $2 < |\Theta| < \infty$. She is endowed with a primitive perception technology that generates a perceived value $\theta'$ of the state. The primitive perception is informative but noisy: the perceived value $\theta'$ equals the true state $\theta$ with a high probability, but mistakes, $\theta' \neq \theta$, occur sometimes. We view the primitive perception technology as a black-box model of a physiological sensor that generates a noisy impression $\theta'$ of the true state $\theta$. The agent can use the sensor repeatedly but is not able to aggregate the information. She conditions the repetition of the sensor’s use on the most recent perception and announces the terminal perception.

We formalise this perception task as follows. The agent makes an announcement $a \in A = \Theta$, where $2 < |\Theta| < \infty$, and receives payoff $u(a, \theta) = u_\theta > 0$ if her announcement is correct, $a = \theta$, and $u(a, \theta) = 0$ if $a \neq \theta$. Each use of the agent’s sensor generates a signal value/perception $\theta'$ in $X = \Theta$, with conditional probabilities $p(\theta' \mid \theta) > 0$. The set $P$ is the singleton $\{p\}$. We make the following assumption.

**Symmetry:** $p(\theta' \mid \theta) = p(\theta \mid \theta')$.

The symmetry assumption leads to a significant simplification of the second-thought-free condition described in Lemma 9 in Appendix. Additionally, we make a simplifying assumption that the agent uses the identity action strategy $\sigma_1$; she announces the state equal to her last perception. We also make the assumption that the optimal termination probabilities $\beta_x$ are positive for all $x \in \Theta$.\(^8\) Let $r^* = r(p, \beta^*, \sigma_1)$ be the optimal feasible choice rule.

**Proposition 3.** The optimal termination probabilities satisfy the system of linear equations,

$$
\sum_{\theta \in \Theta} \beta_\theta^* \frac{p(\tilde{\theta} \mid \theta)}{(\pi_\theta u_\theta p(\theta \mid \theta))^{1/2}} = \sum_{\theta \in \Theta} \beta_{\theta'}^* \frac{p(\tilde{\theta} \mid \theta')}{(\pi_{\theta'} u_{\theta'} p(\theta' \mid \theta'))^{1/2}} \text{ for all } \theta, \theta' \in \Theta. \quad (10)
$$

The proposition implies that the decision rate $f_\theta = \sum_{\theta} \beta_{\theta}^* p(\tilde{\theta} \mid \theta)$ in each state $\theta$ is proportional to $(\pi_\theta u_\theta p(\theta \mid \theta))^{1/2}$ and thus is high in those states that are reliably identified by the primitive experiment and in which the ex ante expected reward for the correct state recognition is high.

### 5.1. Salience

Bordalo et al. (2012) interpret salience as directed attention focus. They quote the popular work by Daniel Kahneman (2011):

‘Our mind has a useful capability to focus on whatever is odd, different or unusual.’

The quote states a causal relation between the two features of the salient phenomena: these are: (i) odd, different or unusual, and because of (i), people benefit from (ii) focusing their attention on such phenomena. Here, we confirm Kahneman’s intuition within our proposed framework. Our

\(^8\) These two assumption are satisfied when $p(\theta \mid \theta)$ is sufficiently close to one for each $\theta$.  

© 2019 Royal Economic Society.
microfoundation of the salience effect is related to the insight emerging in psychological research on visual salience. Itti (2007) conceptualises the visual salience effect as attention allocation to a subset of the visual field that is ‘sufficiently different from its surroundings to be worthy of [one’s] attention’. Similarly, in our model, a payoff state is salient if it stands out sufficiently from similar states to be worthy of the focus of the agent’s information search.

For two states $\theta_1$ and $\theta_2$, we say that $\theta_1$ is more distinct than $\theta_2$ if for each other state $\theta_3 \neq \theta_1, \theta_2, p(\theta_1 \mid \theta_3) < p(\theta_2 \mid \theta_3)$. Suppose for illustration that the perceptual task involves recognition of a color from a set \{azure, indigo, red\}. Intuitively, the red color stands out of this set, and this is captured by the above definition. Assume that the two shades of blue are similar in that the agent’s first impression confuses them in 10% of cases, $p(\text{azure} \mid \text{indigo}) = p(\text{indigo} \mid \text{azure}) = 0.1$, but $p(\theta \mid \text{red}) = p(\text{red} \mid \theta) = 0.01$ for $\theta \in \{\text{azure}, \text{indigo}\}$. Then, the red color is more distinct according to our definition than either of the two blue shades.

We focus on the effect stemming from the agent’s differential ability to perceptually discriminate between the states, and thus we abstract from the differences in the ex ante rewards across states; we assume that $\pi_{\theta}u_{\theta}$ is constant across all states. Additionally, we impose the following assumption:

5.1.1. Sufficient Precision
$p(\theta \mid \theta) > p(\theta' \mid \theta)$ for all $\theta \neq \theta'$.

**Proposition 4.** If state $\theta_1$ is more distinct than state $\theta_2$, then the agent’s terminal perception is biased in favour of the more distinct state $\theta_1$ at the expense of the less distinct state $\theta_2$: $r^*(\theta_1 \mid \theta_2) > r^*(\theta_2 \mid \theta_1)$.

Since the primitive perception technology $p$ is symmetric by assumption, the asymmetry in favour of the distinct state of the optimal terminal perception $r^*$ is driven solely by the optimisation of the termination strategy. To gain the intuition for the salience of the distinct states, consider a state $\theta^*$ that is similar to many other states and an agent who always terminates the process after the first round: $\beta = 1$. This agent is relatively uninformed whenever she forms perception $\theta^*$, since the true state differs from $\theta^*$ with a sizeable probability. The agent with this indistinct perception $\theta^*$ would thus benefit from ‘having a second thought’—i.e., from running the primitive perception formation process once again. The optimal termination strategy involves repeating the primitive process with relatively high probability whenever the agent forms a perception of an indistinct state, and this shifts the terminal perception in favour of the distinct states.

6. Extensions

In the first subsection, we discuss how our model can accommodate agents with more general memory constraints. Subsection 6.2 accommodates agents who discount future payoffs at an exponential rate.

6.1. Sophisticated agents

To demonstrate the flexibility of the general model, we now discuss two specific settings. They feature sophisticated agents with non-trivial memory that can be used to aggregate information

© 2019 Royal Economic Society.
over several observed signal realisations. Perhaps surprisingly, we show that those settings can in fact be interpreted as special cases of our general model that on its face value allows only for trivial memory. We show that such accommodation of non-trivial memory is possible via expansion of the set $\mathcal{P}$ of the primitive experiments. This allows us to establish the generality of the second-thought-free condition.

Moreover, when the state and action spaces are binary, then the setting with sophisticated agents boils down to the simple binary setting as formulated in Section 3, except for the determination of the perceptual-distance parameter $\delta$, which is now endogenously determined by the agent’s ability to process information.

**Example 1** (imperfect information aggregation). This setting relaxes the agent’s inability to aggregate information across the repetitions of her reasoning by endowing her with a finite set of memory states that she can use to represent the signal histories. The setting of this example builds on Hellman and Cover (1970) and Wilson (2014). The agent can repeatedly sample from a single statistical experiment that generates signal realisations from a finite signal space. Additionally, the agent is endowed with a finite set of memory states. After each run of the experiment, the agent randomises between terminating and continuation of the decision process, where in the latter case, she may transition to a new memory state. The termination decisions and the transitions among memory states follow a stationary mixed strategy that conditions on the current memory state and the last observed signal. Once the agent terminates, she maps the last memory state and the last observed signal value to a chosen action. The feasible statistical experiment and the set of memory states specify a set of constructible choice rules, from which the agent chooses the one that maximises her ex ante expected payoff.

The formal specification of this example follows. The agent is endowed with one Blackwell experiment $\mu(x|\theta)$ with a finite signal space $X$ and, additionally, with a finite set $M$ of memory states $m$. After each run of the experiment $\mu$, the agent either terminates or continues with decision-making. If the agent continues, then she transitions from the current memory state to a new memory state and reruns the statistical experiment $\mu(x|\theta)$. That is, the agent selects a (generalisation of the) termination strategy: $\gamma: M \times X \rightarrow \Delta(M \cup \{t\})$, where $\gamma(m'|m,x)$ is the probability that the agent in memory state $m$ who has observed signal realisation $x$ in the last run of the experiment $\mu$ continues with the decision-making and transitions to memory state $m'$, and $\gamma(t|m,x)$ is the probability that such an agent terminates. The terminating agent chooses action $\sigma(m,x)$ that depends both on the current memory state and on the signal realisation observed in the last run of $\mu$. The agent starts the decision-making in the memory state $m_0$. A pair $\gamma, \sigma$ induces a $\theta$-dependent Markov chain over the memory states that eventually terminates with choice $\sigma(m,x)$, where $m$ is the last memory state and $x$ is the last signal realisation observed. Let $p(a|\theta; \gamma, \sigma)$ be the probability that the agent terminates with the choice $a$ in state $\theta$, and let $\mathcal{P}_{iia}$ be the set of all stochastic choice rules $p$ that this agent can construct. She selects the choice rule from $\mathcal{P}_{iia}$ that maximises her ex ante expected payoff.

We now demonstrate that this example is a special case of our baseline model. Consider the baseline model with the signal space $X = A$ and the set of the feasible primitive experiments $\mathcal{P} = \mathcal{P}_{iia}$. The set $\mathcal{R}(\mathcal{P}_{iia}) = \{r(p, \beta, \sigma): p \in \mathcal{P}_{iia}, \beta \in B, \sigma \in S\}$ is then the set of stochastic choice rules that can be constructed as follows. The agent runs any process $p \in \mathcal{P}_{iia}$, and observes
a signal value/action recommendation $a$ with probability $p(a \mid \theta)$. She terminates with probability $\beta_a$, according to the termination strategy $\beta = (\beta_a)_{a \in A}$, and upon the termination chooses an action $a' = \sigma(a)$, where $\sigma \in S$ is any mapping $A \rightarrow A$. She reruns the process $p$ with probability $1 - \beta_a$, observes a new action recommendation generated by $p$, et cetera, until she terminates after a stochastic number of repetitions of the process $p$.

As it turns out, no new choice rules beyond those from $\mathcal{P}_{iia}$ can be constructed by these selective repetitions. This follows because the repetitions of the rule $p \in \mathcal{P}_{iia}$ with the termination strategy $\beta$ can always be replicated with an appropriate choice of a different rule in $\mathcal{P}_{iia}$ that whenever $p$ would terminate with $a$ restarts the process from scratch with probability $1 - \beta_a$.

Formally:

**Lemma 7.** $\mathcal{R}(\mathcal{P}_{iia}) = \mathcal{P}_{iia}$.

According to the lemma, Example 1 is a special case of our baseline model with $\mathcal{P} = \mathcal{P}_{iia}$ and $X = A$, since in such a specification of the baseline model, the set of feasible rules coincides with those in Example 1. In particular, the optimal choice rule $p^* \in \mathcal{P}_{iia}$ solving Example 1 coincides with the optimal rule $r^* \in \mathcal{R}(\mathcal{P}_{iia})$ solving this specification of the baseline model.

The repeated-cognition problem with $\mathcal{P} = \mathcal{P}_{iia}$ is purely formal in that the optimal termination probabilities $\beta^*_x = 1$ for all $x \in X = A$, and thus the agent conducts the optimal process $p^* \in \mathcal{P}_{iia}$ only once and terminates. Nevertheless, the observation that $p^*$ solves the repeated-cognition problem has an important implication.

**Corollary 4.** The choice rule that solves Example 1 (imperfect information aggregation) is second-thought-free.

Wilson (2014) differs from this example mainly in that she assumes exogenous termination probabilities. By adding optimisation over the terminations to the model of Wilson, we gained the partial characterisation of the optimal choice rule with no need to fully solve the problem: one can conclude that the optimal choice rule is second-thought-free without analysing the optimal use of the memory states.

**Example 2 (partial forgetting).** The agent of this example can remember up to a fixed finite number of signal realisations generated by a single statistical experiment. In each round of her decision process, she can discard a subset of the currently remembered signals values, extract a new signal realisation, or terminate, where each of these decisions is determined by a stationary mixed strategy that conditions on the currently remembered stock of the signal values. The statistical experiment and the maximal number of signals that the agent can remember determine the set of stochastic choice rules that she can construct, from which she chooses the rule that maximises her ex ante expected payoff.

We first formalise this example as follows. Let $H$ be the set of signal histories $h$ of length $|h| \leq N$. The agent at a history $h$ can (i) terminate her decision-making, (ii) discard some of the information accumulated, or (iii), if $|h| < N$, acquire a new signal realisation. (i) An agent terminating at $h$ chooses action $\sigma(h)$. (ii) An agent who discards some information transitions to a truncation $h'$ of her current history $h$. $^9$ (iii) An agent who acquires a new signal realisation transitions to a history $hx$, where $x$ is the new signal realisation drawn from $\mu(x \mid \theta)$. The decision-making is governed by a pair of mappings $\gamma : H \times \Theta \rightarrow \Delta(H \cup \{t\})$ and $\sigma : H \rightarrow A$, where

$^9$ A truncation is obtained by deleting one or more last elements in $h$. © 2019 Royal Economic Society.
\( \gamma(h' \mid h, \theta) \) stands for the probability that the agent at history \( h \) in state \( \theta \) continues decision-making and transitions to \( h' \), and \( \gamma(t \mid h, \theta) \) is the probability of termination at history \( h \) in state \( \theta \). The mapping \( \gamma \) is constrained to satisfy (1) \( \gamma(h' \mid h, \theta) \) is independent of \( \theta \) if \( h' \) is a truncation of \( h \), (2) \( \gamma(t \mid h, \theta) \) is independent of \( \theta \), (3) \( \frac{\gamma(hx' \mid h, \theta)}{\gamma(hx \mid h, \theta)} = \frac{\mu_x(\theta)}{\mu_{x'}(\theta)} \), (4) \( \gamma(h' \mid h, \theta) = 0 \) unless \( h' \) is a truncation of \( h \), or \( h' = hx \) for some \( x \in X \) and \( |hx| \leq N \). Constraints 1 and 2 require the agent to condition the decision to discard information or to terminate only on her current history independently of the state. Constraint 3 allows the agent to expand her information set only by running the experiment \( \mu(x \mid \theta) \). Constraint 4 restricts each step of information acquisition to one draw from \( \mu(x \mid \theta) \) or to a partial discarding of the accumulated information. Let \( p(a \mid \theta; \gamma, \sigma) \) be the probability that the agent who employs \( (\gamma, \sigma) \) terminates with action \( a \) in the state \( \theta \). The agent chooses \( \gamma \) and \( \sigma \) to maximise her ex ante expected payoff.

As with the previous example, let \( \mathcal{R}(\mathcal{P}_{pf}) \) be the set of feasible choice rules in our baseline model with the set of feasible primitive experiments \( \mathcal{P} \) identified with \( \mathcal{P}_{pf} \).

**Lemma 8.** \( \mathcal{R}(\mathcal{P}_{pf}) = \mathcal{P}_{pf} \).

Thus, again, the rule \( p^* \in \mathcal{P}_{pf} \) solving this example, and the optimal rule \( r^* \in \mathcal{R}(\mathcal{P}_{pf}) \) coincide, and thus the rule solving the example must be second-thought-free.

**Corollary 5.** The choice rule that solves Example 2 (partial forgetting) is second-thought-free.

Additionally, when the state and action sets are binary, Proposition 2 applies to both examples with \( \overline{d} = \frac{p^*(1|0)p^*(0|0)}{p^*(0|1)p^*(0|1)} \), and thus, relative to the baseline setting in which the agent remembers only one signal, the examples have the same solution except for the determination of the endogenous parameter \( \overline{d} \). Thus, for instance, if the state 1 is a priori more attractive than state 0, then the agent is more likely to make the correct choice in state 1 than in state 0; \( r^*(1 \mid 1) > r^*(0 \mid 0) \). Like in Subsection 4.1, the optimal decision procedure favours the evidence supporting the a priori attractive state.

### 6.2. Impatient Agents

Our baseline model abstracts from the cost of time in that the agent is only concerned with how the repetitions of the signal extraction affect the correlation of the signal with the state. We now incorporate discounting.

We continue to study the baseline model from Section 1, except that the agent discounts future payoffs exponentially with the discount factor \( \delta \in (0, 1) \). To accommodate discounting, we redefine the choice rule induced by the experiment \( p \), the termination strategy \( \beta \) and the action strategy \( \sigma \) as follows.

\[
 r_{\delta}(a \mid \theta; p, \beta, \sigma) = \sum_{t=1}^{\infty} \sum_{x' : \sigma(x_t) = a} \delta^t \rho(x' \mid \theta; p, \beta),
\]

where \( \rho(x' \mid \theta; p, \beta) \) defined in (1) is the conditional probability of the signal history \( x' \). That is, \( r_{\delta}(a \mid \theta; p, \beta, \sigma) \) is the discounted probability of the choice of action \( a \) in the state \( \theta \). When \( \delta = 1 \), then (11) coincides with our baseline definition of the choice rule.
The set of feasible discounted rules is $\mathcal{R}_d(\mathcal{P}) = \{r_\delta(p, \beta, \sigma) : p \in \mathcal{P}, \beta \in B, \sigma \in S\}$. The discounted repeated-cognition problem is to select a feasible rule $r_\delta$ that maximises the expected payoff:

$$\max_{r_\delta \in \mathcal{R}_d(\mathcal{P})} \sum_{\theta \in \Theta, a \in A} \pi_\theta r_\delta(a \mid \theta)u(a, \theta),$$

(12)

where discounting is incorporated in the definition of the feasible rules.

The next result generalises the second-thought-free condition. Let $r_\delta^* = r_\delta(p^*, \beta^*, \sigma^*)$ be the choice rule solving the discounted repeated-cognition problem (12).

**PROPOSITION 5.** If the termination strategy $\beta^*_x \in (0, 1)$ is interior for all $x$ such that $\sigma^*(x) = a$, then

$$\sum_{\theta \in \Theta} \pi_\theta u(a, \theta)r_\delta^*(a \mid \theta) = \delta \sum_{\theta \in \Theta, a' \in A} \pi_\theta u(a', \theta)r_\delta^*(a' \mid \theta)r_\delta^*(a \mid \theta).$$

(13)

The condition has the same interpretation as the second-thought-free condition in the absence of discounting. The left-hand side is the payoff for following the optimal decision process $r_\delta^*$ summed up across all contingencies that terminate with choice of $a$. The right-hand side is the payoff that the agent would get across the same contingencies if she restarted the decision process $r_\delta^*$ instead of the termination.

For illustration, we now revisit the confirmation bias application from Subsection 4.1 with an impatient agent. We find that, unless discounting is too strong, the impatient agent chooses qualitatively the same strategy as the patient one, although the impatient agent speeds up her decision-making by choosing larger termination probabilities.

The setting is as follows. The agent chooses $a \in \{0, 1\}$ and receives $u(a, \theta) = u_0 > 0$ if $a = \theta$, and zero reward otherwise. Action 1 is a priori more attractive than action 0: $\pi_1 u_1 > \pi_0 u_0$. The agent has access to a single primitive experiment $p$ that generates signal values in $X = \{0, 1\}$. The experiment is symmetric with probabilities $p(1 \mid 1) = p(0 \mid 0) = \alpha > 1/2$. We impose a sufficient-informativeness condition that the signal is sufficiently precise relative to the ex ante attractiveness of action 1:

$$\frac{\alpha}{1 - \alpha} > \frac{\pi_1 u_1}{\pi_0 u_0}.$$

**PROPOSITION 6.** The agent chooses the action equal to the last observed signal realisation. She terminates her decision-making immediately after she encounters signal realisation 1: $\beta^*_1 = 1$. When $\delta \in \left(\frac{1}{\alpha + (1 - \alpha)R}, 1\right]$, then the agent who observes $x = 0$ terminates with an interior probability $\beta^*_0 \in (0, 1)$ that decreases in $\delta$. When $\delta \in \left(0, \frac{1}{\alpha + (1 - \alpha)R}\right]$, then the agent terminates immediately: $\beta^*_0 = \beta^*_1 = 1$.

7. Summary

Agents, who cannot comprehend all facts available to them, benefit from selective attention. We show that agents can implement a targeted information search in a process that resembles the natural phenomenon of hesitation. Like a hesitant person, the agent can, conditional on the action contemplated, decide whether she implements the action or whether she will have a second thought, and run the cognition process once more. Such hesitation can be productive, despite consisting of repetitions of the same stochastic cognition process. By conditioning the probability of the repetition on the conclusion of the reasoning, the agent controls the correlation of her terminal conclusion and the payoff state. The optimal decision process arising in our

© 2019 Royal Economic Society.
model exhibits natural hesitation patterns: the agent will have second thoughts—that is, she will repeat her cognition—whenever the expected payoff for the currently favoured choice is inferior to the expected payoff for continuing decision-making. At optimum, the agent terminating the decision-making must be indifferent between terminating with the currently contemplated action, and repeating the process.

In a sense, the condition formalises the concept of a reasonable doubt. Abstracting from many considerations such as information aggregation across the jury members, a jury deciding a trial under common law should be, if using the optimal decision procedure, indifferent between declaring a verdict and announcing a hung jury and initiate retrial.

Let us conclude by reviewing the limitations of our main result. The central assumption—the ability of the agent to freely repeat her decision process—may fail for several reasons. One reason is that the agent may only have access to a limited data set that constrains her to a finite number of repetitions of the primitive decision process, making the optimal termination strategy non-stationary. Another complication arises if the outcomes of distinct runs of the same cognition process are not conditionally independent as assumed in our model; this may arise if some cognition errors are systematic and are likely to emerge in distinct repetitions of the cognition. We conjecture that the second-thought-free condition holds in such a case, with the agent internalising the correlations between the cognition runs.

Appendix A

A.1. Proofs for Section 3

Proof of Lemma 3. Assume that there exists a solution with $\beta_x$ positive for $n > 2$ signals $x \in X$. We show that there exists a solution with $n - 1$ positive signals. The proposition follows from the induction on $n$.

Let us prove the induction step. Fix the primitive experiment $p$ employed by the agent, let $\beta$ be an optimal termination strategy for the given $p$, and let $X'$ be the set of signals with positive $\beta_x$, and write shortly $s(x \mid \theta)$ for the effective experiment $s(x \mid \theta; p, \beta)$ induced by $p$ and $\beta$. Let us abuse notation by letting $s(x) = \sum_\theta \pi_\theta s(x \mid \theta)$ stand for the unconditional effective probability of $x$. For $x \in X'$ let $q_s \in \Delta(\Theta)$ be the posterior belief upon terminating with $x$: $q_s(\theta) = \pi_\theta s(x \mid \theta)/s(x)$.

Since $|X'| > 2$ and the state space $\Theta$ is binary, there exists a signal $x^* \in X'$ such that $q_{x^*}$ is in the convex hull of the posteriors $q_s$, $x \in X' \setminus \{x^*\}$. Let $\mu_x$ be the coefficients that decompose $q_{x^*}$ into $q_s$, $x \in X' \setminus \{x^*\}$. That is, $\mu \in \Delta(X' \setminus \{x^*\})$ such that $q_{x^*}(\theta) = \sum_{x \in X' \setminus \{x^*\}} \mu_x q_s(\theta)$ for all $\theta \in \Theta$.

We will construct an alternative feasible effective experiment $\tilde{s}(x \mid \theta)$ with unconditional probabilities of $x$ denoted by $\tilde{s}(x)$ and the posteriors $\pi_\theta \tilde{s}(x \mid \theta)/\tilde{s}(x)$ denoted by $\tilde{q}_s(\theta)$ such that:

$$
\tilde{s}(x) = \begin{cases} 
  s(x) + s(x^*)\mu_x & \text{if } x \in X' \setminus \{x^*\}, \\
  0 & \text{otherwise},
\end{cases}
$$

(A1)

and

$$
\tilde{q}_s(\theta) = q_s(\theta) \text{ for all } x \in X' \setminus \{x^*\}, \theta \in \Theta.
$$

(A2)

Since the experiment $\tilde{s}$ is more informative than $s$ (in the sense of the Blackwell comparison), there exists a solution with this alternative feasible effective experiment $\tilde{s}$, as needed for the induction step.
It remains to construct $\tilde{s}$. Note that if an effective experiment $s(x \mid \theta; p, \beta) = \frac{\beta_s p(x \mid \theta)}{\sum_{x'} \beta_{x'} p(x' \mid \theta)}$ is induced by some $p$ and $\beta$, then for any vector of probabilities $\tilde{\beta}_x$, the experiment

$$\tilde{s}(x \mid \theta) = \frac{\tilde{\beta}_x s(x \mid \theta; p, \beta)}{\sum_{x' \in X \setminus \{x^s\}} \tilde{\beta}_{x'} s(x' \mid \theta; p, \beta)} = \frac{\tilde{\beta}_x \beta_s p(x \mid \theta)}{\sum_{x' \in X} \tilde{\beta}_{x'} \beta_{x'} p(x' \mid \theta)}$$

is also feasible, since it is induced by $p$ and $\beta' = (\tilde{\beta}_x)_{x \in X}$.

We claim that if $\beta_x \in (0, 1)$ for all $x \in X$, then the resulting $\tilde{s}$ satisfies the properties (A1) and (A2). Let us check:

$$\tilde{s}(x \mid \theta) = \frac{\tilde{\beta}_x s(x \mid \theta)}{\sum_{x' \in X \setminus \{x^s\}} \tilde{\beta}_{x'} s(x' \mid \theta)} = \frac{\tilde{\beta}_x s(x \mid \theta)}{c \left( \sum_{x' \in X \setminus \{x^s\}} s(x' \mid \theta) + \sum_{x' \in X \setminus \{x^s\}} \frac{s(x^s) \mu_{x'}}{s(x')} s(x' \mid \theta) \right)} = \frac{\tilde{\beta}_x s(x \mid \theta)}{c \left( \sum_{x' \in X \setminus \{x^s\}} s(x' \mid \theta) + \sum_{x' \in X \setminus \{x^s\}} \frac{s(x^s) \mu_{x'}}{\pi_{\theta}} q_{x'}(\theta) \right)} = \frac{\tilde{\beta}_x s(x \mid \theta)}{c \left( \sum_{x' \in X \setminus \{x^s\}} s(x' \mid \theta) + s(x^s \mid \theta) \right)} = \frac{\tilde{\beta}_x s(x \mid \theta)}{c} = \left( 1 + \frac{s(x^s) \mu_x}{s(x)} \right)s(x \mid \theta).$$

The property (A1) holds since for all $x \in X \setminus \{x^s\}$:

$$\tilde{s}(x) = \left( 1 + \frac{s(x^s) \mu_x}{s(x)} \right)s(x) = s(x) + s(x^s) \mu_x.$$ 

To establish the property (A2), check that for all $x \in X \setminus \{x^s\}$ and all $\theta \in \Theta$:

$$\tilde{q}_x(\theta) = \frac{\pi_{\theta} \tilde{s}(x \mid \theta)}{\sum_{\theta' \in \Theta} \tilde{s}(x \mid \theta')} = \frac{\pi_{\theta} \left( 1 + \frac{s(x^s) \mu_x}{s(x)} \right)s(x \mid \theta)}{\sum_{\theta' \in \Theta} \left( 1 + \frac{s(x^s) \mu_x}{s(x)} \right)s(x \mid \theta')} = \frac{\pi_{\theta} s(x \mid \theta)}{\sum_{\theta' \in \Theta} s(x \mid \theta')} = q_x(\theta).$$

□
Proof of Lemma 4. For any positive $\beta$,
\[
\frac{\beta_0 p(0 | 0)}{\beta_0 p(0 | 1) + \beta_1 p(1 | 1)} = \frac{\beta_0 p(0 | 0)}{\beta_0 p(0 | 1) + \beta_1 p(1 | 1)} = p(1 | 1) p(0 | 0) = d_p.
\]
Thus, every $r \in \mathcal{R}_{p, \sigma}$ either always selects a same action, or satisfies $\frac{\beta_1 p(1 | 1)}{\beta_0 p(0 | 0)} = d_p.$ Vice versa, if a rule $r'$ satisfies $\frac{\beta_1 p(1 | 1)}{\beta_0 p(0 | 0)} = d_p$, then it belongs to $\mathcal{R}_{p, \sigma}$. To see this, let $r_0$ denote the rule that always selects action $a$. Consider positive $\beta_0$, and note that $r(p, (\beta_0, \beta_1), \sigma)$ is continuous in $\beta_1$, and converges to $r_1$ and $r_0$ as $\beta_1$ approaches 1 and 0. Thus, there exists $\beta$ such that $r'(1 | 1) = r(1 | 1; p, \beta, \sigma).$ Moreover, there is a unique rule $r$ that satisfies $r(1 | 1) = r'(1 | 1)$ and $\frac{\beta_1 p(1 | 1)}{\beta_0 p(0 | 0)} = d_p.$ Thus, $r'$ must be $r(p, \beta, \sigma)$ and hence constructible from $p$. \(\square\)

Proof of Lemma 5. The statement is trivial when $r(p, \beta, \sigma)$ chooses an action $a'$ with probability 1, since then we can set $\beta_a' = 1$ and $\beta_x' = 0$ for $x \neq a'$. Accordingly, assume that both actions are chosen with positive probabilities under the rule $r(p, \beta, \sigma)$ and $\sigma(x) = 1 - x$. For the sake of contradiction, assume that $r(p, \beta, \sigma)$ achieves a higher payoff than all rules constructible with $p$ and $\sigma$. Then, the payoff difference between the rule $r(p, \beta, \sigma)$ and the choice rule that always selects $a = 1$ must be positive:
\[
\pi_0 u_0 r(0 | 0; p, \beta, \sigma) + \pi_1 u_1 r(1 | 1; p, \beta, \sigma) - \pi_1 u_1 = \pi_0 u_0 r(1 | 0; p, \beta, \sigma) + \pi_1 u_1 r(0 | 1; p, \beta, \sigma) - \pi_1 u_1 > 0,
\]
where we have used $u(a | \theta; p, \beta, \sigma) = r(1 - a | \theta; p, \beta, \sigma)$ for the first equality. Similarly, the payoff difference between the rule $r(p, \beta, \sigma)$ and the rule that always selects $a = 0$ must be positive:
\[
\pi_0 u_0 r(0 | 0; p, \beta, \sigma) + \pi_1 u_1 r(1 | 1; p, \beta, \sigma) - \pi_0 u_0 = \pi_0 u_0 r(1 | 0; p, \beta, \sigma) + \pi_1 u_1 r(0 | 1; p, \beta, \sigma) - \pi_0 u_0 = \pi_1 u_1 r(0 | 1; p, \beta, \sigma) - \pi_0 u_0 r(0 | 0; p, \beta, \sigma) > 0.
\]
The last two inequalities imply
\[
\frac{r(1 | 1; p, \beta, \sigma)}{r(1 | 0; p, \beta, \sigma)} < \frac{\pi_0 u_0}{\pi_1 u_1} < \frac{r(0 | 1; p, \beta, \sigma)}{r(0 | 0; p, \beta, \sigma)},
\]
which establishes contradiction because by Lemma 4, the rule $r(a | \theta; p, \beta, \sigma)$ satisfies
\[
r(1 | 1; p, \beta, \sigma) r(0 | 0; p, \beta, \sigma) = p(1 | 1) p(0 | 0).
\]
and therefore it inherits the monotone likelihood ratio property from $p$. \(\square\)

Proof of Lemma 6. Consider the choice rule $r(p, \beta, \sigma)$ constructed from the experiment $p$ with perceptual distance $d_p = d$, and fix the probability $r(0 | 0; p, \beta, \sigma) = \alpha$ of the correct choice

---

10 Rules $r_a$ that always select an action $a$ can be trivially constructed from $p$ and $\sigma$ by using $\beta_a = 1$ and $\beta_x = 0$ for $x \neq a$. © 2019 Royal Economic Society.
in state 0 to a value $\alpha \in (0, 1)$. Then, by Lemma 4, the probability $r(1 \mid 1; p, \beta, \sigma_I)$ of the correct choice in state 1 satisfies
\[ r(1 \mid 1; p, \beta, \sigma_I)\alpha \frac{1 - r(1 \mid 1; p, \beta, \sigma_I)(1 - \alpha)}{d} = 1. \]

For each $\alpha$, the solution for $r(1 \mid 1; p, \beta, \sigma_I)$ of this equation increases in $d$. □

Proof of Proposition 2. The agent’s objective is linear with respect to the choice rule $r(p, \beta, \sigma)$. Thus, the optimal rule is the point of tangency of the set $R_{\pi, \sigma}$ of the feasible rules and of an indifference line; see Figure 2. The slope $\frac{dr(1 \mid 1; \pi, \beta, \sigma)}{dr(1 \mid 1; \pi, \beta, \sigma)}$ is decreasing in $r(1 \mid 1; \pi, \beta, \sigma)$ and attains value $-1/d$ for $r(1 \mid 1; \pi, \beta, \sigma_I) = 0$, and value $-d$ for $r(1 \mid 1; \pi, \beta, \sigma_I) = 1$. Thus, when $R < 1/d$ or $R > d$, then the problem has the corner solution as specified in statements 1 and 2 of the proposition.

When $R \in (1/d, d)$, then the optimal choice rule $r^* = r(\pi, \beta^*, \sigma)$ satisfies the feasibility condition $\frac{r^*(1 \mid 1) - r^*(0 \mid 1)}{r^*(1 \mid 1) - r^*(0 \mid 1)} = d$, the second-thought-free condition (5) (applied to action $a = 1$):
\[ \pi_1u_1r^*(1 \mid 1) = \pi_0u_0r^*(0 \mid 1) + \pi_1u_1r^*(1 \mid 1)r^*(1 \mid 1), \]

and two normalisation conditions $\sum_{\theta} r^*(a \mid \theta) = 1$, for $\theta \in \{0, 1\}$. These four conditions jointly imply the explicit solution for the optimal choice rule in (8). The expression (9) for $\beta^*_1 / \beta^*_0$ follows from (8) and the condition $\frac{r^*(1 \mid 0)}{r^*(0 \mid 0)} = \frac{\beta^*_1[p(1 \mid 0)]}{\beta^*_0[p(0 \mid 0)]}$.

□

A.2. Proofs for Section 5

The next result is an auxiliary lemma used in the proof of Proposition 4.

Lemma 9. The optimal effective choice rule $r^*$ satisfies for any pair of states $\theta, \theta' \in \Theta$:
\[ \pi_\theta u_\theta r^*(\theta \mid \theta')r^*(\theta' \mid \theta') = \pi_\theta u_\theta r^*(\theta \mid \theta)\pi_\theta u_\theta r^*(\theta' \mid \theta'). \quad (A3) \]

Condition (A3) is a strengthening of the second-thought-free condition (5). It requires that the agent who has terminated the decision process with perception $\theta$, and knows that the second run of the process $r^*$ terminates with a value $\theta'$ is indifferent between $\theta$ and $\theta'$. This condition is stronger than the second-thought-free condition (5), since (5) requires (A3) to hold only on average across all $\theta'$. This strengthening holds for the special case of a symmetric experiment $p$.

Proof of Lemma 9. The optimal effective choice rule satisfies the second-thought-free condition (5), equivalent to:
\[ \pi_\theta u_\theta r^*(\theta \mid \theta) = \sum_{\theta' \in \Theta} \pi_{\theta'} u_{\theta'} r^*(\theta \mid \theta')r^*(\theta' \mid \theta') \text{ for all } \theta \in \Theta, \]

which after two algebraic steps gives:
\[ \pi_\theta u_\theta r^*(\theta \mid \theta)(1 - r^*(\theta \mid \theta)) = \sum_{\theta' \neq \theta} \pi_{\theta'} u_{\theta'} r^*(\theta \mid \theta')r^*(\theta' \mid \theta') \text{ for all } \theta \in \Theta, \]
\[ \sum_{\theta' \neq \theta} \pi_{\theta'} u_{\theta'} r^*(\theta \mid \theta')r^*(\theta' \mid \theta') = \sum_{\theta' \neq \theta} \pi_{\theta'} u_{\theta'} r^*(\theta' \mid \theta')r^*(\theta \mid \theta') \text{ for all } \theta \in \Theta. \]

The last system of equations is formally equivalent to the system of balance conditions for a Markov chain. To see this, consider an ergodic Markov chain with transition probabilities from $\theta$
By Lemma 2, we can substitute 
$$\mu \theta$$ 
and thus, for each state \( \theta \), \( \pi_\theta u_\theta r^* (\theta \mid \theta) \) is proportional to the ergodic probability \( \mu (\theta) \) of the state \( \theta \) for the chain with transition probabilities \( r^* (\theta' \mid \theta) \).

Recall that if a Markov chain with transition probabilities \( m(\theta' \mid \theta) \) is reversible, then its stationary distribution \( \mu (\theta) \) satisfies detailed balance conditions

$$
\mu (\theta) m(\theta' \mid \theta) = \mu (\theta') m(\theta \mid \theta') \quad \text{for all } \theta \neq \theta'.
$$

Thus, it suffices to prove that the probabilities \( r^* (\theta' \mid \theta) \) constitute a reversible Markov chain.

Recall that a Markov chain \( m(\theta' \mid \theta) \) is reversible if and only if it satisfies the Kolmogorov criterion, which requires for all sequences of states \( \theta_1, \theta_2, \ldots, \theta_n \),

$$
\frac{m(\theta_2 \mid \theta_1) m(\theta_3 \mid \theta_2) \cdots m(\theta_n \mid \theta_{n-1}) m(\theta_1 \mid \theta_n)}{m(\theta_n \mid \theta_1) m(\theta_{n-1} \mid \theta_n) \cdots m(\theta_2 \mid \theta_1)} = 1. \quad (A4)
$$

The Markov chain with transition probabilities \( p(\theta' \mid \theta) \) given by the primitive experiment \( p \) satisfies the Kolmogorov criterion \( (A4) \) since \( p \) is symmetric by assumption. Finally, for any positive termination strategy \( \beta \), the effective choice rule \( r(\theta' \mid \theta; p, \beta, \sigma_I) \) satisfies the Kolmogorov criterion too. This is because \( r(\theta' \mid \theta; p, \beta, \sigma_I) = \frac{\beta_\theta p(\theta' \mid \theta)}{\sum_\theta \beta_\theta p(\theta' \mid \theta)} \), and when the expressions for \( r(\theta' \mid \theta; p, \beta, \sigma_I) \) are substituted into \( (A4) \), then the terms \( \beta_\theta \) and the denominators cancel out, and hence

$$
\frac{r(\theta_2 \mid \theta_1; p, \beta, \sigma_I) r(\theta_3 \mid \theta_2; p, \beta, \sigma_I) \cdots r(\theta_n \mid \theta_{n-1}; p, \beta, \sigma_I)}{r(\theta_n \mid \theta_1; p, \beta, \sigma_I) r(\theta_{n-1} \mid \theta_n; p, \beta, \sigma_I) \cdots r(\theta_2 \mid \theta_1; p, \beta, \sigma_I)} = \frac{p(\theta_2 \mid \theta_1) p(\theta_3 \mid \theta_2) \cdots p(\theta_n \mid \theta_1)}{p(\theta_n \mid \theta_1) p(\theta_{n-1} \mid \theta_n) \cdots p(\theta_2 \mid \theta_1)} = 1,
$$

as needed. \( \square \)

**Proof of Proposition 3.** Lemma 9 implies for all pairs \( \theta, \theta' \in \Theta \):

$$
\pi_\theta u_\theta r^* (\theta \mid \theta) r^* (\theta' \mid \theta) = r^* (\theta \mid \theta') \pi_\theta u_\theta r^* (\theta' \mid \theta').
$$

By Lemma 2, we can substitute \( r^* (\theta' \mid \theta) = \frac{\beta_\theta^* \beta_{\theta'}^* \pi_\theta u_\theta p(\theta \mid \theta) p(\theta' \mid \theta)}{\sum_\theta \beta_\theta^* \pi_\theta u_\theta p(\theta \mid \theta) p(\theta' \mid \theta)} \), which gives

$$
\frac{\beta_\theta^* \beta_{\theta'}^* \pi_\theta u_\theta p(\theta \mid \theta) p(\theta' \mid \theta)}{\left( \sum_\theta \beta_\theta^* p(\theta \mid \theta) \right)^2} = \frac{\beta_\theta^* \beta_{\theta'}^* p(\theta \mid \theta') \pi_\theta u_\theta p(\theta' \mid \theta')}{\left( \sum_\theta \beta_\theta^* p(\theta \mid \theta') \right)^2}.
$$

Using the symmetry of \( p \) we get

$$
\frac{\sum_\theta \beta_\theta^* p(\theta \mid \theta')}{\sum_\theta \beta_\theta^* p(\theta \mid \theta)} = \left( \frac{\pi_\theta u_\theta p(\theta' \mid \theta')}{\pi_\theta u_\theta p(\theta \mid \theta)} \right)^{1/2}, \quad (A5)
$$

which gives \( (10) \) after rearrangement. \( \square \)

© 1999 Royal Economic Society.
Proof of Proposition 4. To compare \( r^*(\theta_1 \mid \theta_2) \) and \( r^*(\theta_2 \mid \theta_1) \), we write

\[
\frac{r^*(\theta_1 \mid \theta_2)}{r^*(\theta_2 \mid \theta_1)} = \frac{\sum_\beta \hat{p}_\beta \hat{p}(\theta_1 \mid \theta_2)}{\sum_\beta \hat{p}_\beta \hat{p}(\theta_2 \mid \theta_1)} = \left( \frac{\sum_\beta \hat{p}_\beta \hat{p}(\theta_2 \mid \theta_1)}{\sum_\beta \hat{p}_\beta \hat{p}(\theta_1 \mid \theta_2)} \right)^{1/2} = \frac{\hat{p}_\beta \hat{p}(\theta_1 \mid \theta_2)}{\hat{p}_\beta \hat{p}(\theta_2 \mid \theta_1)} = \beta_\beta p^{1/2}(\theta_1 \mid \theta_1),
\]

where we have used (A5) in the second step, and the symmetry of \( p \) and equality of \( \pi_{\beta \theta} \) across \( \theta \) in the last step. Define \( \hat{\beta}_\theta = \beta_\beta p^{1/2}(\theta \mid \theta) \). We need to prove that if \( \theta_1 \) is more distinct than \( \theta_2 \), then \( \hat{\beta}_{\theta_1} > \hat{\beta}_{\theta_2} \).

By (A5), \((\hat{\beta}_\theta)_\theta \) satisfy the system of linear equations:

\[
\sum_{\theta'} D_{\theta \theta'} \hat{\beta}_{\theta'} = 1 \text{ for all } \theta,
\]

where \( D \). We claim that if \( \theta_1 \) is more distinct than \( \theta_2 \), then \( D_{\theta_1 \theta_1} < D_{\theta_2 \theta_2} \) for all \( \theta_1 \neq \theta_2 \). This follows from \( p(\theta_3 \mid \theta_1) < p(\theta_3 \mid \theta_2) \) and from the symmetry of \( p \):

\[
p(\theta_1 \mid \theta_1) = 1 - p(\theta_2 \mid \theta_1) - \sum_{\theta_3 \neq \theta_1, \theta_2} p(\theta_3 \mid \theta_1) > 1 - p(\theta_1 \mid \theta_2) - \sum_{\theta_3 \neq \theta_1, \theta_2} p(\theta_3 \mid \theta_2)
\]

and therefore,

\[
D_{\theta_1 \theta_1} = \frac{p(\theta_1 \mid \theta_3)}{p^{1/2}(\theta_1 \mid \theta_1)p^{1/2}(\theta_3 \mid \theta_3)} < \frac{p(\theta_2 \mid \theta_3)}{p^{1/2}(\theta_2 \mid \theta_2)p^{1/2}(\theta_3 \mid \theta_3)} = D_{\theta_2 \theta_2}.
\]

Thus,

\[
D_{\theta_1 \theta_1} \hat{\beta}_{\theta_1} + D_{\theta_2 \theta_1} \hat{\beta}_{\theta_2} = 1 - \sum_{\theta_3 \neq \theta_1, \theta_2} D_{\theta_1 \theta_1} \hat{\beta}_{\theta_3} > 1 - \sum_{\theta_3 \neq \theta_1, \theta_2} D_{\theta_2 \theta_1} \hat{\beta}_{\theta_3} = D_{\theta_1 \theta_2} \hat{\beta}_{\theta_1} + D_{\theta_2 \theta_1} \hat{\beta}_{\theta_2}.
\]

Using that \( D_{\theta \theta} = 1 \) and \( D_{\theta \theta'} = D_{\theta' \theta} \), we have

\[
\hat{\beta}_{\theta_1} + D_{\theta_2 \theta_1} \hat{\beta}_{\theta_2} > \hat{\beta}_{\theta_2} + D_{\theta_2 \theta_1} \hat{\beta}_{\theta_1}.
\]

The assumption of sufficient precision of \( p \) and symmetry of \( p \) imply that \( D_{\theta_1 \theta_1} < 1 \), and thus \( \hat{\beta}_{\theta_1} > \hat{\beta}_{\theta_2} \), as needed.

\[\square\]

A.3. Proofs of the Results from Section 6

Proof of Lemma 7. All rules feasible in \( \mathcal{P}_{iia} \) are feasible in \( \mathcal{R}(\mathcal{P}_{iia}) \). \( \mathcal{R}(\mathcal{P}_{iia}) \supseteq \mathcal{P}_{iia} \). This is immediate since when \( \beta_a = 1 \) for all \( a \in A \), then \( r(p, \beta, \sigma) = p \) for all \( p \in \mathcal{P}_{iia} \).

It remains to show \( \mathcal{R}(\mathcal{P}_{iia}) \subset \mathcal{P}_{iia} \). Consider \( p(\gamma, \sigma) \in \mathcal{P}_{iia} \) constructed in the setting of Example 1 by the use of the generalised termination strategy \( \gamma(m, x) \), and the action strategy \( \sigma(m, x) \). Recall that \( r(p(\gamma, \sigma), \beta, \bar{\sigma}) \) is the choice rule constructed by repetitions of the rule \( p(\gamma, \sigma) \) according to the termination strategy \( \beta = (\beta_a)_{a \in A} \) and by applying the action strategy \( \bar{\sigma} : A \rightarrow A \) upon the termination. We need to show that there exists \( \gamma' \) and \( \sigma' \) such that \( r(p(\gamma, \sigma), \beta, \bar{\sigma}) = p(\gamma', \sigma') \). This is indeed so when the termination probability \( \gamma'(t \mid m, x) = \gamma(t \mid m, x)\beta_{\sigma(m,x)} \) for \( m \neq m_0 \), the transition probability to the original memory state \( m_0 \) is \( \gamma'(m_0 \mid m, x) = \gamma(m_0 \mid m, x) + \gamma(t \mid m, x) \left( 1 - \beta_{\sigma(m,x)} \right) \), which is the sum of the probabilities that the original process \( \gamma \) transits to \( m_0 \) and that the decision process \( r(p(\gamma, \sigma), \beta, \bar{\sigma}) \) restarts

© 2019 Royal Economic Society.
after termination of \( p(y, \sigma) \). Additionally, for all \( \tilde{m} \neq m_0 \), \( y'(\tilde{m} \mid m, x) = y(\tilde{m} \mid m, x) \). The above choice of \( y' \) implies that the process \( p(y', \sigma') \) replicates the Markov process over the memory states under \( r(p(y, \sigma), \beta, \tilde{\sigma}) \). Finally, to replicate the choices upon terminations, we set the action strategy \( \sigma'(m, x) = \tilde{\sigma}(\sigma(m, x)) \) for all \((m, x)\).

**Proof of Lemma 8.** Again, trivially, \( R(P_{pf}) \supset P_{pf} \), since \( r(p, (1, \ldots, 1), \sigma_I) = p \) for all \( p \in P_{pf} \). Additionally, \( R(P_{pf}) \subset P_{pf} \). This is indeed so because for any \( \beta = (\beta_a)_{a \in A} \) and any \( \tilde{\sigma}: A \rightarrow A, r(p(y, \sigma), \beta, \tilde{\sigma}) = p(y', \sigma') \) where the termination probability \( y'(t \mid h, \theta) = y(t \mid h, \theta) \beta_{\tilde{\sigma}(h)} \), the transition probability to the empty signal history \( \emptyset \) is set to \( y'(\emptyset \mid h, \theta) = y(\emptyset \mid h, \theta) \). Finally, the action strategy is set to \( \sigma'(h) = \tilde{\sigma}(\sigma(h)) \) for all histories \( h \).

**Proof of Proposition 5.** We extend the definition of the effective experiment to the setting with discounting. Let

\[
s_\delta(x \mid \theta; p, \beta) = \sum_t \sum_{x' : x_t = x} \delta^t \rho(x' \mid \theta; p, \beta),
\]

where \( \rho(x' \mid \theta; p, \beta) \) is the probability of the signal history \( x' \) defined in (1). Thus, \( s_\delta(x \mid \theta; p, \beta) \) is the discounted probability that the agent’s last observed signal value is \( x \). It satisfies the recursion:

\[
s_\delta(x \mid \theta; p, \beta) = \beta_x p(x \mid \theta) + \delta \sum_{x' \in X} \left(1 - \beta_x \right) p(x' \mid \theta) s_\delta(x \mid \theta; p, \beta), \quad (A6)
\]

where the first summand is the probability that the decision process terminates with \( x \) in the first round and the second summand is the discounted probability that the process terminates with \( x \) later. Solving (A6) for \( s_\delta \) gives

\[
s_\delta(x \mid \theta; p, \beta) = \frac{\beta_x p(x \mid \theta)}{1 - \delta + \delta \sum_{x' \in X} \beta_{x'} p(x' \mid \theta)}.
\]

The discounted repeated-cognition problem (12) is thus equivalent to

\[
\max_{p \in P, \beta \in B, \sigma \in \mathcal{S}} \sum_{\theta \in \Theta} \sum_{x \in X} \pi_\theta \frac{\beta_x p(x \mid \theta)}{1 - \delta + \delta \sum_{x' \in X} \beta_{x'} p(x' \mid \theta)} u(\sigma(x, \theta)). \quad (A7)
\]

Consider \( x \) with an interior termination probability \( \beta_x^* \in (0, 1) \) and let \( a = \sigma^*(x) \). The first-order condition of the problem (A7) with respect to \( \beta_x \) is:

\[
\sum_{\theta \in \Theta} \pi_\theta \frac{s_\delta(x \mid \theta; p^*, \beta^*)}{\beta_x^*} u(a, \theta) - \delta \sum_{\theta \in \Theta, x' \in X} \pi_\theta s_\delta(x' \mid \theta; p^*, \beta^*) \frac{s_\delta(x \mid \theta; p^*, \beta^*)}{\beta_x^*} u(\sigma^*(x'), \theta) = 0,
\]

\[
\sum_{\theta \in \Theta} \pi_\theta \frac{s_\delta(x \mid \theta; p^*, \beta^*)}{\beta_x^*} u(a, \theta) - \delta \sum_{\theta \in \Theta, a' \in A} \pi_\theta r^*_\sigma(a' \mid \theta; p^*, \beta^*, \sigma^*) \frac{s_\delta(x \mid \theta; p^*, \beta^*)}{\beta_x^*} u(a', \theta) = 0,
\]

where we have summed over all \( x' \) such that \( \sigma^*(x') = a' \) in the second line. Multiplication by \( \beta_x^* \) and summation over all \( x \) such that \( \sigma^*(x) = a \) gives (13).

**Proof of Proposition 6.** Due to the condition that \( \alpha/(1 - \alpha) > R \), any \((\beta, \sigma)\) that leads to a selection of only one action with certainty is dominated by the decision process that terminates after the first round and chooses an action equal to the observed signal value. Thus, both \( \beta_0^* \) and \( \beta_1^* \) are positive, and the action strategy is \( \sigma^*(x) = x \) or \( \sigma^*(x) = 1 - x \). Let us show that the action strategy \( \sigma^* \) must be the identity function \( \sigma_I \).
Assume for contradiction that \( \sigma^*(x) = 1 - x \). The payoff difference between the rule \( r_\delta(p, \beta^*, \sigma^*) \) and the choice rule that always selects \( a = 1 \) must be positive, since the latter is dominated:

\[
\begin{align*}
\pi_0 u_0 r_\delta(0 \mid 0; p, \beta^*, \sigma^*) + \pi_1 u_1 r_\delta(1 \mid 1; p, \beta^*, \sigma^*) - \pi_1 u_1 &= \\
\pi_0 u_0 r_\delta(1 \mid 1; p, \beta^*, \sigma_I) + \pi_1 u_1 r_\delta(0 \mid 1; p, \beta^*, \sigma_I) - \pi_1 u_1 &\geq
\end{align*}
\]

where the first inequality follows from the fact that any discounted choice rule satisfies \( \sum_\delta r_\delta(a \mid \theta; p, \beta, \sigma) \leq 1 \). Similarly, the payoff difference between the rule \( r_\delta(p, \beta^*, \sigma^*) \) and the rule that always selects \( a = 0 \) must be positive:

\[
\begin{align*}
\pi_0 u_0 r_\delta(0 \mid 0; p, \beta^*, \sigma^*) + \pi_1 u_1 r_\delta(1 \mid 1; p, \beta^*, \sigma^*) - \pi_0 u_0 &= \\
\pi_0 u_0 r_\delta(1 \mid 0; p, \beta^*, \sigma_I) + \pi_1 u_1 r_\delta(0 \mid 1; p, \beta^*, \sigma_I) - \pi_0 u_0 &\geq
\end{align*}
\]

The last two inequalities imply:

\[
\begin{align*}
r_\delta(0 \mid 1; p, \beta^*, \sigma_I) > r_\delta(0 \mid 0; p, \beta^*, \sigma_I) > r_\delta(1 \mid 0; p, \beta^*, \sigma_I) > r_\delta(1 \mid 1; p, \beta^*, \sigma_I).
\end{align*}
\]

This establishes contradiction because as shown in the proof of Proposition 5, \( r_\delta(x \mid \theta; p, \beta^*, \sigma_I) = s_\delta(x \mid \theta; p, \beta^*) = \frac{\beta^*_\delta p(x \mid \theta)}{1 - \delta \sum_{\theta'} \beta^*_\delta p(x' \mid \theta)}, \) and thus

\[
\frac{r_\delta(1 \mid 1; p, \beta^*, \sigma_I)r_\delta(0 \mid 0; p, \beta^*, \sigma_I)}{r_\delta(0 \mid 1; p, \beta^*, \sigma_I)r_\delta(1 \mid 0; p, \beta^*, \sigma_I)} = \frac{p(1 \mid 1)p(0 \mid 0)}{p(0 \mid 1)p(1 \mid 0)} > 1.
\]

Further, it must hold that \( \beta^*_0 = 1 \) or \( \beta^*_1 = 1 \). Otherwise, if both \( \beta^*_0 < 1 \) and \( \beta^*_1 < 1 \), then the agent can increase both \( \beta^*_x \) by a same factor. This preserves the conditional action distribution in each state \( \theta \) and increases the decision rates in both states, and thus it is a profitable deviation.

Additionally, it must be that \( \beta^*_1 = 1 \): using the expressions for \( s_\delta(\theta \mid \theta; p, \beta) = r_\delta(\theta \mid \theta; p, \beta, \sigma_I) \), the payoff for \( \sigma_I \) and \( (\beta_0, \beta_1) = (\beta, 1) \) is

\[
\pi_0 u_0 1 - \delta + \delta(\beta \alpha + 1 - \alpha) + \pi_1 u_1 1 - \delta + \delta(\alpha + \beta(1 - \alpha)), \tag{A8}
\]

and payoff for \( \sigma_I \) and \( (\beta_0, \beta_1) = (1, \beta) \) is

\[
\pi_0 u_0 1 - \delta + \delta(\alpha + \beta(1 - \alpha)) + \pi_1 u_1 1 - \delta + \delta(\beta \alpha + 1 - \alpha). \tag{A9}
\]

The assumptions that \( \pi_1 u_1 > \pi_0 u_0 \) and that \( \alpha > 1/2 \) imply that, for any \( \beta \in (0, 1) \), (A8) exceeds (A9), as needed.

It therefore remains to find \( \beta^*_0 \in (0, 1) \). If the optimal value is interior, then it satisfies (13) with \( a = 0 \):

\[
\pi_0 u_0 r_\delta(0 \mid 0; p, \beta^*, \sigma_I) = \delta \left( \pi_0 u_0 r_\delta^2(0 \mid 0; p, \beta^*, \sigma_I) + \pi_1 u_1 r_\delta(1 \mid 1; p, \beta^*, \sigma_I) r_\delta(0 \mid 1; p, \beta^*, \sigma_I) \right).
\]

After the substitution of \( r_\delta(x \mid \theta; p, \beta, \sigma_I) = \frac{\beta^*_\delta p(x \mid \theta)}{1 - \delta \sum_{\theta'} \beta^*_\delta p(x' \mid \theta)} \), this condition simplifies into a quadratic equation for \( \beta^*_0 \). When \( \delta < \frac{1}{\alpha + (1 - \alpha) \delta} \), then this condition does not have an interior
solution and the derivative of the value \((A8)\) with respect to \(\beta_0\) at \(\beta_0 = 1\) is positive. Thus, in this case, the unique \(\beta^*_0\) satisfying the first-order condition is \(\beta^*_0 = 1\).

When \(\delta > \frac{1}{\alpha+1}\), then the condition has an interior solution and the derivative of the value \((A8)\) with respect to \(\beta_0\) at \(\beta_0 = 1\) is negative. Thus, for this range of parameters, the unique \(\beta^*_0\) satisfying the first-order condition is the interior value that solves the quadratic equation, solution of which decreases in \(\delta\). \(\square\)

**PSE and UCL**

*University of Zurich and CERGE-EI*

**References**


© 2019 Royal Economic Society.


© 2019 Royal Economic Society.