Selective Sampling with Information-Storage Constraints

Philippe Jehiel, Jakub Steiner

Selective Sampling with Information-Storage Constraints

Philippe Jehiel
Jakub Steiner

Selective Sampling with Information-Storage Constraints[*]

Philippe Jehiel

PSE and UCL

Jakub Steiner

University of Zurich and CERGE-EI

January 30, 2019

Abstract

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 a simple class of binary problems— include (i) confirmation bias, (ii) speed-accuracy complementarity, (iii) overweighting of rare events, and (iv) salience effect.

keywords: bounded rationality, information processing, stochastic choice, confirmation bias, speed-accuracy complementarity, probability weighting, salience.

JEL codes: D8, D9.

1

1 Introduction

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 analyze, 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 realization 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 realization 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 when the agent cannot aggregate information—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 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 -2 if she invests in the bad state, and receives 0 when she does not invest whatever the state. Both states are a priori equally likely and 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 realizations in sequence but remembers only the last one when making her investment choice. As follows from simple optimization 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 good prefers to rerun the decision process, since the new run will either lead to investing again or will give rise to the signal realization that conflicts with the first observation and will lead to not investing. Since, conditional on two conflicting signals, not investing is preferred, the agent benefits from having second thoughts when the first observed

signal is good.

In the sequel, we interpret the probability of terminating the decision process after receiving a particular signal as a search intensity for the 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 good 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 behavioral stylized facts. The unifying principle of our behavioral insights is the intuition that the agent targets her search towards the type of evidence that would provide her with more informed posteriors under the uniform search. This principle generates *confirmation bias*, 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*; 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 when 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 behavioral application, we show 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.[1] 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 ensuring that the termination strategies are optimally determined.

There is a wide range of studies that propose different modeling of optimization over information structures. Relative to rational inattention models (Sims, 2003), we provide a procedural

---

[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.

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 has a similar flexibility in her choice of information structures as in the sequential-sampling model of Wald (1945) with perfect information aggregation but our approach allows for a simple derivation of the speed-accuracy complementarity which is less immediate to obtain with drift diffusion models (see however Fudenberg et al. (2018)). Our approach is simpler than most of studies based on finite automata (Hellmann and Cover, 1970; Wilson, 2014) in that it focuses solely on the optimal determination of termination probabilities (as a function of the observed signal). This allows us to derive simple necessary conditions that apply to all models in which such termination strategies are optimally determined, which turns out to be all that is needed to derive connections to psychological biases as reported above.

This paper belongs to a growing economic literature that explains behavioral stylized facts as the constrained-optimal behavior 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) endogenize discounting; and Wilson (2014), Compte and Postlewaite (2012), and Leung (2017) establish constrained-optimal ignorance of weakly informative news.[2]

2 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 Θ and A are finite. The agent chooses a Blackwell experiment p, where p is a family of conditional signal distributions $p(x \mid \theta)$ that depend on $\theta \in \Theta$. The experiment generates a signal realization x from a finite space X. The conditional signal distributions are fully mixed: $p(x \mid \theta) > 0$ for all x, θ. We allow the agent to choose among possibly several such experiments and we let $\mathcal{P}$ denote the 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 realization. She chooses a vector $\beta = (\beta_x)_{x \in X} \in B = [0, 1]^{|X|} \setminus \{(0, \ldots, 0)\}$ of termination probabilities β_x for each signal realization x; we call β a *termination strategy*. The agent

---

[2]Somewhat less related is a literature that explores how exogenous errors in learning lead to behavioral biases; see coarse learning in Jehiel (2005) and its application to overoptimism in Jehiel (2018). By contrast, in our approach, the agent optimizes the error distribution given the constraints.
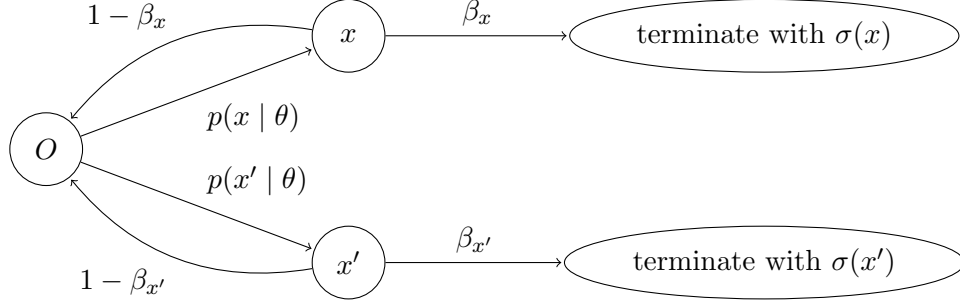
Figure 1: For each (p, β, σ), the decision process is a Markov chain evolving on the agent's states of mind, with transition probabilities that depend on the payoff state θ. The chain begins in the state of mind O and transits to states $x \in X$ with probabilities $p(x \mid \theta)$. The process returns to O with probability $1 - \beta_x$, or terminates with choice of $a = \sigma(x)$ with probability β_x.

runs the experiment p for the first time, receives signal realization x_1 with probability $p(x_1 \mid \theta)$ and terminates the reasoning with probability β_{x_1}. She restarts with the complementary probability $1 - \beta_{x_1}$, and receives a signal realization x_2 from a new run of the process p with probability $p(x_2 \mid \theta)$, terminates with probability β_{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 β_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 realization x, she selects an action $a = \sigma(x)$ according to an *action strategy* $\sigma : X \longrightarrow A$.[3] Let S be the set of all mappings from X to 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 $\mathbf{x}^t = (x_1, \ldots, x_t)$ is

$$\rho\left(\mathbf{x}^t \mid \theta; p, \beta\right) = \beta_{x_t} p(x_t \mid \theta) \prod_{l=0}^{t-1} (1 - \beta_{x_l}) \, p\left(x_l \mid \theta\right). \tag{1}$$

We let

$$r(a \mid \theta; p, \beta, \sigma) = \sum_t \sum_{\mathbf{x}^t : \sigma(x_t) = a} \rho\left(\mathbf{x}^t \mid \theta; p, \beta\right) \tag{2}$$

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

---

[3] We do not allow for mixed action strategies since the optimum can always be achieved with a pure action strategy.

constructed by some p, β, σ.

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

$$\max_{r \in \mathcal{R}(\mathcal{P})} \sum_{\theta \in \Theta, a \in A} \pi_\theta r(a \mid \theta) u(a, \theta). \tag{3}$$

The optimization in (3) can be an outcome of selective pressures that favor 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.

3   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.

3.1   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 π 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,*

$$\mathrm{E}_\alpha[u(a_1, \theta) \mid a_1 = a] \geq \mathrm{E}_\alpha[u(a_2, \theta) \mid a_1 = a], \tag{4}$$

where the expectations are with respect to the random variables θ 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:*

$$\mathrm{E}_\alpha[u(a_1, \theta) \mid a_1 = a] = \mathrm{E}_\alpha[u(a_2, \theta) \mid a_1 = a]. \tag{5}$$

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

$$\mathrm{E}_\alpha\left[u(a_1, \theta)\right] = \mathrm{E}_\alpha\left[\mathrm{E}_\alpha\left[u(a_1, \theta) \mid a_1\right]\right] > \mathrm{E}_\alpha\left[\mathrm{E}_\alpha\left[u(a_2, \theta) \mid a_1\right]\right] = \mathrm{E}_\alpha\left[u(a_2, \theta)\right],$$

which contradicts that a_1 and a_2 are conditionally iid. $\qquad\square$

3.2 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 realization 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 β 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 β_x.

Lemma 2. *An agent who employs a primitive experiment p and a termination strategy β terminates with x in state θ with probability:*

$$s(x \mid \theta; p, \beta) = \frac{\beta_x p(x \mid \theta)}{\sum_{x' \in X} \beta_{x'} p(x' \mid \theta)}. \tag{6}$$

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). $\qquad\square$

7

The lemma implies that $s(p, \beta)$ and hence also $r(p, \beta, \sigma)$ are homogeneous of degree zero with respect to β. Thus, since we abstract from the delay costs, for any optimal termination strategy β^*, $\alpha\beta^*$ for $\alpha \in (0, 1)$ is optimal too, and it generates the same optimal choice rule r^* as β^*. This multiplicity of implementation of the optimal choice rule would disappear in a natural approximation of our model with exponential discount factor approaching 1. Such approximation would select the quickest available decision process that implements the optimal feasible rule r^*; that is, it would impose that $\max_{x \in X} \beta_x = 1$.

Proof of Proposition 1. Using (6), we rewrite the objective as follows.

$$\max_{p \in \mathcal{P}, \beta \in B, \sigma \in S} \sum_{\theta \in \Theta, x \in X} \pi_\theta \frac{\beta_x p(x \mid \theta)}{\sum_{x' \in X} \beta_{x'} p(x' \mid \theta)} u(\sigma(x), \theta). \tag{7}$$

Let rule $r(p^*, \beta^*, \sigma^*)$ solve the repeated-cognition problem. Consider an action a chosen with a positive probability and x such that $\sigma^*(x) = a$ and $\beta_x^* > 0$. The constraint $\beta_x \geq 0$ is not binding for this x, and the first-order condition of (7) with respect to β_x is:

$$\sum_{\theta \in \Theta} \pi_\theta \frac{s(x \mid \theta; p^*, \beta^*)}{\beta_x^*} u(a, \theta) - \sum_{\theta \in \Theta, x' \in X} \pi_\theta s(x' \mid \theta; p^*, \beta^*) \frac{s(x \mid \theta; p^*, \beta^*)}{\beta_x^*} u(\sigma^*(x'), \theta) =$$

$$\sum_{\theta \in \Theta} \pi_\theta \frac{s(x \mid \theta; p^*, \beta^*)}{\beta_x^*} u(a, \theta) - \sum_{\theta \in \Theta, a' \in A} \pi_\theta r(a' \mid \theta; p^*, \beta^*, \sigma^*) \frac{s(x \mid \theta; p^*, \beta^*)}{\beta_x^*} u(a', \theta) \geq 0.$$

Multiplication by β_x^*, summation over all x such that $\sigma^*(x) = a$, and division by $\sum_\theta \pi_\theta r(a \mid \theta; p^*, \beta^*, \sigma^*)$ gives (4). Thus, the terminating agent weakly prefers termination to continuation. Lemma 1 implies the indifference (5). $\qquad \square$

Comment. Our agent can be viewed as having imperfect recall in the sense of Piccione and Rubinstein (1997). Our approach corresponds to their ex ante approach, and the insight of Proposition 1 can be related to the observation in their absent-minded driver example that the ex ante optimal solution is also a (modified) multi-self equilibrium in which the decision problem is viewed as a team composed of multiple selves all sharing the decision-maker's objective.

4 Analytical solution of the binary setting

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, $u(a, \theta) = u_\theta > 0$ if $a = \theta$ and $u(a, \theta) = 0$ otherwise. State θ is drawn from an interior prior $\pi \in \Delta(\Theta)$. The set $\mathcal{P}$ of the feasible statistical experiments is a singleton, and the unique feasible experiment p delivers signals from $X = \{0, 1\}$. (Lemma 5 in Appendix states that the restriction to binary signals is without loss of generality since when $|X| > 2$ then there exists a solution in which the agent terminates with a positive probability for at most two signals values.) We choose without loss of generality the signals labels such that

the likelihood ratio $p(1 \mid \theta)/p(0 \mid \theta)$ increases in θ. Signal 1 is thus relatively indicative of the state 1. We introduce a summary parameter d,

$$d = \frac{p(0 \mid 0)p(1 \mid 1)}{p(0 \mid 1)p(1 \mid 0)}, \tag{8}$$

and refer to d as to the perceptual distance between the two states. This summary statistics is independent of the agent's incentives and of the prior.[4] Finally, a parameter $R = \frac{\pi_1 u_1}{\pi_0 u_0}$ summarizes the incentives and prior; it measures the relative a priori attractiveness of action 1.

Proposition 2. *1. When $R \geq d$ ($R \leq 1/d$), then the agent always chooses action 1 (action 0);*

2. when $R \in (1/d, d)$, then the agent chooses both actions with positive probabilities and

$$r^*(1 \mid 1) = \frac{dR - \sqrt{dR}}{(d-1)R}, \; r^*(0 \mid 0) = \frac{d - \sqrt{dR}}{d - 1}, \tag{9}$$

$$\frac{\beta_1^*}{\beta_0^*} = \frac{dR - \sqrt{dR}}{\sqrt{dR} - R} \frac{p(0 \mid 1)}{p(1 \mid 1)}. \tag{10}$$

Thus, for the binary setting the optimal choice rule exists and is unique. The ratio β_1^*/β_0^* of the termination probabilities is unique too, if both actions are chosen with positive probabilities. 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.
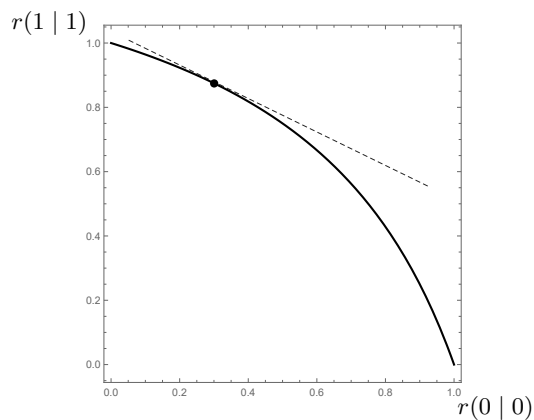
Figure 2: Each point in $[0,1]^2$ corresponds to a choice rule. The depicted curve is a set $\mathcal{R}$ of the choice rules constructible from the experiment p. The objective is linear in the choice rule, and the indifference curves are downward sloping lines. The dashed line is the tangential indifference curve and the dot depicts the solution of the repeated-cognition problem.

---

[4]In the working paper of this paper, we endow the agent with finite set of primitive experiments and show that the agent selects the one that maximizes d.

Proof. The effective experiment $s(x \mid \theta; p, \beta) = s(x \mid \theta)$ satisfies for any $\beta_x > 0$, $x = 0, 1$,

$$\frac{s(1 \mid 1)s(0 \mid 0)}{s(1 \mid 0)s(0 \mid 1)} = \frac{\frac{\beta_1 p(1|1)}{\sum_x \beta_x p(x|1)} \frac{\beta_0 p(0|0)}{\sum_x \beta_x p(x|0)}}{\frac{\beta_1 p(1|0)}{\sum_x \beta_x p(x|0)} \frac{\beta_0 p(0|1)}{\sum_x \beta_x p(x|1)}} = \frac{p(1 \mid 1)p(0 \mid 0)}{p(1 \mid 0)p(0 \mid 1)} = d.$$

Thus, s inherits the increasing likelihood property from p, and hence the identity action strategy $\sigma_I(x) = x$ dominates the strategy $\sigma(x) = 1 - x$. Therefore, the optimal choice rule lies in the compact set $\mathcal{R} = \{r : r(1 \mid 1)r(0 \mid 0) = dr(1 \mid 0)r(0 \mid 1)\}$ depicted on Figure 2. The agent's objective is linear in r and the optimal rule is the point of tangency of $\mathcal{R}$ and of an indifference line. When $R < 1/d$ or $R > d$, then the problem has the corner solution. When $R \in (1/d, d)$, then the optimal rule r^* satisfies the feasibility condition $\frac{r^*(1|1)r^*(0|0)}{r^*(0|1)r^*(1|0)} = d$, the second-thought-free condition (5) (applied to action $a = 1$):

$$\pi_1 u_1 r^*(1 \mid 1) = \pi_0 u_0 r^*(0 \mid 0)r^*(1 \mid 0) + \pi_1 u_1 r^*(1 \mid 1)r^*(1 \mid 1),$$

and two normalization conditions $\sum_a r^*(a \mid \theta) = 1$, for $\theta = 0, 1$. These four conditions jointly imply (9) and (10) follows from (9) and the condition $\frac{r^*(1|\theta)}{r^*(0|\theta)} = \frac{\beta_1^* p(1|\theta)}{\beta_0^* p(0|\theta)}$. $\qquad \square$

5 Behavioral applications

We present three behavioral effects illustrated in the binary setting from Section 4: confirmation bias, speed-accuracy complementarity, and overweighting of rare events. The fourth effect—salience of distinct states—is presented in a setting with multiple states.

5.1 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 memorize 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. (2016). 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, $R \in (1, d)$, and the unique primitive experiment is symmetric, $p(1 \mid 1) = p(0 \mid 0)$, then the agent searches relatively more intensively for signal value 1: $\beta_1^* > \beta_0^*$.*

Proof. Since β_1^*/β_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 (10),

$$\frac{\beta_1^*}{\beta_0^*} = \sqrt{d}\frac{p(0 \mid 1)}{p(1 \mid 1)} = \sqrt{\frac{p(0 \mid 0)p(0 \mid 1)}{p(1 \mid 1)p(1 \mid 0)}} = 1,$$

where the last equality follows from the symmetry of p. $\qquad\square$

To see the connection to confirmation bias, consider an agent whose task is to announce the realized 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 favor 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 β_0^* and repeats the experiment. Although finding the exact optimal value β_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.

Comments. 1) The above insight can receive an alternative political economy interpretation. The two states θ represent left vs right wing policy. Consider a right-wing newspaper that targets the right-wing readers viewed as having a prior belief in favor 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 $p(x \mid \theta)$ to choose as its headline. Our model explains why such journals would target their search toward evidence favoring the right-wing policy.[5]

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 asymmetries arise both in her and our frameworks, the two papers study distinct optimizations. 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.

---

[5]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 (2016) for constrained-optimal media-bias models.

5.2 Speed-accuracy complementarity

Our model generates the speed-accuracy complementarity effect—a stylized 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: $u_0 = u_1 = 1$, $\pi_1 > \pi_0$, considering again a symmetric primitive experiment, $p(1 \mid 1) = p(0 \mid 0) > 1/2$.

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

Corollary 2. *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: $\Pr_\varphi(a = \theta \mid \theta, t) = \Pr_\varphi(a = \theta \mid \theta)$. At optimum, this conditional probability of the correct choice is larger in the a priori more likely 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 θ. A long response time indicates that the agent has repeatedly encountered the surprising signal value 0 and has hesitated to terminate. Hence, conditional on large t, the likelihood of the surprising state becomes high. The longer the agent has hesitated, the more likely it is that she is facing the surprising state in which she is making more mistakes.

Proof. $\beta_1^* > \beta_0^*$ by Corollary (1) since $\pi_1 u_1 > \pi_0 u_0$. We let $f_\theta = \beta_1^* p(1 \mid \theta) + \beta_0^* p(0 \mid \theta)$ denote the probability of termination per each round in state θ. The response time t in the state θ is geometrically distributed with the decision rate f_θ: $\Pr_\varphi(t \mid \theta) = f_\theta(1 - f_\theta)^t$ 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. $\qquad \square$

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 1) The posterior probability of correct recognition is higher when announcing the a priori more likely state and 2) Late announcements are relatively less precise. This is in contrast to the prediction of the traditional Wald model (see Fudenberg et al. (2018) for an elaboration of Wald model in which the stakes attached to the correct recognition are a priori unknown).

5.3 Overweighting of rare events

We continue to consider the state-recognition task; $u_0 = u_1 = 1$. In contrast to previous applications, we assume that the two states $\theta = 0, 1$ are equally likely, but the distribution of the signal values $x = 0, 1$ is asymmetric across states. Specifically, the probability of $x = 1$ in the state θ is $\rho_\theta \in (0, 1)$ and the probability of $x = 0$ is $1 - \rho_\theta$. We assume, without loss of generality, that $\rho_0 < \rho_1 < 1 - \rho_0$.[6] 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. *At the optimum, the agent is biased in favor of the rare event $x = 1$: $\beta_1^* > \beta_0^* > 0$ (and her guess of the state equals the observed signal realization, i.e. $\sigma = \sigma_I$).*

Proof. The belief formation problem studied 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 labeling of the signals satisfies the monotone likelihood property. Since $R = 1 \in (1/d, d)$, Proposition 2 implies that the agent's behavior is stochastic, both β_0^* and β_1^* are positive, and the ratio of the search intensities β_1^*/β_0^* satisfies (10). Since $R = 1$, (10) simplifies to

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

The inequality $\beta_1^* > \beta_0^*$ follows from $\rho_0 < \rho_1 < 1 - \rho_0$. $\qquad \square$

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 datapoint. 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.

---

[6]We can always achieve this by relabeling the states θ and the signals values x, unless $\rho_0 = \rho_1$ or $\rho_0 = 1 - \rho_1$.

5.4 Salience

Bordalo, Gennaioli, and Shleifer (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 microfoundation of the salience effect is related to the insight emerging in psychological research on visual salience. Itti (2007) conceptualizes 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.

As before, the agent faces a perceptual task that requires her to announce a realization of the random state θ. She is endowed with a primitive perception technology that generates a perceived value θ' of the state. The primitive perception is informative but noisy: the perceived value θ' equals the true state θ 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 first impression θ' of the true state θ. 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 formalize this perception task as follows. The agent makes an announcement $a \in A = \Theta$, where $2 < |\Theta| < \infty$, and receives payoff $u(a, \theta) = 1$ if her announcement is correct, $a = \theta$, and $u(a, \theta) = 0$ if $a \neq \theta$. The prior is uniform. Each use of the agent's sensor generates a signal value/perception $\theta' \in X = \Theta$, with conditional probabilities $p(\theta' \mid \theta)$. We make two assumptions on p:

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

Sufficient precision: $p(\theta \mid \theta) > p(\theta' \mid \theta)$ for all $\theta \neq \theta'$.

For two states θ_1 and θ_2, we say that θ_1 is *more distinct* than θ_2 if for each 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 make a simplifying assumption that the agent uses the identity action strategy σ_I; she announces the state equal to her last perception. We also make a regularity assumption that the optimal termination probabilities β_x are positive for all $x \in \Theta$.[7] Let $r^* = r(p, \beta^*, \sigma_I)$ be the optimal

---

[7]This regularity assumption must be satisfied when $p(\theta \mid \theta)$ is sufficiently close to one for each θ, since then any

feasible choice rule.

Proposition 3. *If state θ_1 is more distinct than state θ_2, then the agent's terminal perception is biased in favor of the more distinct state θ_1 at the expense of the less distinct state θ_2:*

$$r^*(\theta_1 \mid \theta_2) > r^*(\theta_2 \mid \theta_1).$$

See Appendix for the proofs omitted in the main text.

Since the primitive perception technology p is symmetric by assumption, the asymmetry in favor of the distinct state of the optimal terminal perception r^* is driven solely by the optimization of the termination strategy. To gain the intuition for the salience of the distinct states, consider a state θ^* 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 θ^*, since the true state differs from θ^* with a sizeable probability. The agent with this indistinct perception θ^* 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 favor 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 over several observed signal realizations. 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 4, except for the determination of the perceptual-distance parameter d, which is now endogenously determined by the agent's ability to process information.

---

strategy for which some β_x is zero is dominated by the strategy that immediately terminates and announces the first impression.

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 Hellmann and Cover (1970) and Wilson (2014). The agent can repeatedly sample from a single statistical experiment that generates signal realizations 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 randomizes 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 the memory states specify a set of constructible choice rules, from which the agent chooses the one that maximizes her ex ante expected payoff.

The formal specification of this example follows. The agent is endowed with one Blackwell experiment $\mu(x \mid \theta)$ with a finite signal space X and, additionally, with a finite set M of the memory states m. After each run of the experiment μ, 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 $\mu(x \mid \theta)$. That is, the agent selects a (generalization of the) termination strategy: $\gamma : M \times X \longrightarrow \Delta(M \cup \{\mathfrak{t}\})$, where $\gamma(m' \mid m, x)$ is the probability that the agent in the memory state m who has observed signal realization x in the last run of the experiment μ continues with the decision-making and transitions to the memory state m', and $\gamma(\mathfrak{t} \mid m, x)$ 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 realization observed in the last run of μ. The agent starts the decision-making in the memory state m_0. A pair γ, σ induces a θ-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 realization observed. Let $p(a \mid \theta; \gamma, \sigma)$ be the probability that the agent terminates with the choice a in state θ, 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 maximizes 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 a set of the 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 β_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 $A \longrightarrow A$. 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 β 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 3. $\mathcal{R}(\mathcal{P}_{iia}) = \mathcal{P}_{iia}$.

According to the lemma, Example 1 is the 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 the 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 optimization over the terminations to the model of Wilson, we gained the partial characterization of the optimal choice rule without fully solving the problem: One can conclude that the optimal choice rule is second-thought-free without analyzing 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 realizations 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 realization, 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 the signals that the agent can remember determine the set of the stochastic choice rules that she can construct, from which she chooses the rule that maximizes her ex ante expected payoff.

We first formalize this example as follows. Let H be the set of the 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 realization. (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.[8] (iii) An agent who acquires a new signal realization transitions to a history hx, where x is the new signal realization drawn from $\mu(x \mid \theta)$. The decision-making is governed by a pair of mappings $\gamma : H \times \Theta \longrightarrow \Delta(H \cup \{\mathfrak{t}\})$ and $\sigma : H \longrightarrow A$, where $\gamma(h' \mid h, \theta)$ stands for the probability that the agent at history h in state θ continues decision-making and transitions to h', and $\gamma(\mathfrak{t} \mid h, \theta)$ is the probability of termination at history h in state θ.

---

[8]A truncation is obtained by deleting one or more last elements in h.

The mapping γ is constrained to satisfy 1. $\gamma(h' \mid h, \theta)$ is independent of θ if h' is a truncation of h, 2. $\gamma(\mathfrak{t} \mid h, \theta)$ is independent of θ, 3. $\frac{\gamma(hx \mid h, \theta)}{\gamma(hx' \mid h, \theta)} = \frac{\mu(x \mid \theta)}{\mu(x' \mid \theta)}$, 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 (γ, σ) terminates with action a in the state θ. The agent chooses γ and σ to maximize her ex ante expected payoff.

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

Lemma 4. $\mathcal{R}(\mathcal{P}_{pf}) = \mathcal{P}_{pf}$.

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 $d = \frac{p^*(1 \mid 1) p^*(0 \mid 0)}{p^*(0 \mid 1) p^*(0 \mid 1)}$, and thus, relative to the simple setting with the single exogenous binary experiment, the examples have the same solution with one additional endogenous parameter d.

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 2, 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 β and the action strategy σ as follows.

$$r_\delta(a \mid \theta; p, \beta, \sigma) = \sum_t \sum_{\mathbf{x}^t : \sigma(x_t) = a} \delta^t \rho\left(\mathbf{x}^t \mid \theta; p, \beta\right), \tag{11}$$

where $\rho\left(\mathbf{x}^t \mid \theta; p, \beta\right)$ defined in (1) is the conditional probability of the signal history $\mathbf{x}^t$. That is, $r_\delta(a \mid \theta; p, \beta, \sigma)$ is the *discounted* probability of the choice of action a in the state θ. When $\delta = 1$, then (11) coincides with our baseline definition of the choice rule.

The set of feasible discounted rules is $\mathcal{R}_\delta(\mathcal{P}) = \{r_\delta(p, \beta, \sigma) : p \in \mathcal{P}, \beta \in B, \sigma \in S\}$. The discounted repeated-cognition problem is to select a feasible rule r_δ that maximizes the expected payoff:

$$\max_{r_\delta \in \mathcal{R}_\delta(\mathcal{P})} \sum_{\theta \in \Theta, a \in A} \pi_\theta r_\delta(a \mid \theta) u(a, \theta), \tag{12}$$
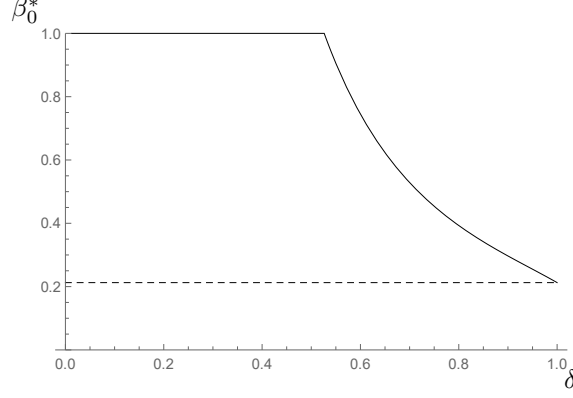
18

Figure 3: Confirmation bias with discounting. Action 1 is a priori more attractive: $\pi_1 u_1 = 10 \times \pi_0 u_0$. The primitive experiment is symmetric: $p(1 \mid 1) = p(0 \mid 0) = .9$. The agent terminates immediately when she observes signal value 1, $\beta_1^* = 1$. When $\delta > .53$, 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 β_0^*/β_1^* from the baseline model without discounting.

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

The next result generalizes 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 4. *If the termination strategy $\beta_x^* \in (0, 1)$ is interior for all x such that $\sigma^*(x) = a$, then*

$$\sum_{\theta \in \Theta} \pi_\theta u(a, \theta) r_\delta^*(a \mid \theta) = \delta \sum_{\theta \in \Theta, a' \in A} \pi_\theta u(a', \theta) r_\delta^*(a' \mid \theta) r_\delta^*(a \mid \theta). \tag{13}$$

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_δ^* 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_δ^* instead of the termination.

For illustration, we now revisit the confirmation bias application from Section 5.1 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 $u(a, \theta) = u_\theta > 0$ 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 $p(1 \mid 1) = p(0 \mid 0) = \alpha > 1/2$.

Proposition 5. *The agent chooses the action equal to the last observed signal realization. She terminates her decision-making immediately after she encounters signal realization 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 interior probability*

19

$\beta_0^* \in (0,1)$ *that decreases in* δ. *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 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 favored 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 formalizes 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 internalizing the correlations between the cognition runs.

References

Bordalo, P., N. Gennaioli, and A. Shleifer (2012). Salience theory of choice under risk. *The Quarterly Journal of Economics 127*(3), 1243–1285.

Calvert, R. L. (1985). The value of biased information: A rational choice model of political advice. *The Journal of Politics 47*(2), 530–555.

Che, Y.-K. and K. Mierendorff (2016). Optimal sequential decision with limited attention. *unpublished, Columbia University*.

Compte, O. and A. Postlewaite (2012). Belief formation. PIER Working Paper 12-027, University of Pennsylvania.

Fryer, R. G., P. Harms, and M. O. Jackson (2016). Updating beliefs when evidence is open to interpretation: Implications for bias and polarization.

Fudenberg, D., P. Strack, and T. Strzalecki (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review 108*(12), 3651–84.

Gabaix, X. and D. Laibson (2017). Myopia and discounting. Technical report, National Bureau of Economic Research.

Gentzkow, M. and J. M. Shapiro (2006). Media bias and reputation. *Journal of Political Economy 114*(2), 280–316.

Hellmann, M. and T. M. Cover (1970). Learning with finite memory. *Annals of Mathematical Statistics 41*, 765–782.

Itti, L. (2007). Visual salience. *Scholarpedia 2*(9), 3327.

Jehiel, P. (2005). Analogy-based expectation equilibrium. *Journal of Economic Theory 123*(2), 81–104.

Jehiel, P. (2018). Investment strategy and selection bias: An equilibrium perspective on overoptimism. *American Economic Review 108*(6), 1582–97.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–291.

Khaw, M. W., Z. Li, and M. Woodford (2017). Risk aversion as a perceptual bias. Technical report, National Bureau of Economic Research.

Leung, T. K. (2017). Limited cognitive ability and selective information processing. *unpublished, Toulouse School of Economics*.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. Number 8. Oxford University Press on Demand.

Meyer, M. A. (1991). Learning from coarse information: Biased contests and career profiles. *The Review of Economic Studies 58*(1), 15–41.

Netzer, N. (2009). Evolution of time preferences and attitudes toward risk. *The American Economic Review 99*(3), 937–955.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology 2*(2), 175.

Oswald, M. E. and S. Grosjean (2004). Confirmation bias. *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory 79*.

Piccione, M. and A. Rubinstein (1997). On the interpretation of decision problems with imperfect recall. *Games and Economic Behavior 20*(1), 3–24.

Rabin, M. and J. L. Schrag (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics 114*(1), 37–82.

Ratcliff, R. and G. McKoon (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation 20*(4), 873–922.

Rayo, L. and G. S. Becker (2007). Evolutionary efficiency and happiness. *Journal of Political Economy 115*(2), 302–337.

Robson, A. J. (2001). The biological basis of economic behavior. *Journal of Economic Literature 39*(1), 11–33.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

Suen, W. (2004). The self-perpetuation of biased beliefs. *The Economic Journal 114*(495), 377–396.

Swensson, R. G. (1972). The elusive tradeoff: Speed vs accuracy in visual discrimination tasks. *Perception & Psychophysics 12*(1), 16–32.

Wald, A. (1945). Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics 16*(2), 117–186.

Wilson, A. (2014). Bounded memory and biases in information processing. *Econometrica 82*(6), 2257–2294.

A Appendix

A.1 Extension of the binary setting from Section 4 to large signal spaces

The next lemma considers the setting from Section 4 with two actions and two payoff states and allow for arbitrary signal space X with $|X| > 2$. There exists a solution in which the agent ignores all but two signal realizations.

Lemma 5. *There exists a solution in which the termination probability β_x is positive for at most two signal values $x \in X$.*

Proof of Lemma 5. Assume that there exists a solution with β_x positive for $n > 2$ signals $x \in X$. We show that then 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 β be an optimal termination strategy for the given p, and let X' be the set of signals with positive β_x, and write shortly $s(x \mid \theta)$ for the effective experiment $s(x \mid \theta; p, \beta)$ induced by p and β. 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 $q_x \in \Delta(\Theta)$ be the posterior belief upon terminating with x: $q_x(\theta) = \pi_\theta s(x \mid \theta)/s(x)$.

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

$$\tilde{s}(x) = \begin{cases} s(x) + s(x^*)\mu_x \text{ if } x \in X' \setminus \{x^*\}, \\ 0 \text{ otherwise,} \end{cases} \tag{14}$$

and

$$\tilde{q}_x(\theta) = q_x(\theta) \text{ for all } x \in X' \setminus \{x^*\}, \theta \in \Theta. \tag{15}$$

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 $s(x \mid \theta; p, \beta) = \frac{\beta_x p(x|\theta)}{\sum_{x'} \beta_{x'} p(x'|\theta)}$ is induced by some p and β, then for any vector of probabilities $\tilde{\beta}_x$, the experiment

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

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

We claim that if

$$\tilde{\beta}_x = \begin{cases} c\left(1 + \frac{s(x^*)\mu_x}{s(x)}\right) & \text{if } x \in X' \setminus \{x^*\}, \\ 0 & \text{otherwise,} \end{cases}$$

where c is a constant such that $\tilde{\beta}_x \in (0,1)$ for all $x \in X$, then the resulting $\tilde{s}$ satisfies the properties (14) and (15). Let us check:

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

The property (14) holds since for all $x \in X' \setminus \{x^*\}$:

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

To establish the property (15), check that for all $x \in X' \setminus \{x^*\}$ and all $\theta \in \Theta$:

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

$\square$

A.2   Proof of the result from Subsection 5.4

The next result is an auxiliary lemma used in the proof of Proposition 3.

Lemma 6. *Suppose that termination probabilities β_x are positive for all $x \in \Theta$. Then, the optimal effective choice rule r^* satisfies for any pair of states $\theta, \theta' \in \Theta$:*

$$r^*(\theta \mid \theta) r^*(\theta' \mid \theta) = r^*(\theta \mid \theta') r^*(\theta' \mid \theta'). \tag{16}$$

Condition (16) is a strengthening of the second-thought-free condition (5). It requires that the agent who has terminated the decision process with perception θ, and knows that the second run of the process r^* terminates with a value θ' is indifferent between θ and θ'. This condition is stronger than the second-thought-free condition (5), since (5) requires (16) to hold only on average across all θ'. This strengthening holds for the special case of this application with a symmetric experiment p.

Proof of Lemma 6. The optimal effective choice rule satisfies the second-thought-free condition (5), equivalent to:

$$r^*(\theta \mid \theta) = \sum_{\theta' \in \Theta} r^*(\theta \mid \theta')r^*(\theta' \mid \theta') \text{ for all } \theta \in \Theta,$$

which after two algebraic steps gives:

$$r^*(\theta \mid \theta)\big(1 - r^*(\theta \mid \theta)\big) = \sum_{\theta' \neq \theta} r^*(\theta \mid \theta')r^*(\theta' \mid \theta') \text{ for all } \theta \in \Theta,$$

$$\sum_{\theta' \neq \theta} r^*(\theta \mid \theta)r^*(\theta' \mid \theta) = \sum_{\theta' \neq \theta} r^*(\theta' \mid \theta')r^*(\theta \mid \theta') \text{ for all } \theta \in \Theta.$$

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 θ to θ' equal to $r^*(\theta' \mid \theta)$. The balance condition for the stationary distribution $\mu(\theta)$ of this chain is

$$\sum_{\theta' \neq \theta} \mu(\theta)r^*(\theta' \mid \theta) = \sum_{\theta' \neq \theta} \mu(\theta')r^*(\theta \mid \theta') \text{ for all } \theta \in \Theta,$$

and thus $r^*(\theta \mid \theta)$ is proportional to the ergodic probability $\mu(\theta)$ of the state θ 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

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

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 $\theta_1, \theta_2, \ldots, \theta_n$,

$$\frac{m(\theta_2 \mid \theta_1)m(\theta_3 \mid \theta_2)\ldots m(\theta_n \mid \theta_{n-1})m(\theta_1 \mid \theta_n)}{m(\theta_n \mid \theta_1)m(\theta_{n-1} \mid \theta_n)\ldots m(\theta_2 \mid \theta_3)m(\theta_1 \mid \theta_2)} = 1. \tag{17}$$

The Markov chain with transition probabilities $p(\theta' \mid \theta)$ given by the primitive experiment p satisfies the Kolmogorov criterion (17) since p is symmetric by assumption. Finally, for any positive termination strategy β, the effective choice rule $r(\theta' \mid \theta; p, \beta, \sigma_I)$ satisfies the Kolmogorov criterion too. This is because $r(\theta' \mid \theta; p, \beta, \sigma_I) = \frac{\beta_{\theta'}p(\theta' \mid \theta)}{\sum_{\tilde{\theta}} \beta_{\tilde{\theta}}p(\tilde{\theta} \mid \theta)}$, and when the expressions for $r(\theta' \mid \theta; p, \beta, \sigma_I)$

are substituted into (17), then the terms $\beta_{\theta'}$ and the denominators cancel out, and hence

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

as needed. $\qquad\square$

Proof of Proposition 3. Lemma 6 implies for all pairs $\theta, \theta' \in \Theta$:

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

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

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

Using symmetry of p we get

$$\frac{\sum_{\tilde{\theta}} \beta_{\tilde{\theta}}^* p(\tilde{\theta} \mid \theta')}{\sum_{\tilde{\theta}} \beta_{\tilde{\theta}}^* p(\tilde{\theta} \mid \theta)} = \frac{p^{1/2}(\theta' \mid \theta')}{p^{1/2}(\theta \mid \theta)}, \tag{18}$$

Equation (18), used below to prove the proposition, has an interesting interpretation on its own. It states that the decision rate $f_\theta = \sum_{\tilde{\theta}} \beta_{\tilde{\theta}}^* p(\tilde{\theta} \mid \theta)$ in each state θ is proportional to $p^{1/2}(\theta \mid \theta)$ and thus high in those states that are reliably identified by the primitive experiment p.

To compare $r^*(\theta_1 \mid \theta_2)$ and $r^*(\theta_2 \mid \theta_1)$, we write

$$\frac{r^*(\theta_1 \mid \theta_2)}{r^*(\theta_2 \mid \theta_1)} = \frac{\frac{\beta_{\theta_1}^* p(\theta_1|\theta_2)}{\sum_{\tilde{\theta}} \beta_{\tilde{\theta}}^* p(\tilde{\theta}|\theta_2)}}{\frac{\beta_{\theta_2}^* p(\theta_2|\theta_1)}{\sum_{\tilde{\theta}} \beta_{\tilde{\theta}}^* p(\tilde{\theta}|\theta_1)}} = \frac{\frac{\beta_{\theta_1}^* p(\theta_1|\theta_2)}{p^{1/2}(\theta_2|\theta_2)}}{\frac{\beta_{\theta_2}^* p(\theta_2|\theta_1)}{p^{1/2}(\theta_1|\theta_1)}} = \frac{\beta_{\theta_1}^* p^{1/2}(\theta_1 \mid \theta_1)}{\beta_{\theta_2}^* p^{1/2}(\theta_2 \mid \theta_2)},$$

where we have used (18) in the second step, and symmetry of p in the last step. Define $\hat{\beta}_\theta = \beta_\theta^* p^{1/2}(\theta \mid \theta)$. We need to prove that if θ_1 is more distinct than θ_2, then $\hat{\beta}_{\theta_1} > \hat{\beta}_{\theta_2}$.

By (18), $\left(\hat{\beta}_\theta\right)_\theta$ satisfy the system of linear equations:

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

where $D_{\theta'\theta} = \frac{p(\theta'|\theta)}{p^{1/2}(\theta'|\theta')p^{1/2}(\theta|\theta)}$. We claim that if θ_1 is more distinct than θ_2, then $D_{\theta_3\theta_1} < D_{\theta_3\theta_2}$ for all $\theta_3 \neq \theta_1, \theta_2$. This follows from $p(\theta_3 \mid \theta_1) < p(\theta_3 \mid \theta_2)$ and from the symmetry of p:

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

and therefore,

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

Thus,

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

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 $D_{\theta_2\theta_1} < 1$, and thus $\hat{\beta}_{\theta_1} > \hat{\beta}_{\theta_2}$, as needed. $\qquad\square$

A.3 Proofs of the results from Section 6

Proof of Lemma 3. All rules feasible in $\mathcal{P}_{iia}$ are feasible in $\mathcal{R}(\mathcal{P}_{iia})$: $\mathcal{R}(\mathcal{P}_{iia}) \supset \mathcal{P}_{iia}$. This is immediate since when $\beta_a = 1$ for all $a \in A$, then $r(p, \beta, \sigma_I) = p$ 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 2 by the use of the generalized termination strategy $\gamma(m, x)$, and the action strategy $\sigma(m, x)$. Recall that $r(p(\gamma, \sigma), \beta, \hat{\sigma})$ 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 $\hat{\sigma} : A \longrightarrow A$ upon the termination. We need to show that there exists γ' and σ' such that $r(p(\gamma, \sigma), \beta, \hat{\sigma}) = p(\gamma', \sigma')$. This is indeed so when the termination probability $\gamma'(\mathsf{t} \mid m, x) = \gamma(\mathsf{t} \mid m, x)\beta_{\sigma(m,x)}$, the transition probability to the original memory state m_0 is $\gamma'(m_0 \mid m, x) = \gamma(m_0 \mid m, x) + \gamma(\mathsf{t} \mid m, x)\big(1 - \beta_{\sigma(m,x)}\big)$, which is the sum of the probabilities that the original process γ transits to m_0 and that the decision process $r(p(\gamma, \sigma), \beta, \hat{\sigma})$ restarts after termination of $p(\gamma, \sigma)$. Additionally, for all $\tilde{m} \neq m_0$, $\gamma'(\tilde{m} \mid m, x) = \gamma(\tilde{m} \mid m, x)$. The above choice of γ' implies that the process $p(\gamma', \sigma')$ replicates the Markov process over the memory states under $r(p(\gamma, \sigma), \beta, \hat{\sigma})$. Finally, to replicate the choices upon terminations, we set the action strategy $\sigma'(m, x) = \hat{\sigma}(\sigma(m, x))$ for all (m, x). $\qquad\square$

Proof of Lemma 4. Again, trivially, $\mathcal{R}(\mathcal{P}_{pf}) \supset \mathcal{P}_{pf}$, since $r(p, (1, \ldots, 1), \sigma_I) = p$ for all $p \in \mathcal{P}_{pf}$. Additionally, $\mathcal{R}(\mathcal{P}_{pf}) \subset \mathcal{P}_{pf}$. This is indeed so because for any $\beta = (\beta_a)_{a \in A}$ and any $\hat{\sigma} : A \longrightarrow A$, $r(p(\gamma, \sigma), \beta, \hat{\sigma}) = p(\gamma', \sigma')$ where the termination probability $\gamma'(\mathsf{t} \mid h, \theta) = \gamma(\mathsf{t} \mid h, \theta)\beta_{\sigma(h)}$, the transition probability to the empty signal history $\emptyset$ is set to $\gamma'(\emptyset \mid h, \theta) = \gamma(\emptyset \mid h, \theta) + \gamma(\mathsf{t} \mid h, \theta)\big(1 - \beta_{\sigma(h)}\big)$, and for all $\tilde{h} \neq \emptyset$, $\gamma'(\tilde{h} \mid h, \theta) = \gamma(\tilde{h} \mid h, \theta)$. Finally, the action strategy is set to $\sigma'(h) = \hat{\sigma}(\sigma(h))$ for all histories h. $\qquad\square$

Proof of Proposition 4. We extend the definition of the effective experiment to the setting with

discounting. Let

$$s_\delta(x \mid \theta; p, \beta) = \sum_t \sum_{\mathbf{x}^t : x_t = x} \delta^t \rho\left(\mathbf{x}^t \mid \theta; p, \beta\right),$$

where $\rho\left(\mathbf{x}^t \mid \theta; p, \beta\right)$ is the probability of the signal history $\mathbf{x}^t$ 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:

$$s_\delta(x \mid \theta; p, \beta) = \beta_x p(x \mid \theta) + \delta \sum_{x' \in X} \left(1 - \beta_{x'} p\left(x' \mid \theta\right)\right) s_\delta(x \mid \theta; p, \beta), \tag{19}$$

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 (19) for s_δ 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

$$\max_{p \in \mathcal{P}, \beta \in B, \sigma \in S} \sum_{\theta \in \Theta, x \in X} \pi_\theta \frac{\beta_x p(x \mid \theta)}{1 - \delta + \delta \sum_{x' \in X} \beta_{x'} p(x' \mid \theta)} u(\sigma(x), \theta). \tag{20}$$

Consider x with an interior termination probability $\beta_x^* \in (0, 1)$ and let $a = \sigma^*(x)$. The first-order condition of the problem (20) with respect to β_x is:

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

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

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

Proof of Proposition 5. Any (β, σ) 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 β_0^* and β_1^* are positive, and the action strategy is $\sigma^*(x) = x$ or $\sigma^*(x) = 1 - x$. Let us show that the action strategy σ^* must be the identity function σ_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:

$$\pi_0 u_0 r_\delta(0 \mid 0; p, \beta^*, \sigma^*) + \pi_1 u_1 r_\delta(1 \mid 1; p, \beta^*, \sigma^*) - \pi_1 u_1 =$$

$$\pi_0 u_0 r_\delta(1 \mid 0; p, \beta^*, \sigma_I) + \pi_1 u_1 r_\delta(0 \mid 1; p, \beta^*, \sigma_I) - \pi_1 u_1 \geq$$

$$\pi_0 u_0 r_\delta(1 \mid 0; p, \beta^*, \sigma_I) - \pi_1 u_1 r_\delta(1 \mid 1; p, \beta^*, \sigma_I) > 0,$$

28

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

$$\pi_0 u_0 r_\delta(0 \mid 0; p, \beta^*, \sigma^*) + \pi_1 u_1 r_\delta(1 \mid 1; p, \beta^*, \sigma^*) - \pi_0 u_0 =$$
$$\pi_0 u_0 r_\delta(1 \mid 0; p, \beta^*, \sigma_I) + \pi_1 u_1 r_\delta(0 \mid 1; p, \beta^*, \sigma_I) - \pi_0 u_0 \geq$$
$$\pi_1 u_1 r_\delta(0 \mid 1; p, \beta^*, \sigma_I) - \pi_0 u_0 r_\delta(0 \mid 0; p, \beta^*, \sigma_I) > 0.$$

The last two inequalities imply:

$$\frac{r_\delta(0 \mid 1; p, \beta^*, \sigma_I)}{r_\delta(0 \mid 0; p, \beta^*, \sigma_I)} > \frac{\pi_0 u_0}{\pi_1 u_1} > \frac{r_\delta(1 \mid 1; p, \beta^*, \sigma_I)}{r_\delta(1 \mid 0; p, \beta^*, \sigma_I)}.$$

This establishes contradiction because as shown in the proof of Proposition 4, $r_\delta(x \mid \theta; p, \beta^*, \sigma_I) = s_\delta(x \mid \theta; p, \beta^*) = \frac{\beta_x^* p(x|\theta)}{1 - \delta + \delta \sum_{x'} \beta_{x'}^* p(x'|\theta)}$, 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 β_x^* by a same factor. This preserves the conditional action distribution in each state θ 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 σ_I and $(\beta_0, \beta_1) = (\beta, 1)$ is

$$\pi_0 u_0 \frac{\beta \alpha}{1 - \delta + \delta(\beta \alpha + 1 - \alpha)} + \pi_1 u_1 \frac{\alpha}{1 - \delta + \delta(\alpha + \beta(1 - \alpha))}, \tag{21}$$

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

$$\pi_0 u_0 \frac{\alpha}{1 - \delta + \delta(\alpha + \beta(1 - \alpha))} + \pi_1 u_1 \frac{\beta \alpha}{1 - \delta + \delta(\beta \alpha + 1 - \alpha)}, \tag{22}$$

The assumptions that $\pi_1 u_1 > \pi_0 u_0$ and that $\alpha > 1/2$ imply that, for any $\beta \in (0, 1)$, (21) exceeds (22), 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 $r_\delta(x \mid \theta; p, \beta, \sigma_I) = \frac{\beta_x p(x|\theta)}{1 - \delta + \delta \sum_{x'} \beta_{x'} p(x'|\theta)}$, this condition simplifies into a quadratic equation for β_0^*. When $\delta < \frac{1}{\alpha + (1-\alpha)R}$, then this condition does not have an interior solution and the derivative of the value (21) with respect to β_0 at $\beta_0 = 1$ is positive. Thus, in this case, the unique β_0^* satisfying the first-order condition is $\beta_0^* = 1$.

When $\delta > \frac{1}{\alpha+(1-\alpha)R}$, then the condition has an interior solution and the derivative of the value (21) with respect to β_0 at $\beta_0 = 1$ is negative. Thus, for this range of parameters, the unique β_0^* satisfying the first-order condition is the interior value that solves the quadratic equation, solution of which decreases in δ. $\qquad\square$