WILL YOU NEVER LEARN? SELF DECEPTION AND BIASES IN INFORMATION PROCESSING[1]

DANIEL GOTTLIEB

Olin Business School, Washington University in St. Louis

dgottlieb@wustl.edu

Abstract

This paper presents a repeated model of selective awareness and studies implications for information processing. A person receives a sequence of signals before making a decision and interprets each signal selectively. In any equilibrium, the person disregards all information after a certain number of observations. As a consequence, learning is always incomplete. Additionally, the equilibrium behavior displays patterns consistent with observed biases in information processing. The person displays a tendency to interpret information in ways that support original beliefs and attaches a disproportionately large weight to initial observations (confirmation bias). She also updates beliefs in the right direction, but in insufficient amount compared to the update derived by Bayes' rule (conservatism bias). The model's implications for learning about one's self control and for the design of incentives to collect information are discussed.

[1] I thank Dilip Abreu, Nageeb Ali, Roland Benabou, Aislinn Bohren, Lucas Maestri, George Mailath, Stephen Morris, Andy Postlewaite, José Scheinkman, John Smith, Jean Tirole, and Muhamet Yildiz for valuable suggestions, and seminar participants at Harvard/MIT, London School of Economics, Princeton, University of Pennsylvania, Washington University at St. Louis, the Wharton School, the Behavioral Economics Annual Meeting (Cornell), and the Murray S. Johnson Memorial Conference (UT Austin) for comments. Kenneth Zhang provided outstanding research assistance.

Contents

1 Introduction

Economists typically model humans as statisticians who collect information in an unbiased manner and make impartial inferences. The psychological evidence, however, suggests that we behave more like unscrupulous statisticians, who collect and interpret information guided less by a concern with accuracy than by a desire to feel competent. Sedikides, Green, and Pinter (2004, pp. 165), for example, describe people as "striving for a positive self-definition or the avoidance of a negative self-definition [...] at the expense of accuracy and truthfulness." Practitioners also claim that biases in information processing distort learning. Montier (2007), for example, argues that "the major reason we don't learn from our mistakes [...] is that we simply don't recognize them as such. We have a gamut of mental devices all set up to protect us from the terrible truth that we regularly make mistakes."

This paper studies how biases in information processing arise and persist when people have selective awareness. I consider an infinitely lived person who receives a sequence of binary signals. Each signal provides information about the state of the world. A state of the world may correspond to the person's skills or some other feature that affects payoffs through anticipatory utility. The acquisition of information ends in each period with some positive probability. Then, the person has to take an action from a set that may be either continuous or discrete. The payoff from each action depends on the state of the world.

In choosing how to interpret each signal, the person faces a trade-off between optimism and improved decision making. I show that, both with discrete and with continuous actions, the gain from optimism decreases to zero at a slower rate than the cost of distorting actions. When the action space is discrete, the person becomes increasingly convinced about which action to take as more signals are interpreted correctly. Then, the probability that each additional signal will affect her choice becomes small and so does the cost of misinterpreting the signal. When the action space is continuous, each individual signal may affect the person's decision; however, small distortions close to the optimum have second-order costs. Hence, the person always rationalizes away negative information after a sufficiently large number of observations. Individuals thus never learn the true state.

In the past decade, many researchers in economics and finance have begun to model agents with biased beliefs, such as optimism or overconfidence. Two questions inevitably arise from these models: "In which domains should we expect these biases to be more prevalent?" and "Should we expect people to learn from their mistakes over time, leading to a progressive elimination of bias?". Models that exogenously assume a bias are unable to address these questions adequately.

This paper draws on a more recent class of models that consider agents for whom biased beliefs arise endogenously, in response to some "need" (e.g. desire for self-esteem, anticipatory utility, motivation). Such needs are balanced – consciously or, more likely, unconsciously – against the costs of making worse decisions through explicit processes of information acquisition, interpretation, or recall. Examples include Akerlof and Dickens's (1982) seminal article on cognitive dissonance, Bodner and Prelec's (1995) model of self-signaling, and Brunnermeier and Parker's (2005) model of optimal expectations. The model considered in this paper has been used to provide theories of personal motivation (Benabou and Tirole, 2002); redistributive policies (Benabou and Tirole, 2006a); political attitudes towards reforms (Levy, 2012); groupthink and ideologies (Benabou, 2008, 2011); endowment, sunk cost effects, and some other deviations from expected utility theory (Gottlieb, 2014); preferences for increasing payments (Smith, 2009a, 2009b); and fear of death as an explanation for puzzles in health and savings behavior (Kopczuk and Slemrod, 2005).[2]

Although most papers in this literature study decisions that are faced repeatedly over time, they all assume that there is a single opportunity to manipulate beliefs. This simplifying assumption precludes the analysis of how individuals would update beliefs if they had not just one, but many signals to process. Furthermore, it is often argued that the Bayesian updating assumption embedded in this framework would lead individuals to eventually learn the truth; therefore, departures from rationality would vanish in the long-run. More specifically, since there is usually a probability bounded away from zero that signals are interpreted correctly in the static equilibrium of these models, standard asymptotic results would imply that beliefs converge to the truth.[3]

This paper challenges that argument. I demonstrate that, in any equilibrium of the repeated model, all information is disregarded after a certain number of observations (i.e., repeatedly playing the static equilibrium is not an equilibrium of the repeated model). Learning is thus always incomplete, and departures from rationality persist even when the decision problem is repeated infinitely many times.[4]

[2] Bernheim and Thomadsen (2005) use a similar model to show why individuals may cooperate in a prisoner's dilemma game. Other papers featuring belief manipulation include Schelling (1985); Kuran (1993); Rabin (1994); Carrillo and Mariotti (2000); Bodner and Prelec (2002); and Karlsson, Loewenstein, and Seppi (2009).

[3] For example, in the models of Benabou and Tirole (2006) and Benabou (2008), there are two stable pure strategy equilibria – one in which information is interpreted realistically and one in which it is ignored – and one unstable mixed-strategy equilibrium. If the individual repeatedly plays either the realistic or the mixed-strategy recollection, she eventually learns the truth and the effects captured in their papers (differences in political ideology, redistribution, labor supply, aggregate income, and popular perceptions of the poor) vanish. Likewise, the predictions from the optimal expectations model would converge to the ones from rational expectations models if individuals observe a large number of signals after initial beliefs are chosen; by Doob's Consistency Theorem, under mild conditions, beliefs asymptotically converge to the truth regardless of the prior distribution chosen. See the Online Appendix for a formalization of this argument and how it relates to each specific model.

[4] This paper is related to the literature on experimentation, which started with Rothschild's (1974) analysis of two-armed bandits. In models in that literature, there is a fixed cost of experimentation. Since the benefits from learning tend to zero with the number of observations, learning eventually stops. Here, there is no exogenous cost of learning (the ex-ante optimal strategy involves complete learning). Instead, there is an endogenous cost of learning that arises from the strategic desire to improve one's self-image. While this endogenous cost also tends to zero as the number of informative periods rises, it does so more slowly than the benefit. Moreover, because the cost of learning is endogenous in my model, the comparative statics are also quite different from what one would get in a standard bandit model. The paper is also related to Ali (2011), which studies conditions for Bayesian individuals to learn their degree of self control (see Section 4). There is also a parallel between

The model predicts a tendency to interpret information in ways that support original beliefs. Individuals with a favorable initial streak of information are beguiled by optimistic beliefs and subsequently disregard negative information; those with a negative initial streak of information are stuck in pessimistic traps, and subsequently disregard positive information. This result is consistent with evidence on the confirmation bias. In the words of Oswald and Grosjean (2004, pp. 79):

> "Confirmation bias" means that information is searched for, interpreted, and remembered in such a way that systematically impedes the possibility that the hypothesis could be rejected – that is, it fosters the immunity of the hypothesis.

Psychologists have extensively studied the confirmation bias.[5] The evidence from this literature suggests that, as people become more convinced of their hypotheses, they disregard conflicting information.

A related finding from psychology is the disproportionate effect of first impressions. Individuals excessively weight initial observations in sequences of exchangeable information. The Markovian equilibria of the model have this property. Individuals interpret initial signals realistically and update their beliefs according to Bayes' rule. However, after a certain number of periods, they discard additional information and, therefore, do not update beliefs. The model also predicts an updating bias known as conservatism. Conservatism posits that individuals update beliefs in the right direction, but by too little relative to Bayesian updating (Edwards, 1968).[6]

The model generates new comparative statics results. It predicts that the ability to learn decreases in the self-image component of the information and increases in the value of information. Both predictions are consistent with a literature in psychology that documents how preferences affect beliefs.[7] For example, biased beliefs seem to be more prevalent for traits and behaviors that individuals regard as important.[8] Bahrick, Hall, and Berger (1996) study distortions in college students' memory of their high school grades. They find that accuracy of recollections declines monotonically in the students' letter grades (from 89% for grades of A to 29% for grades

the learning failure in the model and the one from the social-learning literature (Banerjee, 1992; Bikhchandani, Hirshleifer, and Welch 1992). In models of social learning, equilibria are inefficient because individuals fail to account for the information externality from their actions; here, a self that rationalizes a bad signal away also causes an externality on other selves by making them more suspicious of the reliability of a good signal.

[5]Indeed, Evans (1989, pp. 41) argues that "[c]onfirmation bias is perhaps the best known and most widely accepted notion of inferential error to come out of the literature on human reasoning." See Rabin and Schrag (1999) for a survey of the confirmation bias literature.

[6]Barberis, Shleifer, and Vishny (1998) argue that the conservatism bias may explain the underreaction of stock prices to news.

[7]See Kunda (1990) and Helzer and Dunning (2012) for a summaries of the literature, including its relationship with the confirmation bias. The model also predicts that the ability to learn is decreasing in how much information one expects to receive when the value of information is uniform. I am unaware of any experiments that test this prediction.

[8]See, e.g., MacDonald and Ross (1999), or Sanbonmatsu et al. (1987).

of D). As the authors interpret this finding, "[d]istortions are attributed to reconstructions in a positive, emotionally gratifying direction." Likewise, Greene (1981) argues that a desire to see oneself as an accurate decision-maker in fact biases one's beliefs towards confirming previous decisions.

Relatedly, many studies show that information disclosed before an individual makes a decision is processed differently from information disclosed after the decision has been made. Since information is only valuable when it can affect decisions, the model predicts that belief distortions should be less prevalent when one has to decide which decision to make. Accordingly, psychologists find that individuals who are actively considering which decision to make follow a "deliberative mindset," which is marked by a relatively open-minded processing of information. On the other hand, those who cannot affect the decision follow an "implemental mindset," which is characterized by overly positive self-perceptions (Taylor and Gollwitzer, 1995).[9]

The predictions of the model are also supported by recent studies in experimental economics that compare how individuals process information that bears on their self-esteem with how they process more neutral information. Eil and Rao (2011), for example, document differences in how subjects process information about their own intelligence (as measured by IQ) and beauty, versus how they process neutral information about randomly generated numbers. In the realm of neutral information, they find no significant deviation from Bayes' rule. In the realm of intelligence and beauty, however, they find that good news is processed quite differently from bad news, producing a pattern of confirmation bias. In a similar experiment, Möbius, Niederle, Niehaus, and Rosenblat (2011) test how subjects update beliefs when given information about their scores in an IQ test relative to neutral information. Consistently with the predictions from my model, they find significant evidence of confirmation bias for information about IQ, but not for more neutral information. Moreover, they find much stronger evidence for conservatism when updating information about IQ than for neutral information.[10] Mijovic-Prelec and Prelec (2010) find evidence of self-serving belief manipulations in a categorization test; although subjects were incentivized for accurate predictions of their performance, they change their answers based on their (non-contingent) financial stakes in the outcome.

This paper is related to a theoretical literature that models biases in information processing,

[9]See Gollwitzer and Bayer (1999) for a survey of this literature.

[10]Kuhnen (2012) studies how individuals learn from financial news when they actively invest in certain assets compared to when they passively observe information about those assets. She finds that individuals learn passive information more accurately. Moreover, consistent with the anticipatory utility formulation of my model, she finds that asset payoffs affect how much individuals learn: individuals with payoffs in the gains domain learn significantly better than those with payoffs in the losses domain. Hales (2007) also studies how individuals randomly assigned to different financial positions interpret news. He finds that subjects are more willing to agree with information suggesting that they will make money on their investment and to disagree with information suggesting that they will lose money. Choi and Lou (2010) find evidence of self-serving, asymmetric updating by mutual fund managers, whereas Goetzman and Peles (1997) show that individual investors display a similar asymmetric pattern in their recollections of past portfolio performance. In turn, Wiswall and Zafar (2011) find that college students asymmetrically update beliefs about future earnings in self-serving way.

especially the seminal work of Rabin and Schrag (1999).[11] Rabin and Schrag present a model of confirmatory bias in which individuals misinterpret signals that conflict with their current beliefs with a fixed probability. They show that this updating rule generates overconfidence and that individuals may eventually believe in an incorrect hypothesis with near certainty. My model might be seen as endogeneizing their misrepresentation mechanism. In doing so, it generates predictions about the environments in which we should expect biased beliefs to be more prevalent, and the directions that the bias will take. I do *not* claim that motivated reasoning accounts for all of the evidence identified with the confirmation bias or other violations of Bayes' rule; a large literature finds departures from Bayesian updating in ego-independent environments (c.f., Fischhoff and Beyth-Marom, 1983; Rabin, 1998; or Benjamin, Rabin, and Raymond, 2013). Nevertheless, the evidence contrasting the confirmation bias in ego-related with ego-independent environments suggests that motivated reasoning is an important part of it.

The structure of the paper is as follows. Section 2 presents the general model. In Subsection 2.1, I describe the main results when there is a continuum of actions; Subsection 2.2 considers a discrete action space. Section 3 illustrates the equilibria of the model in the special case of quadratic payoffs and uniform distributions. Section 4 discusses new economic applications, and Section 5 concludes. Several extensions, including the analysis of anticipatory utility, different timings, and multiple non-stationary signals are presented in the appendix.

2 The Model

There is a single individual who must eventually take an action. The optimal action depends on her skills θ, which lie in a compact non-empty interval of the real line $\Theta = [\underline{\theta}, \bar{\theta}]$. The individual does not know her own skills with certainty; her prior beliefs about θ are represented by a thrice

[11] Most models in this literature focus on identifying certain biases and exploring how these biases affect the conclusions of standard economic models. See, e.g., Gennaioli and Shleifer (2010), Madarász (2012), Schwartzstein (2014), and Benjamin, Rabin, and Raymond (2013). This paper is also related to Wilson (2014), which presents a model featuring the same information structure as the one in Section 2 and also examines biases in information processing. Our work is complementary since we focus on different memory limitations; Wilson considers an unbiased but limited memory, consisting of a finite number of states. This restriction precludes the individual from conditioning the action on the whole sequence of signals (or any sufficient statistic), which makes it impossible for the true state to eventually be learned. By contrast, in the model in this paper, the individual can, in principle, condition actions on the whole history of signals. The choice of not doing so arises endogenously through either the desire to improve one's self-image or to enjoy anticipatory utility. Although it is hard to dispute that human memory is limited, unbiased decision makers may sometimes be able to keep a record of their observations (say, by writing them or some sufficient statistic down) or search for evidence if needed. By contrast, the individuals considered in this paper would write down inaccurate observations, interpret them incorrectly, delete their records, or choose not to look at them. Alternatively, Brocas and Carrillo (2012) propose a neuroeconomic model that also generates biases in information processing. Köszegi (2006) studies an individual who derives utility from believing to be able to perform a certain task and shows that the individual may become overconfident. In contrast, the utility of decision makers here are linear in probabilities, and so they are indifferent between irrelevant signals. Preferences for information are determined by the decision makers' memory imperfection.

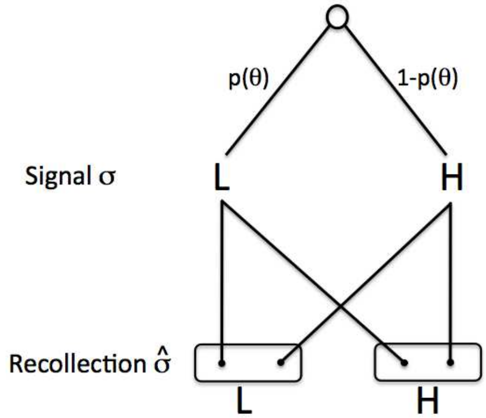

Figure 1: Informational Structure

continuously differentiable density ρ with full support on Θ.

There is an information process that ends with a constant probability $\eta \in (0,1)$ in each period. When it ends, the individual must select an action a from a compact non-empty subset of the real line A.[12] While active, the process generates a new signal in each period. Signals can be either high or low: $\sigma_t \in \{H, L\}$. Conditionally on skills, signals are independent and identically distributed. Let $p(\theta)$ denote the conditional probability of a high signal, where $p : \Theta \to (0,1)$ is a strictly increasing and twice continuously differentiable function. Because p is increasing, a high signal is good news about the individual's skills in the sense of first-order stochastic dominance.[13] These assumptions ensure that Bayesian posteriors converge to the true parameter as the number of signals increase. Thus, a Bayesian individual would eventually learn her true skill.

After observing each signal, the individual decides how to encode it. She may choose to interpret it realistically $\hat{\sigma}_t = \sigma_t$ or to rationalize it as a different signal $\hat{\sigma}_t \neq \sigma_t$ (see Figure 1). I refer to $\hat{\sigma}_t$ as the *recollection* of signal σ_t. Because she remembers her interpretation of each signal but not the signal itself, a period-t history is a vector of recollections up to period t: $h^t = \{\hat{\sigma}_1, \hat{\sigma}_2, ..., \hat{\sigma}_{t-1}\}$. The initial history is the null set: $h^1 = \varnothing$. Let $\mathcal{H}^t := \{L, H\}^{t-1}$ denote the set of period-t histories, and let $\mathcal{H} := \bigcup_{t=1}^{\infty} \mathcal{H}^t$ denote the set of all possible histories. For any $\tau \leq t$, the vector $h^\tau = \{\hat{\sigma}_1, \hat{\sigma}_2, ..., \hat{\sigma}_{\tau-1}\}$ is called a *subhistory* of $h^t = \{\hat{\sigma}_1, \hat{\sigma}_2, ..., \hat{\sigma}_{t-1}\}$. We write $h^\tau \subset h^t$ if h^τ is a subhistory of h^t.

[12]As Wilson (2014) argues, this setup captures an environment in which someone expects to obtain a sequence of information but is unsure about when a decision will have to be made. In Appendix IV, I assume that the person observes a fixed number of signals and takes an action after the last observation. That setup represents an environment in which one is certain about when the decision will have to be made and how much information he or she will acquire before then. The results remain virtually unchanged in both setups (as well as in mixtures between them).

[13]From a statistical perspective, the assumption that p is strictly increasing ensures that θ is identified. If p were not a one-to-one function, it would be impossible for an individual to learn the true parameter θ regardless of the number of observations.

The individual has strictly increasing preferences over her perceived skills θ ("self-image").[14] Without loss of generality, skills can be normalized to be measured in payoff units up to a parameter $\alpha > 0$ that captures the importance of self-image. In each period before the process ends, the individual's payoff equals her expected skills conditional on her recollections: $\alpha E\left[\theta|h^t\right]$. When the information process ends, an individual with skill θ_0 who selects action a gets a payoff from actions $V(a, \theta_0)$, where $V : A \times \Theta \to \mathbb{R}$ is a continuous function. There is no discounting.

A strategy is a function $(\hat{\sigma}_L, \hat{\sigma}_H, a) : \mathcal{H} \to \{L, H\}^2 \times A$ that specifies how each signal is interpreted if the information process does not end and which action is taken if the process ends. That is, $\hat{\sigma}_L(h^t) \in \{L, H\}$ and $\hat{\sigma}_H(h^t) \in \{L, H\}$ specify how the individual interprets low and high signals after history h^t. The functions $\hat{\sigma}_L$ and $\hat{\sigma}_H$ are called *interpretation strategies* and a is called an *action strategy*. Let $\mu(.|h^t)$ denote the individual's posterior beliefs about θ given h^t and let $E_\mu[.|h^t]$ denote the expectation operator with respect to $\mu(.|h^t)$.

Any triple $(\hat{\sigma}_L, \hat{\sigma}_H, \mu)$ induces a probability measure over the space of histories $\mathcal{H}$, which specifies the probability of reaching each history if the information process does not end. For example, if $\hat{\sigma}_L(h^1) = \hat{\sigma}_H(h^1) = H$, then the probability of reaching $h^2 = \{L\}$ is zero, and the probability of reaching $h^2 = \{H\}$ is one. Let $E_{(\hat{\sigma}_L, \hat{\sigma}_H, \mu)}[\cdot|h^t]$ denote the expectation with respect to this measure. Similarly, let $\Pr(h^t|\hat{\sigma}_L, \hat{\sigma}_H)$ denote the probability of reaching history h^t conditional on the interpretation strategy $(\hat{\sigma}_L, \hat{\sigma}_H)$ and prior distribution ρ.

The expected payoff from strategy $(\hat{\sigma}_L, \hat{\sigma}_H, a)$ under belief μ is

$$\sum_{t=0}^{\infty} (1-\eta)^t E_{(\hat{\sigma}_L, \hat{\sigma}_H, \mu)} \left\{ \alpha (1-\eta) \left[E_\mu \left[\theta|h^t\right] + \eta V\left(a\left(h^t\right), \theta\right)\right]\right\}. \tag{1}$$

This payoff function is based on two assumptions that simplify notation but are not important for my results. First, it assumes that the probability of termination η is constant. It is easy, although notationally cumbersome, to generalize my results to time-varying termination probabilities. Similarly, any impatience about the future can be incorporated into the probability of termination. Second, this specification assumes that the game ends after the individual takes the action. The analysis would remain unchanged if we assumed that the individual keeps receiving a payoff equal to her expected skills after the action is taken.[15]

[14] This assumption was previously adopted by Benabou and Tirole (2006b, 2009) and Gottlieb (2014). There are several reasons why people may value believing that they have higher skills. First, they may simply like to think that they have these attributes (see, e.g., Schelling, 1985). Second, people may benefit from having overconfident beliefs in situations in which emotions affect performance (Compte and Postlewaite, 2004). Third, there may be a motivational value of belief manipulation: individuals with time-inconsistent preferences value overconfident beliefs when effort and skills are complementary (Benabou and Tirole, 2002). Preferences over skills can be derived as a reduced form of the models of Benabou and Tirole (2002) and Compte and Postlewaite (2004). Moreover, although it is natural to interpret θ as representing the individual's skills, it can be any payoff-relevant characteristic that is positively correlated with the payoff from the action. The assumption that individuals care directly about self-image can be replaced by anticipatory utility preferences (see Appendix II).

[15] The assumption of separability between the payoff from self views and the payoff from actions is not important for my results. Separability states that payoff from actions are "objective" in the sense that individuals cannot

The individual faces a decision problem with imperfect recall. I follow Piccione and Rubinstein (1997) in modeling such a decision problem as a game between different "selves." In each period before information acquisition ends, a new self receives a signal and chooses how to interpret it. When the process ends, a new self takes an action. The decision is modeled as the Perfect Bayesian Equilibrium (*PBE*) of this multiself game:

Definition 1 *A PBE of the game is a strategy profile* $(\hat{\sigma}_L^*, \hat{\sigma}_H^*, a^*)$ *and posterior beliefs* μ *such that, for all* $h^t \in \mathcal{H}$,

1. *Each interpretation maximizes the expected continuation payoff: For all* $\sigma \in \{L, H\}$,

$$\hat{\sigma}_\sigma^*\left(h^t\right) \in \arg\max_{s\in\{L,H\}} \sum_{\tau=t}^{\infty} (1-\eta)^{\tau-t} E_{\left(\hat{\sigma}_L^*, \hat{\sigma}_H^*, \mu\right)} \left\{ \begin{array}{c} \alpha(1-\eta)\left[E_\mu\left[\theta \middle| \{h^t, s, \hat{\sigma}_{t+1}, ...\hat{\sigma}_\tau\}\right]\right] \\ +\eta V\left(a\left(\{h^t, s, \hat{\sigma}_{t+1}, ...\hat{\sigma}_\tau\}\right), \theta\right) \end{array} \middle| \left(h^t, \sigma\right) \right\};$$

2. *Each action maximizes the expected payoff when the process ends:*

$$a^*\left(h^t\right) \in \arg\max_{a\in A} \left\{E_\mu\left[V(a, \theta) \left|h^t\right.\right]\right\};$$

3. *Posterior beliefs* $\mu(.|h^t)$ *are obtained by Bayes' rule if* $\Pr(h^t|\hat{\sigma}_L^*, \hat{\sigma}_H^*) > 0$.

Conditions 1 and 2 are perfection conditions: 1 requires each self to choose the interpretation strategy that maximizes her expected stream of payoffs given the interpretation and action strategies of other selves; 2 requires actions to maximize the expected payoff in each terminal history. Condition 3 requires beliefs to satisfy Bayes' rule given the equilibrium strategies. We say that beliefs μ are consistent with interpretation strategies $(\hat{\sigma}_L, \hat{\sigma}_H)$ if they satisfy Condition 3, and we say that an action strategy a is consistent with beliefs μ if Condition 2 holds. Given a PBE, we say that history h^t is on the equilibrium path if $\Pr(h^t|\hat{\sigma}_L^*, \hat{\sigma}_H^*) > 0$.

Because in the main text I will only consider pure strategy equilibria, there is no loss of generality in assuming that high signals are always interpreted realistically (up to a relabeling of interpretations).[16] Thus, we can assume that all high signals are interpreted truthfully $\hat{\sigma}_H(h^t) = H$, whereas low signals are either interpreted realistically $\hat{\sigma}_L(h^t) = L$ or rationalized as a high signal $\hat{\sigma}_L(h^t) = H$ (see Figure 2). Because high signals are always assigned to a high recollection, we can omit the subscript L and take the interpretation strategy to be a mapping $\hat{\sigma} : \mathcal{H} \rightarrow \{L, H\}$ that specifies how to interpret a low signal. This reformulation of the model leads to the information processing framework from Benabou and Tirole (2002, 2004, 2006a, 2006b) and Benabou (2008, 2011). It is easier to work with this reformulation because interpretation strategies become a single function, which reduces the set of possible deviations.

manipulate it only by changing their beliefs.

[16]See Appendix I for a formal proof of this claim as well as the formal definition of the PBE of this modified game. Appendix VI considers equilibria in mixed and behavioral strategies.

In equilibrium, the individual knows her interpretation strategy and, by consistency of beliefs, the law of iterated expectations holds on all histories on the equilibrium path. Thus, in equilibrium, misinterpreting information does not affect the expected payoff from skills: $E[\theta] = E[E[\theta|h^t]]$ for all h^t on the equilibrium path. Rationalizing a low signal away raises the payoff conditional on that signal. However, it lowers the payoff conditional on a high signal by making the individual less convinced about its reliability. In expectation, these two effects exactly cancel out. Since more information leads to (weakly) better decision making, the individual would be better off if she could commit not to misinterpret any information. As we will see, interpreting all information realistically is not an equilibrium of the game. Thus, no equilibrium maximizes the individual's ex-ante expected payoff.[17]

Since Condition 3 does not restrict beliefs on histories off the equilibrium path, any belief is allowed in histories that are never reached. Definition 1 is therefore a very weak definition of PBE; it does not even require subgame perfection in games of complete information. Game theorists often impose additional restrictions on posterior beliefs to reduce the multiplicity of equilibria.[18] I will not do so for two reasons. First, since my results hold for all PBE, it suffices to work with this less restrictive definition. Second, the equivalence between the equilibria of the games with the information structures from Figures 1 and 2 typically fails if one imposes additional restrictions on beliefs. Since I want to obtain results that apply to both formulations, I cannot restrict beliefs off the equilibrium path.[19]

Suppose the individual rationalizes every low signal away, assigns the same posterior as her prior distribution to all histories, and chooses a preferred action given her initial beliefs. Because posteriors are not affected by her recollections, both the interpretation strategy and the action strategy are optimal given her beliefs. Moreover, since recollections are uninformative, posterior beliefs are consistent with Bayes' rule. Hence, there always exists a PBE in which the individual rationalizes every low signal away:

Proposition 1 *There exists a PBE in which* $\hat{\sigma}^*(h^t) = H$ *for all* $h^t \in \mathcal{H}$.

[17]This result relies on the linearity of the expected utility function in probabilities. If payoffs were nonlinear functions of probabilities – as, for example, in the model of confidence-enhanced performance of Compte and Postlewaite (2004) or the time-inconsistency model of Benabou and Tirole (2002) – the law of iterated expectations would no longer apply and rationalization could be ex-ante beneficial. The individual's time inconsistency is generated by what Piccione and Rubinstein (1997) refer to as the absentmindedness property.

[18]See, e.g., Fudenberg and Tirole, 1991 pp. 331-333.

[19]For example, in the structure from Figure 2, a low recollection is only possible if a low signal is observed. Therefore, any reasonable selection criterion should require the individual to assign probability one to $\sigma_t = L$ if she recalls $\hat{\sigma}_t = L$. On the other hand, in the information structure from Figure 1, both recollections can be reached after each signal. Since, holding beliefs fixed, the cost choosing high and low recollections is the same for both signals, standard selection criteria (such as the Intuitive or Divinity Criteria) do not rule out any equilibrium. One reasonable refinement in this framework assumes that beliefs remain unchanged in off-path histories. However, this condition does not restrict much the set of PBE. The main new property gained with this refinement is that the comparative statics on the set of equilibria (Proposition 2) hold for all histories (and not only those on the equilibrium path).

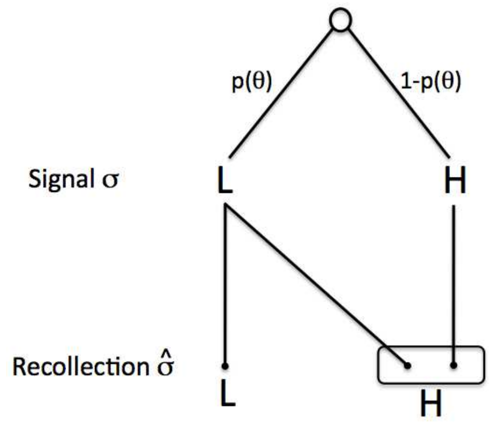

Figure 2: Revised Informational Structure (Benabou-Tirole)

For given posterior beliefs μ and a history h^t, we say that the recollection $\hat{\sigma}_t$ is *informative* if it modifies the posterior:

$$\mu\left(\theta|h^t, \hat{\sigma}_t\right) \neq \mu\left(\theta|h^t\right).$$

Notice that informativeness is defined relative to posterior beliefs and histories. A recollection $\hat{\sigma}_t$ may be informative in some histories and uninformative in others.

For histories on the equilibrium path, belief consistency requires the individual to update beliefs only when low and high signals are interpreted differently. Since high signals are always interpreted as high, this is only possible when low signals are interpreted as low. Hence, a recollection $\hat{\sigma}_t$ after history h^t is informative if and only if $\hat{\sigma}^*\left(h^t\right) = L$. The following lemma states this result formally:

Lemma 1 *Fix an interpretation strategy $\hat{\sigma}^* : \mathcal{H} \to \{L, H\}$ and let μ be posterior beliefs consistent with it. Let $h^t \in \mathcal{H}$ be a history on the equilibrium path. A recollection $\hat{\sigma}_t$ at history h^t is informative if and only if $\hat{\sigma}^*\left(h^t\right) = L$.*

For any history h^t on the equilibrium path, either both high and low recollections will be informative (if $\hat{\sigma}^*\left(h^t\right) = L$), or they will both be uninformative (if $\hat{\sigma}^*\left(h^t\right) = H$). We can, therefore, refer to the informativeness of a period rather than the informativeness of each recollection. We say that history h^t has n *informative periods* if beliefs were modified in n of its subhistories:

$$\#\left\{\left(h^\tau, \hat{\sigma}_\tau\right) \subset h^t : \mu\left(\theta|h^\tau, \hat{\sigma}_\tau\right) \neq \mu\left(\theta|h^\tau\right)\right\} = n.$$

We say that a history h^t has k *successes* in n *informative periods* if, among the n periods in which beliefs were modified, there were k high recollections:

$$\#\left\{\left(h^\tau, H\right) \subset h^t : \mu\left(.|h^\tau, H\right) \neq \mu\left(.|h^\tau\right)\right\} = k.$$

Hence, over such a history, the individual revised her beliefs upwards (in the sense of first-order stochastic dominance) k times and downwards $n-k$ times. The following examples illustrate these definitions:

Example 1 *Let $\hat{\sigma}^*(h^t) = H$ and let posterior beliefs be equal to the prior distribution for all histories. Only high recollections happen with positive probability. Furthermore, for any t, history $h^t = \{H, ..., H\}$ has 0 successes in 0 informative periods.*

Example 2 *Let $\hat{\sigma}^*(h^t) = L$ for all histories and let μ be beliefs consistent with this interpretation strategy. Since every recollection is informative, a history h^t features $k = \#\{\hat{\sigma}_\tau \in h^t : \hat{\sigma}_\tau = H\}$ successes in $n = t-1$ informative periods.*

Example 3 *Let $\hat{\sigma}^*(h^1) = L$ and $\hat{\sigma}^*(h^t) = H$ for all $h^t \neq h^1$. Let μ be beliefs consistent with this interpretation strategy. Then, history $\{H, L, L, L..., L\}$ has 1 success in 1 informative period, and history $\{L, L, L, ..., L\}$ has 0 successes in 1 informative period.*

For a given prior ρ, the number of successes and informative periods fully identifies posterior beliefs along the equilibrium path. More specifically, let h^t be a history on the equilibrium path with k successes in n informative periods. The posterior distribution of θ at h^t has density

$$\mu(\theta|k,n) = \frac{p(\theta)^k [1-p(\theta)]^{n-k} \rho(\theta)}{\int_\Theta p(\theta)^k [1-p(\theta)]^{n-k} \rho(\theta)\, d\theta}.$$

The key difference between the model of selective awareness and other models of non-Bayesian learning is the context-dependence of biases. In order to derive testable implications, we need to understand how changes in the self-image content of signals affect how they are interpreted. There are two ways to derive comparative statics results when there are multiple equilibria. One possibility is to apply a selection criterion. For example, we could select the most efficient equilibrium. A more robust approach, which I pursue here, is to establish set monotonicity (with respect to the inclusion order).

I will say that the equilibrium set is decreasing in the importance of self-image if the set of interpretations and actions on the equilibrium path is decreasing in α. More formally, the equilibrium set is *decreasing in the importance of self-image* if, for any $\alpha_1 > \alpha_0$, whenever $(\hat{\sigma}_1^*, a_1^*, \mu_1^*)$ is the PBE of the game under parameter α_1, there exists a PBE of the game under α_0, $(\hat{\sigma}_0^*, a_0^*, \mu_0^*)$, such that $\hat{\sigma}_0^*(h^t) = \hat{\sigma}_1^*(h^t)$, $a_0^*(h^t) = a_1^*(h^t)$, and $\mu_0^*(\theta|h^t) = \mu_1^*(\theta|h^t)$ for all histories with $\Pr(h^t|\sigma_1^*) > 0$.[20]

Proposition 2 (Comparative Statics) *The equilibrium set is decreasing in the importance of self-image. Moreover:*

[20]Notice that requiring that $\hat{\sigma}_0^*$ and $\hat{\sigma}_1^*$ coincide in all histories on the equilibrium path induced by $\hat{\sigma}_1^*$ implies that the set of histories on the equilibrium paths of both $\hat{\sigma}_0^*$ and $\hat{\sigma}_1^*$ must also coincide.

1. *There exists $\bar{\alpha} \in \mathbb{R}_+$ such that, in any PBE, $\hat{\sigma}^*(h^t) = H$ for all h^t on the equilibrium path whenever $\alpha > \bar{\alpha}$.*

2. *Let $N \in \mathbb{N}$. Suppose $\arg\max_a E[V(a,\theta)|k,n] \nsubseteq \arg\max_a E[V(a,\theta)|k-1,n]$ for all $(k,n) \in \{(k,n) : 1 \leq k \leq n < N\}$. Then, there exists a PBE in which $\hat{\sigma}(h^t) = L$ for all h^t and $t < N$ whenever $\alpha < \underline{\alpha}$ for some $\underline{\alpha} \in \mathbb{R}_{++}$.*

In order to sustain an equilibrium in which a signal is interpreted realistically, we must ensure that the individual will not prefer to interpret a low signal as high. Since the incentive to rationalize low signals away is increasing in the importance of self-image α, the set of equilibria that can be sustained decreases in α. In particular, when the importance of self-image is high enough, the individual cannot commit not to rationalize any low signal away. Then, there is no learning in any equilibrium (Part 1). Conversely, there exists an equilibrium with some learning if the importance of self image is low enough and misinterpreting a signal can reduce the payoff from actions (Part 2).

Remark 1 (Preferences affect Beliefs) *Proposition 2 shows that individuals with the same prior beliefs and subject to the same information may hold systematically different posterior beliefs if they have different payoffs from self-image. In particular, the amount of information that can be learned is decreasing in the importance of self-image. This result is consistent with the evidence described in the introduction.*

The conservatism bias states that individuals update beliefs in the right direction, but by too little relative to the Bayesian update. The individual in this model always displays conservatism. When the equilibrium assigns a low interpretation to a low signal $\hat{\sigma}^*(h^t) = L$, the interpretation fully reveals which signal was observed and beliefs about θ are updated according to Bayes' rule. However, when the equilibrium assigns a high interpretation to a low signal $\hat{\sigma}^*(h^t) = H$, the individual's recollection is uninformative and, therefore, she does not update her beliefs. Hence, selective interpretation introduces additional noise in the individual's recollections of signals, which causes her to update beliefs in the same direction as the Bayesian update conditional on the realized signals, but at a slower rate.

Let θ_t^B denote the expected value of θ obtained by Bayes' rule conditional on the sequence of observed signals $\{\sigma_1, ..., \sigma_{t-1}\}$, and let $\hat{\theta}_t$ denote the expected value of θ calculated according to the individual's beliefs μ (i.e., conditional on the sequence of the individual's interpretations). The following proposition establishes that the individual displays conservatism: Posterior expectations move in the same direction, but are "less variable" (in the sense of second-order stochastic dominance) than the expectations obtained by Bayes' rule.[21]

[21] In Appendix VI, I generalize Proposition 3 to behavioral strategies, leading to a smoother, more realistic version of conservatism.

Proposition 3 (Conservatism) *Let $\hat{\sigma}^*$ be an interpretation strategy and let μ be posterior beliefs consistent with this strategy. For any history h^t such that* $\Pr\left(h^t|\hat{\sigma}^*\right) > 0$,

$$\begin{aligned} \theta_t^B > \theta_{t-1}^B &\implies \hat{\theta}_t \geq \hat{\theta}_{t-1}, \text{ and} \\ \theta_t^B < \theta_{t-1}^B &\implies \hat{\theta}_t \leq \hat{\theta}_{t-1}. \end{aligned}$$

Furthermore, $\hat{\theta}_t$ second-order stochastically dominates θ_t^B.

Consider an equilibrium in which the individual interprets signals realistically at a history with k successes in n informative periods. Deviating and interpreting a low signal as high increases the payoff from self views in the current period by

$$\alpha \left\{ E\left[\theta|k+1, n+1\right] - E\left[\theta|k, n+1\right] \right\}. \tag{2}$$

This is a lower bound on the total gain from the deviation. Since the individual will believe to have observed one more high signal than she really did in all subsequent periods, future payoffs from self image are also increased.

In order to obtain results that hold for any prior distribution, we need an approximation of the conditional expectation $E\left[\theta|k, n\right]$ that is independent of the prior. This will be established using a method originally developed by Laplace (1774), who showed that conditional expectations pile up near the maximum likelihood estimator.[22] Since maximum likelihood estimators do not depend on the prior distribution, this approximation will allow us to obtain uniform results. Let $\mathbf{s} \equiv \{\hat{\sigma}_1, \hat{\sigma}_2, ...\} \in \{L, H\}^\infty$ denote an infinite sequence of informative recollections. The following lemma establishes that we can write the conditional expectation in terms of the maximum likelihood estimator $p^{-1}\left(\frac{k}{n}\right)$ plus terms of higher order:

Lemma 2 *There exist C and $N_\mathbf{s}$ such that, for all $n > N_\mathbf{s}$,*

$$\left| E\left[\theta|k, n\right] - p^{-1}\left(\frac{k}{n}\right) - \xi\left(\frac{k}{n}\right)\frac{1}{n} \right| \leq \frac{C}{n^2} \tag{3}$$

for almost all $\mathbf{s}$ (under the true θ), where

$$\xi(x) := \frac{\sqrt{x(1-x)}}{p'\left(p^{-1}(x)\right)} \left\{ \frac{2\left[p'\left(p^{-1}(x)\right)\right]^3 (1-2x)}{\left[x(1-x)\right]^2} - \frac{3p'\left(p^{-1}(x)\right)p''\left(p^{-1}(x)\right)}{x(1-x)} + \frac{\rho'\left(p^{-1}(x)\right)}{\rho\left(p^{-1}(x)\right)} \right\}.$$

Using Lemma 2, we can then estimate the improvement in self-image from one additional success when the number of informative periods n is large:

[22] Laplace only considered uniform priors. Among the many generalizations of his result, I will follow the approach of Johnson (1970).

Proposition 4 *There exist C and $N_{\mathbf{s}}$ such that, for all $n > N_{\mathbf{s}}$,*

$$\left| E\left[\theta | k+1, n\right] - E\left[\theta | k, n\right] - \frac{1}{p'\left(p^{-1}\left(\frac{k}{n}\right)\right)} \frac{1}{n} \right| \leq \frac{C}{n^2}$$

for almost all $\mathbf{s}$ *(under the true θ).*

Since $p'\left(p^{-1}\left(\frac{k}{n}\right)\right) \rightarrow_{a.s.} p'(\theta) > 0$, the improvement in self image is $O\left(\frac{1}{n}\right)$. I will consider the cost of the deviation for continuous and discrete action spaces separately.

2.1 Continuum of Actions

This subsection obtains a finite bound on the number of informative periods for continuous action spaces. Recall that the action space A is a non-empty compact subset of the real line. I will assume that it is also convex:

Assumption 1. *A is a compact interval of the real line.*

Consider, for instance, a quadratic loss function:

$$V(a, \theta) = -\kappa (a - \theta)^2 . \tag{4}$$

The parameter $\kappa > 0$ captures the importance of taking the correct action. Let $A = \Theta$ be the action space, and let $a_{k,n}$ denote the optimal action in a history with k successes in n informative periods. Under quadratic loss, the optimal action is the mean of the posterior: $a_{k,n} = E[\theta | k, n]$.

Suppose the individual interprets a low signal realistically after history h^t and consider a deviation to $\hat{\sigma}_t = H$. From Proposition 4, the increase in self-image is bounded below by a term of order $\frac{1}{n}$. If the game ends in a history with $\tilde{k}$ successes in $\tilde{n} > n$ informative periods, the individual will choose action $a_{\tilde{k}+1,\tilde{n}}$ instead of $a_{\tilde{k},\tilde{n}}$. The cost of this deviation is

$$E\left[V(a_{\tilde{k},\tilde{n}}, \theta) - V(a_{\tilde{k}+1,\tilde{n}}, \theta) | k, n\right] = \kappa \left\{ E\left[\theta | \tilde{k}+1, \tilde{n}\right] - E\left[\theta | \tilde{k}, \tilde{n}\right] \right\}^2 ,$$

which, by Proposition 4, is of order $\frac{1}{\tilde{n}^2}$. Hence, the cost of the deviation converges to zero faster than the benefit. Therefore, the individual prefers to rationalize low signals away in *any* PBE after a sufficiently informative history.

The result from the quadratic model can be substantially generalized. Assume that the payoff from actions is a sufficiently smooth function, which is concave in actions:

Assumption 2. *V is twice continuously differentiable. For each $\theta \in \Theta$, $\frac{\partial^2 V}{\partial a^2}(a, \theta) < 0$.*

Strict concavity implies that the optimal action for each skill, $a(\theta) := \arg\max_{a \in A} \{V(a, \theta)\}$, is unique. Assume that it is also interior:

Assumption 3. *For almost all θ, $a(\theta)$ is an interior point of A.*

Assumptions 2 and 3 ensure that distortions close to the optimal action have second-order costs. As a result, the cost of distorting self-image is of lower order of magnitude than $\frac{1}{n}$, implying that the individual will prefer to rationalize low signals away in any PBE when n is large. The following theorem establishes this result formally:

Theorem 1 (Confirmation Bias for Continuous Actions) *Suppose Assumptions 1, 2, and 3 are satisfied. There exists $n \in \mathbb{N}$ such that, in any PBE, almost every history on the equilibrium path has at most n informative periods.*

The proof has two main steps. First, we apply an asymptotic expansion to $a_{\tilde{k},\tilde{n}}$ to verify that the distortion in actions from misinterpreting a signal is bounded. Second, we apply a Taylor expansion to the expected payoff from actions around the optimal action $a_{\tilde{k},\tilde{n}}$. Since the first-order term is zero due to the optimality of $a_{\tilde{k},\tilde{n}}$ and the distortion is bounded, only terms with order of magnitude lower than $\frac{1}{n}$ remain.

Theorem 1 shows that the individual's beliefs (almost) never converge to the truth. Therefore, selective awareness imposes a limit to learning. Not only is there always an equilibrium in which no learning ever occurs (Proposition 1); learning is incomplete in *any* equilibrium. By Proposition 2, the amount of learning that can be sustained in equilibrium is decreasing in the importance of self image α.

2.2 Finitely many Actions

This subsection considers the model with a finite action space:

Assumption 4. *A is a finite set.*

The next assumption, which is generic, states that the optimal action is globally unique except at a finite number of skills θ:

Assumption 5. $\arg\max_{a \in A} \{V(a, \theta)\}$ *is a singleton except at a finite set.*

Assumption 5 allows us to partition the type space into a finite number of intervals, with types in the interior of each interval having a strictly preferred action. It is automatically satisfied under the standard assumption that V has either strictly increasing or strictly decreasing differences.

In the continuous-action case, the smoothness of V and the interiority of the solution ensured that small distortions had second-order costs. When actions are finite, this is not true since any distortion that affects actions will necessarily cause a discrete loss. However, for any continuation strategy, the probability that one particular signal affects the optimal action decreases exponentially as the number of informative periods grows. Since the maximal loss from making an incorrect decision is bounded and the probability of affecting the decision decreases exponentially (which converges to zero faster than $1/n$), the benefit from distorting the signal eventually exceeds its cost. At that point, the individual must ignore every additional information in every equilibrium. Thus, as in the continuous-action case, beliefs (almost) never converge to the truth:

Theorem 2 (Confirmation Bias for Finite Actions) *Suppose Assumptions 4 and 5 are satisfied. For almost all θ, there exists n_θ such that, in any PBE, almost all histories on the equilibrium path have at most n_θ informative periods.*

2.3 Markovian Equilibria

A strategy is Markovian if it depends on payoff-relevant information only. In this model, histories affect payoffs only through beliefs about skills. Hence, a Markovian strategy space partitions the set of histories based on beliefs about skills.[23] A Markovian Perfect Bayesian Equilibrium is a PBE in which strategies are Markovian:

Definition 2 *A Markovian Perfect Bayesian Equilibrium (MPBE) is a PBE $(\hat{\sigma}, a, \mu)$ in which, for any histories h^t and $h^{\tau'}$,*

$$\mu\left(\theta | h^t\right) = \mu\left(\theta | h^{\tau'}\right) \implies a\left(h^t\right) = a\left(h^{\tau'}\right) \text{ and } \hat{\sigma}\left(h^t\right) = \hat{\sigma}\left(h^{\tau'}\right).$$

Let $(\hat{\sigma}, a, \mu)$ be an MPBE and let h^t be a history on the equilibrium path with k successes in n informative periods. Suppose the individual interprets a low signal realistically at h^t, i.e., $\hat{\sigma}^*\left(h^t\right) = L$. If the information process does not end at h^t, the individual either observes a high signal at $t+1$, thereby moving to a history with $k+1$ successes in $n+1$ informative periods, or she observes a low signal at $t+1$, moving instead to a history with k successes in $n+1$ informative periods. Therefore, histories in which the individual interprets a low signal realistically are associated with *transient* Markovian states. Conversely, suppose the individual rationalizes a low signal as high in history h^t, i.e., $\hat{\sigma}^*\left(h^t\right) = H$. Because both signals are interpreted equally, a high recollection is not informative and the individual's beliefs remain unchanged: h^t is associated with an *absorbing* Markovian state. Therefore, in an MPBE, once the individual rationalizes a signal away, she will keep rationalizing every future signal away.

In Markovian equilibria, histories on the equilibrium path can be split into two stages. In the first stage, signals are interpreted correctly and beliefs evolve according to Bayes' rule conditional on the observed signals; in the second stage, signals are misinterpreted and beliefs remain unchanged. Consequently, the individual attaches a disproportionately high weight to initial information. The following corollary states this result formally:

[23] This definition follows Maskin and Tirole (2001) in excluding the history length from the state space. Some authors refer to these strategies as stationary Markovian strategies.

Corollary 1 (First Impressions Matter) *Fix an MPBE. For any history h^t on the equilibrium path, there exists $T_{h^t} \leq t$ such that a period $\tau \leq t$ is informative if and only if $\tau < T_{h^t}$. Moreover:*

- *If Assumptions 1-3 are satisfied, there exists $n \in \mathbb{N}$ such that $T_{h^t} \leq n$ for almost every history h^t on the equilibrium path.*
- *If Assumptions 4-5 are satisfied, for almost all θ, there exists $n_\theta \in \mathbb{N}$ such that $T_{h^t} \leq n_\theta$ for almost every history on the equilibrium path.*

Recall that the individual would prefer, ex-ante, to interpret as many signals realistically as possible. Hence, the least efficient PBE is the equilibrium in which all signals are rationalized away, which is Markovian. Moreover, because the game ends with probability η every period, she would prefer to interpret signals realistically as early as possible. Thus, the most efficient PBE is an equilibrium in which the individual interprets signals realistically whenever possible. This is also a Markovian equilibrium.[24] Therefore, Markovian equilibria provide bounds on the amount of inefficiency that may result in any equilibrium of the model. Moreover, if one is willing to adopt a selection criterion that picks an efficient equilibrium, one will necessarily select a Markovian equilibrium.

3 Uniform-Quadratic Model

This section illustrates the theory with a specific formulation of the model. In this formulation, the payoff from actions is determined by the quadratic loss function (4), skills are uniformly distributed, and the probability of a high signal is uniformly distributed on the unit interval:

$$p(\theta) = \frac{\theta - \underline{\theta}}{\bar{\theta} - \underline{\theta}}, \quad \text{and} \quad \theta \sim U[\underline{\theta}, \bar{\theta}].$$

From Bayes' rule, the posterior probability of a high signal and the expected skill conditional on k successes in n informative periods equal

$$E[p|k,n] = \frac{k+1}{n+2}, \quad \text{and} \quad E\left[\theta|k,n\right] = \underline{\theta} + \left(\bar{\theta} - \underline{\theta}\right)\left(\frac{k+1}{n+2}\right). \tag{5}$$

In Proposition 4, we saw that misinterpreting a low signal as high in a history with n informative periods raises expected skills by an order of magnitude $\frac{1}{n}$. Here, we can calculate this effect in closed form:

$$E\left[\theta|k+1,n\right] - E\left[\theta|k,n\right] = \frac{\bar{\theta} - \underline{\theta}}{n+2}.$$

[24] See the Online Appendix for a formal proof.

For general distributions, the benefit from rationalizing low signals away depends on the number of successes k. In the uniform model, however, the effect of misinterpreting a signal on the expectation of skills does not depend on the number of successes. Therefore, the benefit from rationalizing low signals away is the same in all histories with the same number of informative periods.

3.1 Continuous Actions

As in Subsection 2.1, let $A = \Theta$ be the action space, and let $a_{k,n} = E[\theta|k, n]$ denote the optimal action in a history with k successes in n informative periods. Misinterpreting a low signal as high induces future selves in histories with $\tilde{k}$ successes in $\tilde{n}$ informative periods to pick actions as if there were $\tilde{k}+1$ successes instead. The cost of doing so equals

$$E\left\{V\left(E\left[\theta|\tilde{k},\tilde{n}\right],\theta\right) - V\left(E\left[\theta|\tilde{k}+1,\tilde{n}\right]\right)|\tilde{k},\tilde{n}\right\} = \kappa\left(\frac{\bar{\theta}-\underline{\theta}}{\tilde{n}+2}\right)^2.$$

As we saw in Subsection 2.1, the cost of distorting an action close to the optimum is of second order. When the payoff from actions is quadratic, the cost of distorting an action is of second order everywhere (even away from the optimum). Because misinterpreting a signal raises expected skills by $\frac{\bar{\theta}-\underline{\theta}}{n+2}$, the cost of such a distortion is proportional to $\left(\frac{\bar{\theta}-\underline{\theta}}{n+2}\right)^2$, which does not depend on k.

Since both the benefit and the cost of misinterpreting a signal depend on the number of informative signals n but not on the number of successes k, the largest number of realistic interpretations does not depend on k. Let $\lfloor . \rfloor$ denote the floor function. The next proposition, whose proof is in the Online Appendix, obtains a closed-form solution for the maximum number of informative periods:

Proposition 5 (Uniform-Quadratic Continuous) *In any PBE, any strategy on the equilibrium path has at most* $\max\left\{\left\lfloor \eta\kappa\frac{\bar{\theta}-\underline{\theta}}{\alpha} - 2\right\rfloor, 0\right\}$ *informative periods and the bound is tight.*

In particular, an interpretation strategy is part of an MPBE if and only if the individual interprets at most $\bar{n} := \max\left\{\left\lfloor \eta\kappa\frac{\bar{\theta}-\underline{\theta}}{\alpha} - 2\right\rfloor, 0\right\}$ periods realistically and stops learning whenever she interprets a low signal as high. In the MPBE with the highest number of informative periods – the most efficient PBE –, learning ends after a constant number of periods $\bar{n}$ regardless of the history. Since the incentive to rationalize low signals away is increasing in the importance of self-image and decreasing in the cost of making an incorrect decision, the amount of realism that can be sustained in equilibrium decreases in α and increases in κ (consistently with Proposition 2). Figure 3 depicts this MPBE for $\bar{n} = 4$.

Proposition 5 also implies that there is no learning in any PBE if $\eta < \frac{3\alpha}{\kappa(\bar{\theta}-\underline{\theta})}$. Since the expected number of signals is $\frac{1}{\eta}$, the individual does not learn any information in any period

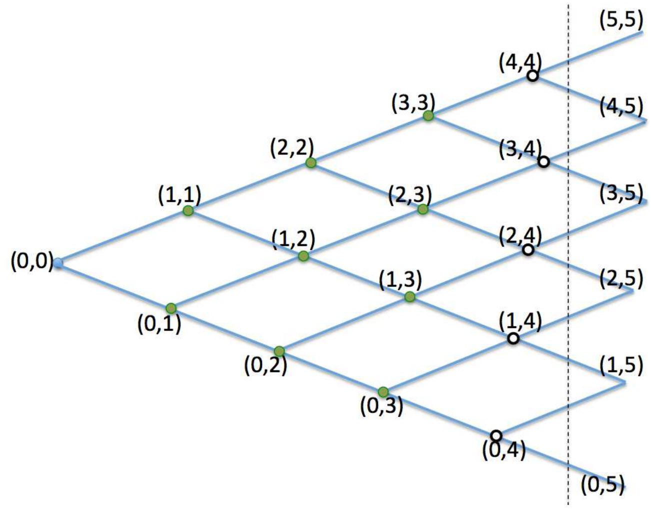

Figure 3: Markovian interpretation strategies in the most efficient MPBE for a uniform-quadratic model with continuous actions in which $\bar{n} = 4$. White circles represent absorbing states.

if she expects information to arrive for a sufficiently long time. When the individual takes the action far enough in the future, a deviation raises self-image for a longer time. As a result, and somewhat paradoxically, the greater the number of signals she expects to receive, the less she is able to learn.

3.2 Binary Actions

In this subsection, we assume that, instead of picking an action in the entire interval of types, the individual must pick one of its endpoints: $A = \{\underline{\theta}, \bar{\theta}\}$. It is optimal to choose the high action in a history with k successes in n informative periods if the proportion of successes exceeds the unconditional mean:

$$a_{k,n} = \begin{cases} \bar{\theta} \text{ if } \frac{k}{n} \geq \frac{1}{2} \\ \underline{\theta} \text{ if } \frac{k}{n} \leq \frac{1}{2} \end{cases} .$$

Proceeding as in the proof of Theorem 2 establishes that, for all $\theta \neq p^{-1}\left(\frac{1}{2}\right)$, there exists n_θ such that, in any PBE, all histories on the equilibrium path have at most n_θ informative periods.[25]

When deciding whether to interpret a low signal as high, the individual balances the gain from having a higher self-image against the cost of possibly taking a worse action. Deviating from a realistic interpretation is only costly if it affects the action taken in some continuation history (i.e., the deviation is "pivotal" in the decision). Hence, a low signal must be rationalized away at h^t if none of the continuation histories following h^t are pivotal. As a result, it is harder to sustain realism in extreme histories than in intermediate ones. Let $\lceil . \rceil$ denote the ceiling function. The following proposition establishes this result formally:

[25]Theorem 2 required $p(\underline{\theta}) > 0$ and $p(\bar{\theta}) < 1$ to justify the asymptotic expansions of the posterior mean and the payoff from actions. Since we are able to compute the posterior mean and the payoff from actions in closed form in the uniform model, we do not need this assumption. The rest of the proof remains unchanged.

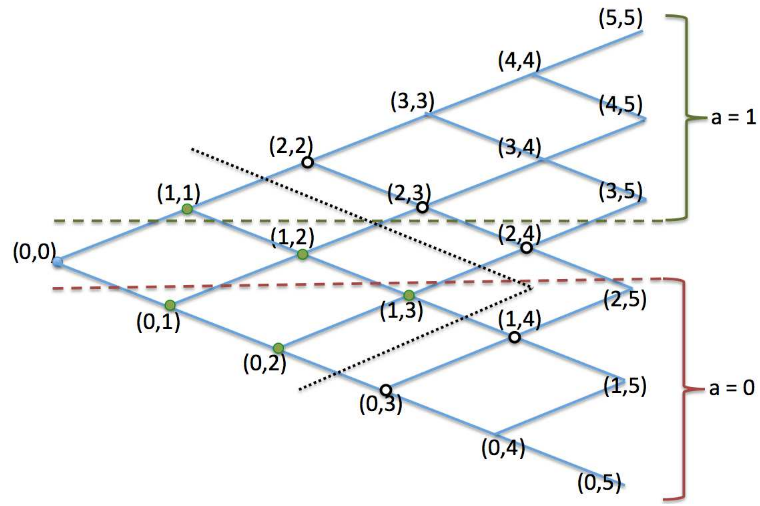

Figure 4: Markovian interpretation strategies in the most efficient MPBE for a uniform-quadratic model with binary actions in which $\bar{n}_\theta = 4$. White circles represent absorbing states.

Proposition 6 (Uniform-Quadratic Binary) *Let $\theta \neq p^{-1}\left(\frac{1}{2}\right)$, fix a PBE, and let n_θ be the largest number of informative periods in this PBE. Then, any history on the equilibrium path has at most $\lceil \frac{n_\theta}{2} \rceil$ successes and at most $\lceil \frac{n_\theta}{2} \rceil + 1$ failures.*

Figure 4 illustrates the result from Proposition 6. The two horizontal lines delineate the regions in which the high and the low actions are strictly preferred ($k > \frac{n}{2}$ and $k < \frac{n}{2}$, respectively), and the states where both actions have the same expected payoff ($k = \frac{n}{2}$). Since the individual must prefer to interpret low signals realistically in any transient state, there must be a path from every transient state to a state in which the individual picks a different action; otherwise, rationalizing a low signal away is costless. In the figure, $(1, 3)$ is the last state in which a signal can affect actions (the individual picks $a = 1$ at state $(2, 4)$ and $a = 0$ at $(1, 4)$). Thus, once a state with 2 successes or 3 failures is reached, there is no cost of rationalizing low signals away and learning stops.

4 Applications

The selective awareness model studied here has many applications in economics and finance. As described in the introduction, versions of the model with a single period of distortion have been applied to a large variety of contexts, including personal motivation (Benabou and Tirole, 2002), redistributive policies (Benabou and Tirole, 2006a), political attitudes towards reforms Levy (2012), groupthink in organizations and contagious exuberance in markets (Benabou, 2008), ideologies (Benabou, 2011), preferences for increasing wages (Smith, 2009a, 2009b), and fear of death as an explanation for puzzles in health and savings behavior (Kopczuk and Slemrod, 2005). A related model, which also assumes that all manipulation happens at a single (ex-ante) period, has been used to explain preference for skewness, overconfidence, and overoptimism (Brunnermeier and Parker, 2005), lack of diversification (Brunnermeier, Gollier, and Parker, 2007), and behavior among individuals at risk for Huntington's disease (Oster, Shoulson, and Dorsey, 2013). This paper complements this literature by showing that these biases persist even if individuals obtain an arbitrarily large amount of feedback about the environment and by deriving the testable predictions from repeated versions of these models.

The model studied here can also be directly applied to study the persistence of optimism. There is a very large literature that studies implications of optimism to economics and finance.[26] Most theoretical models, however, assume that individuals start with an optimistic prior and update their beliefs according to Bayes' rule. These assumptions imply that, under minor conditions, beliefs converge to the truth as the number of observations grows. Empirically, however, most studies suggests that optimism persists with experience.[27] The model presented here is consistent with this empirical pattern. Moreover, and consistently with the evidence described in the introduction, the model predicts that failure to learn is more common for traits and behaviors that individuals regard as important.

In the remainder of this section, I discuss two new applications of the theory.

4.1 Persistence of Naiveté

An important literature in behavioral economics studies individuals with imperfect knowledge of their own self control – c.f. O'Donoghue and Rabin (1999, 2001). Papers in this literature treat beliefs as exogenously given. Consequently, individuals may repeatedly fail to account for their lack of self control.

Ali (2011) studies learning about one's own self control using a model of experimentation. Decisions in each period are made through the conjunction of two separate systems: a (long-run) planner and a (short-run) doer. In each period, the planner decides between exposing herself to temptation or not. If she exposes herself to temptation, the doer chooses between resisting or succumbing to it. The planner does not know the doer's ability to resist temptation and, instead, learns from the doer's choices. If the doer repeatedly succumbs to temptation, the planner eventually decides that the doer is unable to resist temptation and, therefore, decides to no longer expose herself. Thus, the only individuals who continue to expose themselves to temptation are the ones who are able to resist it.

In Appendix III, I embed Benabou and Tirole's (2004) model of willpower in this experimen-

[26] Camerer and Lovallo (1999), for example, experimentally argue that optimism explains the high rate of business failure. Malmendier and Tate (2005) show that CEO optimism is an important predictor of corporate investment distortions.

[27] See, for example, Glaser, Langer, and Weber (2005); Brozynski, Menkhoff, and Schmidt (2006); Deaves, Lüders, and Schröder (2010) and references therein. Glaser and Weber (2007) show that investor optimism is related to biased recollections of past performance.

tation framework.[28] The key difference with respect to Ali's (2011) model is that the planner can selectively interpret the outcome from exposing herself to temptation. In addition to the payoff from the tempting activity, the planner has preferences over her perceived self control. These preferences may be due to self-image concerns or from anticipatory utility.

As in Ali (2001), individuals who believe they have low enough self control stop experimenting and, therefore, fail to learn their true parameters. However, conditionally on exposing oneself to temptation, the effect of distorting an interpretation is $O(\frac{1}{n})$, whereas the cost of doing so decreases exponentially in n (as in Theorem 2). Thus, because of selective awareness, even those who keep exposing themselves to temptation eventually stop learning. Moreover, there is a positive probability that individuals repeatedly succumb to temptation while disregarding their lapses. The model, therefore, provides a motivated-reasoning rationale for individuals who continue to expose themselves to temptation while frequently succumbing to it.

4.2 Incentives to Collect Information

Next, I discuss an application for the theory of organizations. As I show in Appendix II, the anticipatory utility interpretation of the model implies that workers with a higher stake in a company are less capable of processing information objectively. Therefore, tying (either explicitly or implicitly) the compensation of an employee to the performance of the firm reduces the employee's ability to provide accurate evaluations.

The predictions from the model contrast with those from the theory of advocates, which is based on the need to provide incentives to collect information (Dewatripont and Tirole, 1999). In the theory of advocates, assigning a particular cause to an agent and offering a compensation strongly aligned with her defense of that cause is an efficient way to encourage the agent to gather information. In the presence of selective awareness, the opposite is true: incentivizing an agent to defend a particular cause reduces the agent's ability to provide accurate information. In fact, absent moral hazard concerns, the optimal compensation scheme is completely inflexible to the information provided. With moral hazard, the benefit from incentivizing effort has to be weighted against the bias induced by the agent's compensation.

Some researchers have argued that selective awareness is often a much larger problem than the standard conflicts of interest due to moral hazard. For example, starting with Bazerman, Morgan, and Loewenstein (1997), a few papers have documented how selective awareness affects the performances of auditors. As Bazerman, Loewenstein, and Moore (2002) describe it: "Psychological research shows that our desires powerfully influence the way we interpret information, even when we're trying to be objective and impartial. When we are motivated to reach a particular conclusion, we usually do." In the words of Moore et al. (2006):

[28] Benjamin, Rabin, and Raymond (2013) model the persistence of naiveté as a consequence of non-belief in the law of large numbers rather than motivated reasoning.

> Putting the most Machiavellian fringes of professional communities aside, [...] the majority of professionals are unaware of the gradual accumulation of pressures on them to slant their conclusions — a process we characterize as moral seduction. Most professionals feel that their professional decisions are justified and that concerns about conflicts of interest are overblown by ignorant or demagogic outsiders who malign them unfairly. Given what we now know generally about motivated reasoning and self-serving biases in human cognition, and specifically about the incentive and accountability matrix within which auditors work, we should view personal testimonials of auditor independence with skepticism.

The endogenous effects of compensation on beliefs are supported by the recent work of Cheng, Raina, and Xiong (2014), who study beliefs during the housing bubble of 2004-2006. They show that securitization managers, whose future income was directly linked to the performance of the housing market, were actually more likely to buy houses in the period relative to individuals with no private information and, as a result, they obtained a significantly worse performance on their home portfolio.

5 Conclusion

Several papers have recently used the selective awareness framework proposed by Benabou and Tirole to provide explanations for departures from rational decision making. However, it is often argued that the Bayesian updating assumption embedded in this framework combined with the repeated nature of the decisions being modeled would lead individuals to eventually learn the truth and departures from rationality would vanish in the long-run.

This paper formally studied this issue by considering a repeated version of the selective awareness model of Benabou and Tirole (2002, 2004) and Benabou (2008, 2011). It showed that all information is disregarded after a certain number of observations. Therefore, learning is always incomplete, and the departures from rationality presented in the static models in the literature do not disappear even when the decision problem is repeated infinitely many times.

The model predicts a behavior that is consistent with some biases in information processing studied by psychologists. Individuals attribute a disproportionately large weight to initial information. After becoming sufficiently convinced of which action to take, they do not change their beliefs (confirmation bias). They also update beliefs in the right direction, but in insufficient amount compared to the Bayesian updating rule (conservatism bias). The model has implications for learning about one's self control and for the design of incentives for gathering information.

References

Akerlof, G. A. and W. T. Dickens (1982). "The Economic Consequences of Cognitive Dissonance," American Economic Review, 72(3), 307-319.

Ali, N. (2011), "Learning Self-Control," Quarterly Journal of Economics, 126(2), 857-893.

Bahrick, H. P., L. K. Hall, and S. A. Berger (1996). "Accuracy and distortion in memory for high school grades," Psychological Science, 7, 265-271.

Banerjee, A. (1992). "A Simple Model of Herd Behavior," Quarterly Journal of Economics 107(3), 797-817.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992). "A Theory of Fads, Fashion, Custom and Cultural Change as Information Cascades." Journal of Political Economy, 100, 992-1026.

Barberis, N., A. Shleifer, and R. Vishny (1998). "A model of investor sentiment," Journal of Financial Economics, 49 (3), 307-343.

Benabou, R. (2008). "Ideology," Journal of the European Economic Association, 6, 321–352.

Benabou, R. (2013). "Groupthink: Collective Delusions in Organizations and Markets," Review of Economic Studies, 80, 429-462.

Benabou, R. and J. Tirole (2002). "Self-Confidence and Personal Motivation," Quarterly Journal of Economics, 117(3), 871-915.

Benabou, R. and J. Tirole (2004). "Willpower and Personal Rules," Journal of Political Economy, 112 (4), 848-886.

Benabou, R. and J. Tirole (2006a). "Belief in a Just World and Redistributive Politics," Quarterly Journal of Economics, 121(2), 699-746.

Benabou, R. and J. Tirole (2006b). "Incentives and Prosocial Behavior," American Economic Review, 96(5), 1652-1678.

Benabou, R. and J. Tirole (2009). "Over My Dead Body: Bargaining and the Price of Dignity," American Economic Review, 99(2), 459–465.

Benabou, R. and J. Tirole (2011). "Identity, Morals, and Taboos: Beliefs as Assets," Quarterly Journal of Economics, 126, 805-855.

Benjamin, D. J., M. Rabin, and C. Raymond (2013). "A Model of Non-Belief in the Law of Large Numbers," Mimeo., Cornell University, UC Berkeley, and Oxford University.

Bernheim, B. D. and R. Thomadsen (2005). "Memory and Anticipation," Economic Journal, 115, 271–304.

Bodner, R. and D. Prelec. "Self-signaling in a neo-Calvinist model of everyday decision making," in Psychology and Economics, Vol. II. Brocas and J. Carillo (eds.), Oxford University Press, 2002.

Brocas, I. and J. Carrillo, (2012). "From Perception to Action: an Economic Model of Brain Processes," Games and Economic Behavior, 75, 81–103.

Brunnermeier, M. K. and J. A. Parker (2005). "Optimal Expectations," American Economic Review, 95(4): 1092-1118.

Brunnermeier, M. K., C. Gollier, and J. A. Parker (2007). "Optimal Beliefs, Asset Prices, and the Preference for Skewed Returns," NBER Working Paper #12940.

Caplin, A. and J. Leahy (2001). “Psychological Expected Utility Theory and Anticipatory Feelings,” Quarterly Journal of Economics, 116 (1), 55-79.

Carrillo, J., and T. Mariotti (2000). “Strategic Ignorance as a Self-Disciplining Device,” Review of Economic Studies, 66 , 529–544.

Choi, D. and D. Lou (2010). “A Test of the Self-Serving Attribution Bias: Evidence from Mutual Funds,” Mimeo., Hong Kong University and LSE.

Cheng, I.-I., S. Raina, and W. Xiong (2014). “Wall Street and the Housing Bubble,” American Economic Review, forthcoming.

Compte, O. and A. Postlewaite (2004). “Confidence-Enhanced Performance,” American Economic Review, 94, 1536-1557.

Dewatripont, M. and J. Tirole (1999) “Advocates,” Journal of Political Economy, 107(1), 1-39.

Edwards, W. (1968). “Conservatism in human information processing,” In: B. Kleinmutz (Ed.), Formal Representation of Human Judgment. (pp. 17-52). New York: John Wiley and Sons.

Eil, D. and J. M. Rao (2011). “The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself,” American Economic Journal: Microeconomics, 3(2), 114-48.

Evans, J. (1989). Bias in human reasoning: Causes and consequences. Hillsdale, NJ: Erlbaum.

Fudenberg, D. and D. K. Levine. (1992). “Maintaining a Reputation when Strategies are Imperfectly Observed,” Review of Economic Studies 59(3), 561-579.

Fudenberg, D. and J. Tirole (1991). Game Theory. Cambridge, MA: MIT Press.

Gennaioli, N. and A. Shleifer (2010). “What Comes to Mind,” Quarterly Journal of Economics, 125 (4): 1399-1433.

Goetzmann, W. N. and N. Peles (1997). “Cognitive dissonance and mutual fund investors,” Journal of Financial Research, 20(2), 145–158.

Gollwitzer, Peter M., and Ute Bayer. “Deliberative versus implemental mindsets in the control of action.” In: (eds) Dual-process theories in social psychology (1999), 403-422.

Gottlieb, D. (2014). “Imperfect Memory and Choice under Risk,” Forthcoming, Games and Economic Behavior.

Greene, E. (1981). “Whodunit? Memory for evidence in text,” American Journal of Psychology, 94, 479-496.

Hales, J. (2007). “Directional Preferences, Information Processing, and Investors’ Forecasts of Earnings.” Journal of Accounting Research, 45: 607–628.

Johnson, R.A. (1970) “Asymptotic Expansions Associated with Posterior Distributions,” The Annals of Mathematical Statistics, 41(3), 851-864.

Karlsson, N., Loewenstein, G., and Seppi, D. “The ostrich effect: Selective attention to information,” Journal of Risk and Uncertainty 38.2 (2009): 95-115.

Kuhnen, C. M. (2012). “Asymmetric Learning from Financial Information,” Mimeo., Northwestern University.

Kopczuk, W. and J. Slemrod (2005). “Denial of Death and Economic Behavior,” Advances in Theoretical Economics, 5(1), Article 5.

Köszegi, B. (2006). “Ego Utility, Overconfidence, and Task Choice,” Journal of the European Economic Association, 4, 673-707.

Kunda, Z. (1990). “The case for motivated reasoning,” Psychological Bulletin, 108, 480-498.

Kuran, T. (1993). “The Unthinkable and the Unthought,” Rationality and Society, 5, 473-505.

Laplace, P.-S. (1774). “Mémoire sur la probabilité des causes par les événements.” Mémoires presentés à l’Académie Royale des Sciences. English translation by S. M. Stigler (1986), Statistical Science, 1 (3), 364-378.

Levy, R. (2012). “Soothing Politics,” Mimeo., University of Mannheim.

Loewenstein, G. (1987). “Anticipation and the Valuation of Delayed Consumption.” Economic Journal , 387, 97, 666-684.

Lord, Charles G., Lee Ross, and Mark R. Lepper. “Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence.” Journal of Personality and Social Psychology 37.11 (1979): 2098.

MacDonald, T. K., and M. Ross (1999). “Assessing the accuracy of predictions about dating relationships: How and why do lovers’ predictions differ from those made by observers?,” Personality and Social Psychology Bulletin, 25, 1417-1429.

Madarász, K. (2012). “Information Projection: Model and Applications ,” Review of Economic Studies, forthcoming.

Maskin, E. and J. Tirole (2001). “Markov perfect equilibrium: I. Observable Actions,” Journal of Economic Theory, 100(2), 191-219.

Milgrom, P. R. (1981). “Good News and Bad News: Representation theorems and applications,” Bell Journal of Economics (1981): 380-391.

Mijović-Prelec, D. and D. Prelec (2010). “Self-deception as Self-Signalling: a Model and Experimental Evidence,” Philosophical Transactions of the Royal Society B: Biological Sciences 365.1538: 227-240.

Möbius, M. M., M. Niederle, P. Niehaus, and T. S. Rosenblat (2011). “Managing Self-Confidence: Theory and Experimental Evidence,” Mimeo., Iowa State University, Stanford University, and UC San Diego.

Montier, J. (2007). “The Limits to Learning,” In: Behavioural Investing: A Practitioner’s Guide to Applying Behavioural Finance, J. Montier, Wiley Finance.

Moore, D. A., P. E. Tetlock, L. Tanlu, and M. H. Bazerman (2006). “Conflicts of Interest and the Case of Auditor Independence: Moral seduction and Strategic Issue Cycling,” Academy of Management Review 31(1), 10-29.

Oswald, M. E. and S. Grosjean (2004). “Confirmation Bias,” In: R. F. Pohl. Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory. Hove, UK: Psychology Press.

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

Rabin, M. (1994). “Cognitive Dissonance and Social Change,” Journal of Economic Behavior and Organization, 23, 177-194.

Rabin, M. and J. Schrag (1999), “First Impressions Matter: A Model of Confirmatory Bias,” Quarterly Journal of Economics 114(1), 37-82.

Rothschild, M. (1974). “A Two-Armed Bandit Theory of Market Pricing,” Journal of Economic Theory 9(2): 185-202.

Sanbonmatsu, D.M., S. S. Shavitt, S.J. Sherman, and D.R. Roskos-Ewoldsen (1987). “Illusory correlation in the perception of performance by self or a salient other,” Journal of Experimental Social Psychology, 23 (6), 518–543.

Schelling, T. (1985). “The Mind as a Consuming Organ,” In: J. Elster, ed. The Multiple Self. New York: Cambridge University Press.

Schwartzstein, J. (2014). “Selective Attention and Learning,” Journal of the European Economic Association, 12(6), 1423–1452.

Sedikides, C., Green, J. D., and Pinter, B. T. (2004). “Self-protective memory,” In: D. Beike, J. Lampinen, and D. Behrend, eds., The self and memory. Philadelphia: Psychology Press.

Smith, J. (2009a), “Imperfect Memory and the Preference for Increasing Payments,” Journal of Institutional and Theoretical Economics, 2009, 165(4): 684-700.

Smith, J. (2009b), “Cognitive Dissonance and the Overtaking Anomaly: Psychology in the Principal-Agent Relationship,” Journal of Socio-Economics, 2009, 38(4): 684-690.

Taylor, S. E. and P. M. Gollwitzer (1995). “Effects of Mindset on Positive Illusions,” Journal of Personality and Social Psychology, 69(2), 213-226.

Wilson, A. (2014), “Bounded Memory and Biases in Information Processing,” Econometrica, 82 (6), 2257-2294.

Wiswall, M. and Zafar, B. (2011). “Belief Updating among College Students: Evidence from Experimental Variation in Information,” Federal Reserve Bank of New York Staff Reports no. 516.

Appendix

I. Equivalence of Information Structures

Let $\hat{\sigma}_H\left(h^t\right) \in \{L, H\}$ and $\hat{\sigma}_L\left(h^t\right) \in \{L, H\}$ denote the interpretations associated with a high and a low signal after history h^t. In this appendix, we will establish that, up to a relabeling of interpretations $\hat{\sigma}_t$, there is no loss of generality in assuming that the individual always assigns a high interpretation to a high signal.

Because recollections have no intrinsic meaning, for any separating equilibrium (i.e., an equilibrium in which $\hat{\sigma}_H\left(h^t\right) \neq \hat{\sigma}_L\left(h^t\right)$), there exists an equivalent equilibrium that associates the opposite message to each signal. Moreover, for any pooling equilibrium (i.e., an equilibrium in which $\hat{\sigma}_H\left(h^t\right) = \hat{\sigma}_L\left(h^t\right)$), there exists an an equivalent equilibrium that associates the other message to both signals.

In order to deal with this uninteresting multiplicity, I will adopt the following *relabeling conditions*. Whenever we have a separating equilibrium, I will allocate each signal to its own interpretation:

$$\hat{\sigma}_H^*\left(h^t\right) \neq \hat{\sigma}_L^*\left(h^t\right) \implies \hat{\sigma}_H^*\left(h^t\right) = H \text{ and } \hat{\sigma}_L^*\left(h^t\right) = L.$$

Moreover, whenever we have a pooling equilibrium, I will allocate the high recollection to both signals:

$$\hat{\sigma}_H^*\left(h^t\right) = \hat{\sigma}_L^*\left(h^t\right) \implies \hat{\sigma}_H^*\left(h^t\right) = \hat{\sigma}_L^*\left(h^t\right) = H.$$

For notational clarity, I will refer to the games associated with the information structures from Figures 1 and 2 as Game 1 and Game 2. We start with the formal definition of a PBE of Game 2:

Definition 3 *A PBE of the game is a strategy profile $(\hat{\sigma}^*, a^*) : \mathcal{H} \to \{L, H\} \times A$ and posterior beliefs μ such that, for all $h^t \in \mathcal{H}$,*

1. $\hat{\sigma}^*\left(h^t\right) \in \arg\max_{S\in\{L,H\}} \sum_{\tau=t}^{\infty} (1-\eta)^{\tau-t} E_{(\hat{\sigma}^*,\mu)} \left\{ \alpha (1-\eta) E_\mu \left[\theta | h^\tau\right] + \eta V\left(a\left(h^\tau\right), \theta\right) \middle| \left(h^t, S\right) \right\};$
2. $a^*\left(h^t\right) \in \arg\max_{a\in A} \left\{E_\mu\left[V\left(a,\theta\right)|h^t\right]\right\}$; *and*
3. *Posterior beliefs $\mu\left(.|h^t\right)$ are obtained by Bayes' rule if* $\Pr\left(h^t|\hat{\sigma}^*\right) > 0$.

Note that this definition does not require the individual to assign probability 1 to signal $\sigma = L$ upon observing $\hat{\sigma} = L$ when the equilibrium strategy assigns $\hat{\sigma}^* = H$ since the consistency requirement only applies to actions on the equilibrium path. The following proposition, whose proof is presented in the online appendix, states that the PBE of Game 1 satisfying the relabeling conditions are equivalent to the PBE of Game 2:

Proposition 7 *Let $\hat{\sigma}^*$ be an interpretation strategy from Game 2 and let $\hat{\sigma}_L^*\left(h^t\right) = \hat{\sigma}^*\left(h^t\right)$ and $\hat{\sigma}_H^*\left(h^t\right) = H$. Then, $(\hat{\sigma}^*, a^*, \mu)$ is a PBE of Game 2 if and only if $(\hat{\sigma}_L^*, \hat{\sigma}_H^*, \tilde{a}^*, \tilde{\mu})$ is a PBE of Game 1 satisfying the relabeling conditions, for actions $\tilde{a}^*$ and beliefs $\tilde{\mu}$ that coincide with a^* and μ along the equilibrium path.*

It follows directly from the equilibrium definition that a PBE of Game 1 satisfying the relabeling conditions is also a PBE of Game 2. The long but tedious proof of Proposition 7 uses a constructive argument to establish that for any PBE of Game 2, we can find a PBE of Game 1 satisfying the relabeling conditions that coincides with it except for beliefs off the equilibrium path.

II. Anticipatory Utility

In the main text, a state of the world θ represented the individual's skills. Signals were informative about θ and, therefore, provided information about the appropriate action to be taken. In this appendix, I assume instead that the individual has anticipatory utility.[29]

For simplicity, suppose the payoff from actions is

$$V(a,\theta)=\theta-\kappa(\theta-a)^2.$$

As in the standard quadratic loss function, the optimal action equals the mean of the posterior: $a_{k,n}=E[\theta|k,n]$. The additive term θ gives an additional payoff that is independent of actions. The key aspect of this payoff function is that the state of the world θ not only determines the appropriate action; it also provides a payoff independent of the action chosen. For example, suppose θ is the overall state of the economy and a is an investment decision. Then, this payoff function assumes that a high state not only increases the return on new investments, but also increases the return on previously owned assets. (Following the same steps as Theorem 1, it is possible to generalize the results from this appendix for payoff functions satisfying Assumptions 1-3 as well as $\frac{\partial V}{\partial \theta}>0$).

Let μ denote the individual's posterior beliefs about the state of the world θ. When deciding how to interpret a low signal after a history h^t, the current self takes two terms into account. First, she considers the expected payoff from actions $E_\mu\left[V\left(a\left(h^{\tilde{\tau}}\right),y\right)|h^t,L\right]$, where $h^{\tilde{\tau}}$ denotes the (random) history in which the information acquisition process ends. Second, she takes into account the *anticipatory utility* $E_\mu\left[V\left(a\left(h^{\tilde{\tau}}\right),y\right)|h^t,\hat{\sigma}_t\right]$, where $\hat{\sigma}_t$ is her interpretation of the period-t signal. The individual chooses the interpretation $\hat{\sigma}_t\in\{L,H\}$ that maximizes

$$\sum_{s=1}^{\infty}(1-\eta)^s\left\{\alpha E\left[E_\mu\left[V\left(a\left(h^{\tilde{\tau}}\right),y\right)|h^{t+s}\right]|h^t,\hat{\sigma}_t\right]+\eta E_\mu\left[V\left(a\left(h^{\tilde{\tau}}\right),y\right)|h^t,L\right]\right\}, \tag{6}$$

where α captures the relative importance of future anticipatory utility. The PBE definition is analogous to the one from Definition 1, with the substitution of the utility function by (6).

In the model of Section 2, each self balanced the gains from higher self-views with the expected costs of making worse decisions when choosing how to interpret each signal. Then, when the individual was sufficiently confident of which action to take, the self-views effect dominated and she always chose to rationalize low signals away. The anticipatory utility model features a similar trade-off. Each self balances the anticipatory utility gain from believing in a better state of the world with the expected

[29] The concept of anticipatory utility was formally introduced by Loewenstein (1987) and Caplin and Leahy (2001). It is used in the selective awareness models of Benabou (2008, 2011), Benabou and Tirole (2009), and Levy (2012).

cost of making worse decisions. As in the self-views model, when the individual is sufficiently confident of which action to take, the anticipatory utility effect dominates and low signals are rationalized away.

Suppose the individual interprets a low signal realistically after history h^t and consider a deviation to $\hat{\sigma}_t = H$. As shown in Proposition 4, the increase in the expected state is of order $\frac{1}{n}$. If the game ends in a history with $\tilde{k}$ successes in $\tilde{n} > n$ informative periods, the individual will choose action $a_{\tilde{k}+1,\tilde{n}}$ instead of $a_{\tilde{k},\tilde{n}}$. The cost of the deviation is then

$$E\left[V(a_{\tilde{k},\tilde{n}}, \theta) - V(a_{\tilde{k}+1,\tilde{n}}, \theta)|k, n\right] = \kappa \left\{E\left[\theta|\tilde{k}+1, \tilde{n}\right] - E\left[\theta|\tilde{k}, \tilde{n}\right]\right\}^2,$$

which is of order $\frac{1}{\tilde{n}^2}$. Therefore, as in the model of self-views, the individual prefers to rationalize low signals away in *any* PBE after a sufficiently informative history:

Proposition 8 *In the anticipatory utility model, there exists $n_{\kappa,\alpha} \in \mathbb{N}$ such that, in any PBE, any history on the equilibrium path has at most n informative periods. Moreover, $n_{\kappa,\alpha}$ is increasing in κ and decreasing in α.*

III. Persistence of Naiveté

This appendix formalizes the discussion from Section 4 about learning one's time inconsistency. The model embeds Benabou and Tirole's (2004) model of imperfect willpower in Ali's (2011) planner-doer learning framework. A decision maker has an unknown degree of self control θ. His prior beliefs about θ are represented by a thrice continuously differentiable density ρ with full support on $[0, 1]$. There are infinitely many periods, indexed by $n = 1, 2, ...$ In each period, the decision maker must choose between whether or not to expose herself to a tempting activity.

This decision problem is modeled as a game between a long-run player ("planner") and a sequence of identical short-run players with degree of self control θ ("doers"). Each period is composed of three subperiods, indexed by $t = 1, 2, 3$. The timing of each period is as follows:

t=1. The planner chooses between exposing herself to temptation ("risky activity") or avoiding temptation ("safe activity"). Avoiding temptation gives a constant payoff normalized to 1.

t=2. If the planner chose the risky activity at $t = 1$, a doer chooses whether to resist or succumb to temptation:

- Resisting temptation entails a present cost c and an expected future benefit B at the end of the period. The present cost c is drawn from a twice continuously differentiable cumulative distribution distribution G with full support on $[0, \bar{c}]$, where $\bar{c} > B$.
- Succumbing to temptation gives present benefit $b < 1$ and zero future payoff.[30]

[30]Zero is a normalization; the relevant assumption here, which is what makes the activity tempting, is that the future payoff from succumbing to temptation is lower than then future expected payoff from resisting temptation B.

Thus, a doer with self-control parameter θ gets payoff $\theta B - c$ from resisting temptation and b from succumbing to temptation. The planner gets payoff $B - c$ if the doer resists temptation and b if the doer succumbs to temptation.

t=3. If the planner chose the risky activity, she decides how to interpret the doer's action.

In addition to the payoffs from the tempting activity described above, the planner also derives utility from beliefs she has about her self control. Formally, in each period, the planner gets utility $E_{\mu}[u(\theta)]$ from her posterior beliefs μ about θ, where u is a strictly increasing, twice continuously differentiable function. This utility may be due to a direct self-image concern, or due to (unmodeled) anticipatory utility concerns. The planner discounts future payoffs at rate $\beta \in [0,1)$.

If the planner picks the safe activity at t, she does not observe whether the doer would have resisted temptation. The outcome is then $\sigma_t = \text{Safe}$, which is always encoded as $\hat{\sigma}_t = \text{Safe}$. However, if she picks the risky activity at t, the doer either resists temptation ($\sigma_t = H$), or succumbs to temptation ($\sigma_t = L$). She then decides whether to interpret the outcome realistically ($\hat{\sigma}_t = \sigma_t$) or to rationalize it away ($\hat{\sigma}_t \neq \sigma_t$). As before, a period-$t$ history is a vector of recollections up to period t: $h^t = \{\hat{\sigma}_1, \hat{\sigma}_2, ..., \hat{\sigma}_{t-1}\}$.

As in the main text, interpretation strategies can be normalized so that resisting temptation is always interpreted realistically. The interpretation strategy, then, determines whether a lapse $\sigma_t = L$ is interpreted realistically $\hat{\sigma}_t = L$ or rationalized away $\hat{\sigma}_t = H$. Beliefs on the equilibrium path are characterized by the number of times the doer resisted temptation in periods in which the planner would have interpreted a lapse realistically. On the equilibrium path, beliefs remain unchanged in commitment periods. As before, we say a period is informative if its recollection modifies the posterior.

The doer resists temptation if $\theta \geq \frac{b+c}{B}$, which happens with probability

$$\Pr(\sigma_t = H|\theta) = \Pr(\theta B - b \geq c) = G(\theta B - b).$$

Therefore, as in the model in the text, resisting temptation ($\sigma_t = H$) is "good news" about the doer's self control in the sense of first-order stochastic dominance. Conditional on the self control parameter θ, the planner's payoff from the risky activity is

$$\int_0^{\theta B - b} (B - c)\, dG(c) + [1 - G(\theta B - b)]\, b,$$

which is strictly increasing in θ. Since the planner's payoff from the safe activity is constant (1), there exists a threshold level $\bar{\theta}$ below which it is optimal to pick the safe action.

Following the same argument as in Theorem 2, it follows that, in any equilibrium, the planner eventually stops learning:

Proposition 9 (Imperfect Learning about Self Control) *For almost all θ, there exists n_θ such that, in any PBE, almost all histories on the equilibrium path have at most n_θ informative periods.*

The proof of the proposition is on the online appendix. The key idea is that the probability that a deviation affects the optimal experimentation decision converges to zero exponentially whereas the gain from the deviation is $O(1/n)$. Hence, in any path converging to the true parameter, the planner eventually has an incentive to manipulate her recollections into forgetting lapses. Such manipulations, however, prevent her from learning her true self control. Therefore, the planner eventually stops learning about her self control even when she keeps choosing the risky activity.

As in Ali (2011), planners who are convinced to have sufficiently low levels of self control choose to commit and no longer learn. However, even those who do not think they have low self control eventually stop learning. In fact, learning in equilibrium can be very limited. As the following example shows, it is possible to construct equilibria in which the individual frequently succumbs to temptation and nevertheless keeps choosing the risky alternative:

Example 4 *Let (θ, c) be uniformly distributed on $[0,1] \times [0,2]$. Let $B = 4$ and $b = \frac{1}{2}$. The planner's expected payoff from the risky action under the prior distribution is*

$$\frac{\int_0^1 \int_0^{\theta B - b} (\theta B - c)\, dc d\theta + \int_0^1 \int_{\theta B - b}^2 b dc d\theta}{2} = \frac{67}{48} > 1,$$

which exceeds her payoff from the safe action. Moreover, as in the main text, interpreting all signals as high can be supported in equilibrium by setting posteriors equal to the prior. Therefore, there exists a PBE in which the individual always picks the risky action and never updates her beliefs. Moreover, the individual succumbs to temptation with probability $\frac{1}{4}$ in each period under the prior distribution.

As shown in the text, the amount of learning that can be sustained in equilibrium depends on the importance of self views and the cost of succumbing to temptation.

IV. Known Terminal Period

In the main text, we considered a setup in which the information collection process ended randomly. That setup captures an environment in which an individual expects to obtain a sequence of information before making a decision but is unsure about when the decision will have to be made. In this appendix, we consider the polar opposite setup, where information collection ends deterministically after a fixed number of signals. This setup captures situations in which one is certain about when the decision will have to be made and how much information he or she will be able to acquire before then.

Let $T \in \mathbb{N}$ denote the number of signals the individual observes before choosing an action $a \in A$. As in the main text, the individual derives utility from her perceived self image in each period. As in the game with random termination, the individual obtains a payoff from self image in each period and a payoff from actions when the information collection ends. Let $\beta \in [0,1)$ denote the individual's discount factor.[31]

[31]With random termination, the positive probability of termination η ensured that the sum of the discounted expected utility from self views converged. In that setup, any discounting could be introduced through a renormalization of the termination probability. When termination is deterministic, we need to explicitly introduce discounting in order to ensure that the discounted sum of expected utility from self views converges as the number of signals T grows. The results from the model with deterministic termination do not require the discounting of payoff from actions; it is straightforward to generalize them to setups in which payoff from actions are discounted at a rate $\delta \in (0,1]$.

An individual with skill θ and chooses actions $a\left(h^T\right)$ gets expected payoffs

$$E_{(\hat{\sigma}_L,\hat{\sigma}_H,\mu)}\left\{\sum_{t=0}^{T}\beta^t\alpha E_\mu\left(\theta|h^t\right)+\beta^{T+1}V\left(a\left(h^t\right),\theta\right)\right\}. \tag{7}$$

I make the same assumptions on the prior distribution ρ, the conditional probability p, and the payoff from actions V as in Section 2. The equilibrium definition is the same as Definition 3, with the appropriate replacement of the utility function by (7).

When the termination date is deterministic, the individual may become convinced of which action to take before the information collection process ends. At that point, no additional signal would affect the action and, therefore, all future signals would be uninformative. In this case, however, ignoring information is not very interesting as it does not reduce payoffs. In this section, I obtain bounds on the number of informative periods that hold uniformly in the number of signals T. Since they hold for an arbitrarily large number of signals, ignoring information must entail a payoff loss.

Recall the argument behind the incomplete learning results from Theorems 1 and 2. Deviating from a low to a high interpretation raises the posterior mean by a term of order $\frac{1}{n}$. When actions are continuous, distortions close to the optimum have second-order costs. Then, the cost of the deviation has order of magnitude $o\left(\frac{1}{n}\right)$, implying in the existence of a finite bound on the number of informative periods. When actions are discrete, the probability that each signal affects actions decreases exponentially. Then, since the cost from making incorrect decisions is bounded, all signals are also eventually rationalized away. As the proofs of these theorems did not use the assumption that the process ends with a constant probability η, it is immediate to adapt them to the case of a known terminal period:

Corollary 2 *Consider the model with a known terminal period and suppose Assumptions 1, 2, and 3 are satisfied. There exists n such that, in any PBE, almost every history on the equilibrium path has at most n informative periods for any T.*

Corollary 3 *Consider the model with a known terminal period and suppose Assumptions 4 and 5 are satisfied. For almost all θ, there exists n_θ such that, in any PBE, almost all histories on the equilibrium path have at most n_θ informative periods for any T.*

Importantly, the bounds on the number of informative periods in both corollaries (n and n_θ, respectively) are independent of the number of periods T.

In many situations, a person can choose when to stop gathering information and take the action. Since the model predicts that individuals with selective memory disregard information after a certain number of periods, one might expect them to choose to stop gathering information before a Bayesian individual would. In the online appendix, I show that, although this is not true in all equilibria of the model, it is true in Markovian equilibria. Since there always exist equilibria in which the individual ignores any arbitrary number of signals before starting to learn, there may exist equilibria in which the individual with selective memory collects more signals than a Bayesian would.

V. Multiple Signals and Non-Stationary Environments

In the main text, we studied a model with binary signals under arbitrary payoff from actions. The binary signal structure, which is a natural generalization of the framework of Benabou and Tirole (2002, 2004, 2006a, 2006b) and Benabou (2008, 2011), is tractable enough to allow us to obtain results that hold under very general prior distributions, action spaces, and payoff functions. This appendix substantially relaxes the assumption of identically distributed binary signals, but instead focuses on the model with a quadratic loss function.

I establish two main results. First, I generalize the incomplete learning result in the model with continuous actions (Theorem 1). This generalization shows that introducing richer signal structures does not overturn the main conclusion from the model with identically distributed binary signals. Second, I show that models with binary actions and multiple signals can naturally produce a 'belief polarization' phenomenon, where two individuals with different priors update their beliefs in the direction of their priors when given the same information.

For simplicity, I consider the model with a known terminal period (Appendix IV). There are $T+1$ periods. In the last period, the individual chooses an action a from the compact set A. In each of the first T periods, the individual observes the realization of an independent, non-identically distributed signal $\sigma \in \mathcal{S}$, where $\mathcal{S} \equiv \{1, 2, ..., S\}$ denotes the set of possible signals, $S \geq 2$. The probability of signal σ in period t is $p_t(\sigma, \theta) \equiv \Pr(\sigma_t = \sigma | \theta)$.

After observing the signal, the individual chooses how to interpret it. As in Figure 1, each signal can be interpreted as any other signal: $\hat{\sigma}_t \in \{1, ..., S\}$. A history is a vector of interpretations: $h^t = \{\hat{\sigma}_1, \hat{\sigma}_2, ..., \hat{\sigma}_t\}$. Let $\mathcal{H}^t := \{1, 2, ..., S\}^t$ denote the set of period-t histories, and let $\mathcal{H} := \bigcup_{t=1}^{T} \mathcal{H}^t$ denote the set of all possible histories.

An interpretation strategy is a mapping $\hat{\sigma} : \mathcal{H} \to \{1, ..., S\}$, where $\hat{\sigma}(\{\sigma_1, ..., \sigma_t\})$ specifies how signal σ_t is interpreted at history $h^{t-1} = \{\sigma_1, ..., \sigma_{t-1}\}$. An action strategy is a mapping $a : \mathcal{H}^{T+1} \to A$ specifying which action to take when the information process ends. Let $\mu\left(.|h^t\right)$ denote the individual's posterior beliefs about θ given h^t and let $E_\mu\left[.|h^t\right]$ denote the expectation operator with respect to $\mu\left(.|h^t\right)$. The definition of PBE is analogous to the one from Appendix IV, with the replacement of each signal $\sigma \in \{L, H\}$ by $\sigma \in \{1, ..., S\}$. As in Section 3, the payoff from actions is quadratic (4) and actions are continuous $A = \Theta$. Thus, the optimal action equals the posterior mean: $a^*\left(h^t\right) = E_\mu\left[\theta|h^t\right]$.

Let $\mathcal{P}(\mathcal{S})$ denote the power set of $\mathcal{S}$. Note that the prior distribution ρ induces a probability distribution over sets of signals. Let $E_\rho\left[\theta|s^t\right] \equiv E_\rho\left[\theta|h^t \in s^t\right]$ denote the expectation operator associated with it. Let

$$d(t) \equiv \max\left\{\left|E_\rho\left[\theta|s^t s\right] - E_\rho\left[\theta|s^t s'\right]\right| : s^t \in \mathcal{S}^t,\ s, s' \in \mathcal{S}\right\}$$

denote the maximum absolute deviation of posterior means in period t. Notice that d depends only on the prior distribution ρ and the conditional distribution of signals $p_t(\sigma, \theta)$, which are fundamentals of the model.

Definition 4 *The distribution of signals is regular if* $\Pr_\rho\left(\lim_{t\to\infty} d(t) = 0\right) = 1$.

Regularity is a very weak requirement, which states that, with probability one, the effect of a new signal eventually converges to zero as the number of signals goes to infinity. Without regularity, even a Bayesian individual might not be able to learn the true parameter. By Doob's consistency theorem, regularity is always satisfied if signals are identically distributed. More generally, regularity holds as long as the posterior mean follows an active supermartingale (Fudenberg and Levine, 1992). In the uniform binomial model from Section 4, for example, the maximum absolute deviation of posterior means is $d(t) = \frac{1}{t+2}$, which converges to zero everywhere.

Let $h^t h^\tau$ denote the history obtained by concatenating $h^\tau = \{\hat{\sigma}_1^\tau, ..., \hat{\sigma}_{\tau-1}^\tau, \hat{\sigma}_\tau^\tau\}$ after $h^t = \{\hat{\sigma}_1^t, ..., \hat{\sigma}_{t-1}^t, \hat{\sigma}_t^t\}$:

$$h^t h^\tau \equiv \left\{\hat{\sigma}_1^t, ..., \hat{\sigma}_t^t, \hat{\sigma}_1^\tau, ..., \hat{\sigma}_\tau^\tau\right\}. \tag{8}$$

Similarly, let $h^t \hat{\sigma} h^\tau$ denote the successive concatenation of h^t, $\hat{\sigma}$, and h^τ:

$$h^t \hat{\sigma} h^\tau \equiv \left\{\hat{\sigma}_1^t, ..., \hat{\sigma}_t^t, \hat{\sigma}, \hat{\sigma}_1^\tau, ..., \hat{\sigma}_\tau^\tau\right\}. \tag{9}$$

Before presenting the proof of the proposition, I will establish two auxiliary results.

As before, we say that a history h^t is *informative* if $E_\mu\left[\theta|h^t\right] \neq E_\mu\left[\theta|h^t\hat{\sigma}_{t+1}\right]$ for some history $h^t\hat{\sigma}_{t+1}$. We say that h^t is a last history on the equilibrium path if all the continuation histories of h^t on the equilibrium path are uninformative. Formally, h^t is a *last informative history on the equilibrium path* if

- $E\left[\theta|h^t\right] \neq E\left[\theta|h^t\hat{\sigma}_{t+1}\right]$ for some history $h^t\hat{\sigma}_{t+1}$ on the equilibrium path; and
- For all continuation histories of h^t on the equilibrium path ($h^t\hat{\sigma}_{t+1}h^s \supset h^t$), $E\left[\theta|h^t\hat{\sigma}_{t+1}\right] = E\left[\theta|h^t\hat{\sigma}_{t+1}h^s\right]$.

It is straightforward to show that every history on the equilibrium path has a continuation history that is a last informative history on the equilibrium path (see Lemma 10).

The following lemma shows that posterior beliefs in any last informative history on the equilibrium path cannot be "too concentrated":

Lemma 3 *Let $(\hat{\sigma}^*, a^*, \mu^*)$ be a PBE. Let h^t be a last informative history on the equilibrium path. Then, $\left|E_{\mu^*}\left[\theta|h^t\hat{\sigma}_{t+1}\right] - E_{\mu^*}\left[\theta|h^t\hat{s}_{t+1}\right]\right| > \frac{\alpha}{\beta\kappa}$ for all histories $h^t\hat{\sigma}_{t+1}$ and $h^t\hat{s}_{t+1}$ on the equilibrium path.*

Let h^t be a last informative history on the equilibrium path and let n be its number of informative periods. By the previous lemma, $d(n) \geq \frac{\alpha}{\beta\kappa}$.

Let $n^*_{\alpha,\beta,\kappa} \equiv \sup\left\{n \in \mathbb{N} : d(n) \geq \frac{\alpha}{\beta\kappa}\right\}$. Assuming that the distribution of signals is regular, $d(n)$ converges to zero. Therefore, for any $\epsilon > 0$ there exists $\bar{n}_\epsilon$ such that $n > \bar{n}_\epsilon$ implies $d(n) < \epsilon$. Taking $\epsilon = \frac{\alpha}{\beta\kappa}$ establishes that $d(n) < \frac{\alpha}{\beta\kappa}$. Setting $n^*_{\alpha,\beta,\kappa} = \bar{n}_{\frac{\alpha}{\beta\kappa}}$, therefore, establishes that $n^*_{\alpha,\beta,\kappa}$ is finite. Thus, any last informative history on the equilibrium path has at most $n^*_{\alpha,\beta,\kappa}$ informative periods.

Since every informative history on the equilibrium path has a continuation history that is a last informative history on the equilibrium path, it follows that any such history has at most $n^*_{\alpha,\beta,\kappa}$ informative periods. Moreover, any terminal history has at most $n^*_{\alpha,\beta,\kappa} + 1$ informative periods. We have, therefore, established the following result:

Proposition 10 *Suppose the distribution of signals is regular. There exists $n_{\alpha,\beta,\kappa} \in \mathbb{N}$ such that, in any PBE, every history on the equilibrium path has at most $n_{\alpha,\beta,\kappa}$ informative periods.*

Importantly, the maximum number of informative periods $n_{\alpha,\beta,\kappa}$ is not a function of the number of signals T. Therefore, beliefs do not converge as the number of signals grows.

Next, we turn to an example of how belief polarization may naturally occur when actions are binary and there are more than two signals. In their classic work, Lord, Ross, and Lepper (1979) describe the phenomenon as follows:[32]

> Our thesis is that belief polarization will increase, rather than decrease or remain unchanged, when mixed or inconclusive findings are assimilated by proponents of opposite viewpoints. This "polarization hypothesis" can be derived from the simple assumption that data relevant to a belief are not processed impartially. Instead, judgments about the validity, reliability, relevance, and sometimes even the meaning of proffered evidence are biased by the apparent consistency of that evidence with the perceiver's theories and expectations.

For simplicity, consider a one-signal model ($T = 1$). Actions are binary $A = \{0, 1\}$. There are two individuals, A and B. They both have the same payoff from skills and the same quadratic loss function. The two individuals have different prior distributions:

$$\rho_A(\theta) = (1+\delta)\,\theta^{\delta}, \quad \rho_B(\theta) = (1+\delta)\,(1-\theta)^{\delta},$$

where $\delta > 0$ parametrizes the strength of each individual's beliefs.[33] Notice that A has a higher prior than B in the sense of first-order stochastic dominance.

Recall that the optimal action is $a = 1$ if $E[\theta] \geq \frac{1}{2}$ and $a = 0$ if $E[\theta] \leq \frac{1}{2}$. The means of the individuals' prior beliefs are:

$$E_A(\theta) = \frac{1+\delta}{2+\delta} > \frac{1}{2}, \quad E_B(\theta) = \frac{1}{2+\delta} < \frac{1}{2}.$$

Therefore, person A initially favors the high action ($a = 1$) whereas person B initially favors the low action ($a = 0$).

There are three possible signals, H, M, and L, with conditional probabilities

$$\Pr(\sigma = s|\theta) = \begin{cases} \gamma\theta \text{ if } s = H \\ 1-\gamma \text{ if } s = M \\ \gamma(1-\theta) \text{ if } s = L \end{cases},$$

where $\gamma \in (0, 1)$. In this model, a high signal ($s = H$) is good news about θ, a low signal ($s = L$) is bad news about θ, and a medium signal ($s = M$) is neutral in that it does not affect the posterior distribution of θ. While this particular distribution simplifies the calculations, it can be substantially

[32] See Rabin and Schrag (1999) for a detailed description of the literature.

[33] Notice that, if δ is an integer, ρ_A corresponds to the posterior distribution of someone with a uniform prior conditional on δ high signals. Symmetrically, ρ_B corresponds to the posterior distribution of someone with a uniform prior conditional on δ low signals.

generalized. The key assumption is that distribution of signals satisfies the monotone likelihood ratio property: a high signal is a strong indicator of the appropriateness of a high action; a low signal is a strong indicator of the appropriateness of a low action; whereas a medium signal is a weak indicator of which action to take.

As in the main model, there always exists an equilibrium in which the signal is uninformative (Proposition 1). Here, I will study equilibria in which the signal is informative. Because person A chooses a high action if she observes either a high or a medium signal, she cannot commit not to reinterpret a medium signal as high. Hence, there is no equilibrium in which person A interprets high and medium signals differently: the only possible equilibrium in which the signal is informative is the one in which medium and high signals are "pooled" and the low signal "separates." In this equilibrium, person A updates her beliefs upwards after a medium signal. Formally, her posterior distribution after observing a medium signal is

$$\rho_A(\theta|\sigma \in \{M, H\}) = \frac{(1+\delta)\,\theta^{1+\delta} + (1+\delta)\,\theta^{\delta}}{\int (1+\delta)\,\theta^{1+\delta} + (1+\delta)\,\theta^{\delta} d\theta},$$

which dominates the prior distribution $\rho_A(\theta)$ in terms of the monotone likelihood ratio property (and, therefore, in terms of first-order stochastic dominance).

Conversely, person B chooses a low action if she observes either a low or a medium signal and, therefore, cannot commit not to reinterpret a low signal as medium. Hence, the only possible equilibrium for person B in which the signal is informative is the one in which low and medium signals are pooled and the high signal separates. In this equilibrium, person B revises her beliefs downwards after a medium signal since her posterior distribution after observing a medium signal is

$$\rho_B(\theta|\sigma \in \{M, H\}) = \frac{(1+\delta)(1-\theta)^{1+\delta} + (1+\delta)(1-\theta)^{\delta}}{\int (1+\delta)(1-\theta)^{1+\delta} + (1+\delta)(1-\theta)^{\delta} d\theta},$$

which is dominated by the prior distribution $\rho_B(\theta)$ in terms of the monotone likelihood ratio property.

Thus, in the equilibria in which the signal is informative, the intermediate signal induces individuals A and B to update their beliefs in opposite directions. Person A, whose prior distribution dominated by Person B's prior, revises her beliefs upwards. Person B, on the other hand, revises her beliefs downwards after observing the intermediate signal.

As in Section 2, the equilibria in which the signal is informative exist if and only if the importance of self views is small enough or the importance of actions is large enough. As I show in the online appendix, the informative equilibrium for person A exists if and only if

$$\frac{\alpha}{\kappa\beta} \leq \frac{1+\delta}{3+\delta} \times \frac{1}{2-\gamma+\delta}.$$

Person B, the informative equilibrium exists if and only if

$$\frac{\alpha}{\beta\kappa} \leq \frac{\left(\frac{3+\delta}{2+\delta}\right)^2 - \left(\frac{1+\delta+2\gamma}{2+\delta-\gamma} - \frac{1+\delta}{2+\delta}\right)^2}{(3+\delta)\left(2 - \frac{1+\delta+2\gamma}{2+\delta-\gamma}\right)}.$$

More generally, in any equilibrium, the individual cannot interpret two signals differently if they are associated with the same action. Then, for individuals with a low prior distribution, intermediate signals do not affect the optimal action and must, therefore, be pooled with low signals. Those with a high prior pool intermediate signals with high signals. Consequently, individuals with a low prior revise their beliefs downwards after an intermediate signal whereas those with a high prior revise their beliefs upwards.

VI. Behavioral Strategies

So far, we have focused on pure strategy equilibria. This appendix considers strategies in which the individual randomizes. Since this is a game with imperfect recall, mixed strategies are no longer equivalent to behavioral strategies (i.e., Kuhn's Theorem does not hold). A mixed strategy randomizes over pure strategies only at the outset: the individual randomizes over pure strategies before the game starts and follows the realized pure strategy throughout the game. As a result, in this game, any pure strategy in the support of a mixed strategy equilibrium must be a pure strategy equilibrium. Moreover, no mixed strategy equilibrium can dominate all pure strategy equilibria.

With behavioral strategies, the randomization occurs at each history. As Piccione and Rubinstein (1997) show, in decision problems with imperfect recall, behavioral strategy equilibria may dominate all pure strategy equilibria. Accordingly, this appendix studies equilibria in behavioral strategies and establishes two main results. First, it extends the conservatism result from Proposition 3 to behavioral strategy equilibria, allowing for a smoother, more realistic version of conservatism in which updates are partial (in pure strategies, updates are either fully Bayesian or completely uninformative). Second, it establishes a uniform bound on learning for the model with a quadratic payoff from actions. Hence, allowing for behavioral strategies does not overturn the incomplete learning result from the text.

When we allow for behavioral strategies, the equivalence between the decision problem from Figure 1 and the Benabou and Tirole model (Figure 2) no longer holds. When the individual can play a strictly mixed behavioral strategy after observing a high signal, we can no longer relabel the recollection chosen after $\sigma = H$ as $\hat{\sigma} = H$. In this appendix, I consider the Benabou and Tirole model. The same arguments, however, can be applied to the decision problem from Figure 1. For simplicity, I assume a deterministic number of signals T as in Appendix IV.

As in the rest of the paper, the prior density function ρ is a thrice continuously differentiable with full support on $\Theta = [\underline{\theta}, \bar{\theta}]$ and the conditional probability of a high signal $p : \Theta \to (0,1)$ is a strictly increasing and twice continuously differentiable function. A period-t history is a vector of recollections up to period t: $h^t = \{\hat{\sigma}_1, \hat{\sigma}_2, ..., \hat{\sigma}_{t-1}\}$. Let $\mathcal{H}^t := \{L, H\}^{t-1}$ denote the set of period-t histories, and let $\mathcal{H} := \bigcup_{t=1}^{T+1} \mathcal{H}^t$ denote the set of all possible histories.

A behavioral strategy is a pair of mappings $\lambda : \mathcal{H} \to [0,1]$ and $\gamma : \mathcal{H}^{T+1} \to \Delta(\mathbb{R})$, where $\lambda(h^t)$ assigns the probability of playing $\hat{\sigma} = H$ after observing a low signal given history h^t, and $\gamma(h^{T+1})$ assigns the probability of playing each action $a \in \mathbb{R}$ at terminal history h^{T+1}. As usual, the utility function is extended to random interpretations and actions by taking expectations. As in Definition 3, a PBE is a triple (λ, γ, μ) in which interpretation and action strategies maximize the expected payoffs of each self and beliefs satisfy Bayes' rule on histories on the equilibrium path.

Before imposing additional assumptions, we generalize Proposition 3 to equilibria in behavioral strategies. As in the main text, let θ_t^B denote the expected value of θ obtained by Bayes' rule conditional on the sequence of observed signals $\{\sigma_1, ..., \sigma_{t-1}\}$, and let $\hat{\theta}_t$ denote the expected value of θ calculated according to the individual's beliefs μ (i.e., conditional on the sequence of the individual's interpretations). The following proposition, whose proof is in the online appendix, generalizes the conservatism bias for equilibria in behavioral strategies:

Proposition 11 (Conservatism) *Let λ be an interpretation strategy and let μ be posterior beliefs consistent with this strategy. For any history h^t such that* $\Pr\left(h^t|\lambda\right) > 0$,

$$\theta_t^B > \theta_{t-1}^B \implies \hat{\theta}_t \geq \hat{\theta}_{t-1}, \text{ and}$$
$$\theta_t^B < \theta_{t-1}^B \implies \hat{\theta}_t \leq \hat{\theta}_{t-1}.$$

Furthermore, $\hat{\theta}_t$ second-order stochastically dominates θ_t^B.

Next, we turn to the result on the failure of learning. Consider the continuous action model with quadratic payoffs:

$$V(a, \theta) = -\kappa (a - \theta)^2,$$

where $a \in (\underline{\theta}, \bar{\theta})$ and $p(\theta) = \theta$. Let $\gamma\left(h^t\right) = \delta_a$ denote the strategy that assigns probability one to action a. Under quadratic losses, the optimal action corresponds to the mean of the posterior: $\gamma^*\left(h^t\right) = \delta_{E[\theta|h^t]}$. Thus, the action strategy in any PBE is a pure strategy.

With behavioral strategies, however, the Bayesian updates depend not only on whether the signal is informative but also on the specific mixing probability in each history. Thus, we cannot use the asymptotic expansions derived in the text. Nevertheless, with quadratic loss functions, we do not need to derive the specific asymptotic expansions to show that learning is incomplete. With quadratic loss functions, the benefit in self-image is proportional to difference in the posterior mean of θ, whereas the cost of distorting the action is proportional to the square of the difference. Then, if beliefs are sufficiently concentrated, the individual cannot commit not to misinterpret a low signal as high.

Note that, for a fixed prior distribution ρ, any interpretation strategy λ determines the prior probability of reaching each terminal history. Formally, the interpretation strategy λ induces a probability distribution $\mathbb{P}_\lambda$ over the power set of $\mathcal{H}^{T+1}$.

Proposition 12 *For each $T \in \mathbb{N}$, let $(\lambda_T^*, \gamma_T^*, \mu_T^*)$ be a PBE of the game with T signals. Then,*

$$\mathbb{P}_{\lambda_T^*}\left(\lim_{T\to\infty} E_\mu\left[\theta|h^{T+1}\right] = \theta\right) = 0.$$

VII. Proofs

Proof of Proposition 1. For all histories h^t, let $\hat{\sigma}_t^*\left(h^t\right) = H$, $a^*\left(h^t\right) \in \arg\max_a \int_\Theta V(a, \theta)\, d\theta$, and $\mu\left(\theta|h^t\right) = \rho(\theta)$. Since the interpretation strategy does not affect beliefs and actions, it satisfies

Condition 1 from Definition 1. By construction, $a^*\left(h^t\right)$ satisfies Condition 2. Moreover, because there all signals are rationalized as $\hat{\sigma} = H$, consistency requires the posterior distribution to be equal to the prior distribution on the equilibrium path. Hence, Condition 3 holds. ■

Proof of Proposition 2. Before presenting the proof, it is helpful to introduce some notation. Given an action strategy $a : \mathcal{H} \to A$ and posterior beliefs μ, let $\tilde{V}\left(a; h^t, \hat{h}^t\right)$ denote the expected payoff from actions conditional on history h^t when the individual follows the actions prescribed by a in the continuation histories following history $\hat{h}^t$:

$$\tilde{V}\left(a; h^t, \hat{h}^t\right) = \eta \sum_{s=0}^{\infty} (1-\eta)^s E_\mu \left[V\left(a\left(\hat{h}^{t+s}\right), \theta\right) |h^t\right].$$

Let $(\hat{\sigma}^*, a^*, \mu^*)$ be a PBE for the game with parameter α_1 and let $\alpha_0 < \alpha_1$. I will construct an equilibrium of the game with parameter α_0 that coincides with $(\hat{\sigma}^*, a^*, \mu^*)$ for all histories h^t in which $\Pr\left(h^t|\hat{\sigma}^*\right) > 0$.

For all histories on the equilibrium path (under $\hat{\sigma}^*$), pick the same strategies and beliefs as in the original equilibrium. For any history h^t off the equilibrium path (under $\hat{\sigma}^*$), set $\mu_0^*\left(\theta|h^t\right) = \mu_0^*\left(\theta|h^{t-1}\right)$ and choose a_0^* consistent with these beliefs. Then, the individual is indifferent between any interpretation at h^t. In particular, we can pick the same interpretation as the original one $\hat{\sigma}_0^*\left(h^t\right) = \hat{\sigma}^*\left(h^t\right)$. We will show that this is a PBE of the game with parameter α_0.

Let h^t be a history on the equilibrium path and let k and n denote the number of high recollections and informative periods in it. First, we show that we can sustain the same interpretation strategies σ^* for all histories on the equilibrium path. Then, since μ^* is uniquely determined by Bayes' rule given σ^* for histories on the equilibrium path (and it does not depend on α) and $a^*\left(h^t\right) \in \arg\max_a E_\mu\left[V(a,\theta)|h^t\right]$ (which also does not depend on α), the same beliefs μ^* and action strategy a^* can also be sustained along the equilibrium path.

If $\hat{\sigma}^*\left(h^t\right) = L$, Bayesian updating requires beliefs to be updated to the distribution of θ conditional on k high recollections in $n+1$ informative periods. There is no profitable deviation from picking $\hat{\sigma}_{t+1} = L$ if the expected discounted benefit from self image,

$$\alpha \sum_{s=0}^{\infty} (1-\eta)^s \left\{E\left[E_\rho\left[\theta|h^t, H, \hat{\sigma}_{t+1}, \hat{\sigma}_{t+2}, ..., \hat{\sigma}_{t+s}\right] - E_\rho\left[\theta|h^t, L, \hat{\sigma}_{t+1}, \hat{\sigma}_{t+2}, ..., \hat{\sigma}_{t+s}\right]|h^t, L\right]\right\}. \tag{10}$$

does not exceed its expected cost, $\tilde{V}_\mu\left(a^*; \left(h^t, L\right), \left(h^t, L\right)\right) - \tilde{V}_\mu\left(a^*; \left(h^t, L\right), \left(h^t, H\right)\right)$. All histories in the summation above are on the equilibrium path, since they are obtained by following either (h^t, L) or (h^t, H) along the equilibrium interpretation strategy $\hat{\sigma}^*$. Since, by assumption, h^t is on the equilibrium path and $\hat{\sigma}^*\left(h^t\right) = L$, it follows that both (h^t, L) and (h^t, H) are also on the equilibrium path. Because $\sigma_t = H$ is good news and $\sigma_t = L$ is bad news about θ, each term in the summation in (10) is positive. Hence, the benefit from the deviation is increasing in α, while the cost is not a function of it. Since this condition is satisfied for α_1, it must be satisfied for any $\alpha_0 < \alpha_1$.

If $\hat{\sigma}^*\left(h^t\right) = H$, Bayesian updating requires beliefs to remain unchanged after $\hat{\sigma}_{t+1} = H$: $\mu_0\left(\theta|h^t, H\right) = \mu_0\left(\theta|h^t\right)$. Moreover, history (h^t, L) is off the equilibrium path. By construction, $\mu_0^*\left(\theta|h^t, L\right) = \mu_0^*\left(\theta|h^t\right)$

(i.e., beliefs are also unresponsive to $\hat{\sigma}_{t+1} = L$). Thus, $\hat{\sigma}^*\left(h^t\right) = H$ is (weakly) optimal for any α. Furthermore, $a_0^*\left(h^t, H\right)$ is consistent with μ_0^* (by construction).

Now, let h^t be a history off the equilibrium path. Since beliefs are unresponsive, each self is indifferent between all possible interpretation actions. Hence, any $\hat{\sigma}^*(h^t)$ is a best response. By construction, $a_0^*(h^t)$ is consistent with beliefs μ_0^*. Moreover, Bayes' rule does not constrain beliefs μ_0^*. Thus, we have constructed an equilibrium $(\hat{\sigma}_0^*, a_0^*, \mu_0^*)$ of the game with parameter α_0 that coincides with $(\hat{\sigma}^*, a^*, \mu^*)$ on the equilibrium path.

Next, we establish the proofs of the numbered claims:

1. Suppose that there is a PBE in which at least one signal is interpreted correctly, and let t be the first period where this happens. That is, the only period-t history on the equilibrium path is $h^t = (H, H, ..., H)$ and $\hat{\sigma}^*\left(h^t\right) = L$. Consider a deviation to $\hat{\sigma}_{t+1} = H$ in period t. The benefit from this deviation in terms of self image is bounded below by the benefit in the period t:

$$\alpha \left\{E\left[\theta|1,1\right] - E\left[\theta|0,1\right]\right\} > 0.$$

The loss from the deviation is bounded above by

$$\max_{a\in A} E_\rho\left[V\left(a,\theta\right)\right] - \min_{a\in A} E_\rho\left[V\left(a,\theta\right)\right] \geq 0,$$

which exists (since $E\left[V\left(a,\theta\right)\right]$ is a continuous function of a and A is compact). Letting

$$\bar{\alpha} := \frac{\max_{a\in A} E_\rho\left[V\left(a,\theta\right)\right] - \min_{a\in A} E_\rho\left[V\left(a,\theta\right)\right]}{E\left[\theta|1,1\right] - E\left[\theta|0,1\right]} \geq 0$$

concludes the proof.

2. Fix an equilibrium in which all low signals until period $N-1$ are interpreted as low, and all low signals in periods $t \geq N$ are interpreted as high. Consider a deviation at a history h^t to $\hat{\sigma}_t = H$, where $t < N$. Let k denote the number of successes in history h^t (the number of informative periods is $t-1$ since all previous periods were informative). The gain from the deviation is bounded above by

$$\sum_{\tau=t}^{\infty} (1-\eta)^{\tau-t}\, \alpha \left\{E\left[\theta|k+1,t-1\right] - E\left[\theta|k,t-1\right]\right\} = \frac{\alpha}{\eta}\left\{E\left[\theta|k+1,t-1\right] - E\left[\theta|k,t-1\right]\right\}.$$

(This bound is tight if the individual stops learning in the following period. Otherwise, it is not tight because the expectations follow a supermartingale after a deviation: a self who deviates expects posterior means to decrease, on average, in all informative periods). For each (k,n), pick

$$a_{k,n} \in \arg\max_a E\left[V\left(a,\theta\right)|k,n\right] \setminus \arg\max_a E\left[V\left(a,\theta\right)|k-1,n\right].$$

Because $a_{k,n} \notin \arg\max_a E\left[V\left(a,\theta\right)|k-1,n\right]$,

$$v_{k-1,n} := E\left[V\left(a_{k-1,n},\theta\right) - V\left(a_{k,n},\theta\right)|k-1,n\right] > 0.$$

Let $\underline{v}_n = \min\{v_{k,n} : k \in \{0, 1, ..., n\}\}$, which exists and is strictly positive because it is the minimum of a finite set composed of strictly positive elements. The expected cost from the deviation equals

$$\sum_{\tau=t+1}^{N} (1-\eta)^{\tau-t-1} \eta \underline{v}_\tau + \sum_{\tau=N+1}^{\infty} (1-\eta)^{\tau-t} \eta \underline{v}_N$$

$$= (1-\eta)^{N+1-t} \underline{v}_N + \sum_{\tau=t+1}^{N} (1-\eta)^{\tau-t-1} \eta \underline{v}_\tau.$$

The deviation is not profitable as long as

$$\frac{\alpha}{\eta} \{E[\theta|k+1, t-1] - E[\theta|k, t-1]\} \leq (1-\eta)^{N+1-t} \underline{v}_N + \sum_{\tau=t+1}^{N} (1-\eta)^{\tau-t-1} \eta \underline{v}_\tau$$

The left hand side is strictly positive for all $t < N$. Therefore, setting

$$\underline{\alpha} = \frac{(1-\eta)^{N+1-t} \underline{v}_N + \sum_{\tau=t}^{N} (1-\eta)^{\tau-t} \eta \underline{v}_t}{E[\theta|k+1, t-1] - E[\theta|k, t-1]} > 0$$

concludes the proof. ■

Proof of Proposition 3. The proposition will be established by a sequence of lemmata.

Claim. $\sigma_t = H$ implies $\theta_t^B > \theta_{t-1}^B$ and $\sigma_t = L$ implies $\theta_t^B < \theta_{t-1}^B$.

Proof. Let f denote the posterior density conditional on the vector of true signals up to period $t-1$. After a high signal at t, the posterior distribution conditional of the sequence of true signals changes from $\frac{f(\theta)}{\int f(\theta)d\theta}$ to $\frac{p(\theta)f(\theta)}{\int p(\theta)f(\theta)d\theta}$. The posterior at t (i.e., after the high signal) dominates the distribution at $t-1$ (i.e., before that signal) in the sense of monotone likelihood ratio property if, for any $x > y$,

$$\frac{\frac{p(x)f(x)}{\int p(x)f(x)dx}}{\frac{p(y)f(y)}{\int p(y)f(y)dy}} > \frac{\frac{f(x)}{\int f(x)dx}}{\frac{f(y)}{\int f(y)dy}}.$$

Rearranging this expression, we obtain $p(x) > p(y)$, which is true since p is strictly increasing. Then, it follows from Milgrom (1981, Proposition 2) that

$$\theta_t^B = \frac{\int \theta p(\theta) f(\theta) d\theta}{\int p(\theta) f(\theta) d\theta} > \frac{\int \theta f(\theta) d\theta}{\int f(\theta) d\theta} = \theta_{t-1}^B$$

when $\sigma_t = H$. By the exact same argument, $\theta_{t-1}^B > \theta_t^B$ when $\sigma_t = L$. ■

Claim. $\sigma_t = H$ implies $\hat{\theta}_t \geq \hat{\theta}_{t-1}$ and $\sigma_t = L$ implies $\hat{\theta}_t \leq \hat{\theta}_{t-1}$.

Proof. Let g denote the individual's posterior distribution conditional on the recollections up to $t-1$: $\hat{\sigma}_1, ..., \hat{\sigma}_{t-1}$. The individual may observe a low or a high signal, and she may interpret low signals realistically or not.

First, suppose she interprets low signals realistically. If she observes a low signal, the posterior mean changes from $\hat{\theta}_{t-1} = \frac{\int \theta g(\theta) d\theta}{\int g(\theta) d\theta}$ to $\hat{\theta}_t = \frac{\int \theta [1-p(\theta)] g(\theta) d\theta}{\int [1-p(\theta)] g(\theta) d\theta}$. Again, using the monotone likelihood ratio property, establishes that $\hat{\theta}_t < \hat{\theta}_{t-1}$. Analogously, if she observes a high signal, the posterior mean increases from $\hat{\theta}_{t-1} = \frac{\int \theta g(\theta) d\theta}{\int g(\theta) d\theta}$ to $\hat{\theta}_t = \frac{\int \theta p(\theta) g(\theta) d\theta}{\int p(\theta) g(\theta) d\theta} > \hat{\theta}_{t-1}$. Now, suppose she interprets low signals as high. Then, her posterior remains equal to the prior distribution regardless of the observed signal: $\hat{\theta}_{t-1} = \hat{\theta}_t$. ■

Recall that θ_t^B second-order stochastically dominates $\hat{\theta}_t$ if and only if we can express θ_t^B as a mean-preserving spread of $\hat{\theta}_t$. Posteriors remain unchanged in all histories in which low signals are interpreted as high. Let $s^t \equiv \{\sigma_1, ..., \sigma_{t-1}\}$ denote a vector of true signals up to period t. Let the space of all possible vectors of signals and recollections $\Omega \equiv \bigcup_{t=1}^{\infty} \{\mathcal{H}^t \times \mathcal{H}^t\}$ denote the sample space. An element (s^t, h^t) of Ω is, therefore, a vector of histories and recollections with the same length.

Construct the vector $\phi(s^t) \equiv \{\phi_1(s^t), \phi_2(s^t), ..., \phi_{t-1}(s^t)\}$ inductively:

$$\phi_1(s^t) = \begin{cases} H \text{ if } \hat{\sigma}^* (\emptyset) = H \\ \sigma_1 \text{ if } \hat{\sigma}^* (\emptyset) = L \end{cases},$$

$$\phi_2(s^t) = \begin{cases} H \text{ if } \hat{\sigma}^* \left(\left\{\phi_1 \left(s^t\right)\right\}\right) = H \\ \sigma_2 \text{ if } \hat{\sigma}^* \left(\left\{\phi_1 \left(s^t\right)\right\}\right) = L \end{cases},$$

$$...$$

$$\phi_{t-1}(s^t) = \begin{cases} H \text{ if } \hat{\sigma}^* \left(\left\{\phi_1 \left(s^t\right), ..., \phi_{t-2} \left(s^t\right)\right\}\right) = H \\ \sigma_{t-1} \text{ if } \hat{\sigma}^* \left(\left\{\phi_1 \left(s^t\right), ..., \phi_{t-2} \left(s^t\right)\right\}\right) = L \end{cases}.$$

The vector $\phi(s^t)$ is the history of recollections in period t when the individual observed the vector of true signals s^t. Hence, the vector of true signals s^t is a filtration of the vector of recollections $\phi\left(s^t\right) = h^t$.

Let

$$\epsilon\left(s^t\right) \equiv \underbrace{E\left[\theta | s^t\right]}_{\theta_t^B} - \underbrace{E\left[\theta | \phi\left(s^t\right)\right]}_{\hat{\theta}_t}.$$

By the law of iterated expectations and the fact that s^t is a filtration of h^t, we have

$$E\left[\epsilon\left(s^t\right) | h^t\right] = E\left[E\left[\theta | s^t\right] | h^t\right] - E\left[\theta | h^t\right] = E\left[\theta | h^t\right] - E\left[\theta | h^t\right] = 0.$$

Thus, $E\left[\theta | s^t\right] = \theta_t^B$ is a mean-preserving spread of $\hat{\theta}_t$. ■

Proof of Lemma 2. Let $\{x_n\} =: \mathbf{s}$ be a sequence of independent Bernoulli random variables with parameter $p(\theta)$. Each random variable is distributed according to the probability mass function $f(x|\theta) = p(\theta)^x [1 - p(\theta)]^{1-x}$, $x \in \{0, 1\}$. From Johnson (1970, Theorem 3.1)[34], there exist C and

[34]It is immediate to verify that Assumptions 1-9 from Johnson (1970) are satisfied. The theorem then requires that $\int_{\underline{\theta}}^{\overline{\theta}} |\theta| \rho(\theta) d\theta < \infty$, which is true because $\int_{\underline{\theta}}^{\overline{\theta}} |\theta| \rho(\theta) d\theta \leq \bar{\theta} < \infty$.

$N_{\mathbf{s}}$ such that, for all $n > N_{\mathbf{s}}$,

$$\left| E\left[\theta|x_1, x_2, ..., x_n\right] - \hat{\theta} - \frac{1}{b\left(\hat{\theta}, x_1, ..., x_n\right)} \left(6a\left(\hat{\theta}, x_1, ..., x_n\right) + \frac{\rho'\left(\hat{\theta}\right)}{\rho\left(\hat{\theta}\right)} \right) \frac{1}{n} \right| \leq \frac{C}{n^2}, \tag{11}$$

where $\hat{\theta}$ is the maximum likelihood estimator (MLE) of θ,

$$a\left(\theta, x_1, ..., x_n\right) := \frac{1}{6n} \sum_{i=1}^{n} \frac{\partial^3}{\partial\theta^3} \log f\left(x_i|\theta\right), \text{ and} \tag{12}$$

$$b\left(\theta, x_1, ..., x_n\right) := \sqrt{-\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial\theta^2} \log f\left(x_i|\theta\right)}.$$

The MLE of $p\left(\theta\right)$ is $\widehat{p\left(\theta\right)} = \frac{\sum_{i=1}^{n} x_i}{n}$; from the invariance property, $\hat{\theta} = p^{-1}\left(\frac{\sum_{i=1}^{n} x_i}{n}\right)$. Let the random variable $k := \sum_{i=1}^{n} x_i$ denote the number of successes. Using the expressions for f and $\hat{\theta}$ and substituting in equations (12), we obtain:

$$a\left(\hat{\theta}, x_1, ..., x_n\right) = \frac{1}{\left(1 - \frac{k}{n}\right)\frac{k}{n}} \left\{ \frac{\left[p'\left(p^{-1}\left(\frac{k}{n}\right)\right)\right]^3}{3} \left(\frac{1}{\frac{k}{n}} - \frac{1}{1 - \frac{k}{n}} \right) - \frac{p'\left(p^{-1}\left(\frac{k}{n}\right)\right) p''\left(p^{-1}\left(\frac{k}{n}\right)\right)}{2} \right\},$$

and

$$b\left(\hat{\theta}, x_1, ..., x_n\right) = \frac{p'\left(p^{-1}\left(\frac{k}{n}\right)\right)}{\sqrt{\frac{k}{n}\left(1 - \frac{k}{n}\right)}}.$$

Plugging back in equation (11), establishes that, for $n > N_{\mathbf{s}}$,

$$\left| E\left[\theta|x\right] - p^{-1}\left(\frac{k}{n}\right) - \frac{\xi\left(\frac{k}{n}\right)}{n} \right| \leq \frac{C}{n^2},$$

where $\xi\left(x\right) := \frac{\sqrt{x(1-x)}}{p'(p^{-1}(x))} \left\{ \frac{2\left[p'\left(p^{-1}(x)\right)\right]^3 (1-2x)}{[x(1-x)]^2} - \frac{3p'\left(p^{-1}(x)\right)p''\left(p^{-1}(x)\right)}{x(1-x)} + \frac{\rho'\left(p^{-1}(x)\right)}{\rho(p^{-1}(x))} \right\}$. ■

Proof of Proposition 4. Let $\mathbf{s}$ and $\mathbf{s}'$ be two infinite sequences of informative recollections that coincide except at one index (say, j): $s_j = L$, $s'_j = H$, and $s_i = s'_i$ for all $i \neq j$. Since the posterior depends only on the number of successes and informative recollections (k, n), posteriors are the same for any $j \in \mathbb{N}$.

From inequality (3), there exist C_1, $N_{1,\mathbf{s}}$, and $N_{1,\mathbf{s}'}$, such that

$$\left| E\left[\theta|k, n\right] - p^{-1}\left(\frac{k}{n}\right) - \xi\left(\frac{k}{n}\right)\frac{1}{n} \right| \leq \frac{C_1}{n^2}, \text{ and}$$

$$\left| E\left[\theta|k+1, n\right] - p^{-1}\left(\frac{k+1}{n}\right) - \xi\left(\frac{k+1}{n}\right)\frac{1}{n} \right| \leq \frac{C_1}{n^2}$$

for all $n > \max\left\{N_{1,\mathbf{s}}, N_{1,\mathbf{s}'}\right\}$. Notice that $N_{1,\mathbf{s}'}$ can be written as a function of $\mathbf{s}$ only since it does not depend on the index in which $\mathbf{s}'$ and $\mathbf{s}$ differ, j. By the triangle inequality,

$$\left| E\left[\theta|k+1,n\right] - E\left[\theta|k,n\right] - p^{-1}\left(\frac{k+1}{n}\right) + p^{-1}\left(\frac{k}{n}\right) - \frac{\xi\left(\frac{k+1}{n}\right)-\xi\left(\frac{k}{n}\right)}{n} \right| \leq \frac{2C_1}{n^2}. \tag{13}$$

From Taylor's theorem,

$$p^{-1}\left(\frac{k+1}{n}\right) = p^{-1}\left(\frac{k}{n}\right) + \frac{1}{n}\frac{1}{p'\left(p^{-1}\left(\frac{k}{n}\right)\right)} + \frac{1}{2n^2}\frac{p''\left(p^{-1}\left(\frac{k}{n}+\delta\right)\right)}{\left[p'\left(p^{-1}\left(\frac{k}{n}+\delta\right)\right)\right]^3}$$

for $\delta \in [0, \frac{1}{n}]$. Since p is twice continuously differentiable and p' is bounded away from zero, there exist N_2 and C_2 such that

$$\left| p^{-1}\left(\frac{k+1}{n}\right) - p^{-1}\left(\frac{k}{n}\right) - \frac{1}{n}\frac{1}{p'\left(p^{-1}\left(\frac{k}{n}\right)\right)} \right| \leq \frac{C_2}{n^2} \tag{14}$$

for all $n > N_2$ on an almost sure set.

Since $p\left(\theta\right)$ is bounded away from zero and one, ξ is continuously differentiable in its entire domain. Applying Taylor's theorem, we obtain

$$\frac{\xi\left(\frac{k+1}{n}\right)-\xi\left(\frac{k}{n}\right)}{n} = \frac{1}{n^2}\xi'\left(\frac{k}{n}+\epsilon\right)$$

for some $\epsilon \in [0, \frac{1}{n}]$.
By the uniform strong law, $\frac{k}{n}$ converges almost surely to $p\left(\theta\right)$. Since ξ is continuously differentiable, $\xi'\left(\frac{k}{n}+\delta_0\right)$ converges uniformly almost surely to $\xi'\left(\theta\right)$. Thus, there exist constants N_3 and C_3 such that

$$\left| \frac{\xi\left(\frac{k+1}{n}\right)-\xi\left(\frac{k}{n}\right)}{n} \right| \leq \frac{C_3}{n^2} \tag{15}$$

for all $n > N_3$.
Combining (14) and (15) with inequality (13) (using the triangle inequality), we obtain, for all $n > N_{\mathbf{s}}$,

$$\left| E\left[\theta|k+1,n\right] - E\left[\theta|k,n\right] - \frac{1}{n}\frac{1}{p'\left(p^{-1}\left(\frac{k}{n}\right)\right)} \right| \leq \frac{C}{n^2}$$

on an almost sure set, where $N_{\mathbf{s}} := \max\left\{N_{1,\mathbf{s}}, N_{1,\mathbf{s}'}, N_2, N_3\right\}$ and $C := 2C_1 + C_2 + C_3$. ■

Proof of Theorem 1

Before presenting the proof of the theorem, it is helpful to simplify the notation. Note that any distribution over skills θ induces a distribution over signal structures $p\left(\theta\right)$. For notational simplicity, I will work with the distribution over signal structures $\pi = p\left(\theta\right)$ rather than the distribution over skills θ. This is without loss of generality here because $p\left(\theta\right)$ is a strictly increasing function (this is a standard maneuver in probability theory, known as the "change of variables" or the "push out" method). For any $\pi \in p\left(\Theta\right)$, let $\mu\left(\pi\right) := \rho\left(p^{-1}\left(\pi\right)\right)$ denote the prior distribution over the space of probabilities of a high signal π.

The posterior distribution of θ,

$$\rho(\theta|k,n) = \frac{p(\theta)^k [1-p(\theta)]^{n-k} \rho(\theta)}{\int_\Theta p(\theta)^k [1-p(\theta)]^{n-k} \rho(\theta)\, d\theta},$$

can be written in terms of the posterior distribution of π,

$$\mu(\pi|k,n) = \frac{\pi^k (1-\pi)^{n-k} \mu(\pi)}{\int_\Theta \pi^k (1-\pi)^{n-k} \mu(\pi)\, d\pi}.$$

Let $W(a,\pi) := V\left(a, p^{-1}(\pi)\right)$ denote the payoff from action a when the individual has the skill parameter $\theta = p^{-1}(\pi)$.

Let $\gamma := \frac{k}{n}$, denote the proportion of successes, which is also the maximum likelihood estimator of π. I abuse of notation by letting

$$a(\pi) := \arg\max_{a\in A} W(a,\pi)$$

denote the optimal action for "probability type" π. Let

$$\tilde{a}_n(\gamma) := \arg\max_{a\in A} \int W(a,\pi)\,\mu(\pi|\gamma n, n)\, d\theta, \tag{16}$$

denote the optimal action conditional on the proportion γ of successes in n observations. It is nonempty since the objective function is continuous and A is compact; it is unique because the objective function is strictly concave and A is convex. The following lemma shows that $\tilde{a}_n(\gamma)$ converges pointwise to $a(\gamma)$.

Lemma 4 $\lim_{n\to\infty} \tilde{a}_n(\gamma) = a(\gamma)$ *for all* $\gamma \in [\underline{\pi}, \bar{\pi}]$.

Proof. Since $\mu(\pi|\gamma n, n)$ converges to a mass point at γ and $W(a,\cdot)$ is a continuously differentiable (and, therefore, absolutely continuous) function,

$$\lim_{n\to\infty} \int W(a,\pi)\,\mu(\pi|\gamma n,n)\, d\pi = W(a,\gamma).$$

Suppose, in order to obtain a contradiction, that $\{\tilde{a}_n(\gamma)\}_{n=1}^\infty$ does not converge to $a(\gamma)$. Then, there exists $\epsilon > 0$ and subsequence $\{\tilde{a}_{n_i}(\gamma)\}_{i=1}^\infty$ such that

$$|\tilde{a}_{n_i}(\gamma) - a(\gamma)| \geq \epsilon \tag{17}$$

for all n_i. Since $\{\tilde{a}_{n_i}(\gamma)\}$ is in the compact set $A \in \mathbb{R}$, it follows from the Bolzano-Weierstrass theorem that it has a converging subsequence $\left\{\tilde{a}_{n_{i_j}}\right\}$; let a^* to denote its limit. Because $\tilde{a}_{n_{i_j}}(\gamma)$ maximizes $\int W(a,\pi)\,\mu\left(\pi|\gamma n_{i_j}, n_{i_j}\right) d\theta$, we must have

$$\int W\left(\tilde{a}_{n_{i_j}}(\gamma),\pi\right) \mu\left(\pi|\gamma n_{i_j}, n_{i_j}\right) d\pi \geq \int W(a(\gamma),\pi)\,\mu\left(\pi|\gamma n_{i_j}, n_{i_j}\right) d\pi \quad \forall n_{i_j}.$$

Since

$$\lim_{n\to\infty}\int W(a,\pi)\,\mu(\pi|\gamma n,n)\,d\pi = W(a,\pi)\,,$$

and $\lim_{n_{i_k}\to\infty}\tilde{a}_{n_{i_k}}(\gamma)=a^*$, it follows that

$$W(a^*,\gamma)\geq W(a(\gamma),\gamma)\,.$$

Thus, a^* maximizes $W(a,\pi)$. Since, by Assumption 2, $a(\gamma)$ is the unique maximizer of $W(a,\gamma)$, we must have $a^*=a(\gamma)$, which contradicts (17). ■

Substituting the posterior distribution in (16) and noting that the denominator is positive and it is not a function of a, yields

$$\tilde{a}_n(\gamma)=\arg\max_{a\in A}\mathcal{V}_n(a,\gamma)\,,$$

where $\mathcal{V}_n(a,\gamma)\equiv\int W(a,\pi)\left[\pi^{\gamma}(1-\pi)^{1-\gamma}\right]^n\mu(\pi)\,d\pi$. Since $a(\gamma)$ is in the interior of A and $\tilde{a_n}(\gamma)$ converges to $a(\gamma)$, $a_n(\gamma)$ must also be in the interior of A for large n. Thus, it must satisfy the following first order condition:

$$\frac{\partial\mathcal{V}_n}{\partial a}(a_n(\gamma),\gamma)=0.$$

Since $\frac{\partial\mathcal{V}_n}{\partial a}$ is strictly decreasing and continuously differentiable, the implicit function theorem ensures that $a_n(\gamma)$ is a continuously differentiable function and

$$a_n'(\gamma)=-\frac{\frac{\partial^2\mathcal{V}_n}{\partial a\partial\gamma}(a_n(\gamma),\gamma)}{\frac{\partial^2\mathcal{V}_n}{\partial a^2}(a_n(\gamma),\gamma)}. \tag{18}$$

Our next lemma shows that $\{a_n'(\gamma)\}$ is bounded:

Lemma 5 *For each γ, there exist constants N_γ and K_γ such that $|a_n'(\gamma)|<K_\gamma$ for all $n>N_\gamma$.*

Proof. The second derivatives are

$$\frac{\partial^2\mathcal{V}_n}{\partial a\partial\gamma}(a,\gamma)=n\int\ln\left(\frac{\pi}{1-\pi}\right)\frac{\partial W}{\partial a}(a,\pi)\left[\pi^{\gamma}(1-\pi)^{1-\gamma}\right]^n\mu(\pi)\,d\pi,$$

and

$$\frac{\partial^2\mathcal{V}_n}{\partial a^2}(a,\gamma)=\int\frac{\partial^2 W}{\partial a^2}(a,\pi)\left[\pi^{\gamma}(1-\pi)^{1-\gamma}\right]^n\mu(\pi)\,d\pi.$$

Substituting these expressions in (19), yields

$$a_n'(\gamma)=-\frac{n\int\ln\left(\frac{\pi}{1-\pi}\right)\frac{\partial W}{\partial a}(a_n(\gamma),\pi)\left[\pi^{\gamma}(1-\pi)^{1-\gamma}\right]^n\mu(\pi)\,d\pi}{\int\frac{\partial^2 W}{\partial a^2}(a_n(\gamma),\pi)\left[\pi^{\gamma}(1-\pi)^{1-\gamma}\right]^n\mu(\pi)\,d\pi}. \tag{19}$$

Recall that $a_n(\gamma)$ converges pointwise to $a(\gamma)$. Since $\frac{\partial^2 W}{\partial a^2}$ and μ are continuous and π is bounded away from 0 and 1, this expression has the same limit as

$$-\frac{n\int \ln\left(\frac{\pi}{1-\pi}\right)\frac{\partial W}{\partial a}\left(a\left(\gamma\right),\pi\right)\left[\pi^{\gamma}\left(1-\pi\right)^{1-\gamma}\right]^{n}\mu\left(\pi\right)d\pi}{\int \frac{\partial^{2}W}{\partial a^{2}}\left(a\left(\gamma\right),\pi\right)\left[\pi^{\gamma}\left(1-\pi\right)^{1-\gamma}\right]^{n}\mu\left(\pi\right)d\pi}.$$

Since $a\left(\gamma\right)$ maximizes $W\left(a,\gamma\right)$ and $a\left(\gamma\right)$ is interior, it satisfies the first-order condition: $\frac{\partial W}{\partial a}\left(a\left(\gamma\right),\gamma\right)=0$. Applying a Taylor expansion to $\frac{\partial W}{\partial a}\left(a\left(\gamma\right),\pi\right)$ at $\pi=\gamma$, we obtain

$$\frac{\partial W}{\partial a}\left(a\left(\gamma\right),\pi\right)=\underbrace{\frac{\partial W}{\partial a}\left(a\left(\gamma\right),\gamma\right)}_{0}+\frac{1}{2}\frac{\partial^{2}W}{\partial a^{2}}\left(a\left(\gamma\right),\gamma\right)\left(\pi-\gamma\right)+R\left(\pi-\gamma\right),$$

where $\lim_{h\to 0}\frac{R(h)}{h}=0$. Substituting in (19), gives

$$a_{n}'(\gamma)=-\frac{n\int \ln\left(\frac{\pi}{1-\pi}\right)\left[\frac{1}{2}\frac{\partial^{2}W}{\partial a^{2}}\left(a\left(\gamma\right),\gamma\right)+R\left(\pi-\gamma\right)\right]\left(\pi-\gamma\right)\left[\pi^{\gamma}\left(1-\pi\right)^{1-\gamma}\right]^{n}\mu\left(\pi\right)d\pi}{\int \frac{\partial^{2}W}{\partial a^{2}}\left(a\left(\gamma\right),\pi\right)\left[\pi^{\gamma}\left(1-\pi\right)^{1-\gamma}\right]^{n}\mu\left(\pi\right)d\pi}$$

$$=-\frac{n\int \ln\left(\frac{\pi}{1-\pi}\right)\left[\frac{1}{2}\frac{\partial^{2}W}{\partial a^{2}}\left(a\left(\gamma\right),\gamma\right)+R\left(\pi-\gamma\right)\right]\left(\pi-\gamma\right)\left[\pi^{\gamma}\left(1-\pi\right)^{1-\gamma}\right]^{n}\mu\left(\pi\right)d\pi}{\int \left[\pi^{\gamma}\left(1-\pi\right)^{1-\gamma}\right]^{n}\mu\left(\pi\right)d\pi}$$

$$\times\frac{\int \left[\pi^{\gamma}\left(1-\pi\right)^{1-\gamma}\right]^{n}\mu\left(\pi\right)d\pi}{\int \frac{\partial^{2}W}{\partial a^{2}}\left(a\left(\gamma\right),\pi\right)\left[\pi^{\gamma}\left(1-\pi\right)^{1-\gamma}\right]^{n}\mu\left(\pi\right)d\pi}.$$

The second term,

$$\frac{\int \left[\pi^{\gamma}\left(1-\pi\right)^{1-\gamma}\right]^{n}\mu\left(\pi\right)d\pi}{\int \frac{\partial^{2}W}{\partial a^{2}}\left(a\left(\gamma\right),\pi\right)\left[\pi^{\gamma}\left(1-\pi\right)^{1-\gamma}\right]^{n}\mu\left(\pi\right)d\pi}=\frac{1}{E\left[\frac{\partial^{2}W}{\partial a^{2}}\left(a\left(\gamma\right),\pi\right)|\gamma,n\right]},$$

converges in probability to $\frac{1}{\frac{\partial^{2}W}{\partial a^{2}}(a(\gamma),\gamma)}$ since W is twice continuously differentiable and $E\left[\frac{\partial^{2}W}{\partial a^{2}}\left(a\left(\gamma\right),\pi\right)|\gamma,n\right]$ converges in probability to $\frac{\partial^{2}W}{\partial a^{2}}\left(a\left(\gamma\right),\gamma\right)<0$.

The first term equals

$$-nE\left[\ln\left(\frac{\pi}{1-\pi}\right)\frac{1}{2}\frac{\partial^{2}W}{\partial a^{2}}\left(a\left(\gamma\right),\gamma\right)\left(\pi-\gamma\right)|\gamma,n\right]-nE\left[\ln\left(\frac{\pi}{1-\pi}\right)R\left(\pi-\gamma\right)\left(\pi-\gamma\right)|\gamma,n\right]$$

$$=-\frac{n}{2}\frac{\partial^{2}W}{\partial a^{2}}\left(a\left(\gamma\right),\gamma\right)E\left[\ln\left(\frac{\pi}{1-\pi}\right)\left(\pi-\gamma\right)|\gamma,n\right]-nE\left[\ln\left(\frac{\pi}{1-\pi}\right)\frac{R\left(\pi-\gamma\right)}{\pi-\gamma}\left(\pi-\gamma\right)^{2}|\gamma,n\right].$$

Since π is bounded away from 0 and 1, standard asymptotic results (see, e.g., Johnson (1970)) establish that

$$E\left[\ln\left(\frac{\pi}{1-\pi}\right)\left(\pi-\gamma\right)|\gamma,n\right]=O_{\gamma}\left(\frac{1}{n}\right)\tag{20}$$

and

$$E\left[(\pi-\gamma)^2|\gamma,n\right]=O_\gamma\left(\frac{1}{n}\right). \tag{21}$$

From (20), $-\frac{n}{2}\frac{\partial^2 W}{\partial a^2}(a(\gamma),\gamma)\,E\left[\ln\left(\frac{\pi}{1-\pi}\right)(\pi-\gamma)|\gamma,n\right]=O_\gamma(1)$. Moreover, since $\lim_{h\to 0}\frac{R(h)}{h}=0$, it then follows that

$$E\left[\underbrace{\ln\left(\frac{\pi}{1-\pi}\right)}_{\to_p \ln\left(\frac{\gamma}{1-\gamma}\right)}\underbrace{n(\pi-\gamma)^2}_{O(1)}\underbrace{\frac{R(\pi-\gamma)}{\pi-\gamma}}_{\to_p 0}|\gamma,n\right]\to_p 0.$$

Thus, $a_n'(\gamma)$ is $O_\gamma(1)$. ■

Proof of the theorem. In order to obtain a contradiction, suppose that $\hat{\sigma}\left(h^t\right)=L$ for some history h^t with n informative recollections and consider a deviation to $\hat{\sigma}=H$. The gain from self-image is bounded below by $\alpha\left\{E[\theta|k+1,n]-E[\theta|k,n]\right\}$, which, as established in Lemma 4, is of order $\frac{1}{n}$. The deviation will lead the individual to choose action $\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}+\frac{1}{n}\right)$ if the game ends after a history associated with state $\left(\tilde{k},\tilde{n}\right)=(\tilde{n}\tilde{\gamma},\tilde{n})$. Therefore, the expected cost of distorting actions equals

$$\sum_{\substack{\tilde{\gamma}\geq k/\tilde{n}\\ \tilde{n}\geq n}}\int\left\{\left[W\left(\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right),\pi\right)-W\left(\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}+\frac{1}{\tilde{n}}\right),\pi\right)\right]\mu\left(\pi|\tilde{n}\tilde{\gamma},\tilde{n}\right)\right\}d\pi\Phi\left(\tilde{n}\tilde{\gamma},\tilde{n}|k,n\right), \tag{22}$$

where $\Phi\left(\tilde{k},\tilde{n}|k,n\right)$ denotes the probability that the game ends at a history associated to $\left(\tilde{k},\tilde{n}\right)$, given the strategies being played in the fixed PBE. We will establish that the expression in (22) decreases to 0 at a rate faster than $\frac{1}{n}$.

Recall the expression for the expected payoff from action a conditional on $(\tilde{\gamma},\tilde{n})$:

$$\mathcal{V}_n(a,\gamma)\equiv\int W(a,\pi)\left[\pi^\gamma(1-\pi)^{1-\gamma}\right]^n\mu(\pi)\,d\pi.$$

From Taylor's theorem,

$$\begin{aligned}\mathcal{V}_{\tilde{n}}\left(\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}+\tfrac{1}{\tilde{n}}\right),\tilde{\gamma}\right) &= \\ &\mathcal{V}_{\tilde{n}}\left(\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right),\tilde{\gamma}\right)+\tfrac{\partial\mathcal{V}_{\tilde{n}}}{\partial a}\left(\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right),\tilde{\gamma}\right)\times\left[\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}+\tfrac{1}{\tilde{n}}\right)-\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right)\right] \\ &+r\left(\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}+\tfrac{1}{\tilde{n}}\right)-\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right)\right),\end{aligned} \tag{23}$$

where $\lim_{h\to 0}\frac{r(h)}{h}=0$. For large enough $\tilde{n}$, $\tilde{a}_{\tilde{n}}(\tilde{\gamma})$ must be interior. Therefore, it must satisfy the first-order condition: $\frac{\partial\mathcal{V}_{\tilde{n}}}{\partial a}\left(\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right),\tilde{\gamma}\right)=0$. Substitute in equation (23) and divide both sides by $\frac{1}{n}$:

$$\frac{\mathcal{V}_{\tilde{n}}\left(\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right),\tilde{\gamma}\right)-\mathcal{V}_{\tilde{n}}\left(\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}+\frac{1}{n}\right),\tilde{\gamma}\right)}{\frac{1}{n}}=\frac{r\left(\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}+\frac{1}{n}\right)-\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right)\right)}{\frac{1}{n}}.$$

In order to show that the cost has a lower order of magnitude than $\frac{1}{n}$, we need to verify that

$$\lim_{\tilde{n}\to\infty} \frac{r\left(\tilde{a}_{\tilde{n}}\left(\tilde{\gamma}'\right) - \tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right)\right)}{\frac{1}{\tilde{n}}} = 0.$$

If $\tilde{a}_{\tilde{n}}\left(\tilde{\gamma} + \frac{1}{n}\right) = \tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right)$, the result is immediate; otherwise, we can multiply and divide the term on the right by $\tilde{a}_{\tilde{n}}\left(\tilde{\gamma} + \frac{1}{n}\right) - \tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right)$:

$$\frac{r\left(\tilde{a}_{\tilde{n}}\left(\tilde{\gamma} + \frac{1}{n}\right) - \tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right)\right)}{\tilde{a}_{\tilde{n}}\left(\tilde{\gamma} + \frac{1}{n}\right) - \tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right)} \frac{\tilde{a}_{\tilde{n}}\left(\tilde{\gamma} + \frac{1}{n}\right) - \tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right)}{\frac{1}{n}}.$$

From the definition of r and the fact that $\lim_{\tilde{n}\to\infty} \tilde{a}_{\tilde{n}}\left(\tilde{\gamma} + \frac{1}{\tilde{n}}\right) - \tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right) = 0$, the first term converges to zero. Thus, it suffices to show that there exists K and N such that $\tilde{n} > N$

$$\frac{\left|\tilde{a}_{\tilde{n}}\left(\tilde{\gamma} + \frac{1}{\tilde{n}}\right) - \tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right)\right|}{\frac{1}{\tilde{n}}} < K_{\tilde{\gamma}}.$$

But, from the mean value theorem,

$$\frac{\left|\tilde{a}_{\tilde{n}}\left(\tilde{\gamma} + \frac{1}{\tilde{n}}\right) - \tilde{a}_{\tilde{n}}\left(\tilde{\gamma}\right)\right|}{\frac{1}{\tilde{n}}} = \left|\tilde{a}'_{n}\left(\gamma^*\right)\right|,$$

for some $\gamma^* \in [\tilde{\gamma}, \tilde{\gamma} + \frac{1}{\tilde{n}}]$. Hence, from the previous lemma, such $K_{\tilde{\gamma}}$ exists. ■

Proof of Theorem 2

The proof will proceed by a series of lemmata. Let $\rho(\theta|k,n)$, $\tilde{a}_n\left(\gamma\right)$, and $a\left(\gamma\right)$ be as defined in the proof of Theorem 1. Because V and p^{-1} are continuous and $a\left(\gamma\right)$ is a singleton except at a finite number of points, we can partition the type space in a finite number of intervals and the optimal action is unique and constant in the interior of all those intervals. More precisely, there exists a finite set $\{\pi_1, \pi_2, ..., \pi_M\}$ with $\pi_1 = p\left(\underline{\theta}\right) < \pi_2 < ... < \pi_M = p\left(\bar{\theta}\right)$ such that $a\left(\gamma\right) = a\left(\gamma'\right)$ is a singleton for all $\gamma, \gamma' \in (\pi_i, \pi_{i+1})$, $i = 1, ..., N-1$.

Note that $\gamma = \frac{k}{n}$ is the maximum likelihood estimator (MLE) of $\pi = p\left(\theta\right)$. The first lemma shows that when n large enough, the action taken when the individual observes a proportion γ of successes, $\tilde{a}_n(\gamma)$, coincides with the action she would take if she knew with certainty that her type was $\pi = \gamma$:

Lemma 6 *There exists a constant N such that, for all γ in which $a\left(\gamma\right)$ is a singleton, $\tilde{a}_n(\gamma) = a\left(\gamma\right)$ for all $n > N$.*

Proof. First, we establish that there exists N_γ with this property. Let $a(\gamma)$ be unique. Then, for all $a \neq a\left(\gamma\right)$, $V(a(\gamma), p^{-1}\left(\gamma\right)) > V(a, p^{-1}\left(\gamma\right))$. Note that, for all $a \neq a(\gamma)$, we have

$$\lim_{n\to\infty} \left\{E[V(a(\gamma), p^{-1}\left(\pi\right))|\gamma n, n] - E[V(a, p^{-1}\left(\pi\right))|\gamma n, n]\right\} = V(a(\gamma), p^{-1}\left(\gamma\right)) - V(a, p^{-1}\left(\gamma\right)) > 0.$$

Thus, for any $\epsilon > 0$, there exists $n_{\gamma,\epsilon}$ such that

$$-\epsilon < \left\{E[V(a(\gamma), p^{-1}(\pi))|\gamma n, n] - E[V(a, p^{-1}(\pi))|\gamma n, n]\right\} - \left[V(a(\gamma), p^{-1}(\gamma)) - V(a, p^{-1}(\gamma))\right] < \epsilon$$

for $n > n_{\gamma,\epsilon}$. Taking $\epsilon = \min_a \left\{V(a(\gamma), p^{-1}(\gamma)) - V(a, p^{-1}(\gamma))\right\} > 0$ (which exists and is strictly positive because A is finite), it follows that there exists N_γ such that

$$E[V(a(\gamma), p^{-1}(\pi))|\gamma n, n] > E[V(a, p^{-1}(\pi))|\gamma n, n]$$

for all $n > N_\gamma$. Thus, $a(\gamma)$ also maximizes $E[V(a, \theta)|\gamma, n]$ for all such n.

Next, we show that we can take N_γ independent of γ. Let N be the supremum of N_γ for $\gamma \notin \{\pi_1, \pi_2, ..., \pi_M\}$. Suppose, in order to obtain a contradiction, that $N = +\infty$. Then, by the definition of the supremum, for any n, there exists $\gamma_n \neq \{\pi_1, \pi_2, ..., \pi_M\}$ such that $\tilde{a}_n(\gamma_n) \neq a(\gamma_n)$. Therefore, the sequence $\{\gamma_n\}$ is such that $\tilde{a}_n(\gamma_n) \neq a(\gamma_n)$ for all n. Since the sequence $\{\gamma_n\}$ is bounded, it has a convergent subsequence $\{\gamma_{n_j}\}$; let $\gamma^* \in [\underline{\pi}, \bar{\pi}]$ denote its limit. Then, we have a sequence $\{\gamma_{n_j}\} \to \gamma^*$ such that $\tilde{a}_n(\gamma_{n_j}) \neq a(\gamma_{n_j})$ for all n_j, which, because A is finite, implies $\lim_{n\to\infty} \tilde{a}_n(\gamma^*) \neq a(\gamma^*)$. This contradicts our previous result, which shows that there exists N_{γ^*} such that $\tilde{a}_n(\gamma^*) \neq a(\gamma^*)$ for all $n > N_{\gamma^*}$. ■

Next, we note that the posterior probability of a high signal can be approximated by its MLE $\frac{k}{n}$:

Lemma 7 *There exist D and N such that for any $n > N$:*

$$\left| E[\pi|k, n] - \frac{k}{n} + \frac{1}{n}\left[1 - 2\frac{k}{n} + \frac{k}{n}\left(1 - \frac{k}{n}\right)\frac{\rho'\left(\frac{k}{n}\right)}{\rho\left(\frac{k}{n}\right)}\right]\right| \leq \frac{D}{n^2}.$$

Proof. Follows directly from Hald (1967, pp. 360). ■

We will also use the fact that the MLE of π is asymptotically normally distributed:

Lemma 8 $\frac{\sqrt{n}(\gamma-\pi)}{\sqrt{\gamma(1-\gamma)}} \to_D N(0, 1)$, *where* $\gamma = \frac{k}{n}$.

Proof. See Hald (1967). ■

Proof of the theorem. Fix a type $\pi \notin \{\pi_1, \pi_2, ..., \pi_M\}$. In order to obtain a contradiction, suppose that the claim from the theorem is not true. Then, there must exist a sequence of recollections $h^\infty := \{\hat{\sigma}_t\}_{t\in\mathbb{N}}$ such that:

1. Every finite restriction of h^∞ is on the equilibrium path: $\Pr(\hat{\sigma}_1, ..., \hat{\sigma}_t|\sigma^*) > 0$, and

2. h^∞ has an infinite number of informative periods: $\#\{t : \sigma^*(\hat{\sigma}_1, ..., \hat{\sigma}_{t-1}) = L\} = +\infty$.

Let $\{\hat{\sigma}_{t_n}\}_{n=0}^\infty$ be the subsequence of h^∞ containing its informative periods. Define the random variable $x_{t_n} \equiv \mathbf{1}(\hat{\sigma}_{t_n} = H)$, where $\mathbf{1}(.)$ denotes the indicator function; let $k_n = \sum_{i=0}^n x_{t_i}$ denote the number of successes in the subsequence of informative periods up to n.

Let $h^{t_n} := (\hat{\sigma}_1, ..., \hat{\sigma}_{t_n})$ be the history in the informative period t_n. By construction, h^{t_n} has k_n successes in n informative periods. Since h^{t_n} is informative, $\hat{\sigma}^*(h^{t_n}) = L$ so that both (h^{t_n}, L) and

(h^{t_n}, H) are on the equilibrium path. Consider a deviation at h^{t_n} to $\hat{\sigma} = H$. There are two possibilities. Either $\hat{\sigma}_{t_n+1} = H$ so that we remain on h^∞ after the deviation, or $\hat{\sigma}_{t_n+1} = L$ so we leave h^∞ after the deviation. We will consider them separately.

First, suppose $\hat{\sigma}_{t_n+1} = H$ for n large enough. Then, from Lemma 6, the deviation only affects the action to be chosen if the game ends at a state $(\tilde{k}, \tilde{n})$ in which

$$\frac{\tilde{k}}{\tilde{n}} < \pi_i < \frac{\tilde{k}+1}{\tilde{n}}, \tag{24}$$

for $i = 1, ..., M$. From Lemma 8, $\frac{\sqrt{\tilde{n}}(\gamma-\pi)}{\sqrt{\gamma(1-\gamma)}} \to_D N(0,1)$. Thus, the probability of a signal satisfying condition 24 converges to

$$\Phi\left(\sqrt{\frac{\tilde{n}}{\gamma(1-\gamma)}}(\pi_i - \pi)\right) - \Phi\left(\sqrt{\frac{\tilde{n}}{\gamma(1-\gamma)}}\left(\pi_i - \pi - \frac{1}{\tilde{n}}\right)\right),$$

where Φ is the c.d.f. of the Normal distribution. The cost of making an incorrect decision is bounded above by

$$\bar{V} := \max_{\theta,a}\{V(a,\theta)\} - \min_{a,\theta}\{V(a,\theta)\},$$

which is finite because V is continuous and $\Theta \times A$ is compact. Thus, the expected cost from the deviation in terms of actions is bounded above by the maximum cost of making an incorrect decision times the probability of affecting the decision:

$$\bar{V}\sum_{i=1}^{M}\left\{\Phi\left(\sqrt{\frac{\tilde{n}}{\frac{\tilde{k}}{\tilde{n}}\left(1-\frac{\tilde{k}}{\tilde{n}}\right)}}(\pi_i - \pi)\right) - \Phi\left(\sqrt{\frac{\tilde{n}}{\frac{\tilde{k}}{\tilde{n}}\left(1-\frac{\tilde{k}}{\tilde{n}}\right)}}\left(\pi_i - \pi - \frac{1}{\tilde{n}}\right)\right)\right\},$$

which converges to zero *exponentially* for all $\pi \notin \{\pi_1, \pi_2, ..., \pi_M\}$. Therefore, the expected cost has a lower order of magnitude than the benefit in self image, which is $O(1/n)$. Hence, there is always a profitable deviation if there exists n large enough for which $\hat{\sigma}_{t_n+1} = H$.

Next, let us consider the other case. That is, suppose there exists $\bar{n}$ such that $\hat{\sigma}_{t_n+1} = L$ for all $n > \bar{n}$. Then, among the informative periods in h^∞, there is an infinite number of low signals and only a finite number (at most $\bar{n}$) of high signals. Hence, beliefs on h^∞ converge to $\underline{\pi} := p(\underline{\theta})$. The previous argument does not work in this case because, following a deviation to $\hat{\sigma} = H$, histories no longer belong to the sequence h^∞. As a result, the individual may stop interpreting signals realistically: posterior beliefs will depend on the individual's continuation strategies after such a deviation. We need a bound on the cost of the deviation that holds uniformly among all possible continuation strategies. We will obtain such uniform bound and show that it converges to zero at a much faster rate than $\frac{1}{n}$.

First, suppose $a(p^{-1}(\underline{\pi}))$ is a singleton. Then, there exists a neighborhood of $\underline{\pi}$ where $a(p^{-1}(\pi)) = a(p^{-1}(\underline{\pi}))$. The probability that the deviation affects actions is bounded above by following a strategy that interprets signals realistically until we reach $\gamma = \frac{\tilde{k}}{\tilde{n}} \geq \pi_1$ and then rationalizes all future signals away. This is an upper bound both because it is maximizes the probability that we affect decisions among all possible continuation strategies but also because it ignores the probability that the game ends (so we may never be able to reach such a state). I will show that the probability determined by this upper bound is $O\left(\frac{1}{n^{n-2k}}\right)$.

As shown previously, $\Pr(\sigma = H|k,n) = E[\pi|k,n] = \frac{k}{n} + O\left(\frac{1}{n^2}\right)$. Since beliefs on h^∞ converge to $\underline{\pi} < \pi_1$ and n is large enough, there is no loss of generality in assuming that $\frac{k}{n} < \pi_1$. Therefore, we need to calculate the probability of obtaining S successes and F failures such that $\frac{k+S}{n+S+F} \geq \pi_1$.

$$k + S \geq \pi_1 (n + S + F) \iff S \geq \frac{\pi_1 (n+F) - k}{1 - \pi_1}.$$

Letting $\lceil \cdot \rceil$ denote the ceiling function, we need $S = \left\lceil \frac{\pi_1(n+F)-k}{1-\pi_1} \right\rceil$ and F failures, for $F = 0, 1, 2, \ldots$.

Note that the order of successes and failures does not matter for their conditional probabilities (up to terms of lower order); for any vector $\{\sigma_{t+1}, \sigma_{t+2}, \ldots, \sigma_{t+S+F}\}$ with S successes and F failures,

$$\Pr(\sigma_{t+1}, \ldots, \sigma_{t+S+F}|k,n) = \frac{k}{n}\frac{k+1}{n+1}\frac{k+2}{n+2}\cdots\frac{k+S-1}{n+S-1}\left(1-\frac{k+S}{n+S}\right)\left(1-\frac{k+S}{n+S+1}\right)\cdots\left(1-\frac{k+S}{n+S+F-1}\right)$$

$$= \frac{(n-1)!}{(n+S+F-1)!}\frac{(k+S-1)!}{(k-1)!}\frac{(n-k+F-1)!}{(n-k-1)!}.$$

To simplify notation, let $r := \frac{\pi_1}{1-\pi_1}$, so that $S = \lceil r(n+F) - (1+r)k \rceil$. Summing for $F = 0, 1, 2, \ldots.$, we obtain the probability of reaching $(\tilde{k}, \tilde{n})$ such that $\frac{\tilde{k}}{\tilde{n}} > \pi_1$ is bounded above by

$$\sum_{F=0}^{\infty} \frac{(n-1)!}{(n+\lceil r(n+F)-(1+r)k\rceil+F-1)!}\frac{(k+\lceil r(n+F)-(1+r)k\rceil-1)!}{(k-1)!}\frac{(n-k+F-1)!}{(n-k-1)!}$$

$$= \frac{(n-1)!}{(k-1)!(n-k-1)!}\sum_{F=0}^{\infty}\frac{(n-k+F-1)!(k+\lceil r(n+F)-(1+r)k\rceil-1)!}{(n+\lceil r(n+F)-(1+r)k\rceil+F-1)!}. \tag{25}$$

Notice that

$$\frac{(n-1)!}{(k-1)!(n-k-1)!} = k \times \binom{n-1}{k} = k \times \prod_{i=1}^{k}\frac{n-1-(k-i)}{i}$$

which is $O\left(n^k\right)$. Moreover, after some algebraic manipulations, we can rewrite the inverse of the terms inside the summation as

$$\frac{(n+\lceil r(n+F)-(1+r)k\rceil+F-1)!}{(n-k+F-1)! \times (k+\lceil r(n+F)-(1+r)k\rceil-1)!}$$

$$= (n+\lceil r(n+F)-(1+r)k\rceil+F-1) \times \binom{n+\lceil r(n+F)-(1+r)k\rceil+F-2}{n-k+F-1}$$

$$= (n+\lceil r(n+F)-(1+r)k\rceil+F-1)\prod_{i=1}^{n-k+F-1}\frac{K-1+i+\lceil r(n+F)-(1+r)k\rceil}{i},$$

which is $O\left(n^{n-k+F}\right)$. Thus, there exist constants C_0, C_1, and $\bar{n}$ such that

$$\frac{(n-1)!}{(k-1)!(n-k-1)!} < \frac{C_0}{n^k}, \quad \text{and}$$

$$\frac{(n-k+F-1)!\,(k+\lceil r\,(n+F)-(1+r)\,k\rceil-1)!}{(n+\lceil r\,(n+F)-(1+r)\,k\rceil+F-1)!}<\frac{C_1 r^F}{n^{n-k+F}}$$

for all $n>\bar{n}$. Combining both conditions, we obtain

$$\frac{(n-1)!}{(k-1)!\,(n-k-1)!}\sum_{F=0}^{\infty}\frac{(n-k+F-1)!\,(k+\lceil r\,(n+F)-(1+r)\,k\rceil-1)!}{(n+\lceil r\,(n+F)-(1+r)\,k\rceil+F-1)!}$$

$$<\frac{C_0}{n^k}\sum_{F=0}^{\infty}\frac{C_1 r^F}{n^{n-k+F}}=\frac{C_0 C_1}{n^{n-2k}}\sum_{F=0}^{\infty}\frac{r^F}{n^F}=\left(\frac{n}{n-r}\right)\frac{C_0 C_1}{n^{n-2k}}.$$

Since $\frac{n}{n-r}\to 1$, it follows that the expression in equation (25) is $O\left(\frac{1}{n^{n-2k}}\right)$. Hence, the expected cost from the deviation – which is bounded above by the probability of affecting actions times the maximum cost $\bar{V}$ – is also $O\left(\frac{1}{n^{n-2k}}\right)$. Since the benefit from the deviation is $O\left(\frac{1}{n}\right)$, there is always a profitable deviation for n large enough.

Finally, suppose that $a\left(p^{-1}\left(\underline{\pi}\right)\right)$ is not a singleton. Then, because it is a singleton only in a finite number of points, there exists $\pi_1>\underline{\pi}$ such that $a\left(p^{-1}\left(\pi\right)\right)=:a_1$ is a singleton for all $\pi\in\left(\underline{\pi},\pi_1\right)$. Moreover, because posteriors have full support and converge to $\underline{\pi}$, it follows that there exists n such that $a_n\left(p^{-1}\left(\frac{k}{n}\right)\right)$ converges to a_1. Then, the result is obtained by following the exact same steps as when $a\left(p^{-1}\left(\underline{\pi}\right)\right)$ is a singleton. ■

Proof of Lemma 3.

Let h^t be a last informative history on the equilibrium path. By Bayes' rule, there must exist at least two histories on the equilibrium path with different conditional expectations. With no loss of generality, let

$$E\left[\theta|h^t\hat{\sigma}'_{t+1}\right]>E\left[\theta|h^t\hat{\sigma}_{t+1}\right].$$

In order for this to be an equilibrium, no signal σ_{t+1} that is prescribed to choose $\hat{\sigma}_{t+1}$ can profit by deviating to $\hat{\sigma}'_{t+1}$. Such a deviation raises payoffs from self views by

$$\alpha\left(\frac{1-\beta^{T-t+1}}{1-\beta}\right)\left\{E\left[\theta|h^t\hat{\sigma}'_{t+1}\right]-E\left[\theta|h^t\hat{\sigma}_{t+1}\right]\right\}.$$

Deviating from $\hat{\sigma}_{t+1}$ to $\hat{\sigma}'_{t+1}$ reduces the payoff from actions by

$$\beta^{T-t+1}\kappa\left\{E\left[\theta|h^t\hat{\sigma}'_{t+1}\right]-E\left[\theta|h^t\hat{\sigma}_{t+1}\right]\right\}\left\{E\left[\theta|h^t\hat{\sigma}_{t+1}\right]+E\left[\theta|h^t\hat{\sigma}'_{t+1}\right]-2E\left[\theta|h^t,\sigma_{t+1}\right]\right\},$$

where $E\left[\theta|h^t,\sigma_{t+1}\right]$ denotes the mean conditional on history h^t and the true signal σ_{t+1} in period $t+1$.

Thus, the deviation is profitable if

$$\frac{\alpha}{\kappa}\left(\frac{1-\beta^{T-t+1}}{1-\beta}\right)\geq\beta^{T-t+1}\left\{E\left[\theta|h^t\hat{\sigma}_{t+1}\right]+E\left[\theta|h^t\hat{\sigma}'_{t+1}\right]-2E\left[\theta|h^t,\sigma_{t+1}\right]\right\}.$$

Because $\frac{1-\beta^{T-t+1}}{1-\beta}\geq 1$ and $\beta\geq\beta^{T-t+1}$, a sufficient condition for the deviation to be profitable is

$$\frac{\alpha}{\beta\kappa}\geq E\left[\theta|h^t\hat{\sigma}_{t+1}\right]+E\left[\theta|h^t\hat{\sigma}'_{t+1}\right]-2E\left[\theta|h^t,\sigma_{t+1}\right]. \tag{26}$$

Consistency of beliefs requires the posterior mean conditional on recollection $\hat{\sigma}_{t+1}$ to lie in the convex hull of posterior means conditional on all signals that are interpreted as $\hat{\sigma}_{t+1}$. Thus, there exists a signal σ_{t+1} interpreted as $\hat{\sigma}_{t+1}$ (i.e., $\hat{\sigma}^*\left(h^t\sigma_{t+1}\right) = \hat{\sigma}_{t+1}$) such that

$$E\left[\theta|h^t\hat{\sigma}_{t+1}\right] \leq E\left[\theta|h^t, \sigma_{t+1}\right].$$

After some algebraic manipulations, we obtain:

$$E\left[\theta|h^t\hat{\sigma}'_{t+1}\right] - E\left[\theta|h^t\hat{\sigma}_{t+1}\right] \geq E\left[\theta|h^t\hat{\sigma}'_{t+1}\right] + E\left[\theta|h^t\hat{\sigma}_{t+1}\right] - 2E\left[\theta|h^t, \sigma_{t+1}\right].$$

Substituting in equation (26), we obtain the sufficient condition for the deviation to be profitable:

$$\frac{\alpha}{\beta\kappa} \geq E\left[\theta|\left\{h^t, \hat{\sigma}'_{t+1}\right\}\right] - E\left[\theta|\left\{h^t, \hat{\sigma}_{t+1}\right\}\right],$$

concluding the proof ■

Proof of Proposition 12. Recall that $h^t h^\tau$ denotes the history obtained by concatenating h^τ and h^t, whereas $h^t\hat{\sigma}h^\tau$ denotes the concatenation of h^τ, $\hat{\sigma}$, and h^τ – see (8) and (9). We say that h^t is a last history on the equilibrium path if all of its continuation histories on the equilibrium path are uninformative. Formally, h^t is a *last informative history on the equilibrium path* if

- $E\left[\theta|\left\{h^tH\right\}\right] \neq E\left[\theta|\left\{h^tL\right\}\right]$, and
- For all continuation histories $\left\{h^t\hat{\sigma}_t h^s\right\} \supset h^t$ on the equilibrium path, $E\left[\theta|\left\{h^t\hat{\sigma}_t\right\}\right] = E\left[\theta|\left\{h^t\hat{\sigma}_t h^s\right\}\right].$

The following lemma establishes that beliefs in a last informative history on the equilibrium path cannot be "too concentrated."

Lemma 9 *Let h^t be a last informative history on the equilibrium path. Then, $E\left[\theta|\left\{h^tH\right\}\right] - E\left[\theta|\left\{h^tL\right\}\right] > \frac{\alpha}{\beta\kappa}$.*

Proof. Let h^{T+1} be a history on the equilibrium path. If h^{T+1} has no informative periods, the statement is trivially true. Assume that h^{T+1} has at least one informative period and let $t \leq T+1$ denote its last informative period (i.e., t is the last period in which the individual interprets a low signal realistically with positive probability: $\lambda\left(h^t\right) < 1$).

In order for this to be an equilibrium, the individual cannot strictly prefer to play $\hat{\sigma} = H$ at h^t. Since all future periods are uninformative, the benefit from playing $\hat{\sigma} = H$ instead of $\hat{\sigma} = L$ is

$$\alpha \sum_{s=t}^{T} \beta^{s-t}\left\{E\left[\theta|\left\{h^t, H\right\}\right] - E\left[\theta|\left\{h^t, L\right\}\right]\right\} = \alpha\left(\frac{1-\beta^{T-t+1}}{1-\beta}\right)\left\{E\left[\theta|\left\{h^t, H\right\}\right] - E\left[\theta|h^t, \sigma_t = L\right]\right\},$$

where I used the fact that $E\left[\theta|\left\{h^t, L\right\}\right] = E\left[\theta|h^t, \sigma_t = L\right]$. The cost of playing $\hat{\sigma} = H$ instead of $\hat{\sigma} = L$ in terms of choosing a suboptimal action is

$$\beta^{T+1-t}E\left[V\left(E\left[\theta|\left\{h^t, L\right\}\right], \theta\right) - V\left(E\left[\theta|\left\{h^t, H\right\}\right], \theta\right)|h^t, \sigma_t = L\right]$$

$$= \beta^{T+1-t}\kappa \left\{ E\left[\theta | \left\{h^t, H\right\}\right] - E\left[\theta | h^t, \sigma_t = L\right] \right\}^2 .$$

Thus, the condition for $\hat{\sigma} = H$ not to be strictly preferred to $\hat{\sigma} = L$ is

$$\alpha \left(\frac{1 - \beta^{T-t+1}}{1-\beta} \right) \left\{ E\left[\theta | \left\{h^t, H\right\}\right] - E\left[\theta | h^t, \sigma_t = L\right] \right\} \leq \beta^{T+1-t}\kappa \left\{ E\left[\theta | \left\{h^t, H\right\}\right] - E\left[\theta | h^t, \sigma_t = L\right] \right\}^2 .$$

Since the individual plays L with strictly positive probability, Bayesian updating gives

$$E\left[\theta | h^t, \sigma_t = L\right] < E\left[\theta | h^t\right] < E\left[\theta | \left\{h^t, H\right\}\right] \leq E\left[\theta | h^t, \sigma_t = H\right] .$$

Then, because $\left(\frac{1-\beta^{T-t+1}}{1-\beta} \right) > 1 > \beta \geq \beta^{T+1-t}$, a necessary condition for the individual to play L with positive probability at h^t is

$$\frac{\alpha}{\beta\kappa} < E\left[\theta | \left\{h^t, H\right\}\right] - E\left[\theta | h^t, \sigma_t = L\right] ,$$

concluding the proof. ■

The proof proceeds by backward induction. The following lemma establishes the appropriate starting point.

Lemma 10 *Let h^t be an informative history on the equilibrium path. Then, h^t has a continuation history on the equilibrium path $h^{t+s} = \{h^t, \hat{\sigma}_t, ..., \hat{\sigma}_{t+s-1}\}$ with the following properties:*

1. *h^{t+s} is a last informative history on the equilibrium path; and*
2. *any continuation history of $h^{t+s-1} = \{h^t, \hat{\sigma}_t, ..., \hat{\sigma}_{t+s-2}\}$ has no informative signals in periods $\tau > t + s$.*

Proof. Let h^t be an informative history on the equilibrium path. If $t = T$, if follows that both $\left\{h^T, H\right\}$ and $\left\{h^T, L\right\}$ are on the equilibrium path. Since these are terminal histories, they must also be a last informative history on the equilibrium path.

If $t < T$, the histories $\left\{h^t, H\right\}$ and $\left\{h^t, L\right\}$ must also be on the equilibrium path. There are two alternatives: either (i) both of them rationalize away all future signals on the equilibrium path, or (ii) one of them has a continuation history on the equilibrium path in which a one low signal is interpreted as low with positive probability.

If alternative (i) is true, the result from the lemma is immediate. Suppose, therefore, that alternative (ii) holds. Let h^{t_1} ($t_1 \geq t + 1$) denote the continuation history of either $\left\{h^t, H\right\}$ or $\left\{h^t, L\right\}$ on the equilibrium path in which a low signal is interpreted as low with positive probability ($\lambda\left(h^{t_1}\right) < 1$). Then, both $\left\{h^{t_1}, H\right\}$ and $\left\{h^{t_1}, L\right\}$ are also on the equilibrium path. Again, we have two alternatives: either (i) all continuation histories of both of them are uninformative, or (ii) one of them has a continuation history h^{t_2} on the equilibrium path in which $\lambda\left(h^{t_2}\right) < 1$, where $t_2 \geq t_1 + 1$. Again, the result from the lemma is immediate if (i) holds.

Proceeding by induction, it follows that either we reach an informative history in which all future signals are rationalized away, or we can construct a sequence of continuation histories $\{h^{t_n}\}_{n\in\mathbb{N}}$ with

$t_n \geq t_{n-1}+1$ in which all histories are informative. However, this contradicts the fact that history any terminal-period history must be uninformative (there are no new signals after period T) and, therefore, such sequence cannot have more than T elements. ■

Let h^t be a last informative history on the equilibrium path such that any continuation history of h^{t-1} has no informative signals in periods greater than t. Since h^t is informative, we must have either $\lambda\left(\left\{h^{t-1},H\right\}\right)<1$ or $\lambda\left(\left\{h^{t-1},L\right\}\right)<1$. There are three possible cases:

1. $\lambda\left(\left\{h^{t-1},H\right\}\right)<1$ and $\lambda\left(\left\{h^{t-1},L\right\}\right)<1$,
2. $\lambda\left(\left\{h^{t-1},H\right\}\right)<1$ and $\lambda\left(\left\{h^{t-1},L\right\}\right)=1$, and
3. $\lambda\left(\left\{h^{t-1},H\right\}\right)=1$ and $\lambda\left(\left\{h^{t-1},L\right\}\right)<1$.

As I show below, there exists a strictly positive bound (independent of T) on the difference in posterior means in all informative histories in each of these cases. The following lemmata will be useful throughout our analysis:

Lemma 11 *Suppose h^{t-1} is an informative history on the equilibrium path. Then*

$$E\left[\theta|\left\{h^{t-1},H\right\}\right]-E\left[\theta|\left\{h^{t-1},L\right\}\right]\geq E_{\hat{\sigma}}\left\{E\left[\theta|\left\{h^{t-1},H,\hat{\sigma}\right\}\right]-E\left[\theta|\left\{h^{t-1},L,\hat{\sigma}\right\}\right]\right\},$$

with strict inequality if and only if $\lambda\left(h^{t-1},H\right)<1$.

Proof. To simplify notation, I will omit h^{t-1} from the conditional distributions in the expressions below. Thus $\{H\}$ and $\{L\}$ will refer to histories $\{h^{t-1},H\}$ and $\{h^{t-1},L\}$, respectively.

If $\hat{\sigma}=H$ is played with probability 1 after history $\{H\}$, it follows that $E\left[\theta|\left\{H,\hat{\sigma}\right\}\right]=E\left[\theta|\left\{H\right\}\right]$. Moreover, since beliefs conditional on $\sigma=L$ are the same as beliefs conditional on $\hat{\sigma}=L$, by the law of iterated expectations, we have

$$E_{\hat{\sigma}}\left\{E\left[\theta|\left\{L\hat{\sigma}\right\}\right]\right\}=E\left[\theta|L\right]. \tag{27}$$

Thus, the condition holds with equality in this case. Suppose $\hat{\sigma}=H$ is played with a probability less than one after history $\{H\}$. Note that, in this case, $E_{\hat{\sigma}}\left\{E\left[\theta|\left\{H\hat{\sigma}\right\}\right]\right\}$ equals

$$\begin{aligned}
&\Pr\left(\hat{\sigma}_t=H|L\right)E\left[\theta|\left\{HH\right\}\right]+\Pr\left(\hat{\sigma}_t=L|L\right)E\left[\theta|\left\{HL\right\}\right]\\
=\;&E\left[\theta|\left\{HL\right\}\right]+\Pr\left(\hat{\sigma}_t=H|H\right)E\left[\theta|\left\{HH\right\}-E\left[\theta|\left\{HL\right\}\right]\right]\\
&+\left[\Pr\left(\hat{\sigma}_t=H|L\right)-\Pr\left(\hat{\sigma}_t=H|H\right)\right]E\left[\theta|\left\{HH\right\}-E\left[\theta|\left\{HL\right\}\right]\right].
\end{aligned}$$

By the law of iterated expectations,

$$E\left[\theta|\left\{HL\right\}\right]+\Pr\left(\hat{\sigma}_t=H|H\right)E\left[\theta|\left\{HH\right\}-E\left[\theta|\left\{HL\right\}\right]\right]=E\left[\theta|H\right].$$

Substituting back in the previous equation, gives

$$E_{\hat{\sigma}}\left\{E\left[\theta|\left\{H\hat{\sigma}\right\}\right]\right\}=E\left[\theta|H\right]-\left[\Pr\left(\hat{\sigma}_t=H|H\right)-\Pr\left(\hat{\sigma}_t=H|L\right)\right]\left\{E\left[\theta|\left\{HH\right\}\right]-E\left[\theta|\left\{HL\right\}\right]\right\}<E\left[\theta|H\right]. \tag{28}$$

The result then follows from expressions (27) and (28). ■

Lemma 12 $E\left(\theta|\left\{h^{t-1}H\right\}\right)-E\left[\theta|\left\{h^{t-1}L\right\}\right]<\frac{E\left(\theta|h^{t-1}\right)-E\left[\theta|\left\{h^{t-1}L\right\}\right]}{\underline{\theta}}$ *for any* h^{t-1} *on the equilibrium path.*

Proof. I will omit h^{t-1} from the conditional distributions in the expressions below. Thus $\{H\}$ and $\{L\}$ will refer to histories $\{h^{t-1},H\}$ and $\{h^{t-1},L\}$, respectively. Bayesian updating gives:

$$E\left(\theta|\hat{\sigma}_2=H\right)=\alpha E\left[\theta|\sigma=H\right]+\left(1-\alpha\right)E\left[\theta|\sigma=L\right],$$

and

$$E\left(\theta\right)=\pi E\left[\theta|\sigma=H\right]+\left(1-\pi\right)E\left[\theta|\sigma=L\right],$$

where $\pi\equiv\Pr\left(\sigma=H\right)$ and $\underline{\theta}<\pi<\alpha\leq 1$. Thus,

$$\frac{E\left(\theta|\hat{\sigma}_2=H\right)-E\left[\theta|\sigma=L\right]}{\alpha}=E\left[\theta|\sigma=H\right]-E\left[\theta|\sigma=L\right],$$

$$\frac{E\left(\theta\right)-E\left[\theta|\sigma=L\right]}{\pi}=E\left[\theta|\sigma=H\right]-E\left[\theta|\sigma=L\right],$$

Combining both equalities, we obtain:

$$E\left(\theta|\hat{\sigma}_2=H\right)-E\left[\theta|\sigma=L\right]=\frac{\alpha}{\pi}\left\{E\left(\theta\right)-E\left[\theta|\sigma=L\right]\right\}<\frac{E\left(\theta\right)-E\left[\theta|\sigma=L\right]}{\underline{\theta}},$$

where the inequality uses $\alpha<1$ and $\pi>\underline{\theta}\therefore\frac{1}{\pi}<\frac{1}{\underline{\theta}}$. ■

Case 1.

In case 1, there are three informative histories to consider: h^{t-1}, $\left\{h^{t-1},L\right\}$, and $\left\{h^{t-1},H\right\}$. Lemma (1) gives:

$$E\left(\theta|\left\{h^{t-1},H,H\right\}\right)-E\left(\theta|\left\{h^{t-1},H,L\right\}\right)>\frac{\alpha}{\beta\kappa},\text{ and}$$

$$E\left(\theta|\left\{h^{t-1},L,H\right\}\right)-E\left(\theta|\left\{h^{t-1},L,L\right\}\right)>\frac{\alpha}{\beta\kappa}.$$

For history h^{t-1}, the result follows from the Lemma 11, which concludes the analysis of case 1.

In my analysis of cases 2 and 3, I will use the following decomposition of the payoff from actions:

$$E\left\{V\left(E\left[\theta|h^t\right]+\delta,\theta\right)|h^t\right\}=-\kappa\left[\mathrm{Var}\left(\theta|h^t\right)+\delta^2\right]\quad\forall\delta\in\mathbb{R},$$

which is obtained by straightforward algebraic manipulations of the quadratic payoff function. This decomposition shows that the payoff from actions depends additively on the variance of beliefs $(-\kappa\mathrm{Var}(\theta|h^t))$ and the square of the distortion (δ^2).

Case 2.

In case 2, all continuation histories of $\{h^{t-1}, L\}$ interpret low signals as high with probability one and, therefore, beliefs remain constant after reaching $\{h^{t-1}, L\}$. Since a low signal is interpreted realistically with positive probability at $\{h^{t-1}, H\}$, beliefs conditional on $\{h^{t-1}, H, L\}$ differ from those conditional on $\{h^{t-1}, H, H\}$. However, beliefs remain constant after reaching both $\{h^{t-1}, H, L\}$ and $\{h^{t-1}, H, H\}$. Thus, we need to obtain bounds for the two informative histories: h^{t-1} and $\{h^{t-1}, H\}$. For $\{h^{t-1}, H\}$, the result follows directly from Lemma (1):

$$E\left[\theta|\left\{h^{t-1}HH\right\}\right] - E\left[\theta|\left\{h^{t-1}HL\right\}\right] > \frac{\alpha}{\beta\kappa}. \tag{29}$$

Below, we verify that it also holds for h^{t-1}.

To simplify notation, I will omit h^{t-1} from the conditional distributions in the expressions below. Hence, I will write $\Pr(\hat{\sigma}_t = H|\hat{\sigma})$ for $\Pr\left(\hat{\sigma}_t = H|\left\{h^{t-1}, \hat{\sigma}\right\}\right)$, $E(\theta|\hat{\sigma})$ for $E\left(\theta|\left\{h^{t-1}, \hat{\sigma}\right\}\right)$, and so on.

The expected payoff from playing $\hat{\sigma} = L$ after observing a low signal at h^{t-1} is

$$\alpha\left(\frac{1-\beta^{T-t+1}}{1-\beta}\right) E[\theta|L] - \kappa\beta^{T-t+1}\mathrm{Var}(\theta|L). \tag{30}$$

The expected payoff from playing $\hat{\sigma} = H$ after observing a low signal at h^{t-1} is

$$\alpha\left\{E[\theta|H] + \beta\frac{1-\beta^{T-t}}{1-\beta}\left[\Pr(\hat{\sigma}_t = H|L)\, E[\theta|\{HH\}] + \Pr(\hat{\sigma}_t = L|L)\, E[\theta|\{HL\}]\right]\right\}$$

$$-\kappa\beta^{T-t+1}\left\{\begin{array}{l}\Pr(\hat{\sigma}_t = H|L)\left\{[E(\theta|\{HH\}) - E(\theta|\{LH\})]^2 + \mathrm{Var}(\theta|\{LH\})\right\} \\ +\Pr(\hat{\sigma}_t = L|L)\left\{[E(\theta|\{HL\}) - E(\theta|\{LL\})]^2 + \mathrm{Var}(\theta|\{LL\})\right\}\end{array}\right\},$$

where the second line uses the payoff from actions decomposition.

Notice that, by the law of iterated expectations,

$$\Pr(\hat{\sigma}_t = H|L)\, E[\theta|\{HH\}] + \Pr(\hat{\sigma}_t = L|L)\, E[\theta|\{HL\}]$$

$$= E[\theta|H] - [\Pr(H|H) - \Pr(H|L)]\{E[\theta|\{HH\}] - E[\theta|\{HL\}]\}.$$

Thus, beliefs after playing $\hat{\sigma} = H$ in an informative period follow a supermartingale: $E_{\hat{\sigma}}[E(\theta|H\hat{\sigma})|H] < E[\theta|H]$. Intuitively, the current-period self expects the beliefs of future selves to, on average, be revised downwards after a deviation. Moreover, by the law of total variance,

$$E_{\hat{\sigma}_t}[\mathrm{Var}(\theta|\{L\hat{\sigma}_t\})|L] = \mathrm{Var}(\theta|L) - \mathrm{Var}_{\hat{\sigma}_t}[E(\theta|\{L\hat{\sigma}_t\})|L].$$

Using the definition of variance, we can show that

$$\mathrm{Var}_{\hat{\sigma}_t}[E(\theta|\{L\hat{\sigma}_t\})|L] = \Pr(\hat{\sigma}_t = H|L)\Pr(\hat{\sigma}_t = L|L)[E(\theta|\{LH\}) - E(\theta|\{LL\})]^2.$$

Substituting back on the expected payoff from playing $\hat{\sigma} = H$ gives

$$\alpha E[\theta|H]\left(\frac{1-\beta^{T-t+1}}{1-\beta}\right) - \alpha\frac{\beta-\beta^{T-t+1}}{1-\beta}\left[\Pr(\hat{\sigma}_t = H|H) - \Pr(\hat{\sigma}_t = H|L)\right]\left\{E[\theta|\{HH\}] - E[\theta|\{HL\}]\right\}$$

$$-\kappa\beta^{T-t+1}\left\{\begin{array}{c}\Pr(\hat{\sigma}_t = H|L)\left\{[E(\theta|\{HH\}) - E(\theta|\{LH\})]^2\right\} \\ +\Pr(\hat{\sigma}_t = L|L)\left\{[E(\theta|\{HL\}) - E(\theta|\{LL\})]^2\right\} \\ +\mathrm{Var}(\theta|L) - \Pr(\hat{\sigma}_t = H|L)\Pr(\hat{\sigma}_t = L|L)[E(\theta|\{LH\}) - E(\theta|\{LL\})]^2\end{array}\right\}. \tag{31}$$

Recall that, in equilibrium, the individual plays L at h^{t-1} with positive probability. Thus, the payoff from H cannot exceed the payoff from playing L. Using the expressions for these payoffs from (30) and (31), we can write the equilibrium condition as

$$\alpha\left(\frac{1-\beta^{T-t+1}}{1-\beta}\right)\{E[\theta|H] - E[\theta|L]\} - \alpha\frac{\beta-\beta^{T-t+1}}{1-\beta}\left[\Pr(\hat{\sigma}_t = H|H) - \Pr(\hat{\sigma}_t = H|L)\right]\{E[\theta|\{HH\}] - E[\theta|\{HL\}]\}$$

$$\leq \kappa\beta^{T-t+1}\left\{\begin{array}{c}\overbrace{\Pr(\hat{\sigma}_t = H|L)[E(\theta|\{HH\}) - E(\theta|\{LH\})]^2 + \Pr(\hat{\sigma}_t = L|L)[E(\theta|\{HL\}) - E(\theta|\{LL\})]^2}^{(\text{Bias})^2} \\ -\underbrace{\Pr(\hat{\sigma}_t = H|L)\Pr(\hat{\sigma}_t = L|L)[E(\theta|\{LH\}) - E(\theta|\{LL\})]^2}_{\text{Value of Information}}\end{array}\right\}. \tag{32}$$

The expression on the first line are the self-image gain from playing $\hat{\sigma} = H$ instead of $\hat{\sigma} = L$. The terms terms on the second line are the expected loss from the bias in posterior beliefs minus the informational gain from the deviation. Next, we establish that the benefit from the deviation is $O([E(\theta|\{H\}) - E(\theta|\{L\})])$ whereas the cost is $O\left(\left[E(\theta|\{H\}) - E(\theta|\{L\})^2\right]\right)$.

Lemma 13 $E(\theta|\{HH\}) - E(\theta|\{HL\}) \leq \frac{1}{\underline{\theta}} \times [E(\theta|\{H\}) - E(\theta|\{L\})]$.

Proof. For notational simplicity, let $p \equiv \Pr(\hat{\sigma}_t = H|\hat{\sigma}_{t-1} = H)$. Then,

$$E[\theta|\{H\}] = pE[\theta|\{HH\}] + (1-p)E[\theta|\{HL\}]$$

$$\therefore [E(\theta|\{HH\}) - E[\theta|\{HL\}]] = \frac{E[\theta|\{H\}] - E[\theta|\{HL\}]}{p} < \frac{E[\theta|\{H\}] - E[\theta|\{L\}]}{\underline{\theta}},$$

where the inequality uses $\Pr(\hat{\sigma}_t = H|\hat{\sigma}_{t-1} = H) > \underline{\theta}$ and $E[\theta|\{HL\}] > E[\theta|\{L\}]$. ■

Lemma 14 $E(\theta|\{HL\}) - E(\theta|\{LL\}) \leq \frac{1}{1-\lambda(h^{t-1})} \times \frac{1-\underline{\theta}}{(1-\bar{\theta})^2} \times [E(\theta|\{H\}) - E(\theta|\{L\})]$.

Proof. For notational simplicity, let $p \equiv \Pr(\hat{\sigma}_{t-1} = H|\hat{\sigma}_t = L)$ and note that

$$E[\theta|\hat{\sigma}_t = L] = E[\theta|\hat{\sigma}_{t-1} = L] = E[\theta|\sigma_{t-1} = L].$$

Then, using the definition of conditional expectations, we have

$$E[\theta|\{L\}] = pE[\theta|\{HL\}] + (1-p)E[\theta|\{LL\}]$$

$$\therefore E\left[\theta|\{HL\}\right]-E\left[\theta|\{L\}\right]=(1-p)\left\{E\left[\theta|\{HL\}\right]-E\left[\theta|\{LL\}\right]\right\}.$$

Using the fact that $E\left[\theta|\{HL\}\right]<E\left[\theta|\{H\}\right]$ and $1-p$

$$\therefore \left\{E\left[\theta|\{HL\}\right]-E\left[\theta|\{LL\}\right]\right\}<\frac{E\left[\theta|\{H\}\right]-E\left[\theta|\{L\}\right]}{1-p}.$$

But. by Bayes' rule,

$$1-p=\Pr\left(\hat{\sigma}_{t-1}=L|\hat{\sigma}_{t}=L\right)=(1-\lambda_{t-1})\frac{E\left[(1-\theta)^{2}\right]}{(1-E\left[\theta\right])},$$

concluding the proof. ■

Lemma 15 $E\left(\theta|\{HH\}\right)-E\left(\theta|\{LH\}\right)\leq\frac{1}{\underline{\theta}^{2}}\times\left[E\left(\theta|\{H\}\right)-E\left(\theta|\{L\}\right)\right]$.

Proof. Let $p\equiv\Pr\left(\hat{\sigma}_{t-1}=H|\hat{\sigma}_{t}=H\right)$. Then, by the definition of conditional expectations,

$$E\left(\theta|\hat{\sigma}_{t}=H\right)=pE\left(\theta|\{HH\}\right)+(1-p)\,E\left(\theta|\{LH\}\right).$$

Rearrange this expression as:

$$E\left(\theta|\{HH\}\right)-E\left(\theta|\{LH\}\right)=\frac{E\left(\theta|\hat{\sigma}_{t}=H\right)-E\left(\theta|\{LH\}\right)}{p}<\frac{E\left(\theta|\hat{\sigma}_{t}=H\right)-E\left(\theta|\{L\}\right)}{p},\tag{33}$$

where the inequality uses the fact that $E\left(\theta|\{LH\}\right)>E\left(\theta|\{L\}\right)$. By Bayes' rule, we have, for $s\in\{t,t-1\}$,

$$E\left(\theta|\hat{\sigma}_{s}=H\right)=\Pr\left(\sigma_{s}=H|\hat{\sigma}_{s}=H\right)E\left(\theta|\sigma_{s}=H\right)+\left[1-\Pr\left(\sigma_{s}=H|\hat{\sigma}_{s}=H\right)\right]E\left(\theta|\sigma_{s}=L\right).$$

Using this expression for both $s=t$ and $s=t-1$, we obtain:

$$E\left(\theta|\hat{\sigma}_{t}=H\right)-E\left(\theta|\sigma_{t}=L\right)=\frac{\Pr\left(\sigma_{t}=H|\hat{\sigma}_{t}=H\right)}{\Pr\left(\sigma_{t-1}=H|\hat{\sigma}_{t-1}=H\right)}\left[E\left(\theta|\hat{\sigma}_{t-1}=H\right)-E\left(\theta|\sigma_{t-1}=L\right)\right].$$

Plugging back in (33), yields

$$E\left(\theta|\{HH\}\right)-E\left(\theta|\{LH\}\right)<\frac{\Pr\left(\sigma_{t}=H|\hat{\sigma}_{t}=H\right)\left[E\left(\theta|\hat{\sigma}_{t-1}=H\right)-E\left(\theta|\sigma_{t-1}=L\right)\right]}{\Pr\left(\hat{\sigma}_{t-1}=H|\hat{\sigma}_{t}=H\right)\Pr\left(\sigma_{t-1}=H|\hat{\sigma}_{t-1}=H\right)}.\tag{34}$$

Note that

$$\Pr\left(\sigma_{s}=H|\hat{\sigma}_{s}=H\right)=\frac{\Pr\left(\sigma_{s}=H\right)}{\Pr\left(\sigma_{s}=H\right)+\lambda_{s}\left[1-\Pr\left(\sigma_{s}=H\right)\right]}\in\left(\frac{\underline{\theta}}{\underline{\theta}+(1-\underline{\theta})\,\lambda},1\right],$$

where λ_s denotes the probability of playing $\hat{\sigma}=H$ at history h^{s} and we used the fact that $p(\theta)=\theta>\underline{\theta}$. Moreover,

$$\Pr\left(\hat{\sigma}_{t-1}=H|\hat{\sigma}_{t}=H\right)>\underline{\theta},$$

since the lowest probability of a high interpretation is obtained if all low signals are interpreted realistically and, in that case, the probability of a high interpretation is bounded below by $\underline{\theta}$. Combining both bounds, we obtain

$$\frac{\Pr\left(\sigma_t = H|\hat{\sigma}_t = H\right)}{\Pr\left(\hat{\sigma}_{t-1} = H|\hat{\sigma}_t = H\right) \times \Pr\left(\sigma_{t-1} = H|\hat{\sigma}_{t-1} = H\right)} < \frac{\underline{\theta} + \left(1 - \underline{\theta}\right)\lambda_{t-1}}{\underline{\theta}^2} < \frac{1}{\underline{\theta}^2}.$$

Substitute back in (34) yields

$$E\left(\theta|\left\{HH\right\}\right) - E\left(\theta|\left\{LH\right\}\right) < \frac{\left[E\left(\theta|H\right) - E\left(\theta|L\right)\right]}{\underline{\theta}^2},$$

which concludes the proof. ■

For notational simplicity, let $\Delta \equiv E\left[\theta|H\right] - E\left[\theta|L\right]$. Using Lemmata (13)-(15) and the no-deviation condition (32), we obtain the following necessary condition for the equilibrium to exist:

$$\alpha\left(\frac{1-\beta^{T-t+1}}{1-\beta}\right)\Delta < \alpha\frac{\beta-\beta^{T-t+1}}{1-\beta}\left[\Pr\left(\hat{\sigma}_t = H|H\right) - \Pr\left(\hat{\sigma}_t = H|L\right)\right]\frac{\Delta}{\underline{\theta}}$$
$$+\kappa\beta^{T-t+1}\left\{\frac{\Pr\left(\hat{\sigma}_t = H|L\right)}{\underline{\theta}^4} + \frac{\Pr\left(\hat{\sigma}_t = L|L\right)}{\left[1-\lambda(h^{t-1})\right]^2}\frac{\left(1-\underline{\theta}\right)^2}{\left(1-\bar{\theta}\right)^4}\right\}\Delta^2. \tag{35}$$

By Bayes' rule,

$$\Pr\left(\hat{\sigma}_t = H|H\right) = \Pr\left(\sigma_t = H|H\right) + \Pr\left(\sigma_t = L|H\right)\left[1-\lambda\left(h^t\right)\right], \tag{36}$$

and

$$\Pr\left(\hat{\sigma}_t = H|L\right) = \Pr\left(\sigma_t = H|L\right) + \Pr\left(\sigma_t = L|L\right)\left[1-\lambda\left(h^t\right)\right]. \tag{37}$$

Combining expressions (36) and (37), we obtain

$$\begin{array}{ll}\Pr\left(\hat{\sigma}_t = H|H\right) - \Pr\left(\hat{\sigma}_t = H|L\right) & = \left[\Pr\left(\sigma_t = H|H\right) - \Pr\left(\sigma_t = H|L\right)\right]\lambda\left(h^t\right) \\ & = \left[E\left(\theta|H\right) - E\left(\theta|L\right)\right]\lambda\left(h^t\right) \leq \Delta.\end{array},$$

where the inequality uses the definition of Δ and the fact that $\lambda(h^t) \leq 1$. Substituting back in inequality (35), we obtain the necessary condition for the equilibrium to exist:

$$\alpha\left(\frac{1-\beta^{T-t+1}}{1-\beta}\right) < \left\{\frac{\alpha}{\underline{\theta}}\frac{\beta-\beta^{T-t+1}}{1-\beta} + \kappa\beta^{T-t+1}\left[\frac{\Pr\left(\hat{\sigma}_t = H|L\right)}{\underline{\theta}^4} + \frac{\Pr\left(\hat{\sigma}_t = L|L\right)}{\left(1-\lambda(h^{t-1})\right)^2}\frac{\left(1-\underline{\theta}\right)^2}{\left(1-\bar{\theta}\right)^4}\right]\right\}\Delta.$$

There are two possibilities depending on whether $1-\lambda(h^{t-1})$ is uniformly bounded away from zero. Formally, let a game be parameterized by the number of signals T. Suppose that there exists a constant $K < 1$ such that, for any number of signals $T \in \mathbb{N}$, $\lambda(h^{t-1}) < K$ in any history h^{t-1} satisfying the conditions of Case 2 (i.e., $\left\{h^{t-1}, H\right\}$ is a last informative history and all continuation histories of $\left\{h^{t-1}, L\right\}$ are uninformative).

Then, the previous inequality implies that, for any such history in a game with T signals,

$$\alpha\left(\frac{1-\beta^{T-t+1}}{1-\beta}\right)<\left\{\frac{\alpha}{\underline{\theta}}\frac{\beta-\beta^{T-t+1}}{1-\beta}+\kappa\beta^{T-t+1}\left[\frac{\Pr\left(\hat{\sigma}_t=H|L\right)}{\underline{\theta}^4}+\frac{\Pr\left(\hat{\sigma}_t=L|L\right)}{\left(1-K\right)^2}\frac{\left(1-\underline{\theta}\right)^2}{\left(1-\bar{\theta}\right)^4}\right]\right\}\Delta, \quad (38)$$

where, as before, we omit h^{t-1} for notational simplicity. Note that $\frac{1-\beta^{T-t+1}}{1-\beta}\geq 1$, $\frac{\beta-\beta^{T-t+1}}{1-\beta}\leq\frac{\beta}{1-\beta}$, $\Pr\left(\hat{\sigma}_t=H|L\right)<1$ and $\frac{\Pr(\hat{\sigma}_t=L|L)}{(1-K)^2}<\frac{\underline{\theta}}{(1-K)^2}$. Thus, a necessary condition for the equilibrium to exist is

$$\Delta>\frac{\alpha}{\frac{\alpha}{\underline{\theta}}\frac{\beta}{1-\beta}+\kappa\beta\left[\frac{1}{\underline{\theta}^4}+\frac{\underline{\theta}}{(1-K)^2}\frac{\left(1-\underline{\theta}\right)^2}{\left(1-\bar{\theta}\right)^4}\right]},$$

which is a uniform bound on Δ. Thus, when a uniform bound on $1-\lambda(h^{t-1})$ exists, there also exists a uniform bound on $\Delta=E\left[\theta|\left\{H\right\}\right]-E\left[\theta|\left\{L\right\}\right]$.

Suppose, in order to obtain a contradiction, that no bound on $\Delta=E\left[\theta|\left\{H\right\}\right]-E\left[\theta|\left\{L\right\}\right]$ holds uniformly across all histories h^{t-1} satisfying the conditions on Case 2 for all number of signals $T\in\mathbb{N}$. Then, for any $\epsilon>0$, there exists a game (parameterized by T), and a history h^{t-1} satisfying the conditions of Case 2 such that

$$E\left[\theta|\left\{h^{t-1}H\right\}\right]-E\left[\theta|\left\{h^{t-1}L\right\}\right]<\epsilon.$$

Taking $\epsilon=\frac{1}{N}$, we can construct a sequence $\left\{\Delta_N\right\}_{N\in\mathbb{N}}$ such that

$$\Delta_N=E\left[\theta|\left\{h_N^{t-1}H\right\}\right]-E\left[\theta|\left\{h_N^{t-1}L\right\}\right]<\frac{1}{N}$$

for some Case-2 history h_N^{t-1} on the equilibrium path. Since $\Delta_N>0$ for all N, it follows from the Squeeze Theorem that $\left\{\Delta_N\right\}_{N\in\mathbb{N}}$ converges to zero. By inequality (38), this requires $\left\{\lambda(h_N^{t-1})\right\}_{N\in\mathbb{N}}$ converges to zero. Thus,

$$\lim_{N\to\infty}E\left[\theta|\left\{h_N^{t-1}H\right\}\right]-E\left[\theta|\left\{h_N^{t-1}L\right\}\right]=\lim_{N\to\infty}E\left[\theta|h_N^{t-1}\right]-E\left[\theta|\left\{h_N^{t-1}L\right\}\right]=0. \quad (39)$$

However, by inequality (29), we must have

$$E\left[\theta|\left\{h_N^{t-1}HH\right\}\right]-E\left[\theta|\left\{h_N^{t-1}HL\right\}\right]>\frac{\alpha}{\beta\kappa}$$

for all h_N^{t-1} (since h_N^{t-1} satisfies the conditions of Case 2, history $\left\{h^{t-1},H\right\}$ must be informative). Because $\lambda(h_N^{t-1})$ converges to zero, we have

$$\lim_{N\to\infty}\left[E\left(\theta|\left\{h_N^{t-1}HH\right\}\right)-E\left(\theta|\left\{h_N^{t-1}H\right\}\right)\right]=0,$$

and

$$\lim_{N\to\infty}\left[E\left(\theta|\left\{h_N^{t-1}HL\right\}\right)-E\left(\theta|\left\{h_N^{t-1}L\right\}\right)\right]=0.$$

Thus, there exists $\bar{N}$ such that, for all $N > \bar{N}$,

$$E\left[\theta|\left\{h_N^{t-1}H\right\}\right] - E\left[\theta|\left\{h_N^{t-1}L\right\}\right] > \frac{\alpha}{\beta\kappa}.$$

Then, using Lemma 12, we obtain

$$E\left(\theta|h_N^{t-1}\right) - E\left[\theta|\left\{h_N^{t-1}L\right\}\right] > \frac{\alpha\underline{\theta}}{\beta\kappa}$$

for all $N > \bar{N}$, which contradicts equation (39). Thus, we have obtained a bound on $E\left(\theta|\left\{h^{t-1}H\right\}\right) - E\left[\theta|\left\{h^{t-1}L\right\}\right]$ for all Case-2 histories that is uniform across equilibria and number of signals T.

Case 3.

In case 3, there are two informative histories for which we need to verify the bounds on the mean of the posterior: h^{t-1} and $\left\{h^{t-1}, L\right\}$. The result for history $\left\{h^{t-1}, L\right\}$ follows straight from the first lemma:

$$E\left[\theta|\left\{h^{t-1}LH\right\}\right] - E\left[\theta|\left\{h^{t-1}LL\right\}\right] > \frac{\alpha}{\beta\kappa}. \tag{40}$$

As in our analysis of case 2, I will omit h^{t-1} from the continuation histories for notational simplicity. Thus, in the expressions above, I will refer to history $\{h^{t-1}, L\}$ as $\{L\}$ and I will refer to history $\{h^{t-1}, H\}$ as $\{H\}$.

At h^{t-1}, the payoff from playing $\hat{\sigma} = H$ is

$$\alpha\frac{1-\beta^{T+1-(t-1)}}{1-\beta}E\left[\theta|\{H\}\right] - \kappa\beta^{T+1-(t-1)}\left\{\mathrm{Var}\left(\theta|\{L\}\right) + \left[E\left(\theta|\{H\}\right) - E\left(\theta|\{L\}\right)\right]^2\right\}.$$

The payoff from playing $\hat{\sigma} = L$ is

$$\alpha\frac{1-\beta^{T+1-(t-1)}}{1-\beta}E\left[\theta|\left\{L\right\}\right] - \kappa\beta^{T+1-(t-1)}\left[\Pr\left(\hat{\sigma}_t = H|L\right)\mathrm{Var}\left(\theta|\{LH\}\right) + \Pr\left(\hat{\sigma}_t = L|L\right)\mathrm{Var}\left(\theta|\{LL\}\right)\right]. \tag{41}$$

Let $\Delta_1 \equiv E\left(\theta|H\right) - E\left(\theta|L\right)$ and $\Delta_2 \equiv E\left[\theta|\left\{LH\right\}\right] - E\left[\theta|\left\{LL\right\}\right]$. By the Law of Total Variance,

$$\mathrm{Var}(\theta|L) - E_{\hat{\sigma}}\left[\mathrm{Var}\left(\theta|\{L\hat{\sigma}\}\right)\right] = \mathrm{Var}_{\hat{\sigma}}\left[E\left(\theta|\{L\hat{\sigma}\}\right)\right].$$

Using the definition of variance, we obtain

$$\begin{aligned}\mathrm{Var}_{\hat{\sigma}}\left[E\left(\theta|\left\{L,\hat{\sigma}\right\}\right)|L\right] &= \Pr\left(\hat{\sigma} = H|L\right)\left[E\left(\theta|\left\{LH\right\}\right) - \left(E\left(\theta|\left\{LL\right\}\right) + \Pr\left(\hat{\sigma} = H|L\right)\Delta_2\right)\right]^2 \\ &\quad + \Pr\left(\hat{\sigma} = L|L\right)\left[E\left(\theta|\left\{LL\right\}\right) - \left(E\left(\theta|\left\{LL\right\}\right) + \Pr\left(\hat{\sigma} = H|L\right)\Delta_2\right)\right]^2 \\ &= \Pr\left(\hat{\sigma} = H|L\right)\Pr\left(\hat{\sigma} = L|L\right)\Delta_2^2.\end{aligned}$$

Hence,

$$\Pr\left(\hat{\sigma}_t = H|L\right)\mathrm{Var}\left(\theta|\{LH\}\right) + \Pr\left(\hat{\sigma}_t = L|L\right)\mathrm{Var}\left(\theta|\{LL\}\right) = \mathrm{Var}\left(\theta|L\right) - \Pr\left(\hat{\sigma} = H|L\right)\Pr\left(\hat{\sigma} = L|L\right)\Delta_2^2. \tag{42}$$

Substituting (42) in equation (41), we can write the payoff from playing $\hat{\sigma} = L$ as

$$\alpha \frac{1 - \beta^{T+2-t}}{1 - \beta} E\left[\theta | \{L\}\right] - \kappa \beta^{T+2-t} \left[\mathrm{Var}\left(\theta | \{L\}\right) - \Pr\left(\hat{\sigma} = H | \{L\}\right) \Pr\left(\hat{\sigma} = L | \{L\}\right) \Delta_2^2\right].$$

Since, the equilibrium prescribed playing $\hat{\sigma} = L$ with positive probability at history h^{t-1}, the payoff from playing $\hat{\sigma} = H$ cannot exceed the payoff from playing $\hat{\sigma} = L$:

$$\alpha \frac{1 - \beta^{T+2-t}}{1 - \beta} \Delta_1 \leq \kappa \beta^{T+2-t} \left\{\Delta_1^2 + \Pr\left(\hat{\sigma} = H | \{L\}\right) \Pr\left(\hat{\sigma} = L | \{L\}\right) \Delta_2^2\right\}. \tag{43}$$

Note that, by the law of iterated expectations,

$$E\left[\theta | \{L\}\right] = \Pr\left(\hat{\sigma} = H | L\right) E\left[\theta | \{LH\}\right] + \Pr\left(\hat{\sigma} = L | L\right) E\left[\theta | \{LL\}\right],$$

which can be rearranged as

$$\Pr\left(\hat{\sigma} = L | L\right) \left\{E\left[\theta | \{LH\}\right] - E\left[\theta | \{LL\}\right]\right\} = E\left[\theta | \{LH\}\right] - E\left[\theta | \{L\}\right] < E\left[\theta | \{H\}\right] - E\left[\theta | \{L\}\right].$$

Thus,

$$\left[E\left[\theta | \{LH\}\right] - E\left[\theta | \{LL\}\right]\right] < \frac{E\left[\theta | \{H\}\right] - E\left[\theta | \{L\}\right]}{\Pr(\hat{\sigma} = L | \{L\})} = \frac{\Delta_1}{\Pr(\hat{\sigma} = L | \{L\})}.$$

Moreover, $\frac{1-\beta^{T+2-t}}{1-\beta} > 1 + \beta$, and $\beta^{T+2-t} < \beta^2$. Substituting back in (43), we obtain the following necessary condition for an equilibrium:

$$\alpha (1 + \beta) \leq \kappa \beta^2 \left[1 + \frac{\Pr\left(\hat{\sigma} = H | \{L\}\right)}{\Pr(\hat{\sigma} = L | \{L\})}\right] \Delta_1.$$

Recall also that

$$\Pr\left(\hat{\sigma}_t = H | L\right) = \Pr\left(\sigma_t = H | L\right) + \Pr\left(\sigma_t = L | L\right) \lambda\left(h^t\right) < 1,$$

$$\Pr\left(\hat{\sigma}_t = L | L\right) = \Pr\left(\sigma_t = L | L\right) \left[1 - \lambda\left(h^t\right)\right] > \underline{\theta} \left[1 - \lambda\left(h^t\right)\right].$$

Thus, the following condition is necessary for an equilibrium:

$$\alpha (1 + \beta) \leq \kappa \beta^2 \left(1 + \frac{1}{\underline{\theta}\left[1 - \lambda\left(h^t\right)\right]}\right) \Delta_1. \tag{44}$$

As in Case 2, there are two possibilities depending on whether $\lambda\left(h^t\right)$ is or is not uniformly bounded away from one. Suppose that there exists a constant $K < 1$ such that, for any number of signals $T \in \mathbb{N}$, $\lambda(h^t) \leq K$ in any history h^t satisfying the conditions of Case 3. Then, (44) yields

$$\frac{\alpha}{\kappa \beta^2} \frac{1 + \beta}{1 + \frac{1}{\underline{\theta}(1-K)}} \leq \Delta_1,$$

which is a uniform bound on $\Delta_1 \equiv E\left(\theta | H\right) - E\left(\theta | L\right)$.

Suppose, in order to obtain a contradiction, that no bound on $\Delta_1 \equiv E\left(\theta | H\right) - E\left(\theta | L\right)$ holds uniformly across all histories h^{t-1} satisfying the conditions on Case 3 for all number of signals $T \in \mathbb{N}$. Then, for any $\epsilon > 0$, there exists a game (parameterized by T), and a history h^{t-1} satisfying the conditions of Case 3 such that

$$E\left[\theta \middle| \left\{h^{t-1}H\right\}\right] - E\left[\theta \middle| \left\{h^{t-1}L\right\}\right] < \epsilon.$$

Taking $\epsilon = \frac{1}{N}$, we can construct a sequence $\{\Delta_{1_N}\}_{N\in\mathbb{N}}$ such that

$$\Delta_{1_N} = E\left[\theta \middle| \left\{h_N^{t-1}H\right\}\right] - E\left[\theta \middle| \left\{h_N^{t-1}L\right\}\right] < \frac{1}{N}$$

for a Case-3 history h_N^{t-1} on the equilibrium path.

Since $\Delta_{1_N} > 0$ for all N, by the Squeeze Theorem, $\{\Delta_{1_N}\}_{N\in\mathbb{N}}$ must converge to zero. By inequality (38), this requires $\left\{\lambda(h_N^{t-1})\right\}_{N\in\mathbb{N}}$ converges to zero. Recall that, from inequality (44), the following condition is necessary for an equilibrium:

$$\alpha\left(1+\beta\right)\Delta_1 \leq \kappa\beta^2\left\{\Delta_1^2 + \Pr\left(\hat{\sigma} = L \middle| \left\{h_N^t\right\}\right)\Delta_2^2\right\}. \tag{45}$$

Notice that $\Pr\left(\hat{\sigma}_t = L|L\right) = \Pr\left(\sigma_t = L|L\right)\left[1-\lambda\left(h^t\right)\right] < \bar{\theta}\left[1-\lambda\left(h^t\right)\right]$ converges to zero as $\lambda\left(h^t\right)$ approaches 1. Then, since Δ_2^2 is bounded, condition (45) implies that there exists $\bar{N}$ such that $N > \bar{N}$ implies

$$\frac{\alpha\left(1+\beta\right)}{\kappa\beta^2} \leq \Delta_{1_N},$$

which contradicts the fact that $\{\Delta_{1_N}\}_{N\in\mathbb{N}}$ converges to zero. Thus, there exists a uniform bound on Δ_1 that holds across all Case-3 histories in any equilibrium and for any number of signals T.

Proceeding by backward induction establishes that there is a uniform bound on the distance in means $E[\theta|\left\{h^{t-s}, H\right\}] - E[\theta|\left\{h^{t-s}, L\right\}]$ for all histories s periods away from the last informative period.