Paying (for) Attention: The Impact of Information Processing Costs on Bayesian Inference*

Scott Duke Kominers[†] Xiaosheng Mu[‡] Alexander Peysakhovich[§]

January 24, 2019

Abstract

Human information processing is often modeled as costless Bayesian inference. However, research in psychology shows that attention is a computationally costly and potentially limited resource. We thus study Bayesian agents for whom computing posterior beliefs is costly; such agents face a tradeoff between economizing on attention costs and having more accurate beliefs. We show that even small processing costs can lead to significant departures from the standard costless processing model. There exist situations in which beliefs can cycle persistently and never converge. In addition, when updating is costly, agents are more sensitive to signals about rare events than to signals about common events. Thus, these individuals can permanently overestimate the likelihood of rare events. There is a commonly held assumption in economics that individuals will converge to correct beliefs/optimal behavior given sufficient experience. Our results contribute to a growing literature in psychology, neuroscience, and behavioral economics suggesting that this assumption is both theoretically and empirically fragile.

*The authors appreciate the helpful comments of Daniel Benjamin, Alessandro Bonatti, Roland Fryer, Drew Fudenberg, Martin Nowak, Larry Samuelson, Andrei Shleifer, Jakub Steiner, Juuso Toikka, and seminar audiences at Harvard and MIT. Kominers appreciates the hospitality of the Harvard Program for Evolutionary Dynamics and Microsoft Research New England, which hosted him during parts of this research, and gratefully acknowledges the support of NSF Grants CCF-1216095 and SES-1459912, the Harvard Milton Fund, an AMS–Simons travel grant, the Ng Fund and the Mathematics in Economics Research Fund of the Harvard Center of Mathematical Sciences and Applications, and the Human Capital and Economic Opportunity Working Group sponsored by the Institute for New Economic Thinking. Some of this research was conducted while Kominers was a Research Scholar at the Becker Friedman Institute for Research in Economics at the University of Chicago and Peysakhovich was a Research Scientist at the Harvard Program for Evolutionary Dynamics.

[†]Harvard Business School, Department of Economics, Center of Mathematical Sciences and Applications, and Center for Research on Computation and Society, Harvard University, and National Bureau of Economic Research.

[‡]Cowles Foundation and Columbia University.

[§]Human Cooperation Lab, Yale University.

1 Introduction

Modeling human information processing as Bayesian is standard across many fields. Cognitive scientists have modeled visual perception as a form of Bayesian inference (Knill and Richards (1996)); psychologists have argued that babies are "intuitive Bayesians" (Gopnik (2010)); neuroscientists have attempted to investigate how the brain would perform Bayesian inferences (Doya (2007)); and economists have used Bayesian models to justify many assumptions on human behavior (Epstein and Le Breton (1993)).

The Bayesian actor model assumes not only an ability to update beliefs, but an ability to do so costlessly. In this paper, we examine how Bayesian behavior changes if the costless processing assumption is relaxed; we find that this small departure from standard Bayesian information processing models can lead to significant changes in behavior.

We consider an agent who is uncertain about the way the world works. The agent is a Bayesian, and begins with a prior probability distribution over possible models of the world. The agent receives signals about the true model of the world, and can choose whether to internalize signal information by updating his prior via Bayesian inference. Updating is costly, but internalizing signal information is valuable because it helps the agent solve asset allocation problems that arise each period.[1]

In the single-period version of our model, we show that the agent's optimal strategy can be implemented using an elegant algorithm: Upon receiving a signal, he performs a Bayesian update in a low-dimensional space to compute a "decision value" that determines whether the signal provides enough new information to be worth internalizing. If the signal's decision value exceeds the cost of updating, then the agent internalizes the signal, performing the full Bayesian update (which is useful for any future decision that may arise).

Building on this characterization, we show that the agent can reject a large proportion of signals even if updating costs are very small. Moreover, the agent is specially attentive to signals suggesting that rare events are more likely than the prior suggests; in other words, individuals optimally attend more closely to signals suggesting that rare events are more common than they seem (e.g., hearing about airplane crashes) than to signals that rare events are, in fact, very rare (e.g., government reports on the general safety of airplanes).

We proceed to study the dynamic version of our model, where agents repeatedly decide whether or not to update their beliefs based on the signal received. We show that in dynamic environments, agents can exhibit several further behaviors that depart from the traditional costless baseline. Agents can exhibit permanent belief cycling: Even as they receive infinite information, their beliefs need not converge. More starkly, there are situations where agents' beliefs remain bounded away from the truth with probability 1.

Although agents in our model exhibit seemingly irrational behaviors, these behaviors are not a result of irrationality on the part of agents. Indeed, agents in our model are

[1]We focus on the asset allocation problem because it is important in many economic applications. Asset allocation also has two nice technical properties: First, the expected payoff is continuous in the agent's belief; and second, better-informed agents have strictly higher ex-ante expected payoffs.

always rationally optimizing, *given the constraint that Bayesian inference is computationally costly*.

There is ample psychological evidence that individuals display seemingly irrational patterns of belief updating. For example, Edwards (1982) reports situations in which individuals update more conservatively than Bayesian updating would predict. Bar-Hillel (1980) argues that individuals do not pay enough attention to base rates. Finally, Kahneman and Tversky (1979) and Prelec (1998) show that individuals appear to overweight rare events and underweight common events. In the presence of updating costs, some signal structures can make agents look conservative, while others can make agents look like they overweight the probability of rare events. Thus costly Bayesian updating provides a principled and unified formulation that can explain why certain biases may exist in some environments and not others.

Our model also links to the psychology and neuroscience of attention. The optimal strategy under costly Bayesian updating (described above) can be implemented in a way that resembles the "early selection theory of attention" (Pashler and Sutherland (1998)). First, a fast, parallel system analyzes the signals received and determines which ones are important enough for internalization. Important signals are then passed on to a slower, serial system for further processing. This theory explains why attention exhibits "cocktail party effects" (Bronkhorst (2000)): Important ambient signals such as hearing one's name, are readily attended to, while most incoming information is rejected/lost. Our model also explains "flashbulb memories," the fact that certain episodes are recorded much more vividly in memory than others (Brown and Kulik (1977)).

1.1 Related Literature

Alternative constraints on belief processing have been considered. Wilson (2014) characterizes optimal (Markovian) memory protocols for agents that have a finite number of memory modules. In the two-state setting she considers, an agent optimally associates each memory module with an interval of beliefs and—after observing a signal—transitions to the module whose corresponding interval contains the posterior. Although Wilson (2014) assumes that the agent forgets about the realized history of signals, she allows the agent to fully introspect on the source of his beliefs by Bayesian updating over all possible signal histories; in our dynamic model, this channel of introspection is shut down.[2] Like us, Wilson (2014) finds that agents should ignore signals that are mildly informative given the prior (see also Compte and Postlewaite (2012)). Consequently, information processing is "non-commutative" and first impressions matter. Contrary to her model, however, agents facing information processing costs as in our model do not succumb to confirmation bias. Rather, they regard evidence opposite to the prior as having higher decision value (see the discussion on belief cycling later in this paper).

Steiner and Stewart (2006) study the optimal design of probability perception when there is noisy information at the time of decision making.[3] Steiner and Stewart (2006)

[2]Like in the work of Lipman (1999), we model bounded rationality as a "lack of logical omniscience."

[3]Other models of belief manipulation include those of Bénabou and Tirole (2002), Bénabou and Tirole (2004), Compte and Postlewaite (2004), Brunnermeier and Parker (2005) and Gossner and Steiner (2016).

show that the well-known biases of overweighting small probabilities and underweighting large probabilities arise as a second-best solution to correct for the winner's curse effect. In our framework, agents also face friction in information processing. Yet, rather than holding the friction fixed and analyzing *ex ante* perception, we allow agents to repeatedly interact with the friction by choosing which signals to incorporate at a cost and which ones to ignore. Although our prediction that agents pay more attention to rare events bears resemblance to the finding of Steiner and Stewart (2006), in our context this is because rare events entail higher decision values.

Other models assume that agents have information processing constraints either because they must use one of a finite number of learning algorithms (Brock and Hommes (1997)) or because the amount of information they can process is limited (Sims (2003), Gabaix (2014)). However, any fixed upper bound on information processing becomes non-binding as beliefs approach the extreme, so that a rationally inattentive agent always updates at extreme beliefs. Thus the model of rational inattention cannot produce belief traps of the type we find.

Schwartzstein (2014) considers a model in which agents are Bayesian, yet selectively attend only to aspects they believe are important. With a sufficiently strong (and wrong) prior, a selectively attentive agent may "persistently fail to attend to information even if it would greatly improve forecast accuracy." The bad, self-confirming outcome Schwartzstein (2014) finds largely depends on the fact that the agent internalizes information via a model that is sometimes misspecified. This does not occur in our framework because our agents face processing costs instead of model misspecification.[4]

Fryer et al. (2015) introduce a special signal (consisting of a pair of opposite signals) that the agent remembers as favoring his current bias and then updates upon; this leads to double updating and confirmatory bias. In our setting, by contrast, perfectly mixed signals are simply ignored because they have no decision value.[5]

Entropy formulas related to those we find have also appeared in the literature on costly information acquisition. Those papers' similarity to ours is due to log utility being prominent; see, for example, Lemma 2 of Cabrales et al. (2013b). However, in costly information models like that of Cabrales et al. (2013b),[6] the cost of information is paid *ex ante* (as an information acquisition cost), whereas in our model the agent sees the signal first and then decides whether or not to pay to update.

2 Baseline Setup

We consider an agent who must allocate 1 unit of assets across possible outcomes $\Omega = \{\omega_1, \omega_2, \ldots, \omega_k\}$. The true outcome is generated by a *distribution*, denoted by θ^*—the **(true) model of the world**. Our agent's **beliefs** about the world are based on a set

[4]For further work on information processing under misspecified models, see Benjamin, Rabin, and Raymond (2016) and Benjamin, Bodoh-Creed, and Rabin (2016).

[5]Also related are the work on memory and recall, for example Mullainathan (2002), Gennaioli and Shleifer (2010), and Bordalo et al. (2015).

[6]See also Cabrales et al. (2013a) and Bergemann and Bonatti (2015).

$\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ of **possible models of the world**, which contains the true model, θ^*. We assume that the agent does not know the true model θ^* but starts with prior beliefs $p_0 \in \Delta(\Theta)$ over models of the world putting positive weight on θ^* (i.e. $p_0(\theta^*) > 0$).[7] Thus, learning about the true probability distribution θ^* helps the agent make better decisions.

When an outcome ω_i actually occurs, the agent gets utility $u(\alpha_i)$, where α_i is the amount of assets he allocated to outcome ω_i. We use the asset allocation problem as our benchmark because it is simple, has economic meaning in many contexts, and implies a direct value for information. Before allocating his asset, the agent receives a **signal** s drawn from the set $S = \{s_1, s_2, \ldots, s_m\}$. Given a model of the world θ, the distribution of the signal is denoted by $\lambda(\theta)$, so that the probability of any signal s is $\lambda(\theta)[s]$. We assume signal and outcome realizations are conditionally independent from each other given θ. We also assume the information structure can distinguish between different models, so that for any $\theta \neq \theta'$, there exists s with $\lambda(\theta)[s] \neq \lambda(\theta')[s]$.

The final component of our model is our departure from the standard Bayesian model: *Performing Bayesian updating is costly*. After seeing a signal s, the agent chooses whether to internalize the signal by updating his prior. Updating requires payment of a (fixed) cost c, and results in the replacement of the prior beliefs p_0 with the posterior beliefs $\overline{p}(s)$. If the agent does not update, then he pays no cost but the signal information is lost.

We consider first the case with a one-shot decision; this can also be thought of the case in which the agent acts in each period but has a discount rate of $\delta = 0$. After presenting our results for the single-period case, we explain how our findings change in a dynamic world with $\delta > 0$.

3 Single-Period Costly Updating

For a distribution p on models of the world $\theta \in \Theta$, we let $p^\Omega \in \Delta(\Omega)$ be the induced probability distribution on outcomes, defined by

$$p^\Omega[\omega] = \sum_{\theta \in \Theta} p(\theta)\theta(\omega).$$

We can then define the expected utility of an agent who has beliefs p over models of the world and allocates assets according to $\alpha \in \Delta(\Omega)$:

$$EU_p(\alpha) = \sum_{i=1}^{k} p^\Omega[\omega_i] \cdot u(\alpha_i).$$

The optimal action α^* is defined by

$$\alpha^*(p) = \mathrm{argmax}_{\alpha \in \Delta(\Omega)} EU_p(\alpha),$$

which leads to our key concept of "decision value" below.

[7]We use the standard notation $\Delta(X)$ for the set of probability distributions over set X.

Definition 1. *Given prior* p_0*, the* ***decision value*** *of a signal realization* s *is*

$$EU_{\overline{p}}(\alpha^*(\overline{p})) - EU_{\overline{p}}(\alpha^*(p_0)),$$

where $\overline{p} = \overline{p}(s)$ *represents the posterior beliefs given* s*.*

To interpret this formula, note that when the agent sees signal s, he faces two options: On one hand, he can choose to internalize the signal and update beliefs to $\overline{p}$; this allows him to allocate according to $\alpha^*(\overline{p})$ and obtain payoff $EU_{\overline{p}}(\alpha^*(\overline{p}))$ modulo the updating cost. On the other hand, he can alternatively choose to discard the signal information; in this case he knows that at the time of decision his *future self* will allocate according to $\alpha^*(p_0)$, so the expected payoff for his *current self* is given by $EU_{\overline{p}}(\alpha^*(p_0))$. Thus the agent optimally chooses to internalize a signal if and only if its decision value exceeds c. We emphasize that unlike models of costly information acquisition, the decision value in our setting depends on the signal realization (recall that $\overline{p} = \overline{p}(s)$).

To simplify the rest of our analysis, from here onwards we assume that the agent has **log-utility**: that is, $u(\alpha_i) = \log(\alpha_i)$. This functional form assumption drives the functional forms of our results, but not our qualitative findings.

We now show that under log-utility, there is an elegant characterization of the agent's optimal updating strategy in terms of KL divergence.[8]

Proposition 1. *Given prior* p_0*, signal* s*, and* log*-utility, the agent updates if and only if*

$$D(\overline{p}(s)^{\Omega} \mid\mid p_0^{\Omega}) > c,$$

where $D(\overline{p}(s)^{\Omega} \mid\mid p_0^{\Omega})$ *is the Kullback-Leibler (KL) divergence of the prior from the posterior, and* c *is the processing cost.*

Proof. In the asset allocation problem with log utility, the first-order conditions imply

$$c^*(p) = p^{\Omega}.$$

If the agent uses prior beliefs p_0 to make the allocation decision, he ends up with an expected utility of

$$\sum_{\omega} \overline{p}^{\Omega}(\omega) \log(p_0^{\Omega}(\omega)),$$

where $\bar{p}$ represents his posterior beliefs.

However, if the agent uses $\bar{p}$ to make his allocation decision, he has expected utility

$$\sum_{\omega} \overline{p}^{\Omega}(\omega) \log(\overline{p}^{\Omega}(\omega)).$$

[8]The *Kullback-Leibler (KL) divergence* of distribution Q from distribution P over a common discrete sample space Ω is defined as (Kullback and Leibler (1951))

$$D(P \mid\mid Q) = \sum_{\omega \in \Omega} P(\omega) \log\left(\frac{P(\omega)}{Q(\omega)}\right).$$

Thus the decision value of the signal is given by

$$\sum_{\omega} \bar{p}^{\Omega}(\omega)(\log(\bar{p}^{\Omega}(\omega)) - \log(p_0^{\Omega}(\omega))) = \sum_{\omega} \bar{p}^{\Omega}(\omega) \cdot \log\left(\frac{\bar{p}^{\Omega}(\omega)}{p_0^{\Omega}(\omega)}\right),$$

which is exactly the KL divergence of the prior from the posterior. $\square$

3.1 Discussion

Proposition 1 tells us the exact conditions under which updating is useful, but computing a decision value itself requires Bayesian updating. How can the agent solve this problem? The key here is that the KL divergence is computed in terms of *marginal* distributions over Ω. Thus, the agent only needs to do a "partial" update to find the decision value. To further illustrate this point, below we argue that Proposition 1 implies that the agent can use a two-step algorithm that is more efficient than full Bayesian updating, so long as updating in a larger space is more computationally intensive.

First, consider the complete decision space. A Bayesian agent has a distribution $\mu \in \Delta(\Theta \times S \times \Omega)$ that describes the joint probability of any model, signal and outcome combination. To fully update this distribution given a signal s, the decision-maker must make some number of computations, which we denote ℓ_1.

As a different strategy, the agent may instead look at the distribution ν that is the projection of μ onto the space $S \times \Omega$. This space is of lower-dimension than $\Theta \times S \times \Omega$; thus, updating in $S \times \Omega$ given a signal s is simpler, having processing cost $\ell_2 < \ell_1$. Proposition 1 suggests that in order to determine whether the signal s is worth updating fully, it suffices for the agent to partially update to ν and compute the decision value.

In a world where all signals are highly informative and worth internalizing, "gatekeeper" computation of ν is a waste of resources. However, suppose that only a fraction $q < 1$ of signals are worthwhile. The agent pays expected updating cost of $\ell_2 + q \cdot \ell_1$ from following the two-step algorithm, versus the cost of ℓ_1 if he always updates. Thus, if q is small or if ℓ_1 is large relative to ℓ_2—e.g., if there are many possible models but few outcome possibilities—then the gatekeeper strategy we propose conserves on computational cost while still allowing the agent to internalize important information.[9]

We mention that the gatekeeper solution is exactly the kind of system set up by the early selection theory of attention in psychology (Pashler and Sutherland (1998)). In the models posited in that literature, agents are endowed with two information processing systems: A fast one that can process information in parallel, but only at a cursory level, and a slower one that can look at information in more depth, but must do so serially. In such models, the parallel system screens incoming stimuli. After screening, signals deemed unimportant are ignored, while those deemed important are passed "upstairs" for further processing. For our agent, the faster "gatekeeper" system can compute the decision value and, if need be, push the signal to the slower serial system for internalization.

[9] A second, more psychological (or reduced-form) interpretation of Proposition 1 is that the decision value of a signal is more easily accessible than are the signal's implications about the likelihoods of different models of the world.

Of course, one may ask: Why not simply always update in the lower-dimensional space $S \times \Omega$ and forget the existence of $\Theta \times S \times \Omega$? This is because just forming beliefs over $S \times \Omega$ is sufficient in a dynamic setting. Specifically, if the agent only maintains beliefs about s and ω over multiple updates, then he treats signals and outcomes as independent, but in fact they are only independent conditional on the model θ. Consequently, such an agent will learn the wrong distribution in the long run, even in the absence of updating costs.

We now move on to examples showing how agents with costly updating differ in behavior, sometimes greatly, from their costless Bayesian counterparts.

3.2 Example 1: Quantifying Costs

Proposition 1 shows that the addition of updating costs may cause agents to discard some information. We now show, by numerical example, that this effect distorts behavior away from the standard (rational, with perfect updating) baseline model even when processing costs are relatively small.

We consider an example with two possible models of the world, θ_1 and θ_2, and two possible outcomes, ω_1 and ω_2. The probability of outcome ω_1 under θ_1 is .9, the probability of outcome ω_1 under θ_2 is .1. The agent has a prior that puts weight p_0 on θ_1 and $1 - p_0$ on θ_2.[10] In this setup, any possible signal s can be parametrized by an **accuracy level**

$$r \equiv \frac{\lambda(\theta_1)[s]}{\lambda(\theta_1)[s] + \lambda(\theta_2)[s]}.$$

A signal with accuracy level $r = 1$ ($r = 0$) is a signal that θ_1 (θ_2, respectively) is definitely the true model of the world. We now characterize the agent's decision to accept a signal with accuracy level r as a function of his prior.

Given a signal of accuracy r, we can compute the agent's decision value. First, we determine the prior distribution on outcomes:

$$p_0^{\Omega}(\omega_1) = p_0(.9) + (1 - p_0)(.1).$$

Similarly, the posterior distribution on outcomes implied by the signal is

$$\overline{p}^{\Omega}(\omega_1) = \frac{p_0 r}{p_0 r + (1 - p_0)(1 - r)}(.9) + \left(1 - \frac{p_0 r}{p_0 r + (1 - p_0)(1 - r)}\right)(.1).$$

By Proposition 1, the agent should choose to update iff (omitting the superscript Ω)

$$\overline{p}(\omega_1) \log\left(\frac{\overline{p}(\omega_1)}{p_0(\omega_1)}\right) + (1 - \overline{p}(\omega_1)) \log\left(\frac{1 - \overline{p}(\omega_1)}{1 - p_0(\omega_1)}\right) > c.$$

Clearly, there is a threshold cost (depending on the prior) above which the agent discards the signal and below which the agent updates. To interpret the magnitude of

[10] Here and hereafter, when a distribution p is over two objects, we abuse notation slightly by using the distribution notation p to represent the probability of one of the objects.

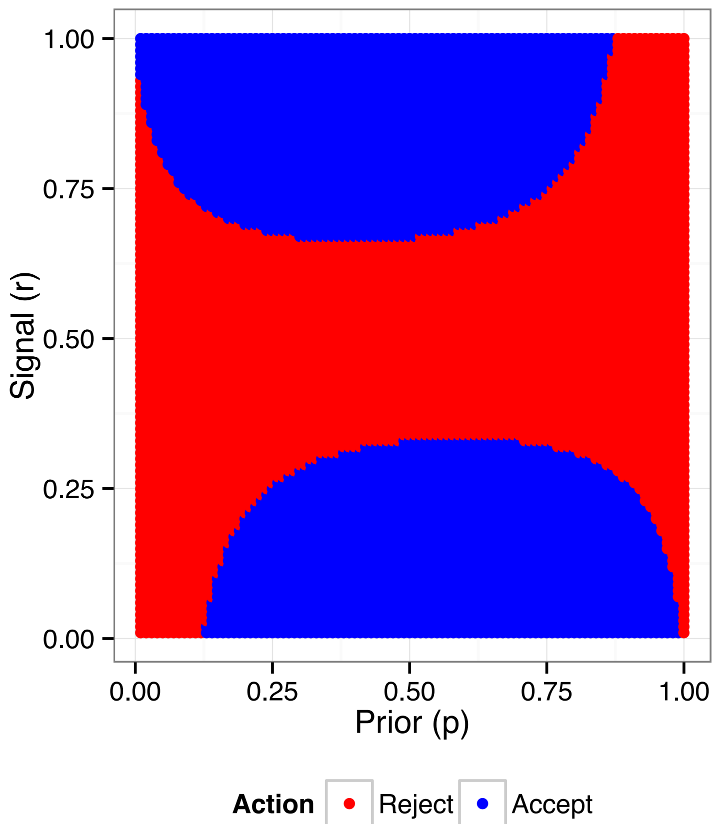

Figure 1: Combinations of prior p_0 (horizontal axis) and signal accuracy r (vertical axis) that lead the agent to accept (blue) or ignore (red) the signal. Here $c = \frac{1}{10} \times .16$.

this threshold, we will perform a unit conversion. We scale c to be a fraction of the total expected utility gain between a completely informed agent (who has expected utility of $.9\log(.9) + .1\log(.1)$) and a completely uninformed agent (who has expected utility of $\log(.5)$). That is, c is written as a fraction of the term

$$(.9\log(.9) + .1\log(.1)) - \log(.5) \approx .16.$$

Figure 1 shows how an agent with prior beliefs p_0 reacts to signals of accuracy r.

Note that even if c is just 10% of the informational value of the problem, the agent will, for any prior, reject signals that have accuracy in the interval $[\frac{1}{3}, \frac{2}{3}]$. Thus costly updating has a large distortionary effect on behavior, relative to the costless updating baseline.

3.3 Example 2: Attention to Rare Events

Individuals tend to overestimate the probability of rare events like plane crashes and terrorist attacks (see, e.g., Barberis (2013)). In the behavioral economics literature, this overestimation has been modeled as accidental distortion of probabilities (Kahneman and Tversky (1979)). Our model, however, suggests a different explanation: When updating is costly, agents' beliefs tend to gravitate towards models of the world in which rare events are more common than they actually are.

To see this, consider an example with agents who live in a big city. The agents faces two outcomes: $\Omega = \{\rho \text{ (robbed)}, \sigma \text{ (safe)}\}$. Correspondingly, there are two possible models of the world: $\Theta = \{\theta^{\mathsf{s}}, \theta^{\mathsf{d}}\}$. We suppose that $\theta^{\mathsf{s}}(\sigma) = .99$ and $\theta^{\mathsf{d}}(\sigma) = .8$, so that robbery is always rare, but much rarer under θ^{s} than under θ^{d}. We can interpret θ^{s} as a model in which the city is "safe"; θ^{d} is a model in which the city is "dangerous."

Suppose there are two agents who have the same updating cost c. One agent, whom we call Dirk, starts out thinking that the city is dangerous, and the other, whom we call Stephen, starts out thinking that the city is safe. Both agents are equally certain of their (differing) models of the world: Dirk puts prior probability $p_0 > .5$ on θ^{d}, while Stephen puts prior probability p_0 on θ^{s}.

We examine what happens when both agents receive equally strong signals suggesting that their beliefs are wrong. Formally, Stephen receives a signal s^{d} that has accuracy level $r < .5$, and Dirk receives a signal s^{s} that has accuracy $1 - r$.

Our next result shows that quite generally, the decision value of Dirk's signal is higher than that of Stephen's.

Proposition 2. *For a fixed $p_0 > .5$, suppose that $r \in (0, .5)$ is such that the posterior beliefs of Stephen put probability at least $1 - p_0$ on the model θ^{s}. Then the decision value of the signal s^{d} for Stephen is strictly higher than the decision value of the signal s^{s} for Dirk.*[11]

Under costly updating, agents who believe that the city is safe—that is, agents who think that rare, dangerous events are very rare—tend to be very responsive to information that suggests otherwise. Meanwhile, agents who believe that rare events are relatively common are not as responsive to equally strong information suggesting that rare events are, in fact, rare.

Proof of Proposition 2. We use two technical lemmata, whose proofs are deferred to the appendix. For $p, q \in (0, 1)$, we abuse notation slightly and write $D(p||q)$ for the relative entropy between two Bernoulli distributions with success probabilities p and q.

First, we observe that signals that move the agent's beliefs *away from* the extreme are more valuable than equally-strong signals that push beliefs *toward* the extreme.

Lemma 1. *For all $.5 \le p < q \le 1$, $D(p||q) > D(q||p)$.*

Second, we establish a monotonicity result on the value of signals that change beliefs by a fixed amount.

[11] Note that Proposition 2 is only true if the signal is not strong enough to completely overturn the prior. To see why r cannot be too small, consider the above setup with $p_0 = .6$ and $r = .1$, so that Stephen's posterior puts probability $\frac{1}{7}$ on θ^{s}. His posterior estimate of the probability of the safe outcome is about .827, compared to the prior estimate .914. From Proposition 1, we compute that Stephen's decision value is about .0166. Similarly we can compute that the decision value for Dirk is .0183, higher than Stephen's.

Lemma 2. *For any $\Delta > 0$ and any p, q satisfying $.5 \le p < q \le 1 - \Delta$, $D(p \,||\, p + \Delta) < D(q \,||\, q + \Delta)$.*

Now, let $p_0^{\mathsf{s}} = .99p + .8(1-p)$ be Stephen's prior probability on the safe outcome and let $\bar{p}^{\mathsf{s}} = \frac{.99pr + .8(1-p)(1-r)}{pr + (1-p)(1-r)}$ be the corresponding posterior probability following a signal with accuracy r. Similarly, we let $p_0^{\mathsf{d}} = .8p + .99(1-p)$ and $\bar{p}^{\mathsf{d}} = \frac{.8pr + .99(1-p)(1-r)}{pr + (1-p)(1-r)}$ be Dirk's prior and posterior estimates respectively.

For any $r < .5$ we have $.5 < \bar{p}^{\mathsf{s}} < p_0^{\mathsf{s}}$, $.5 < p_0^{\mathsf{d}} < \bar{p}^{\mathsf{d}}$, and, importantly, $\Delta \equiv p_0^{\mathsf{s}} - \bar{p}^{\mathsf{s}} = \bar{p}^{\mathsf{d}} - p_0^{\mathsf{d}}$. The condition that r is not too small further ensures that $\bar{p}^{\mathsf{d}} \le p_0^{\mathsf{s}}$, and so $p_0^{\mathsf{d}} \le \bar{p}^{\mathsf{s}}$. Hence, Lemma 1 and Lemma 2 together imply that

$$D(\bar{p}^{\mathsf{d}}||p_0^{\mathsf{d}}) < D(p_0^{\mathsf{d}}||\bar{p}^{\mathsf{d}}) = D(p_0^{\mathsf{d}}||p_0^{\mathsf{d}} + \Delta) \le D(\bar{p}^{\mathsf{s}}||\bar{p}^{\mathsf{s}} + \Delta) = D(\bar{p}^{\mathsf{s}}||p_0^{\mathsf{s}}); \tag{1}$$

the expression (1) says exactly that the decision value for Stephen is higher than that for Dirk. □

4 Belief Dynamics

We now explicitly consider the dynamics of our model. Suppose that the true model of the world remains constant throughout. In this setting, the agent starts each period $t = 1, 2, \ldots$ with a prior p_{t-1}, receives a signal s drawn from probability distribution $\lambda(\theta^*)$, and chooses whether to internalize s. He then faces the allocation decision outlined in our static model, with beliefs $p_t = \bar{p}_{t-1}(s)$ if s is internalized, and with beliefs $p_t = p_{t-1}$ if not. The agent then proceeds to the next period, with new prior p_t. We assume that the agent discounts future payoffs at rate $\delta \in [0, 1)$.

When $\delta > 0$, our dynamic model here assumes that the agent takes into account future information values (which in turn depends on future updating decisions) in deciding whether to internalize signals. This kind of sophistication makes it difficult to characterize the decision value of a signal, unlike in the static/myopic case of Proposition 1. In our formal analysis, we show an upper bound on the total value of a signal by assuming that *the agent can internalize all future signals at no cost*. We then compare that bound to the updating cost and establish conditions under which the latter is larger. Consequently, our results below about the *existence* of belief traps and belief cycling extend without change regardless of what the agent believes his future selves will do.[12] That said, it is certainly of interest to fully determine the optimal dynamic policy for a sophisticated agent. We leave such a characterization for future work.

As a benchmark, we note that when there is no updating cost (i.e., $c = 0$), the agent learns the true model θ^* with certainty: $p_t(\theta^*) \xrightarrow{a.s.} 1$.[13] Next, we examine the (dynamic) implications of updating costs.

[12]Unlike in the model of Wilson (2014), we do not allow the agent to reflect on how he reaches his current beliefs and perform delayed updating on signals that (he infers) have been ignored in the past.

[13]This result is standard, so we omit the proof.

Definition 2. *A **belief trap** is a probability distribution $p \in \Delta(\Theta)$ such that an agent with prior beliefs p will not internalize any signal $s \in S$.*

Note that an agent who starts a period in a belief trap will maintain his beliefs forever, irrespective of the signals he receives. When updating is costless, the only belief trap is one in which the agent knows the true model with certainty. However, we now show that under general conditions, strong beliefs about *any* possible model of the world become belief traps in the presence of a positive updating cost.

Definition 3. *An information structure has full support if $\lambda(\theta)[s] > 0$ for all θ and s.*

Proposition 3. *If $c > 0, \delta < 1$, and the information structure has full support, then there exists $\epsilon > 0$ such that any belief p_0 with $p_0(\theta) > 1 - \epsilon$ for some θ is a belief trap.*

When $\delta = 0$, so that our agent is myopic, the characterization in Proposition 1 applies. We can then directly show that when the prior beliefs are (close to) extreme, no signal generates a decision value exceeding c.

When $\delta > 0$, Proposition 1 no longer holds because the updating decision in a given period has repercussions for the future. In the absence of an analytic characterization, we instead give an upper bound on potential future benefits of updating and show they are smaller than c. The method of proof is to choose ϵ much smaller than what is required in the case of $\delta = 0$, so that any notable utility gain only happens at least T periods into the future. For large T, these benefits are again negligible due to discounting. Details are left to the appendix.

While Proposition 3 shows that extreme beliefs are stable under costly updating, it does *not* ensure that belief traps are attractors. We find that when updating is costly, beliefs may not even converge—they can exhibit permanent cycling with probability 1.

Proposition 4. *Fix $c > 0$ and $\delta < 1$. There exists an information structure and an open interval $B \subset (0, 1)$ such that if $p_0 \in B$, then almost surely $\lim_{t\to\infty} p_t$ is undefined.*

The proof, which we present in the appendix, is by construction: For simplicity, we restrict to the case of two possible outcomes $\Omega = \{0, 1\}$, two models $\Theta = \{a, b\}$ (with a being the true model) and two signals $S = \{A, B\}$. The table below shows the conditional probabilities of outcomes given models and of signals given models.

OUTCOMES/MODELS	a	b
0	.8	.2
1	.2	.8

SIGNALS/MODELS	a	b
A	.85	.15
B	.15	.85

.

Now consider an agent who repeatedly faces this information structure and makes decisions. For illustrate we suppose $\delta = 0$, so Proposition 1 is sufficient to tell us whether an agent will update given a signal. Figure 2 shows the decision value of accepting either signal as a function of the prior. Observe that the value of accepting signal A falls quickly as the agent comes to believe that a is very likely; thus, there is a point that the agent stops accepting signal A. For example, if $c = .07$, then the agent will not accept signal

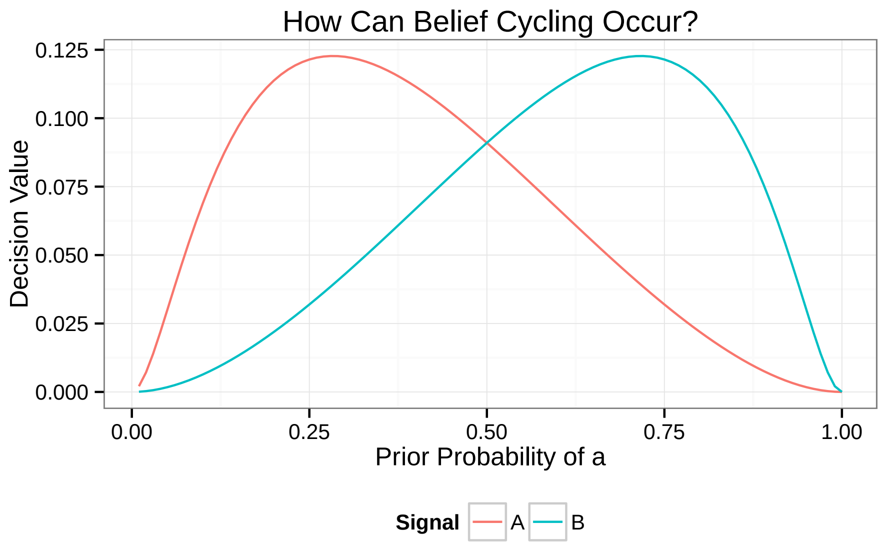

Figure 2: Sufficient conditions for agents beliefs to cycle under costly updating.

A if he believes model a has probability at least .59. At that threshold belief, even if the agent received and accepted signal A, his posterior probability estimate would only be about .891. At that posterior belief, he would no longer accept signal A but he would still accept signal B. This means the agent's probability estimates of model a would never be higher than .891. The same argument applies symmetrically for signal B. Thus, there is no way for the agent to get into the region where he stops accepting both signals; hence, the agent will exhibit persistent belief cycling.

Now, we have shown that it is possible for beliefs to get stuck or to cycle forever. Our final result shows that there exist situations in which the agent's beliefs will converge to a point away from the true model (unless his prior beliefs are very close to the truth).

Proposition 5. *Fix $\delta < 1$, $p_0 < 1$ and $\eta > 0$. There exists an information structure such that an agent with prior p_0, discount factor δ and any small positive updating cost exhibits the following belief dynamics:*

$$\text{prob}\left(\lim_{t\to\infty} p_t(\theta^*) < \eta\right) = 1.$$

Again, here we just sketch the intuition behind the proof: Consider an information structure with two models and two signals as follows:[14]

SIGNALS/MODELS	a	b
A	$1-\epsilon$	$1-L\epsilon$
B	ϵ	$L\epsilon$

[14] In this construction, we assume for simplicity that the outcome is identical to the model.

Suppose that ϵ is very small and L is very large. In this case, signal A is weakly informative that the true model is a; hence, its decision value is almost 0, and for any *fixed* updating cost it can be shown that the agent will never accept signal A. At the same time, λ is large, so signal B is very informative that the true model is, in fact, b. Thus, even if the true model is a, the agent will accept only signal B (unless he begins at a prior where he accepts no signals). In a dynamic world, this means the agent will converge to believing b very strongly, until the decision value of signal B also falls below c.

However, we highlight that the proposition asserts an information structure that applies to *every* c. As we fix the information structure and prior and consider very small c, the preceding argument fails because the agent will (initially) accept signal A. Despite this, we can still show that at any belief, *signal* B *is accepted whenever signal* A *is*. This property again implies that the agent will converge to the wrong belief b in the long run under costly updating.

We mention that because Proposition 5 applies to any small updating cost, it suggests a surprising discontinuity as we move away from the costless baseline. As far as we are aware, such discontinuity does not occur in other forms of learning biases: For example, Bohren and Hauser (2017) show that under model misspecification (and no updating cost), agents with approximately correct models almost surely learn asymptotically.

5 Conclusion

We have shown that the addition of processing costs to Bayesian inference leads rational agents to adopt strategies of "pre-screening" signals before fully internalizing them. This causes individuals to ignore some information, and can have distortionary effects on behavior even when processing costs are small. Additionally, costly processing produces a form of probability overweighting, leading individuals to react more strongly to information suggesting that unexpected events are common than to equivalently strong information suggesting that events believed to be likely are actually rare. Finally, individuals for whom processing is costly can sometimes converge to having arbitrarily strong, but wrong, beliefs even in a world of free-flowing information.

There is a large set of psychological findings on biases in human probability judgment.[15] However, oftentimes the observed biases can appear to be somewhat contradictory—e.g., some strands of research show that individuals do not update enough (i.e., they are too conservative) while others show that individuals can update too much (e.g., ignore base rates or otherwise overweight the likelihood of a rare event). Our model can help to reconcile the range of these findings: When Bayesian updating is costly, individuals can be maximally conservative in some situations (as when caught in a belief trap), and can be over-attentive (overweighting rare events) in others.

[15] Recent contributions in this direction include Haruvy and Erev (2002), Barberis (2013), Fudenberg and Peysakhovich (2014), Peysakhovich and Rand (2015), and Peysakhovich and Karmarkar (2016).

References

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica 44*(3), 211–233.

Barberis, N. (2013). The psychology of tail events: Progress and challenges. *American Economic Review 103*(3), 611–616.

Bénabou, R. and J. Tirole (2002). Self-confidence and personal motivation. *Quarterly Journal of Economics 117*(3), 871–915.

Bénabou, R. and J. Tirole (2004). Willpower and personal rules. *Journal of Political Economy 112*(4), 848–886.

Benjamin, D. J., A. Bodoh-Creed, and M. Rabin (2016). Base-rate neglect: Model and applications. Working Paper.

Benjamin, D. J., M. Rabin, and C. Raymond (2016). A model of non-belief in the law of large numbers. *Journal of the European Economic Association 14*(2), 515–544.

Bergemann, D. and A. Bonatti (2015). Selling cookies. *American Economic Journal: Microeconomics 7*(3), 259–294.

Bohren, J. A. and D. N. Hauser (2017). Bounded rationality and learning: A framework and a robustness result. Working Paper.

Bordalo, P., N. Gennaioli, and A. Shleifer (2015). Memory, attention and choice. Working Paper.

Brock, W. A. and C. H. Hommes (1997). A rational route to randomness. *Econometrica: Journal of the Econometric Society*, 1059–1095.

Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica united with Acustica 86*(1), 117–128.

Brown, R. and J. Kulik (1977). Flashbulb memories. *Cognition 5*(1), 73–99.

Brunnermeier, M. K. and J. A. Parker (2005). Optimal expectations. *American Economic Review 95*(4), 1092–1118.

Cabrales, A., O. Gossner, and R. Serrano (2013a). The appeal of information transactions. Working Paper.

Cabrales, A., O. Gossner, and R. Serrano (2013b). Entropy and the value of information for investors. *American Economic Review 103*(1), 360–377.

Compte, O. and A. Postlewaite (2004). Confidence-enhanced performance. *American Economic Review 94*(5), 1536–1557.

Compte, O. and A. Postlewaite (2012). Belief formation. Working Paper.

Cover, T. M. and J. A. Thomas (2012). *Elements of Information Theory*. John Wiley & Sons.

Doya, K. (2007). *Bayesian Brain: Probabilistic Approaches to Neural Coding*. MIT Press.

Edwards, W. (1982). Conservatism in human information processing. In A. T. Daniel Kahneman, Paul Slovic (Ed.), *Judgment under Uncertainty: Heuristics and Biases*, pp. 359–369. Cambridge University Press.

Epstein, L. G. and M. Le Breton (1993). Dynamically consistent beliefs must be Bayesian. *Journal of Economic Theory 61*(1), 1–22.

Fryer, R. G., P. Harms, and M. O. Jackson (2015). Updating beliefs with ambiguous evidence: Implications for polarization. Working Paper.

Fudenberg, D. and A. Peysakhovich (2014). Recency, records, and recaps: The effect of feedback on behavior in a simple decision problem. In *Proceedings of the 15th ACM conference on Economics and Computation*, pp. 971–986.

Gabaix, X. (2014). A sparsity-based model of bounded rationality. *Quarterly Journal of Economics 129*(4), 1661–1710.

Gennaioli, N. and A. Shleifer (2010). What comes to mind. *Quarterly Journal of Economics 125*(4), 1399–1433.

Gopnik, A. (2010). How babies think. *Scientific American 303*(1), 76–81.

Gossner, O. and J. Steiner (2016). Optimal illusion of control and related perception biases. Working Paper.

Haruvy, E. and I. Erev (2002). On the application and interpretation of learning models. In *Experimental Business Research*, pp. 285–300. Springer.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–292.

Knill, D. C. and W. Richards (1996). *Perception as Bayesian Inference*. Cambridge University Press.

Kullback, S. and R. A. Leibler (1951). On information and sufficiency. *Annals of Mathematical Statistics 22*(1), 79–86.

Lipman, B. L. (1999). Decision theory without logical omniscience: Toward an axiomatic framework for bounded rationality. *Review of Economic Studies 66*(2), 339–361.

Mullainathan, S. (2002). A memory-based model of bounded rationality. *Quarterly Journal of Economics 117*(3), 735–774.

Pashler, H. E. and S. Sutherland (1998). *The Psychology of Attention*. MIT Press.

Peysakhovich, A. and U. R. Karmarkar (2016). Asymmetric effects of favorable and unfavorable information on decision making under ambiguity. *Management Science 62*(8), 2163–2178.

Peysakhovich, A. and D. G. Rand (2015). Habits of virtue: Creating norms of cooperation and defection in the laboratory. *Management Science 62*(3), 631–647.

Prelec, D. (1998). The probability weighting function. *Econometrica 66*(3), 497–528.

Rudin, W. (1964). *Principles of Mathematical Analysis*. McGraw-Hill.

Schwartzstein, J. (2014). Selective attention and learning. *Journal of the European Economic Association 12*(6), 1423–1452.

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

Steiner, J. and C. Stewart (2006). Perceiving prospects properly. *American Economic Review 106*(7), 1601–1631.

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

A Proofs Omitted from the Main Text

Proof of Lemma 1

We need to show that

$$p \log\left(\frac{p}{q}\right) + (1-p)\log\left(\frac{1-p}{1-q}\right) > q \log\left(\frac{q}{p}\right) + (1-q)\log\left(\frac{1-q}{1-p}\right),$$

or equivalently, that

$$(2-p-q)\log\left(\frac{1-p}{1-q}\right) - (p+q)\log\left(\frac{q}{p}\right) > 0. \tag{2}$$

The left-hand-side of (2) vanishes when $p = q$, so it suffices to show that the derivative of that left-hand-side with respect to q is strictly positive. This reduces to verifying

$$\frac{q-p}{q(1-q)} - \log\left(\frac{q}{p}\right) - \log\left(\frac{1-p}{1-q}\right) > 0. \tag{3}$$

Now, we observe that the left-hand-side of (3) vanishes when $p = q$ and its derivative with respect to p is $\frac{1}{p(1-p)} - \frac{1}{q(1-q)} < 0$. Thus, we have (3) and, consequently, (2); the lemma follows.

Proof of Lemma 2

We fix Δ and show that the expression

$$D(p \mid\mid p+\Delta) = p \log\left(\frac{p}{p+\Delta}\right) + (1-p)\log\left(\frac{1-p}{1-p-\Delta}\right) \tag{4}$$

strictly increases in p. For the derivative of (4) to be positive, we need to check that

$$\frac{\Delta}{(p+\Delta)(1-p-\Delta)} > \log\left(\frac{(p+\Delta)(1-p)}{p(1-p-\Delta)}\right). \tag{5}$$

At $p = .5$, (5) reduces to the easy-to-verify inequality $\log \frac{1+x}{1-x} < \frac{2x}{1-x^2}$. Taking the derivative of (5) with respect to p again, it remains to show that

$$\frac{\Delta(2p+2\Delta-1)}{(p+\Delta)^2(1-p-\Delta)^2} > \frac{\Delta(2p+\Delta-1)}{p(1-p)(p+\Delta)(1-p-\Delta)};$$

this is true whenever $p \geq .5$, and the lemma follows.

Proof of Proposition 3

As the information structure has full support, there exists $\gamma > 0$ such that $\lambda(\theta)[s] \geq \gamma$ for every θ and s. Let $M = \frac{1}{\gamma}$. We observe a simple lemma.

Lemma 3. *Whenever an agent's prior belief satisfies $p_0(\theta) > 1-\epsilon$ for some θ and $\epsilon > 0$, his posterior belief must be such that $p_1(\theta) > 1 - M\epsilon$.*

Proof. For the proof, it suffices to note that the posterior belief of any model $\theta' \neq \theta$ can increase by a factor of at most M relative to the prior. □

Now, we use the preceding lemma to provide a lower bound on the utility the agent expects if he commits not to update. We assume without loss of generality that all outcomes are possible given the prior p_0, and denote $N \equiv |\Omega| < \infty$.

In each period t, a non-updating agent forgets the signals he has seen and starts with the same prior p_0. Because he rejects all signals, the amount he allocates to outcome $\omega \in \Omega$ is always $p_0^\Omega(\omega)$. His expected payoff in period t is thus

$$\pi_t \equiv \sum_{\omega\in\Omega} p_1^\Omega(\omega) \cdot \log(p_0^\Omega(\omega)), \tag{6}$$

where p_1^Ω is the posterior distribution on outcomes conditional on the signal realized in period t. Lemma 3 implies that for any θ, $p_1(\theta)$ and $p_0(\theta)$ differ by at most $M\epsilon$ and by at most a multiple of M. It follows that for any ω, $p_1^\Omega(\omega)$ and $p_0^\Omega(\omega)$ satisfy the same properties. So we have from (6) that

$$\begin{aligned}\pi_t &\geq \sum_{\omega\in\Omega} p_0^\Omega(\omega)\cdot\log(p_0^\Omega(\omega)) - \sum_{\omega\in\Omega} |p_1^\Omega(\omega) - p_0^\Omega(\omega)||\log(p_0^\Omega(\omega))| \\ &\geq -H(p_0^\Omega) - NM|\epsilon\log\epsilon|\end{aligned}$$

where $H(\cdot)$ is the entropy function (Cover and Thomas (2012)).[16]

Our non-updating agent thus ensures a total payoff of

$$\sum_{t\geq 1} \delta^{t-1}\pi_t \geq \frac{1}{1-\delta}\left(-H(p_0^\Omega) - NM|\epsilon \log \epsilon|\right). \tag{7}$$

We can also give an upper bound to the agent's total payoff assuming that updating is costless and that the agent can freely choose to update. For this let $q_t \in \Delta(\Theta)$ be the posterior that the agent *knows* about θ after seeing the signal in period t, and let p_t be the posterior that he *retains*. Regardless of the retention decision, $q_t(\theta) \sim p_{t-1}(\theta)\cdot\lambda(\theta)[s^t]$ (where s^t denotes the signal realized in period t). However $p_t = q_t$ only when the agent accepts the signal; $p_t = p_{t-1}$ otherwise.

The agent allocates according to p_t at the end of period t, for an expected payoff of

$$\overline{\pi}_t \equiv \sum_{\omega\in\Omega} p_t^\Omega(\omega)\cdot\log(q_t^\Omega(\omega)) \leq -H(q_t^\Omega) \tag{8}$$

where the inequality is due to KL-divergence being non-negative. From Lemma 3 and induction we have $p_t(\theta), q_t(\theta) \geq 1 - M^t\epsilon$ for each t. Thus the total variation distance between q_t and p_0 is at most $M^t\epsilon$, which also bounds $||q_t^\Omega - p_0^\Omega||$.

Now take T sufficiently large such that $\frac{\delta^T}{1-\delta}H(p_0^\Omega) < \frac{c}{3}$. For this T, we take ϵ sufficiently small such that $\frac{1-\delta^T}{1-\delta}|H(p_0^\Omega) - H(q^\Omega)| < \frac{c}{3}$ whenever $||p_0^\Omega - q^\Omega|| \leq M^T\epsilon$. This can be done independently of p_0^Ω because the entropy function is continuous in the total variation norm, hence uniformly continuous (Rudin (1964, Thm. 4.19)). Making ϵ even smaller if necessary, we may assume that $\frac{1}{1-\delta}NM|\epsilon\log\epsilon| < \frac{c}{3}$.

From (8), the non-negativeness of $H(\cdot)$ and the preceding assumptions, we deduce that the total payoff for a freely updating agent is bounded above by

$$\begin{aligned}
\sum_{t\geq 1}\delta^{t-1}\overline{\pi}_t &\leq \sum_{t\geq 1}\delta^{t-1}(-H(q_t^\Omega)) \\
&\leq \sum_{1\leq t\leq T}\delta^{t-1}(-H(q_t^\Omega)) \\
&< \sum_{1\leq t\leq T}\delta^{t-1}(-H(p_0^\Omega)) + \frac{c}{3} \\
&< \frac{1}{1-\delta}(-H(p_0^\Omega)) + \frac{2c}{3} \\
&\leq \sum_{t\geq 1}\delta^{t-1}\pi_t + c
\end{aligned} \tag{9}$$

where the last step follows from (7).

Now, (9) shows that even when updating is *costless*, the agent cannot improve her *total* payoff by c relative to a counterfactual in which he commits not to update. As a corollary, we see that when the updating cost is c, it is never worthwhile to pay the updating cost of c in any period (given any signal); this completes the proof of the proposition.

[16]The latter step follows from considering the cases $p_0^\Omega(\omega) \geq \epsilon$ and $p_0^\Omega(\omega) < \epsilon$ separately.

Proof of Proposition 4

We will first consider a myopic agent facing the following information structure:

	θ^*	θ'
s^*	$\frac{1}{r+1}$	$\frac{r}{r+1}$
s'	$\frac{r}{r+1}$	$\frac{1}{r+1}$

with some small positive r. Assume throughout the proof that outcomes are simply generated by the point-mass distribution at the model.

For $p \in (0,1)$, define a pair of functions:

$$F_r(p) = \frac{p}{p + (1-p) \cdot r};$$
$$G_r(p) = \frac{p \cdot r}{p \cdot r + (1-p)}.$$

If an agent assigns prior probability p to model θ^*, his posterior becomes $F_r(p)$ given the signal s^* and $G_r(p)$ given the signal s'. Notice that F_r and G_r are inverse functions, corresponding to the fact that after an agent sees s^* followed by s', his posterior equals his prior. It follows that starting with any initial belief p_0, the agent's posterior beliefs all belong to the countable set $\{F_r^{(k)}(p)\}_{k=-\infty}^{\infty}$.

The decision value for the agent upon seeing s^* is

$$\begin{aligned} V_{F,r}(p) &:= D(F_r(p) \,||\, p) \\ &= \frac{p}{p+(1-p)r} \cdot \log\left(\frac{1}{p+(1-p)r}\right) + \frac{(1-p)r}{p+(1-p)r} \cdot \log\left(\frac{r}{p+(1-p)r}\right) \\ &= -\log(p+(1-p)r) + \frac{(1-p)r}{p+(1-p)r} \cdot \log r. \end{aligned} \tag{10}$$

Similarly the decision value of s' is

$$V_{G,r}(p) := -\log(pr + 1 - p) + \frac{pr}{pr+1-p} \cdot \log r. \tag{11}$$

The following lemma establishes the cutoff structure for accepting signals:

Lemma 4. *For all $r \in (0,1)$, when $p \geq \frac{1}{2}$ $V_{F,r}(p)$ is strictly decreasing in p while $V_{G,r}(p)$ is first increasing then decreasing. Moreover, $V_{G,r}(p) > V_{F,r}(p)$ for every $p \in (\frac{1}{2}, 1)$. In words, the decision value of a "reinforcement" signal decreases in the prior probability, and it is always smaller than the decision value of a "counteracting" signal.*

Proof. We omit the subscript reference to r and calculate that

$$V_F'(p) = -\frac{1-r}{p+(1-p)r} - \frac{r \log r}{(p+(1-p)r)^2};$$
$$V_G'(p) = \frac{1-r}{pr+1-p} + \frac{r \log r}{(pr + (1-p))^2}.$$

For $V_F'(p) < 0$ we need $(1-r)(p+(1-p)r) > -r\log r$. As $p \geq \frac{1}{2}$, we only need to check $\frac{1-r^2}{2} > -r\log r$. Differentiating twice, we see this inequality indeed holds for all $r < 1$. For a similar reason, we see that $V_G'(p)$ is first positive and then negative.

To show that $V_G(p) > V_F(p)$ for $p \in (\frac{1}{2}, 1)$, we notice first that the two functions are equal at both endpoints of the interval. It then suffices to show that $V_G'(p) - V_F'(p)$ is first positive then negative, crossing 0 exactly once. We calculate that

$$
\begin{aligned}
&V_G'(p) - V_F'(p) = \\
&(1-r)\left(\frac{1}{pr+1-p} + \frac{1}{p+(1-p)r}\right) - (-r\log r)\left(\frac{1}{(pr+1-p)^2} + \frac{1}{(p+(1-p)r)^2}\right).
\end{aligned}
$$

Now, the ratio between $\frac{1}{(pr+1-p)^2} + \frac{1}{(p+(1-p)r)^2}$ and $\frac{1}{pr+1-p} + \frac{1}{p+(1-p)r}$ equals

$$
\frac{1}{1+r} \cdot \left(\frac{pr+1-p}{p+(1-p)r} + \frac{p+(1-p)r}{pr+1-p}\right),
$$

which is increasing in p. It follows that $V_G(p) > V_F(p)$, and we have the lemma. $\square$

We now define a pair of cutoffs

$$
\begin{aligned}
p_{F,r}(c) &= \max_p\{V_{F,r}(p) \geq c\}; \\
p_{G,r}(c) &= \max_p\{V_{G,r}(p) \geq c\}.
\end{aligned}
\tag{12}
$$

For every cost $c \in (0, \log 2)$ and sufficiently small r, we have $V_F(\frac{1}{2}) = V_G(\frac{1}{2}) > c$. Lemma 4 then implies that $\frac{1}{2} < p_{F,r}(c) < p_{G,r}(c)$. We will frequently omit the reference to c to ease notation.

Using Lemma 4 and the symmetry of the information structure, we deduce that the interval $[0,1]$ is divided into five subintervals : $[0, 1-p_{G,r}], [1-p_{G,r}, 1-p_{F,r}], [1-p_{F,r}, p_{F,r}], [p_{F,r}, p_{G,r}], [p_{G,r}, 1]$. If the agent's prior p is in the first or last subinterval, he is in a belief trap and rejects both signals. If p is in the central subinterval, both signals are accepted. If $p_{F,r} < p < p_{G,r}$, the agent accepts the signal s' and rejects s^*. Similarly he accepts s^* and rejects s' if $1-p_{G,r} < p < 1-p_{F,r}$.

To complete the argument for cycling, we have to show that starting with any initial prior $p_0 \in (1-p_{G,r}, p_{G,r})$, the agent's posterior never jumps out of this interval to one of the two belief traps. For this it suffices that the agent's belief never moves from the central subinterval to the rightmost (or likewise leftmost) subinterval. Using the notation we have introduced, we need to check $F_r(p_{F,r}) < p_{G,r}$. But this follows directly from Lemma 1: We have $c = D(F_r(p_{F,r})||p_{F,r}) < D(p_{F,r} \,||\, F_r(p_{F,r})) = D(G_r(F_r(p_{F,r})) \,||\, F_r(p_{F,r}))$. By definition (12), we have $p_{G,r} > F_r(p_{F,r})$. This proves the existence of belief cycling for myopic agents.

We can extend the argument to cover patient agents. Hold c and r fixed and consider the following modified information structure parametrized by ϵ:

	θ^*	θ'
s^*	$\frac{1}{r+1} \cdot \epsilon$	$\frac{r}{r+1} \cdot \epsilon$
s'	$\frac{r}{r+1} \cdot \epsilon$	$\frac{1}{r+1} \cdot \epsilon$
s^0	$1-\epsilon$	$1-\epsilon$

Notice that s^0 is totally uninformative, while conditional on $s \neq s^0$ the information structure is the same as before.

Because s^0 happens with overwhelming probability, the effective discount factor is close to 0. So under this information structure the patient agent acts as if he is myopic. The argument above can thus be applied with little change.[17]

Proof of Proposition 5

We first consider a myopic agent whose information structure is given by

	θ^*	θ'
s^*	$1-\epsilon$	$1-L\epsilon$
s'	ϵ	$L\epsilon$

with some $L > 1$ and $\epsilon > 0$ sufficiently small. As before, assume that the outcome is equal to the model (that is, if the model is a the outcome 1 has probability 1 and vice versa for b).

Define $r^* = \frac{1}{L}$ and $r' = \frac{1-L\epsilon}{1-\epsilon}$. Then using notation from the proof of Proposition 4, we know that the signal s^* takes the agent's belief from p to $F_{r'}(p)$ when he updates, while s' renders his belief $G_{r^*}(p)$. Note that $r^* < 1$ is independent of ϵ, while $r' \to 1_-$ as $\epsilon \to 0$.

Let us show that for r' sufficiently close to 1,

$$V_{G,r^*}(p) > V_{F,r'}(p), \forall p \in (0,1). \tag{13}$$

Intuitively, because the signal s^* is very weakly information of θ^*, its decision value is small uniformly over p. The remaining technical difficulty is to establish the above inequality for p close to 0 or 1. For this we calculate the derivatives of V_F and V_G near these endpoints. From Lemma 4, $V'_{G,r^*}(0) = 1 - r^* + r^* \log r^*$, which is a fixed positive number. On the other hand $V'_{F,r'}(0) = -\frac{1-r'+\log r'}{r'} \to 0$ as $r' \to 1$. Thus $V'_{F,r'}(0) < V'_{G,r^*}(0)$ and similarly $V'_{F,r'}(1) > V'_{G,r^*}(1)$ for ϵ sufficiently small, establishing (13).

Moreover, from the expression for $V'_{G,r^*}(p)$ in Lemma 4 we see that this derivative becomes uniformly (over p) close to zero as $r' \to 1$. Since $\frac{F_{r'}(p)}{p}$ becomes uniformly close to 1, we can use a similar argument to deduce

$$V_{G,r^*}(F_{r'}(p)) > V_{F,r'}(p), \forall p \in (0,1). \tag{14}$$

[17] Even though the region of accepting any signal may no longer be an interval for any positive ϵ, we can leverage the fact that posteriors take only finitely many values. Thus, the places where the interval property breaks down have vanishingly small length as $\epsilon \to 0$, and we can find an open interval of priors from which these problematic beliefs will never be reached. The argument then goes through.

Now take any $c < \min\{V_{G,r^*}(p_0), V_{G,r^*}(\eta)\}$ (recall r^* is fixed). As in the proof of Proposition 4, the interval [0,1] is divided into five subintervals L, CL, C, CR, R. In the leftmost subinterval L and the rightmost subinterval R, the agent rejects both signals. He accepts both signals in the central subinterval C, which may be empty. Finally he accepts s' but rejects s^* in *both* CL and CR, due to (13). The choice of c ensures that the initial belief p_0 does not belong to R, and (14) implies that the belief will never enter R (as p cannot jump from C to R). It follows that the agent's belief will almost surely enter L, after a finite number of consecutive s' signals. By the choice of c, the final belief is lower than η as desired.

Finally we can extend this argument to a patient agent by making ϵ even smaller, just as we did with the other proofs. A key observation in this setting is that a smaller ϵ not only makes signal s^* less informative, but it also makes any s' signal more valuable because the agent won't get another one for a long time. We omit the technical details.