Updating Beliefs with Ambiguous Evidence: Implications for Polarization*

Roland G. Fryer, Jr., Philipp Harms and Matthew O. Jackson†

Draft: June 2013

Abstract

We introduce and analyze a model in which agents observe sequences of signals about the state of the world, some of which are ambiguous and open to interpretation. Instead of using Bayes' rule on the whole sequence, our decision makers use Bayes' rule in an iterative way: first to interpret each signal and then to form a posterior on the whole sequence of interpreted signals. This technique is computationally efficient, but loses some information since only the interpretation of the signals is retained and not the full signal. We show that such rules are optimal if agents sufficiently discount the future; while if they are very patient then a time-varying random interpretation rule becomes optimal. One of our main contributions is showing that the model provides a formal foundation for *why* agents who observe exactly the same stream of information can end up becoming increasingly polarized in their posteriors.

JEL classification numbers: D10, D80, J15, J71, I30

Keywords: beliefs, polarization, learning, updating, Bayesian updating, biases, discrimination, decision making

*This project benefited greatly from extensive discussions with Richard Holden. We are also grateful to Ben Golub, Lawrence Katz, Matthew Rabin, and Andrei Shleifer for helpful comments and suggestions. The usual caveat applies.

†Fryer and Harms are at the Department of Economics, Harvard University, and NBER, 1875 Cambridge Street, Cambridge MA, 02138 (rfryer@fas.harvard.edu and pharms@edlabs.harvard.edu); and Jackson is at the Department of Economics, Stanford University, Stanford CA, 94305-6072, USA, the Santa Fe Institute (jacksonm@stanford.edu), and CIFAR. We gratefully acknowledge financial support from the NSF grants SES-0961481 and SES-1155302, and ARO MURI Award No. W911NF-12-1-0509.

"It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so. – Mark Twain

1 Introduction

Some argue that the world is becoming more polarized (Sunstein 2009, Brooks 2012). Consider, for instance, Americans' views on global warming. In a 2003 Gallup poll, 68 percent of self-identified Democrats believed that temperature changes over the past century could be attributed to human activity, relative to 52 percent of Republicans (Saad 2013). By 2013, these percentages had diverged to 78 percent and 39 percent.[1] In a similar vein, there were large racial differences in reactions to the O.J. Simpson murder trial. Four days after the verdict was announced, 73 percent of whites but only 27 percent of blacks surveyed in a Gallup/CNN/USA Today poll believed that Simpson was guilty of murdering Nicole Brown Simpson and Ronald Goldman (Urschel 1995).

Psychologists have long recognized humans' propensity to over-interpret evidence that reinforces their beliefs while disregarding contradictory information.[2] A particularly striking example is Darley and Gross (1983). Subjects in their experiment are asked to grade an essay written by a fictional student. Before reading the essay, however, graders view one of two videos of the student playing in his or her home. The first video shows the student in a poor, inner-city neighborhood, while the second depicts a middle-class suburban environment. Subjects shown the first video give the essay significantly lower grades. However, a subset of each group was also shown an identical video of the student in class, providing both correct and incorrect answers to a teacher's questions. Respondents who saw this second video diverged even further in their grading of the essay. Subjects who viewed the low-quality (high-quality) neighborhood and the class video gave the essay a lower (higher) grade than those who just saw the neighborhood video. The showing of the same video to the subjects caused them to further diverge in their grades.

The importance of polarizing beliefs in part derives from the conflict they induce. In influential papers, Esteban and Ray (1994) and Duclos, Esteban, and Ray (2004) derive measures of societal polarization and methods to estimate them. Later work has linked increased polarization to conflict both theoretically (Esteban and Ray 2011) and empirically (Esteban, Mayoral, and Ray 2013). These papers motivate understanding the mechanisms that can lead opinions to diverge even in the face of common information, since such divergence can

[1] Beliefs that the effects of global warming "have already begun to happen" and that "global warming will pose a serious threat to you or your way of life in your lifetime" show similar trajectories for both groups (Saad 2013).

[2] See, e.g., Lord, Ross, and Lepper (1979). A summary of many of these studies can be found in Table 1, and we review some of this evidence in more detail below.

result in society-level disruptions, intolerance, and discrimination.

In this paper, we introduce a model that provides a simple foundation for why the above-described polarizations in beliefs should be observed from rational agents. In our model, agents interpret information according to Bayes' rule as it is received and store the interpreted signal in memory rather than the full information, and this simple modification of Bayesian updating can lead to increasing and extreme polarization. In particular, there are two possible states of nature A, B and an agent observes a series of signals a, b that are correlated with the state of nature. Some of the signals are ambiguous and come as ab and must be interpreted. The difference from standard Bayesian agents with unbounded memory, is that we assume that an agent does not store a sequence such as $a, b, ab, ab, a, ab, b, b \ldots$ in memory, but, due to limited memory, interprets the ab signals as they are seen according to Bayes' rule (or some other rule), and then stores the interpretation in memory. So, if the agent started by believing that A was more likely, then the sequence would be interpreted and stored in memory as $a, b, a, a, a, a, b, b \ldots$. By storing the full ambiguous sequence $a, b, ab, ab, a, ab, b, b \ldots$ the agent would actually see more evidence for b than a, while interpreting signals according to Bayes' rule as they are received and then storing the interpretation leads to $a, b, a, a, a, a, b, b \ldots$ and more evidence for a than b.

This can lead people with slightly different priors to end up with increasingly different posteriors and they interpret the same string of ab's very differently, which then reinforces their differences in beliefs. As we show, if a nontrivial fraction of experiences are open to interpretation, then two agents who have differing priors and who see *exactly* the same sequence of evidence can, with positive probability, end up polarized with one agent's posterior tending to place probability one on A and the other tending to place probability one on B.

With the lens of our model, the two neighborhood videos in Darley and Gross (1983) created different priors on the students' ability. This then affected the way that respondents interpreted the second video of students answering questions. A student with a lower prior on the talent of the student would put more emphasis on the incorrect answers in the second video, and a student with a higher prior would put more emphasis on the correct answers. This then leads to the observed polarization of beliefs when grading the essays. Bayes' rule with unbounded memory would predict the opposite – seeing the same second video should lead beliefs to get closer if anything. With respect to the O.J. Simpson murder trial, people interpreted the same trial as either another example of an unfair legal system or another example of murder. The data are open to interpretation, and based on their differing past experiences and beliefs prior to the trial people reached different conclusions.

The deviation of our model from fully Bayesian updating is minimal: Bayes' rule is still used at all steps and agents are rational, they are simply constrained by limited memory and must interpret ambiguous signals before storing them and then base their posteriors on the memories rather than all of the original data. Some of our results examine optimal ways to interpret ambiguous signals. We show that the rule above, based on iterative use of Bayes' rule is optimal if the agent is not too patient, and so cares enough about making correct decisions in the short run. In contrast, if the agent is very patient, then the agent may wish to randomize interpretations and may do so in a manner that depends on the current belief.

The paper proceeds as follows. Section 2 provides a brief literature review. Section 3 describes a framework for updating beliefs when evidence is unclear and section 4 uses this insight to understand potential mechanisms that drive polarization. The final section concludes. An appendix contains the proofs of all formal results as well as additional technical details.

2 A Brief Review of the Literature

Rabin and Schrag (1999) provide a first decision-making model of confirmation bias. In each period agents observe a noisy signal about the state of the world, at which point they update their beliefs. Agents' perceptions are clouded by bias. In their model, signals that are believed to be less likely are misinterpreted with an exogenous probability. Given the exogenous mistakes introduced into the model, agents can converge to incorrect beliefs if they misinterpret contradictory evidence sufficiently frequently. The model does not clarify the mechanism behind the misinterpretation. Our model provides a foundation for the interpretation and storage of ambiguous information that can be thought of as providing a "why" behind the polarization.

In an important paper, Wilson (2003) introduced a model of limited memory that provides insight into how restrictions on the updating process motivated by the psychology literature can yield biases in beliefs (see also Mullainathan (2002) and Glaeser and Sunstein (2013)). Her agent is endowed with a finite number of "memory states," and the updating process is a set of rules for how to transition between states after observing a period's signal. The agent's goal is to know the correct state of the world when the sequence randomly terminates. The optimal transition rule turns out to be fairly simple. First, agents rank their memory states from 1 to N, with higher numbers indicating an increased likelihood that the "good" state of the world is correct. Following a good (bad) signal, the agent shifts to the next higher (lower) state until she reaches one of the extremes states ($n = 1$ or $n = N$). Wilson shows that optimizing agents with sufficient patience (probability of continuation) will leave the extreme state with very low probability. Thus, with some probability, agents will reach a state in which they ignore large amounts of seemingly contradictory information. This can yield polarization under several scenarios. For instance, two people who receive the same set of signals but start at different memory states can end up in different places for long periods of time, as they can get temporarily "stuck" at one of the extreme states. More generally, Wilson proves that as long as agents do not have identical priors and do not start at one of the extreme states, their beliefs can differ for long periods of time with positive probability.

Our paper makes at least three contributions to the existing literatures. First, our iterative process of Bayesian updating with limited memory provides a microfoundation for why agents become increasingly polarized when faced with the same data, and generates patterns and "anomalies" described in Mullainathan (2002) and Glaeser and Sunstein (2013). Second, our polarization results are different from those established in Wilson (2003) in that our agents can polarize forever, and become increasingly polarized when observing a common stream of information, while Wilson's model is ergodic: agents follow a similar irreducible and aperiodic Markov chain, simply starting in different states. That can allow them to hold different beliefs on some random occasions, but leads them to the same ergodic distribution. Third, our model is simple, transparent, and can be easily applied to alternative settings.

3 A Model of Updating Beliefs when Experiences are Open to Interpretation

There are two possible states of nature: $\omega \in \{A, B\}$.

An agent observes a sequence of signals, s_t, one at each date $t \in \{1, 2, \ldots\}$. The signals lie in the set $\{a, b, ab\}$. A signal a is evidence that the state is A, a signal b is evidence that the state is B, and the signals ab are ambiguous and open to interpretation.

In particular, signals are independent over time, conditional upon the state. With a probability $\pi < 1$ the signal is $s_t = ab$ independent of the state. With probability $1 - \pi$ the signal s_t is either a or b. If the state is $\omega = A$ and the signal is not ab, then the probability that $s_t = a$ is $p > 1/2$. Likewise, if the state is $\omega = B$ and the signal is not ab, then the probability that $s_t = b$ is $p > 1/2$.

For example, A might be a world with many racists and B a world with few racists. In A a fraction $p > 1/2$ of interactions are racist, while in state B a fraction $p > 1/2$ of interactions are non-racist. Or it might be that there are a sequence of studies regarding the impact of climate change or the efficacy of the death penalty that come out, p of which are valid and $1 - p$ of which are corrupt or flawed in state A and the reverse in state B.

$\lambda_0 \in (0, 1)$ is the agent's prior that $\omega = A$.

Throughout our analysis we assume that the prior, λ_0, is not 0 or 1, as otherwise learning is precluded. Similarly, we maintain the assumption that $\pi < 1$ as otherwise no information is ever revealed.

It follows directly from a standard law of large numbers argument (e.g., Doob's (1949) consistency theorem), that the agent's posterior beliefs converge to place weight one on the true state ω almost surely.

OBSERVATION **1** *A Bayesian-updating agent who forms beliefs conditional upon the full sequence of signals observed through each time has a posterior that converges to place probability 1 on the correct state almost surely.*

The signals ab are uninformative as they occur with a frequency π *regardless of the state.* Thus, a Bayesian updater who remembers all of the signals in their entirety effectively ignores interactions that are open to interpretation.

3.1 Interpreting Signals and Limited Memory

An important aspect of the example described in the introduction is that agents have limited memory. Otherwise, there would be no need to interpret signals – agents would store the full "scene" of every interaction in memory and be fully Bayesian.

The key element of our model is that at each date an agent can only store one bit of data. When the agent sees an ambiguous signal $s_t = ab$, the agent *must* interpret and store it as either a or b. An agent has limited memory and cannot remember all the possible interpretations of all social interactions and thus, must store in memory a strict subset of the information involved in the social interaction. The *interpretation* of ambiguous signals $s_t = ab$ as clear signals a or b is based on the agent's experiences through time t.

The only limitation our agent faces is being forced to interpret ambiguous signals and store them in memory. The agent does this in a "rational" manner, interpreting the signals according to their maximum likelihood. As we show below, this will be an optimal storage technique if the agent is faced with making decisions in the short run that depend on how they interpret the situations they face.

3.1.1 An Example of Polarized Beliefs

To illustrate this process and the polarization that it induces, let us consider a specific example. For instance, suppose that the true state is $\omega = A$, the probability that an unambiguous signal matches the state is $p = 2/3$, and the fraction of ambiguous signals is $\pi = 1/2$. In such a case, the sequence of signals might look like:

$$a, ab, ab, b, a, ab, a, ab, ab, b, ab, ab, \ldots$$

In particular, consider a case such that the agent interprets $s_t = ab$ based on what is most likely under her current belief λ_{t-1}: the agent interprets $s_t = ab$ as a if $\lambda_{t-1} > 1/2$ and as b if $\lambda_{t-1} < 1/2$.

Suppose that the agent's prior is $\lambda_0 = 3/4$. When faced with $s_1 = a$, the agent's posterior λ_1 becomes 6/7. So, the agent updates beliefs according to Bayes' rule given the interpreted (remembered) signals.[3] Then when seeing $s_2 = ab$ the agent stores the signal as a, and ends up with a posterior of $\lambda_2 = 5/6$. In this case, the agent would store the sequence as

$$a, a, a, b, a, a, a, a, a, b, a, a, \ldots$$

and the posterior at the end of this sequence would already be very close to 1.

In contrast, consider another agent whose prior is $\lambda_0 = 1/4$. When faced with $s_1 = a$, the agent's posterior λ_1 becomes 2/5. Then when seeing $s_2 = ab$ the agent stores the signal as b, and ends up with a posterior of $\lambda_2 = 1/4$. In this case, the agent would store the sequence as

$$a, b, b, b, a, b, a, b, b, b, b, b, \ldots$$

and the posterior at the end of this sequence would be very close to 0.

Two agents observing exactly the same sequence with different (and non-extreme) priors, come to have increasingly different posteriors. Their posteriors are

3/4,	6/7,	12/13,	24/25,	12/1,	24/25,	48/49,	96/97,	192/193,	384/385,	192/193,	192/193...
1/4,	2/5,	1/4,	1/7,	1/13,	1/7,	1/13,	1/7,	1/13,	1/25,	1/49,	1/97...

3.2 Optimal Interpretations in the Face of Choosing an Action

In the above example, we considered situations where the agent interprets signals according to which state is more likely, which we refer to as *maximum likelihood storage*. There is a tradeoff: if an agent stores a signal in the way that is viewed as most likely given the current beliefs that can help in reacting to the current situation at hand, but then the agent can bias long-term learning.

To explore this tradeoff, consider a situation in which the agent must take an action based on the current interaction. For instance, in the context of our racist example, at each date the agent must react appropriately to either being faced with a racist or not.[4] In the case of an ambiguous encounter, the agent must take an action. The key assumption is that if the agent decides that it is best to respond to $s_t = ab$ at time t in a manner consistent with a signal of a, then the agent must store the signal as an a. The agent cannot treat the signal in one manner for actions and then remember it differently. This is in line with our

[3] In this case $\lambda_t = P(A|s_t = a, \lambda_{t-1}) = 2\lambda_{t-1}/(1 + \lambda_{t-1})$, and $\lambda_t = P(A|s_t = b, \lambda_{t-1}) = \lambda_{t-1}/(2 - \lambda_{t-1})$

[4] Note that it is the particular situation/signal that must be appropriately interpreted, not the state. The probability of the state is useful in making a decision, but it is the particular encounters that the agent must make decisions about.

limited memory assumption.

To be more explicit, suppose the agent has to call out her interpretation a or b of the unclear signal she receives. She gets payoff of $u_t = 1$ if the current situation is correctly identified and $u_t = 0$ when a mistake is made.[5] In particular, if the agent calls out a, then $u_t = 1$ with probability p if the state is A and probability $(1-p)$ if the state is B. Similarly, if the signal is ambiguous and the agent calls out b, then $u_t = 1$ with probability $(1-p)$ if the state is A and probability p if the state is B.

To represent the full set of possible strategies that an agent might have for making interpretations (including strategies for agents with unbounded memories), we first define histories. Let a history $h_t = (s_1, i_1; \ldots ; s_t, i_t) \in \{\{a, b, ab\} \times \{a, b\}\}^t$ be a list of raw signals and their interpretations seen through time t. Let $H_t = \{\{a, b, ab\} \times \{a, b\}\}^t$ be all the histories of length t and $H = \cup_t H_t$ be the set of all finite histories.

A *strategy* for the agent is a function σ that can depend on the history and the agent's beliefs and generates a probability that the current signal is interpreted as a: $\sigma : H \times [0, 1] \to [0, 1]$.[6] In particular, $\sigma(h_{t-1}, \lambda_0)$ is the probability that the agent interprets an ambiguous signal $s_t = ab$ as a conditional upon the history h_{t-1} and the initial prior belief λ_0.

A *limited-memory strategy* is a strategy that depends only on interpreted and not on raw signals.[7]

An agent's expected payoffs can be written as:

$$U(\sigma, \delta, \lambda_0) = E\left(\sum_{t=1}^{\infty} \delta^t u_t(\sigma(h_{t-1}, \lambda_0)) \middle| \lambda_0\right).$$

The optimal strategy in the case of unconstrained memory is that of a fully rational Bayesian. It can be written solely as a function of the posterior belief, as the history that led to the posterior is irrelevant.

Our limited-memory strategies exhibit interesting history dependencies. For instance, a sequence can be reshuffled and will not affect Bayesian updating with unbounded memory. However, with bounded memory, the order of observation becomes important. Seeing a sequence where the a's all appear early tilts the prior towards the state A which then affects the interpretation of ab's towards a's. In contrast, reshuffling a sequence towards having the b's early has the opposite affect. For instance, in our example in Section $a, ab, ab, b, a, ab, a, ab, ab, b, ab, ab,$ was interpreted as $a, a, a, b, a, a, a, a, a, b, a, a, \ldots$ by anyone starting with a prior above 1/2. Suppose we reorder that original sequence to be

[5]Given the binary setting, the normalization is without loss of generality.

[6]Given that we allow the strategy to depend on the history, it is irrelevant whether we allow it to depend on the current posterior or original prior, as either can be deduced from the other given the history.

[7]A strategy is limited-memory if $\sigma(h_t) = \sigma(h'_t)$ whenever the even entries (the interpreted signals, i_t's) of h_t and h'_t coincide.

$b, b, ab, ab, ab, ab, ab, ab, a, a, ab, a,$. With the same prior in favor of A, but sufficiently close to $1/2$, the interpretation would instead flip to be $b, b, b, b, b, b, b, b, a, a, b, a,$, and end up pushing the beliefs towards B. So, the order in which a sequence of signals appear can now be (enormously) consequential.

An optimal limited-memory strategy is one that adjusts the probability of interpreting a signal with the posterior, but appears difficult to derive in a closed form. Nonetheless, we can find strategies that approximate the optimal strategy when the agent is sufficiently patient or impatient. To this aim, we will compare several classes of strategies: (a) the optimal strategy, (b) ones that depend only on the time, and (c) ones that involve randomizing in a simple manner based on the posterior.

In the latter class of strategies, the agent randomizes in interpreting ambiguous signals, but with a fixed probability that leans towards the posterior. Let $\gamma \in [0, 1]$ be a parameter such that the agent follows the posterior with probability γ and goes against the posterior with probability $1 - \gamma$. Under such a strategy, at time t with posterior λ_{t-1}, the agent interprets the unclear signal as a with probability $\gamma 1_{\lambda_{t-1} \geq .5} + (1 - \gamma) 1_{\lambda_{t-1} < .5}$ and b with the remaining probability. We denote this strategy by σ^{γ}.

The special case of $\gamma = 1$ corresponds to maximum likelihood interpretation. As we show now, maximum likelihood is an approximately optimal method of interpretation if the agent cares relatively more about the short run than the long run.

Proposition 1 *If the agent's discount factor is small enough* or *the prior belief is close enough to either 0 or 1, then the maximum likelihood strategy is approximately optimal. That is, for any λ_0 and $\varepsilon > 0$, there exist $\overline{\delta}$ such that if $\delta < \overline{\delta}$, then $U(\sigma^1, \delta, \lambda_0) \geq U(\sigma, \delta, \lambda_0) - \varepsilon$ for all strategies σ. Moreover, the same statement holds for any δ for λ_0 that are close enough to 0 or 1.*

All proofs appear in the appendix.

The intuition is straightforward. The underlying tension is between correctly calling the state in the short run, and interpreting signals over the long run. When discount rates are sufficiently high, it is best to make correct decisions in the short run. If the agent begins with a strong prior in one direction or the other (λ_0 near either 0 or 1), then maximum likelihood is optimal for most discount factors as the agent does not expect to learn much.

At the other extreme, by setting $\gamma = 1/2$, then it is clear that the agent will learn the state with probability one in the long run. However, that is at the expense of making incorrect decisions at many dates. More generally, we can state the following proposition.

Let us say that there is *long-run learning* under the randomized-interpretation strategy σ^{γ} for some γ if the resulting beliefs λ_t converge to the true state almost surely.

PROPOSITION **2** *If $\pi < \frac{p-1/2}{p}$, then ambiguous signals are infrequent enough so that there is long-run learning under the randomized-interpretation strategy σ^γ for any γ. So, consider the case in which $\pi > \frac{p-1/2}{p}$.*

(a) If $\gamma < \frac{p-1/2+\pi(1-p)}{\pi}$ then there is long run learning under the randomized-interpretation strategy σ^γ.

(b) If $\gamma > \frac{p-1/2+\pi(1-p)}{\pi}$ then there is a positive probability that beliefs λ_t converge to the wrong state under the randomized-interpretation strategy σ^γ.

3.3 Approximately Optimal Strategies: Two-Step Rules

Proposition 2 shows that long run learning under a randomized-interpretation strategy only occurs if the randomization is sufficiently high (γ is sufficiently low). This can be very costly in the long run, as although the posterior converges, it happens because the agent is randomly interpreting ambiguous signals, even when the agent is almost certain of how they should be interpreted.

The optimal strategy should adjust the randomization with the belief: as the agent becomes increasingly sure of the true state, the agent should become increasingly confident in categorizing ambiguous signals, and use less randomization. The fully optimal strategy is difficult to characterize as it is the solution to an infinitely nested set of dynamic equations and we have not found a closed-form. Nonetheless, we can easily find a strategy that approximates the optimal strategy when the agent is sufficiently patient.

Consider a T-period *two-step* rule, defined as follows. For T periods the agent uses $\gamma = 1/2$, and after T periods the agent uses $\gamma = 1$. Let σ^T denote such a strategy. We will show that with sufficient patience, such a strategy is approximately optimal.

As a strong benchmark, let σ^{FI} denote the *full-information* strategy, where the agent actually knows whether the state is A or B, and so when seeing ab then always calls a if $\omega = A$ (or b if $\omega = B$). In this case, the expected utility is independent of λ_0 and can be written as a function of the strategy and the discount factor alone: $U(\sigma^{FI}, \delta)$. This is a very stringent benchmark as it presumes information that the agent would never know even under the best circumstances. Even so, we can show that a simple two-step rule can approximate this benchmark.

PROPOSITION **3** *Sufficiently patient agents can get arbitrarily close to the full information payoff by using a variation on the two-step rule. That is, for any ϵ and λ_0, there exist T and $\overline{\delta}$ such that if $\delta > \overline{\delta}$, then $U(\sigma^T, \delta, \lambda_0) > U(\sigma^{FI}, \delta) - \epsilon$.*

Thus, with sufficient patience a very simple strategy approaches the full information benchmark.[8]

Putting these two results together, how a subjective degree of belief should change "rationally" to account for evidence which is open to interpretation depends, in important ways, on how patient a decision maker is. If they are sufficiently patient, a strategy that entails randomization in the presence of unclear evidence for finite time and then following their maximum likelihood estimate thereafter approximates the full information Bayesian outcome.[9]

In stark contrast, if agents sufficiently discount the future (or, equivalently, don't expect a large number of similar interactions in the future), they interpret ambiguous evidence as the state which has the highest maximum likelihood, given their prior belief. This leads to more informed (and optimal, though possibly mistaken) decisions in the short-run, but the potential of not learning over the longer-run.

An important remark is that we have specified two-step rules in terms of time. An alternative method is to do the following. If an agent's beliefs λ_t place at least weight ε on both states ($\lambda_t \in [\varepsilon, 1-\varepsilon]$), then randomize with equal probability on interpretations, and otherwise use the maximum likelihood method. This strategy does not require any attention to calendar time, and also will approximate the full information strategy for small enough ε.

4 Implications for Polarization of Beliefs

As we have already seen in Section 3.1.1, two agents can have very different interpretations of exactly the same events and that can lead them to update differently, and become increasingly polarized in their beliefs. The next result generalizes that observation.

PROPOSITION 4 *Suppose that a nontrivial fraction of experiences are open to interpretation so that $\pi > \frac{p-1/2}{p}$. Consider two interpretative agents 1 and 2 who both use the maximum likelihood rule but have differing priors: agent 1's prior is that A is more likely (so 1 has a prior $\lambda_0 > 1/2$) and agent 2's prior is that B is more likely (so 2 has a prior $\lambda_0 < 1/2$). Let the two agents see exactly the same sequence of signals. With a positive probability that tends to 1 in π the two agents will end up polarized with 1's posterior tending to 1 and 2's posterior tending to 0. With a positive probability tending to 0 in π the two agents will end up with the same (possibly incorrect) posterior tending to either 0 or 1.*

[8]Proposition 3 also clearly holds for a slightly different strategy: σ^x which is defined by setting $\gamma_t = 1/2$ until either $\lambda_t > 1-x$ or $\lambda_t < x$, and then setting $\gamma_t = 1$ forever after. Instead of holding for a large enough T and δ, the proposition then holds for a small enough x and large enough δ.

[9]This result highlights a familiar tradeoff between exploitation (being correct in the short run) and experimentation (long-run learning) typical in the multi-armed bandit literature (Berry and Fristedt 1985).

The proof is based on the observation that when $\pi = 1$ and $\lambda_0 > 1/2$, then all signals are interpreted as a under the maximum likelihood storage rule. Moreover, the law of the belief process depends continuously on π.

4.1 Evidence Consistent with Polarization of Beliefs

Table 5 summarizes a variety of studies in which identical information given to subjects in a variety of experimental settings resulted in increased polarization of beliefs – individuals expressing more confidence in their initial beliefs. The seminal paper in this literature is Lord, Ross, and Lepper (1979), who provided experimental subjects with two articles on the deterrent effects of capital punishment. The first article offered evidence that capital punishment has a deterrent effect on crime, while the second argued there was no relationship. The authors observe both biased assimilation (subjects rate the article reinforcing their viewpoint as more convincing) and polarization (subjects express greater confidence in their original beliefs).

Several other studies have produced similar results using opinions of nuclear power (Plous 1991), homosexual stereotypes (Munro and Ditto 1997), perceptions of fictional brands (Russo, Meloy, and Medvec 1998), theories of the assassination of John F. Kennedy (McHoskey 2002), the perceived safety of nanotechnology (Kahan et al 2007), and the accuracy of statements made by contemporary politicians (Nyahn and Reifler 2010).

Nyhan, Reifler, and Ubel (2013) provide a recent example relating to political beliefs regarding health care reform. They conduct an experiment to determine if more aggressive media fact-checking can correct the (false) belief that the Affordable Care Act would create "death panels." Participants from an opt-in Internet panel were randomly assigned to either a control group in which they read an article on Sarah Palin's claims about "death panels" or a treatment group in which the article also contained corrective information refuting Palin.

Consistent with the maximum likelihood storage rule, Nyhan, Reifler, and Ubel (2013) find that the treatment reduced belief in death panels and strong opposition to the Affordable Care Act among those who viewed Palin unfavorably and those who view her favorably but have low political knowledge. However, identical information served to *strengthen* beliefs in death panels among politically knowledgeable Palin supporters.

5 Conclusion

Polarization of beliefs has been documented by both economists and psychologists. To date, however, there has been little understanding of the underlying mechanisms that lead to such polarization, and especially why increasing polarization occurs even though agents are faced with access to the same information. To fill this void, we illuminate a simple idea: when evidence is open to interpretation, then a straightforward – and constrained optimal – iterative application of Bayes' rule can lead individuals to polarize when their information sets are identical.

Beyond polarization, our adaptation of Bayes' rule has potentially important implications for other information-based models such as discrimination. For instance, using our updating rule, it is straightforward to show that statistical discrimination can persists in a world with infinite signals. This is not true in standard models the rely on a single signal (e.g. Coate and Loury 1993). Moreover, policies designed to counteract discrimination likely have to be more complicated in an environment in which employers are forced to interpret unclear signals before hiring decisions are made.

Bayes' rule is one of the most often used results in the applied sciences – expressing how a belief should change, rationally, to account for evidence. In many applications, from economics to ecology, evidence is open to interpretation and our model can be used to form posteriors.

References

[1] Berry, Donald A. and Bert Fristedt. 1985. *Bandit Problems: Sequential Allocation of Experiments.* Monographs on Statistics and Applied Probability. London; New York. Chapman/Hall.

[2] Brooks, David. 2012, June 1. "The Segmentation Century." *The New York Times*, p. A27. Retrieved from `http://www.nytimes.com/2012/06/01/opinion/brooks-the-segmentation-century.html?_r=0`

[3] Coate, Steven, and Glenn Loury. 1993. "Will Affirmative Action Policies Eliminate Negative Stereotypes." *American Economic Review*, 83(5): 1220-40

[4] Darley, John M. and Paget H. Gross. 1983. "A Hypothesis-Confirming Bias in Labeling Effects." *Journal of Personality and Social Psychology*, 44(1): 20-33.

[5] Doob, J.L. "Application of the theory of martingales." 1949. In *Le Calcul des Probabilités et ses Applications, Colloques Internationaux du Centre National de la Recherche Scientifique*, 13: 2327.

[6] Duclos, Jean-Yves, Joan Esteban, and Debraj Ray. 2004. "Polarization: Concepts, Measurement, Estimation," *Econometrica* 72(6): 1737-1772.

[7] Esteban, Joan, Laura Mayoral, and Debraj Ray. 2013. "Ethnicity and Conflict: An Empirical Study." *American Economic Review*, forthcoming.

[8] Esteban, Joan, and Debraj Ray. 1994. "On the Measurement of Polarization." *Econometrica*, 62(4): 819-851.

[9] Esteban, Joan, and Debraj Ray. 2011. "Linking Conflict to Inequality and Polarization." *American Economic Review*, 101(4): 1345-74.

[10] Glaeser, Ed, and Cass Sunstein. "Why Does Balanced News Produce Unbalanced Views?" NBER Working Paper No. 18975.

[11] Hoeffding, Wassily. 1963. "Probability Inequalities for Sums of Bounded Random Variables." *Journal of the American Statistical Association*, 58(301): 13-30.

[12] Jacod, Jean, and Albert Shiryaev. 2003. *Limit Theorems for Stochastic Processes.* Berlin: Springer-Verlag.

[13] Lord, Charles, Lee Ross and Mark Lepper. 1979. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence." *Journal of Personality and Social Psychology*, 37(11): 2098-2109.

[14] Kahan, Dan M., Paul Slovic, Donald Braman, John Gastil, and Geoffrey L. Cohen. 2007. "Affect , Values, and Nanotechnology Risk Perceptions: An Experimental Investigation." GWU Legal Studies Research Paper No. 261; Yale Law School, Public Law Working Paper No. 155; GWU Law School Public Law Research Paper No. 261; 2nd Annual Conference on Empirical Legal Studies Paper. Available at SSRN: `http://ssrn.com/abstract=968652`

[15] McHoskey, John W. 2002. "Case Closed? On the John F. Kennedy Assassination: Biased Assimilation of Evidence and Attitude Polarization." *Basic and Applied Social Psychology*, 17(3): 395-409.

[16] Mullainathan, Sendhil. 2002. "A Memory-Based Model of Bounded Rationality." *Quarterly Journal of Economics*, 117(3): 735-774.

[17] Munro, Geoffrey D. and Peter H. Ditto. 1997. "Biased Assimilation, Attitude Polarization, and Affect in Reactions to Stereotype-Relevant Scientific Information." *Personality and Social Psychology Bulletin*, 23(6): 636-653.

[18] Nyhan, Brendan and Jason Reifler. 2010. "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior*, 32:303-330.

[19] Nyhan, Brendan, Jason Reifler, and Peter Ubel. "The Hazards of Correcting Myths about Health Care Reform." *Medical Care* 51(2): 127-132.

[20] Plous, Scott. 1991. "Biases in the Assimilation of Technological Breakdowns: Do Accidents Make Us Safer?" *Journal of Applied Social Psychology*, 21(13): 1058-1082.

[21] Russo, J. Edward, Margaret G. Meloy, and Victoria Husted Medvec. 1998. "Predecisional Distortion of Product Information." *Journal of Marketing Research*, 35(4): 438-452.

[22] Saad, Lydia. 2013, April 9. "Republican Skepticism Toward Global Warming Eases." Gallup Politics, `http://www.gallup.com/poll/161714/republican-skepticism-global-warming-eases.aspx`. Retrieved April 27, 2013.

[23] Sunstein, Cass. 2009. *Going to Extremes: How Like Minds Unite and Divide.* Oxford University Press.

[24] Urschel, Joe. 1995, October 9. "Poll: A Nation More Divided." *USA Today*, p. 5A. Retrieved from LexisNexis Academic database, April 22, 2013.

[25] Wilson, Andrea. 2003. "Bounded Memory and Biases in Information Processing." Unpublished manuscript, Princeton University.

Table 1: Summary of Belief Divergence Results

Lord, Ross, and Lepper (1979)	Experimental subjects were provided with evidence for and against the detterant effect of the death penalty. Subjects of all beliefs report that the article matching their baseline is more convincing, and students became more confident in their original position.
Darley and Gross (1983)	Subjects were asked to grade an identical essay after hearing the child described as from a poor, inner-city neighborhood or a middle-class suburban neighborhood. Subjects gave lower grades to the inner-city author, despite no change in the text of the essay. Subjects who viewed a video of the child answering quiz questions before grading the test displayed even greater divergence.
Plous (1991)	Subjects with varying opinions on nuclear energy and deterrence were provided with articles on the Three Mile Island disaster and a narrowly-averted accidental missile launch. Subjects of all viewpoints expressed increased confidence in their original viewpoints after reading the articles.
Munro and Ditto (1997)	Subjects with high and low levels of prejudice towards homosexuals were presented with two fictional studies on the empirical prevalence of a homosexual stereotype. Follow-up interviews revealed evidence of both biased assimilation and attitude polarization.
Russo, Meloy, and Medvec (1998)	Experimenters sequentially provided subjects with information on two fictional brands. In later stages, once participants have formed preferences, neutral information causes subjects to identify more strongly with their preferred brand.
McHoskey (2002)	Students were randomly selected to review either information supporting the claim that Lee Harvey Oswald acted alone in assassinating John F. Kennedy and or information pointing to a larger conspiracy. Students with extreme opinions intensify their positions when presenting with information supporting their beliefs and relax their beliefs to a lesser degree when confronted with contradictory evidence.
Kahan et al (2007)	Subjects were surveyed on their beliefs about the safety of nanotechnology after half were randomly provided with factual information about risks and benefits. Those who were exposed to information displayed greater polarization than those who were not.
Nyhan and Reifler (2010), Nyhan, Reifler, and Ubel (2013)	These studies detail a series of five experiments in which participants are asked to assess the validity of a false or misleading statement by a politician. In each case, the additional information leads the most-committed members of the targeted subgroup to intensify their misperceptions, rather than weakening them.

6 Appendix

Proof of proposition 1.

Let us first show for any λ_0 and $\varepsilon > 0$, there exist $\overline{\delta}$ such that if $\delta \leq \overline{\delta}$, then $U(\sigma^1, \delta, \lambda_0) \geq U(\sigma, \delta, \lambda_0) - \varepsilon$ for all strategies σ. Recall that

$$U(\sigma, \delta, \lambda_0) = E\left(\sum_{t=1}^{\infty} \delta^t u_t(\sigma(h_{t-1}, \lambda_0)) \middle| \lambda_0 \right).$$

Write

$$U(\sigma, \delta, \lambda_0) = E\left(u_1(\sigma(\emptyset, \lambda_0))|\lambda_0\right) + E\left(\sum_{t=2}^{\infty} \delta^t u_t(\sigma(h_{t-1}, \lambda_0)) \middle| \lambda_0 \right).$$

The basic idea is that as $\delta \to 0$, the future does not matter and the decision maker only needs to maximize the current period's payoff which amounts to choosing the most likely interpretation. Note that $E\left(\sum_{t=2}^{\infty} \delta^t u_t(\sigma(h_{t-1}, \lambda_0))|\lambda_0\right)$ lies in the interval $[-\frac{\delta}{1-\delta}, \frac{\delta}{1-\delta}]$ lies within $[-\varepsilon, \varepsilon]$ if $\delta \leq \overline{\delta} = \frac{\varepsilon}{1+\varepsilon}$. Thus, if $\delta \leq \overline{\delta}$, then

$$U(\sigma^1, \delta, \lambda_0) - U(\sigma, \delta, \lambda_0) \leq E\left(u_1(\sigma^1(\emptyset, \lambda_0))\middle|\lambda_0\right) E\left(u_1(\sigma(\emptyset, \lambda_0))|\lambda_0\right) - \varepsilon. \tag{1}$$

Next, note that

$$E\left(u_1(\sigma(\emptyset, \lambda_0))|\lambda_0\right) = E[p \Pr[i_1 = \omega] + (1-p) \Pr[i_1 \neq \omega]|\lambda_0].$$

Since $p > 1/2$, the maximizing solution is to set i_1 to match the most likely state ω given λ_0, and so σ^1 is optimal for the first period optimization. This implies that

$$E\left(u_1(\sigma^1(\emptyset, \lambda_0))\middle|\lambda_0\right) \geq E\left(u_1(\sigma(\emptyset, \lambda_0))|\lambda_0\right).$$

Thus, from (3) it follows that if $\delta \leq \overline{\delta}$, than

$$U(\sigma^1, \delta, \lambda_0) \geq U(\sigma, \delta, \lambda_0) - \varepsilon. \tag{2}$$

Next, let us now show that it is possible to choose $\overline{\delta}$ such that it approaches 1 as λ_0 approaches 0 or 1. For any δ, there exists $T(\delta)$ such that the expected sum of discounted utilities past time $T(\delta)$ amounts to less than $\varepsilon/2$ and so the utility is captured in the first $T(\delta)$ periods:[10]

$$U(\sigma, \delta, \lambda_0) \geq E\left(\sum_{t=1}^{T(\delta)} \delta^t u_t(\sigma(h_{t-1}, \lambda_0)) \middle| \lambda_0 \right) - \varepsilon/2.$$

[10] An upper bound is to set $\frac{\delta^T}{1-\delta} = \varepsilon/2$.

Next, note that the expression

$$E\left(\sum_{t=1}^{T(\delta)}\delta^t u_t(\sigma^1(h_{t-1},\lambda_0))\right)$$

is continuous in λ_0 including the extreme points of $\lambda_0 \in \{0,1\}$ for any given δ. Note also that σ^1 is the optimal strategy if $\lambda_0 = 1$, since then the expected payoff in any given period (independent of the history) is simply the probability that the interpretation is A times p plus the probability that the interpretation is the interpretation is B times $1-p$. This is maximized by setting the interpretation to A. Similarly if If $\lambda_0 = 0$ the optimal strategy is to interpret things as B in any given period. Thus, maximum likelihood storage rule, σ^1, is optimal for $\lambda_0 \in \{0,1\}$. Given the continuity, it is within $\varepsilon/2$ optimal for any λ_0 close enough to 1 or 0. So, for any δ we can find λ_0 close enough to 1 or 0 for which

$$U(\sigma^1,\delta,\lambda_0) \geq U(\sigma,\delta,\lambda_0) - \varepsilon. \tag{3}$$

which completes the proof. ■

Proof of Proposition 2. We first state an auxiliary result, from Hoeffding (1963), that is useful in proving Proposition 2.

Lemma 1 **(Hoeffding's inequality)** *If $X_1,\ldots,X_t$ are independent and $a_i \leq X_i \leq b_i$ for $i = 1,2,\ldots,t$, then for $\delta > 0$,*

$$P\left(\sum_{i=1}^{t}\big(X_i - E(X_i)\big) \geq t\epsilon\right) \leq e^{-2t^2\epsilon^2/\sum_{i=1}^{t}(b_i-a_i)}.$$

Let $n(\lambda)$ be the number of b interpreted signals minus the number of a interpreted signals needed to reach the frontier where $\lambda_t = 1/2$ starting from $\lambda_0 = \lambda$, i.e.,

$$n(\lambda) = \left\lfloor \frac{\log\left(\frac{\lambda}{1-\lambda}\right)}{\log\left(\frac{p}{1-p}\right)} \right\rfloor.$$

The $\lfloor\cdot\rfloor$ reflects starting from a prior below 1/2, and otherwise it would be rounded up.

The process $n_t = n(\lambda_t)$ is a random walk in the integers such that n_t is increased by 1 every time there is an interpreted a signal and decreased by 1 every time there is an interpreted b signal. The conditional laws given the states A and B are denoted by P_A and P_B, respectively, and E_A and E_B are the corresponding expectations.

(a) First, note that if $(1-\pi)p + \pi(1-\gamma) > 1/2$ (which is rewritten as $\gamma < \frac{1/2-(1-p)(1-\pi)}{\pi}$), then even if all of the unclear signals are incorrectly interpreted, the majority of signals will still match the true state. Therefore, if the true state is A, then the increments $\Delta n_t = n_{t+1} - n_t$ are positive in expectation, i.e., $E_A(\Delta n_t) > 0$. Moreover, they have bounded first and second moments. It follows from the strong law of large numbers that $(n_t - E_A(n_t))/t$ converges to zero P_A-a.s., which implies that $n_t \to \infty$ P_A-a.s. and $\lambda_t \to 1$ P_A-a.s. The P_B-a.s. convergence of λ_t to zero is proven in a similar same way.

Note that this is automatically satisfied if $\pi < (p-1/2)/p$ for any γ, which establishes the first sentence of the proposition.

(b) Now suppose that $(1-\pi)p+\pi(1-\gamma) < 1/2$ and assume that the true state is B. We claim that P_B assigns positive probability to the event $\lambda_t \to 1$, which coincides with the event $n_t \to \infty$. First, we note that n_t reaches any preset level with positive probability if t is large enough. Therefore, it is sufficient to prove the proposition for n_0 large. Whenever n_t is positive, it is more likely to increase than to decrease, i.e., $P_B(\Delta n_t = 1) \equiv z > 1/2$. As long as this is the case, $E_B(\Delta n_t) = 2z - 1 > 0$ and Hoeffding's inequality states that for any $\epsilon > 0$,

$$P_B\big(n_t - n_0 \le (2z - 1 - \epsilon)t\big) \le e^{-t\epsilon^2/2}.$$

Setting $\epsilon = (2z-1)/2$ leads to the bound

$$P_B\big(n_t \le (z - /2)t\big) \le P_B\big(n_t - n_0 \le (z - 1/2)t\big) \le e^{-t(z-1/2)^2/2}.$$

When n_0 is large, n_t cannot immediately fall below $(z-1/2)t$. More specifically, this is impossible for $t \le \lfloor n_0/(z+1/2) \rfloor$. It follows that

$$P_B\big(\forall t : n_t > (z - 1/2)t\big) = 1 - P_B\big(\exists t : n_t \le (z-1/2)t\big) \ge 1 - \sum_{t > \lfloor n_0/(z+1/2)\rfloor} e^{-t(z-1/2)^2/2}.$$

The last expression is positive if n_0 is large enough. This proves that

$$P_B\big(\lim_{t\to\infty} \lambda_t = 1\big) = P_B\big(\lim_{t\to\infty} n_t = \infty\big) > 0.$$

It can be shown in a similar way that P_A assigns positive probability to the event $\lambda_t \to 0$.

■

Proof of Proposition 3. In the limit, the belief process places a.s. weight 1 on the true state of Nature when the $\gamma = 1/2$ rule is used, as shown in proposition 2 (and could also be deduced from Levy's 0-1 law and Martingale convergence of beliefs). Therefore, the belief process remains eventually on the correct side of the frontier. Formally, under the $\gamma = 1/2$ rule, the random time

$$S = \inf\{t \geq 0 : \forall s \geq t : 1_{\lambda_t > 1/2} = \omega\}$$

is P_ω-a.s. finite for any $\omega \in \{A, B\}$. Therefore, for any T,

$$\begin{aligned} U(\sigma^T, \delta, \lambda) &\geq E\left(1_{S \leq T} \sum_{t=T}^{\infty} \delta^t u_t(\sigma^T(h_{t-1}, \lambda_{t-1}))\right) = E\left(1_{S \leq T} \sum_{t=T}^{\infty} \delta^t u_t(\sigma^{FI}(\omega))\right) \\ &> E\left(\sum_{t=T}^{\infty} \delta^t u_t(\sigma^{FI}(\omega))\right) - \epsilon/2 > E\left(\sum_{t=0}^{\infty} \delta^t u_t(\sigma^{FI}(\omega))\right) - \epsilon = U(\sigma^{FI}, \delta) - \epsilon. \end{aligned}$$

In the above equation, the first relation holds because some non-negative terms are dropped, the second relation holds because the agent calls out the right state past time S, the third relation holds for large enough T because $P(S \leq T) \to 1$ as $T \to \infty$, and the fourth relation holds for δ close enough to 1 because then the first T stages do not matter relative to the rest. ■

Proof of Proposition 4. The argument in the proof of proposition 2b shows that when $\pi < 1$, then the values 0 and 1 occur with positive probability as the limit of the belief process λ_t as $t \to \infty$. It remains to show that $\lambda_t \to 1$ with probability tending to one as $\pi \to 1$ if $\lambda_0 > 1/2$. So let us assume $\lambda_0 > 1/2$. Obviously, in the extreme case that $\pi = 1$, all signals are interpreted as a and therefore, the probability that $\lambda_t \to 1$ equals 1. This probability depends continuously on the parameter π by Lemma 2. ■

LEMMA **2** *Under any strategy, the belief process Λ is continuous in the total variation norm with respect to the parameters p and π.*

Proof. It is equivalent to show the proposition for the point process n_t defined in the proof of Proposition 2 instead of the process λ_t. Let $p^k \to p$, $\pi^k \to \pi$, and let P^k and P be the corresponding laws of the process n_t. Furthermore, let $\mathcal{F}_t$ be the sigma algebra generated by $n_0, \ldots, n_t$ and P_t^k the restriction of P^k to $\mathcal{F}_t$. It follows directly from the Chapman-Kolmogorov equations or from Jacod and Shiryaev (2003, corollary V.4.39a) applied to the point process $(n_t - n_0 + t)/2$ that the total variation of the signed measure $P_t^k - P_t$ tends to zero, i.e.,

$$\|P_t^k - P_t\| = \sup\{|P_t^k(\phi) - P_t(\phi)| : \phi \ \mathcal{F}_t\text{-measurable function on } \Omega \text{ with } |\phi| \leq 1\} \to 0.$$

The restriction that ϕ is $\mathcal{F}_t$-measurable can be removed by an approximation argument: for any $\mathcal{F}_t$-measurable function ϕ_t, one has

$$\begin{aligned} |P^k(\phi) - P(\phi)| &\leq |P^k(\phi) - P^k(\phi_t)| + |P^k(\phi_t) - P(\phi_t)| + |P(\phi_t) - P(\phi)| \\ &\leq |P^k(\phi) - P^k(\phi_t)| + \|P^k - P\| + |P(\phi_t) - P(\phi)|. \end{aligned} \tag{4}$$

Setting k large enough, $\|P_t^k - P_t\|$ can be made smaller than $\epsilon/3$. Then ϕ_t can be set equal to the $\mathcal{F}_t$-conditional expectation of ϕ under the measure $(P^k + P)/2$. It follows that $\phi_t \to \phi$ a.s. under P^k and P. By the dominated convergence theorem, the first and third term in the right-hand side of equation (4) are smaller than $\epsilon/3$ when t is large enough. It follows that (4) is arbitrarily small for large enough values of k. Thus P^k converges to P in the total variation norm. ■