Coarse Bayesian Updating

Alexander M. Jakobsen*

August 2022

Abstract

Studies have shown that the standard law of belief updating—Bayes' rule—is descriptively invalid in a variety of settings. In this paper, I show that a wide range of non-Bayesian updating behavior can be captured by a small modification of Bayes' rule. My framework—*Coarse Bayesian updating*—posits that agents apply subjective criteria to select among posteriors they (also subjectively) deem feasible. I characterize the model axiomatically, show that it accommodates much of the evidence on non-Bayesian updating, and derive its main implications in static and dynamic settings. Each axiom expresses a property of Bayes' rule but, conditional on the others, stops just short of making the agent fully Bayesian. Finally, I apply the model to a standard setting of decision under risk, leading to a close relationship with the Blackwell information ordering and comparative measures of cognitive sophistication and bias.

1 Introduction

Bayesian updating plays a central role in economic theory. A number of experimental findings, however, document settings in which real behavior cannot be reconciled with Bayes' rule. For example, individuals may under-react to new information or even ignore it altogether; others over-react by falsely extrapolating or, more generally, engaging in pattern-seeking behavior. "Motivated" reasoning, among other mechanisms, may lead individuals to under-react to some signals but over-react to others. Still others may be Bayesian except when information is too extreme or unexpected. Such heterogeneity, both within and between individuals, poses an interesting challenge to the Bayesian paradigm and raises several

*Kellogg MEDS, Northwestern University; alexander.jakobsen@kellogg.northwestern.edu. I am grateful to Yoram Halevy, Shaowei Ke, Dimitri Migrow, Rob Oxoby, Scott Taylor, and Benjamin Young for helpful conversations and feedback. I also thank seminar audiences at UC Davis, Edinburgh, Kellogg MEDS, Michigan, Montreal, UNSW, UTS, Toronto, and UWO as well as participants at BSE Summer Forum (2022), CETC (McGill, 2019), RUD (PSE, 2019), and SAET 2019 for their comments. Alaina Olson provided excellent research assistance. This research was supported by a SSHRC Insight Development Grant.

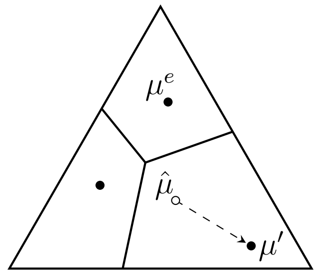


Figure 1: Coarse Bayesian updating. In this example, there are three feasible beliefs (solid dots). The point μ^e is the prior. After observing a signal, the agent determines which cell of the partition contains the Bayesian posterior $\hat{\mu}$, then adopts the representative of that cell (in this case, μ') as his new belief.

questions. Which aspects of Bayes' rule are (or are not) consistent with the wide variety of observed behavior? Do different kinds of mistakes or biases require distinct models of behavior? How "far" must one depart from Bayes' rule to accommodate the evidence, and what are the implications for economic analysis?

In this paper, I approach these questions by introducing and analyzing a generalization of Bayesian updating—*Coarse Bayesian updating*—capturing several non-Bayesian phenomena. I provide an axiomatic foundation for the model where each of three axioms imposes a high degree of Bayesian rationality but, conditional on the others, stops just short of making the agent fully Bayesian. Nonetheless, the framework permits a wide range of biases and other violations of Bayes' rule, including those above. Thus, Coarse Bayesian updating is a small, if not minimal, departure from Bayes' rule that can account for a variety of biases and individual heterogeneity in updating behavior. The model also employs standard primitives, making it amenable to economic applications; I illustrate this by deriving its main implications in a general setting of decision under risk.

Intuitively, Coarse Bayesian updating hinges on a single key property: the agent simplifies the world by considering only a subset of the probability space. Given this restriction, the agent applies subjective criteria to switch among beliefs in that set. More precisely, a Coarse Bayesian agent is characterized by (i) a partition of the probability simplex over a state space, and (ii) a representative distribution for each cell of the partition, one of which is the prior. After observing a signal, the agent determines which cell contains the Bayesian posterior and adopts the representative of that cell as posterior belief (see Figure 1).

The parameters of the model—cells and their representative points—are subjective characteristics of the individual: two Coarse Bayesians may differ in their sets of feasible beliefs, their partitions, or both. In contrast to the canonical framework of Savage (1954), then, Coarse Bayesians exhibit subjectivity not only in their prior beliefs, but in their criteria for revising those beliefs. Consequently, different Coarse Bayesians may exhibit over-reaction, under-reaction, or other biases depending on the signal, the shape of the partition, and the positions of representative points within their cells. There are several ways of interpreting this behavior, such as categorical thinking or signal distortion—I discuss these, and other, interpretations in section 2.

The first result provides a simple characterization of the updating procedure. I take as primitive a finite, exogenous state space and an updating rule specifying an individual's beliefs at every possible signal. In my framework, signals represent messages that can be generated by stochastic information structures. Thus, a signal is a profile of numbers representing likelihoods of the associated message being generated in different states. By employing such primitives, the model is readily adaptable to standard economic or game-theoretic settings.

The characterization involves three testable axioms on the updating rule, each capturing some feature of standard Bayesian behavior. First, *Homogeneity* states that beliefs are invariant to scalar transformations of signals: like Bayes' rule, Coarse Bayesian updating rules only depend on the likelihood ratios of the observed signal. Second, *Cognizance* states that if two signals result in the same belief, then so does a "garbled" signal indicating that one of those signals was generated. A natural interpretation of this axiom is that the agent understands, or is cognizant of, his own updating procedure: if he is uncertain about which of two signals was generated, but recognizes that each would lead to the same posterior belief, then he adopts that belief. Finally, *Confirmation* states that if a signal exactly supports (or confirms) some feasible belief, then the updating rule associates that belief to the given signal. Theorem 1 establishes that an updating rule has a Coarse Bayesian representation if and only if it is Homogeneous, Cognizant, and Confirmatory; moreover, the associated partition, representative elements, and prior are unique.

Next, Proposition 1 establishes that, under mild assumptions, strengthening any of the axioms to an if-and-only-if form forces the agent to be Bayesian. For example, *Homogeneity* states that if two signals have the same likelihood ratios, then they induce the same beliefs. The proposition implies that if one adds a fourth axiom, "two signals have the same likelihood ratios if they induce the same beliefs" (the converse to *Homogeneity*), then the agent must be Bayesian—the added responsiveness to information implied by this converse statement closes the gap between Bayesian and Coarse Bayesian behavior. The same property holds for *Cognizance* and *Confirmation*: adding the converse statement to either axiom makes the agent Bayesian. This is the sense in which Coarse Bayesian updating is, intuitively, a "small" departure from Bayes' rule, even though it is sufficiently flexible to capture a wide range of non-Bayesian updating behavior.

Section 3 explores the main implications of Coarse Bayesian updating and examines connections to related models and evidence. In section 3.1, I discuss evidence on biased belief updating and demonstrate how Coarse Bayesian updating can generate such behavior. In addition to under- and over-reaction, I show how Coarse Bayesians may exhibit "motivated" belief updating, limited perception, extreme-belief aversion, or susceptibility to logical fallacies. Section 3.2 examines the relationship to "paradigm shifts," including the Hypothesis-Testing model of Ortoleva (2012). In particular, I examine whether Coarse Bayesians can be represented as Bayesians with second-order priors. I show that such models, dubbed *Maximum-Likelihood* updating rules, intersect the class of Coarse Bayesian rules but that neither class subsumes the other—unless there are exactly two states, in which case every Coarse Bayesian rule can be expressed as a Maximum-Likelihood rule. Finally, section 3.3 explores some basic properties of the model in dynamic settings. I consider two categories of dynamic updating rules: *pooled* rules and *sequential* rules. Pooled rules incorporate, at every time period, the full history of signal realizations; consequently, pooled rules satisfy strong forms of path-independence and have simple convergence properties. Sequential rules, however, involve signal-by-signal updating, introducing various degrees of path dependence and more nuanced convergence properties.

Section 4 applies the model to a standard setting of decision under risk. I analyze how Coarse Bayesians value information (Blackwell experiments) when faced with menus of actions with state-dependent payoffs. I show that a Coarse Bayesian's ex-ante value of information can be expressed in a familiar posterior-separable form, then establish that, unlike Bayesians, Coarse Bayesians typically exhibit violations of the Blackwell (1951) information ordering—they need not assign higher ex-ante value to more informative experiments. I characterize the set of menus in which a given Coarse Bayesian agent always benefits from Blackwell improvements, and show that the connection runs much deeper: two Coarse Bayesians are identical—same cells, same representative points—if and only if they benefit from the same Blackwell improvements. Thus, the parameters of the model can be uniquely identified from the agent's menu-contingent rankings of Blackwell-comparable experiments.

In section 4.2, I examine how a Coarse Bayesian's welfare changes as he becomes "more Bayesian." I consider three such orderings. First, an agent is *more sophisticated* if he employs a finer partition. I show that more-sophisticated agents are characterized by heightened responsiveness to information, as captured by ex-ante value of information. Second, one agent is *more biased* than another if his updating rule exhibits larger distortions away from Bayesian posteriors. I show that greater bias is characterized by greater susceptibility to harmful exploitation in that worst-case losses, relative to a Bayesian, increase as bias increases. Importantly, neither greater sophistication nor lower bias imply that the agent is better off at all menus or signal realizations. The final result shows that such welfare enhancements require the agent to be perfectly Bayesian on a larger set of signal realizations, giving rise to a third ordering that jointly refines the sophistication and bias orderings.

Throughout the paper, my focus is on the general class of Coarse Bayesian representations and their properties. In particular, I do not take a stance on where partitions or representative elements "come from," viewing them instead as subjective (but identifiable) characteristics of an individual, much like subjective prior beliefs. There are several ways to go about endogenizing the parameters by adding assumptions about the decision problem(s) agents expect to face, the signaling structure, and costs or constraints on the fineness of the updating rule (for example, a bound on the number of cells in the partition). The results of section 4.2 suggest a slightly different approach may be valuable: rather than solving for an optimal updating rule in the context of a specific environment, one may prefer a more robust objective—characterized by the bias ordering, for example—accommodating uncertainty about the environment. I discuss this at the end of section 4.2.

1.1 Related Literature

Economists and psychologists have developed a large body of research documenting systematic violations of Bayesian updating; early contributions include Kahneman and Tversky (1972), Tversky and Kahneman (1974), and Grether (1980). As seen in the surveys of Camerer (1995), Rabin (1998), and Benjamin (2019), there is substantial variation in both experimental protocols[1] and the patterns of behavior displayed by subjects. For example, under-reaction is quite common, but by no means an established law of behavior—over-reaction occurs as well; there is mixed evidence for asymmetric processing of ego-relevant information—subjects may or may not respond differently to good news than they do to bad news; and numerous studies document individual heterogeneity—some subjects are more Bayesian than others (see Benjamin (2019) for a survey and meta-analysis of the literature).

Motivated by this evidence, several authors have developed models to better understand the mechanisms behind, and consequences of, non-Bayesian updating. Models focusing on implications of biased updating are typically cast in simplified frameworks (eg, two states of the world; particular protocols or functional form assumptions), or involve non-standard elements like ambiguous signals or framing effects. See, among others, Barberis et al. (1998), Fryer et al. (2019), Gennaioli and Shleifer (2010), Rabin and Schrag (1999), and Mullainathan et al. (2008). My emphasis, particularly in sections 3 and 4, is on implications that are reasonably independent of any particular application. As such, I employ standard primitives (a finite state space; stochastic information structures; general decision problems) that can

[1] For example, studies differ in whether subjects observe individual signals or larger samples/sequences of evidence; whether prior beliefs are objectively induced or subjectively formed by participants; whether choices are incentivized with monetary rewards; and how problems and information are framed.

be adapted to any economic model.

Decision theorists have developed axiomatic approaches to non-Bayesian updating. Kovach (2020), for example, develops a model of conservative updating. Epstein (2006) provides a model of non-Bayesian updating accommodating under-reaction, over-reaction, and other biases; Epstein et al. (2008) extend this model to an infinite-horizon setting. Zhao (2022) axiomatizes an updating rule for signals indicating that one event is more likely than another. Like these authors, I take a general approach and characterize behavior axiomatically. My model is not targeted toward a specific bias or application, but provides a general framework that can accommodate a variety of non-Bayesian behavior.

Coarse Bayesian updating resembles, to a degree, the well-known representativeness heuristic of Kahneman and Tversky (1972), wherein an individual "evaluates the probability of an uncertain event, or a sample, by the degree to which it is: (i) similar in essential properties to its parent population; and (ii) reflects the salient features of the process by which it is generated" (Kahneman and Tversky, 1972, p. 431). One might interpret Coarse Bayesian representations—cells and their representative points—as a way of formalizing the representativeness heuristic by providing an agent's subjective assessment of "similarity," "essential properties," or "salient features." There are at least two problems with this. First, the representativeness heuristic requires agents to ignore base rate (prior) information, which is inconsistent with Coarse Bayesian behavior: the prior is directly relevant to a Coarse Bayesian because the agent adopts the representative of the cell containing the Bayesian posterior. Second, the Coarse Bayesian framework accommodates behavior that is at odds, intuitively, with the representativeness heuristic. For example, Coarse Bayesian updating permits agents to be perfectly Bayesian as long as they "notice" the signal (see section 3.1). Despite the freedom afforded by the definition above, it would be a stretch to categorize such behavior as an instance of the representativeness heuristic when other explanations, like limited attention, seem more appropriate—especially since this behavior incorporates prior information in a substantive way. In section 2, I offer other interpretations of Coarse Bayesian behavior that avoid these difficulties.

Three studies are especially relevant to Coarse Bayesian updating. First, the *hypothesis testing* model introduced and axiomatized by Ortoleva (2012) posits that agents apply Bayes' rule except when news is sufficiently "surprising," in which case a maximum-likelihood criterion is applied using a second-order prior. Specifically, an agent applies Bayes' rule if the prior probability of the signal exceeds a threshold $\varepsilon \geq 0$; otherwise, the agent updates a second-order prior via Bayes' rule and selects a belief of maximal probability under the new second-order beliefs. In section 3, I show that Coarse Bayesian updating can accommodate similar behavior and compare Coarse Bayesian rules to a general class of Maximum-Likelihood updating rules. I show that Coarse Bayesian rules can be expressed as Maximum-Likelihood rules if there are only two states but that, in general, neither category subsumes the other. Notably, Maximum-Likelihood rules may violate the Confirmation property—perfect evidence of a candidate belief does not guarantee that that belief is selected.

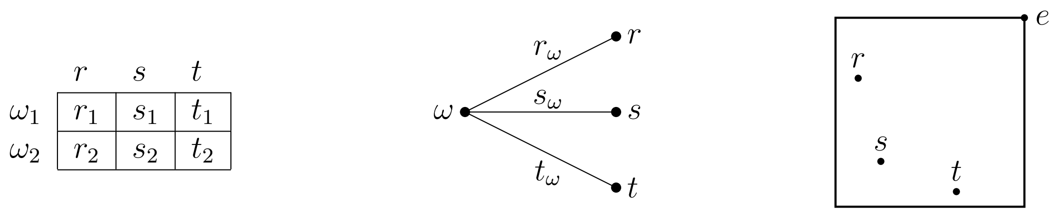
Second, Wilson (2014) studies optimal updating rules for a boundedly rational agent facing a binary decision problem and a stochastic sequence of signals. There are two states, and the agent has limited memory: only K memory states are available. In an optimal protocol, each memory state is associated with a convex set of posterior beliefs and a representative distribution for that set; if an interim Bayesian posterior belongs to some cell, then the representative of that cell is adopted as the agent's belief. Thus, the optimal protocol emerging from Wilson's model can be represented as a (dynamic) Coarse Bayesian updating procedure. Naturally, the parameters of this representation—cells and their representative points—depend on features of the environment like the signal structure, the stakes of the decision problem, and the bound K. Like Bayesian updating, Coarse Bayesian updating procedures do not depend on any factors other than the informational content of realized signals. I do not require Coarse Bayesian representations to be optimal in any sense, nor do I impose cognitive bounds such as a restriction on the number of cells. This allows my model to capture documented behavior (for example, Bayesian updating except when signals are too "extreme"—see section 3) that is inconsistent with Wilson's model.

Third, a working paper, Mullainathan (2002), develops a model of categorical thinking. Agents in this model follow a procedure similar to Coarse Bayesian updating where feasible posteriors represent categories and the mapping from Bayesian posteriors to categories is determined by a partition of the simplex. A key difference is that Mullainathan's partition is derived from the set of feasible posteriors: given a set of feasible posteriors, an optimality condition similar in spirit to maximization of a likelihood function is used to select a posterior. The resulting partition has convex cells, as in a Coarse Bayesian representation, but cells need not contain their representative elements. In other words, behavior in this model need not satisfy Confirmation—see appendix B for a concrete example.

2 Model

Let $\Omega = \{1, \ldots, N\}$ denote a finite set of $N \geq 2$ states and Δ the set of probability distributions over Ω. A distribution $\hat{\mu} \in \Delta$ assigns probability $\hat{\mu}_\omega$ to state $\omega \in \Omega$. Let Δ^0 denote the interior of Δ.

An **experiment** is a matrix with N rows, finitely many columns, and entries in $[0, 1]$ such that each row is a probability distribution and each column has a nonzero entry. Columns



(a) Matrix Form (b) Tree Form (c) Signal Space

Figure 2: Three representations of an experiment $\sigma = [r, s, t]$.

represent messages that might be generated, and rows state-contingent probability distributions over messages. Let $\mathcal{E}$ denote the set of all experiments, with generic element σ.

As in Jakobsen (2021), a **signal** is a profile $s = (s_\omega)_{\omega\in\Omega} \in [0,1]^\Omega$ such that $s_\omega \neq 0$ for at least one state ω. Let S denote the set of all signals. A signal s represents a column (message) of some experiment, and s_ω the likelihood of the message being generated in state ω. The notation $s \in \sigma$ indicates that s is a column of σ. I reserve e to denote an **uninformative signal**; that is, $e \in S$ and $e_\omega = 1$ for all $\omega \in \Omega$. Note that e qualifies as an experiment. Using the notation of signals, an experiment can be viewed as a collection (matrix) of signals, of state-contingent distributions over signals, or of points in S that sum to e; see Figure 2.

For profiles $v = (v_\omega)_{\omega\in\Omega}$ and $w = (w_\omega)_{\omega\in\Omega}$ of real numbers, let $vw := (v_\omega w_\omega)_{\omega\in\Omega}$ denote the profile formed by multiplying v and w component-wise. Similarly, if $w_\omega > 0$ for all ω, let $v/w := (v_\omega/w_\omega)_{\omega\in\Omega}$. The dot product of v and w is given by $v \cdot w := \sum_{\omega\in\Omega} v_\omega w_\omega$. The notation $v \approx w$ indicates that $v = \lambda w$ for some $\lambda > 0$, where $\lambda w := (\lambda w_\omega)_{\omega\in\Omega}$ is the scalar product of λ with w. The standard Euclidean norm of v is denoted $\|v\|$.

For $\hat{\mu} \in \Delta$ and $s \in S$ where $s \cdot \hat{\mu} \neq 0$, let $B(\hat{\mu}|s) := \frac{s\hat{\mu}}{s\cdot\hat{\mu}} \in \Delta$ denote the **Bayesian posterior** of $\hat{\mu}$ at s. Finally, an **updating rule** is a function $\mu : S \to \Delta$ assigning probability distributions $\mu^s := \mu(s) \in \Delta$ to signals $s \in S$. I assume μ^e, the **prior**, has full support.

2.1 Coarse Bayesian Representations

Consider an agent whose behavior is summarized by an updating rule $\mu : S \to \Delta$ where μ^e has full support. The interpretation is that μ^s is the agent's posterior belief conditional on observing signal s.[2] Almost any updating behavior can be expressed as an updating rule. In this section, I show that Coarse Bayesian updating is characterized by three axioms on μ. Each axiom imposes some amount of Bayesian rationality by expressing a property of Bayes'

[2]Note that updating rules condition beliefs on signal realizations s but not on experiments σ. In practice, a signal must be generated by an experiment, in which case one may wish to denote posterior beliefs by $\mu^{(\sigma,s)}$ where $s \in \sigma$. Like Bayesian updating, however, Coarse Bayesian updating depends on s but not the other columns of σ. To simplify the notation, I have omitted the underlying experiment(s) σ.

rule, and is falsifiable with data in the form of an updating rule. The axioms also lead to a simple comparison with Bayes' rule, capturing the sense in which Coarse Bayesian updating is a "small" departure from Bayesian updating.

Axiom 1 (Homogeneity). If $s \approx t$, then $\mu^s = \mu^t$.

Homogeneity requires the agent's analysis of a signal s to depend only on the likelihood ratios $s_\omega / s_{\omega'}$. This is a key feature of standard Bayesian updating, and implies the agent is not susceptible to certain types of framing effects. For example, whether information is stated in terms of frequencies or likelihoods has no effect on the agent's cognitive process.

By Homogeneity, the notation μ^s can be extended to all non-zero profiles $\tilde{s}$ such that $\tilde{s}_\omega \geq 0$ for all ω because such profiles can be scaled by a factor $\lambda > 0$ to yield a signal $\lambda \tilde{s} \in S$. This will be convenient for expressing the remaining axioms.

Axiom 2 (Cognizance). If $\mu^s = \mu^t$, then $\mu^{s+t} = \mu^s$.

Cognizance states that if signals s and t result in the same posterior belief, then the agent adopts that belief if he knows that either s or t has realized. This interpretation stems from the fact that $s + t$ is a "garbled" signal indicating that either s or t was generated.[3] Thus, an interpretation of Cognizance is that *the agent understands his own updating rule*: if he knows that one of two signals was generated and realizes that either one would lead to the same posterior belief—that is, if he is cognizant of his own updating procedure—then he ought to adopt that belief.

Although Cognizance is mainly motivated by normative considerations, it is also potentially important in applications. For example, section 4 studies how Coarse Bayesians value information. This involves ex-ante rankings of information structures that rely on correct forecasts of updating behavior. For such exercises to make sense, an assumption like Cognizance is required.

Axiom 3 (Confirmation). If $t \approx \mu^s / \mu^e$, then $\mu^t = \mu^s$.

To understand Confirmation, observe that for any s, μ^s is a feasible posterior because it is in the range of the updating rule. Moreover, any signal $t \approx \mu^s / \mu^e$ satisfies $B(\mu^e | t) = \mu^s$, making t *perfect evidence* of μ^s. So, Confirmation states that if the agent considers some signal to be sufficient evidence of some μ^*, then he also considers perfect evidence of μ^* to be sufficient evidence of μ^*. Although quite intuitive and normatively appealing, Confirmation is not satisfied by some closely-related models—see section 3.2 and appendix B.

[3] For example, if $s, t \in \sigma$, then there is a garbling matrix M such that $\sigma' = \sigma M$ collapses columns s and t to a single column $s + t$ without altering any other columns of σ.

Theorem 1. *An updating rule μ is Homogeneous, Cognizant, and Confirmatory if and only if there is a partition $\mathcal{P}$ of Δ and a profile $\mu^{\mathcal{P}} = (\mu^P)_{P\in\mathcal{P}}$ of distributions such that*

(i) each cell $P \in \mathcal{P}$ is convex,

(ii) $\mu^P \in P$ for all $P \in \mathcal{P}$, and

(iii) for all $s \in S$, $B(\mu^e|s) \in P$ implies $\mu^s = \mu^P$.

*Such a pair $\langle \mathcal{P}, \mu^{\mathcal{P}}\rangle$ is a **Coarse Bayesian Representation** of μ. If $\langle \mathcal{Q}, \mu^{\mathcal{Q}}\rangle$ is another Coarse Bayesian Representation of μ, then $\mathcal{P} = \mathcal{Q}$ and $(\mu^P)_{P\in\mathcal{P}} = (\mu^Q)_{Q\in\mathcal{Q}}$.*

Proof. First, observe that if $\alpha, \beta \geq 0$ and $s, t, \alpha s + \beta t \in S$, then

$$\begin{aligned} B(\mu^e|\alpha s + \beta t) &= \frac{(\alpha s + \beta t)\mu^e}{(\alpha s + \beta t)\cdot \mu^e} \\ &= \frac{\alpha s \cdot \mu^e}{(\alpha s + \beta t)\cdot \mu^e}\frac{s\mu^e}{s\cdot \mu^e} + \frac{\beta t \cdot \mu^e}{(\alpha s + \beta t)\cdot \mu^e}\frac{t\mu^e}{t\cdot \mu^e} \\ &= \frac{\alpha s \cdot \mu^e}{(\alpha s + \beta t)\cdot \mu^e}B(\mu^e|s) + \frac{\beta t \cdot \mu^e}{(\alpha s + \beta t)\cdot \mu^e}B(\mu^e|t). \end{aligned} \tag{1}$$

Thus, $B(\mu^e|\alpha s + \beta t)$ is a convex combination of $B(\mu^e|s)$ and $B(\mu^e|t)$; the weight attached to $B(\mu^e|s)$ is the prior probability of signal αs given that either αs or βt is generated. It is now straightforward to verify that if μ has a Coarse Bayesian Representation, then Axioms 1–3 are satisfied (Axiom 2 follows from equation (1) and convexity of cells $P \in \mathcal{P}$).

For the converse, we construct a Coarse Bayesian Representation as follows. First, note that Homogeneity and Cognizance imply μ is **Convex**: if $\mu^s = \mu^t$ and $\alpha \in [0, 1]$, then $\mu^{\alpha s+(1-\alpha)t} = \mu^s$. It follows that μ is measurable with respect to a partition of S into convex cones. That is, there is a partition $\mathcal{C}$ of S such that (i) $\mu^s = \mu^t$ if and only if there exists $C \in \mathcal{C}$ such that $s, t \in C$, and (ii) every $C \in \mathcal{C}$ is a convex cone: if $s, t \in C$ and $\alpha, \beta \geq 0$ such that $\alpha s + \beta t \in S$, then $\alpha s + \beta t \in C$. Every $C \in \mathcal{C}$ can be identified with a subset of Δ by letting $P^C := \{B(\mu^e|s) : s \in C\}$. Each set P^C is convex by equation (1) and the fact that sets $C \in \mathcal{C}$ are convex cones. In addition, $\mathcal{P} := \{P^C : C \in \mathcal{C}\}$ is a partition of Δ because $B(\mu^e|s) = B(\mu^e|t)$ if and only if $s \approx t$, forcing s and t to belong to the same cone $C \in \mathcal{C}$. For each $P \in \mathcal{P}$, let μ^P denote the unique distribution $\hat{\mu}$ such that $\mu^s = \hat{\mu}$ for all $s \in C$, where $P = P^C$. Cognizance implies $\mu^s = \mu^P \in P$ whenever $B(\mu^e|s) \in P \in \mathcal{P}$. Uniqueness of $\langle \mathcal{P}, \mu^{\mathcal{P}}\rangle$ follows from uniqueness of $\mathcal{C}$. □

Theorem 1 formalizes the concept of a Coarse Bayesian Representation and establishes that an updating rule has such a representation if and only if it is Homogeneous, Cognizant, and Confirmatory. Each of these testable axioms imposes a degree of Bayesian rationality on the agent by expressing some feature of Bayes' rule—indeed, each axiom is satisfied by a standard Bayesian. As we shall see, Coarse Bayesian updating nonetheless accommodates a variety of behavioral biases and other violations of Bayes' rule.

Coarse Bayesians partition the probability simplex, assign a representative point to each cell, and adopt the representative of a cell as posterior if the Bayesian posterior belongs to that cell. Why might an agent behave this way? Below, I offer four interpretations of the model, some of which may be more appropriate than others depending on the application.

1. *Competing Theories.* Here, the agent simplifies the world by considering a set of feasible theories (representative points), sets criteria (the partition) for switching between them, and analyzes signals to the extent necessary to determine whether a change is justified. The agent is only interested in whether the evidence satisfies his "standard of proof" for a given theory, so he does not necessarily have to fully compute the Bayesian posterior.

2. *Limited Computation.* An agent might attempt to compute the Bayesian posterior but be unable to point-identify it. Consequently, the agent lumps several posteriors together with a single point, making the representation a simplifying heuristic or approximation to Bayes' rule. Since different agents can employ different partitions or representative points, they may disagree on what constitutes a hard problem or a good approximation.

3. *Signal Distortions.* Here, to update beliefs, the agent mentally transforms signals before applying Bayes' rule. Thus, apparent deviations from Bayes' rule are the result of imperfect perception or attention—not necessarily computational constraints. Theorem 2 below formalizes the concept of Signal Distortion Representations and establishes their equivalence to Coarse Bayesian Representations in static settings. In dynamic settings, however, the distinction matters (see section 3.3).

4. *Categorical Thinking.* Here the agent reasons about categories of beliefs, each represented by a cell of the partition. This way, a cell represents distributions that share some properties of interest, and its representative point is a natural example (or "archetype") of a distribution with those properties. When information arrives, the agent determines which category applies and adopts its archetype as posterior. The key difference between this and the competing-theories interpretation is that here the cells (not their representative points) are focal: the agent is primarily interested in whether the true distribution belongs to a given category, and uses representative points to envision the category.

In each case, the parameters of the representation are subjective *characteristics of the individual*: agents may differ in their priors, partitions, or representative points. In the same way that standard Bayesian theories are agnostic about the source of one's prior beliefs, my model does not take a stance on how partitions or representative points are formed. Rather, Theorem 1 characterizes Coarse Bayesian behavior in terms of observable primitives (the updating rule) and establishes that all parameters can be uniquely identified from those primitives—with or without additional assumptions about how they might be derived.[4]

The next result provides a simple comparison between Bayesian and Coarse Bayesian behavior. To proceed, an additional definition is required; some subsequent results in the paper also utilize this definition.

Definition 1. Given $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$, a cell $P \in \mathcal{P}$ is **regular** if it has full dimension in Δ and its representative μ^P belongs to the relative interior of P. If every cell $P \in \mathcal{P}$ is regular, then $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ is **regular**.

Proposition 1. *Suppose μ is non-constant and has a Coarse Bayesian Representation $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ where every non-singleton cell of $\mathcal{P}$ is regular. Then μ is Bayesian (that is, $\mu^s = B(\mu^e|s)$ for all $s \in S$) if and only if any of the following three conditions hold:*

(i) $\mu^s = \mu^t$ *implies* $s \approx t$*;*

(ii) $\mu^{s+t} = \mu^s$ *implies* $\mu^s = \mu^t$*;*

(iii) $\mu^t = \mu^s$ *implies* $t \approx \mu^s/\mu^e$.

Proposition 1 states that, under mild regularity assumptions, strengthening *any* of Axioms A1–A3 to an if-and-only-if form forces a Coarse Bayesian agent to be perfectly Bayesian. Statement (i), the converse to Homogeneity, makes the agent highly responsive to changes to information: different likelihood ratios lead to different posterior beliefs. Statement (ii), the converse to Cognizance, requires that if the agent is unaffected by the knowledge that t may have been generated instead of s, then s and t must lead to the same beliefs. Finally, statement (iii), the converse to Confirmation, asserts that if t leads to the same posterior as s, then t must be perfect evidence of μ^s. These statements are consistent with Bayesian updating, and the proposition implies that if a Coarse Bayesian agent satisfies any of them, then the agent actually satisfies all three and behaves like a standard Bayesian.[5]

A key takeaway from Proposition 1 is that the "wedge" between Bayesian and Coarse Bayesian updating is fairly small. As we shall see, Coarse Bayesian Representations nonetheless permit many documented departures from Bayes' rule. Thus, one can accommodate a

[4]See also the discussion at the end of section 4.2 regarding approaches to endogenizing the parameters.

[5]The regularity assumptions of Proposition 1 are only needed to establish that statement (ii) forces the agent to be Bayesian—statements (i) and (iii) each make *any* Coarse Bayesian perfectly Bayesian.

variety of non-Bayesian behavior without abandoning tenets of Bayesian rationality (namely, those specified by Axioms 1–3) that, combined, almost make the agent perfectly Bayesian.

I conclude this section by providing an alternative representation of Coarse Bayesian behavior and a brief discussion of some limitations of the model.

Theorem 2. *An updating rule μ has a Coarse Bayesian Representation if and only if there is a function $d : S \to S$ such that*

(i) $s \approx t$ *implies* $d(s) \approx d(t)$,

(ii) $d(s) \approx d(t)$ *implies* $d(\lambda s + (1-\lambda)t) \approx d(s)$ *for all* $\lambda \in [0, 1]$,

(iii) $d(d(s)) = d(s)$ *for all* s,

and $\mu^s = B(\mu^e|d(s))$ *for all* s*. The function* d *is a* ***Signal Distortion Representation*** *of* μ*. If* d' *is another such representation, then* $d'(s) \approx d(s)$ *for all* $s \in S$*.*

Signal Distortion Representations formalize the signal distortion interpretation of Coarse Bayesian behavior, replacing the parameters $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ with a function d satisfying three properties analogous to Axioms A1–A3; in particular, an agent who receives signal s applies Bayes' rule to a distorted signal $d(s)$. Theorem 2 establishes that Coarse Bayesian and Signal Distortion behavior is equivalent in static settings; however, as shown in section 3.3, this equivalence fails in dynamic settings.

Naturally, Coarse Bayesian updating is not without its limitations; I discuss some of them below.

1. *Only the realized signal matters.* More precisely, Homogeneity requires that only the likelihood ratios of s can affect posterior beliefs. This rules out sensitivity to the way information is framed, as well as the possibility that extraneous features of the environment might impact beliefs.

2. *Beliefs are represented by probability distributions.* For example, the conjunction fallacy (illustrated by the well-known "Linda" problem of Tversky and Kahneman (1983)) occurs when subjects state that an event E is less likely than some conjunction $E \cap F$. Such beliefs cannot be represented by probability distributions and therefore fall outside the scope of the model.

3. *Updating is typically discontinuous in* s*.* In particular, jumps can occur when perturbations to a signal make the Bayesian posterior cross over a cell boundary. If continuity is an essential conceptual feature of some pattern of behavior—rather than a convenient

technical assumption—then Coarse Bayesian updating procedures will, at most, provide approximations to that behavior.

4. *Convex cells.* This convexity is driven by Cognizance and can be discarded by dropping that axiom. However, as explained above, Cognizance is potentially important in applications because it means agents correctly forecast their own updating behavior.

3 Models, Evidence, and Implications

Coarse Bayesian updating is related to a number of other models of non-Bayesian updating and accommodates a variety of experimental findings. In this section, I examine these relationships and explore the main implications of the model. Sections 3.1–3.3 are independent of each other and can be read in any order; section 4 is also independent of this section.

3.1 Bias, Asymmetry, and Perception

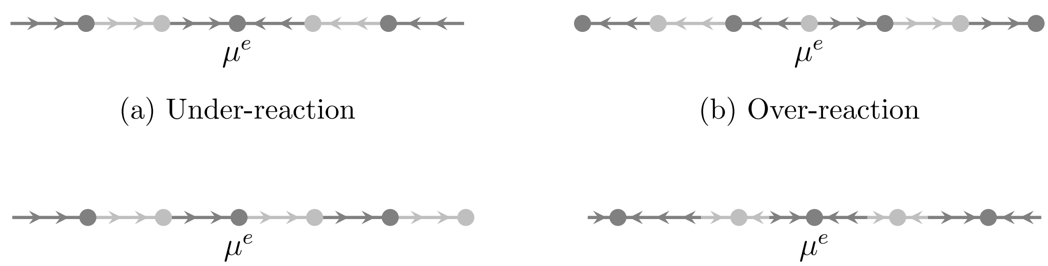
In this section, I discuss a variety biases, mistakes, and other non-Bayesian phenomena and how the Coarse Bayesian framework can accommodate them.

1. Under-reaction, Over-reaction, and Asymmetric Updating. Conservative updating, or under-reaction to information, is a well-documented behavior violating Bayes' rule.[6] Benjamin (2019), for example, conducts a review and meta-analysis of the experimental literature and finds that under-reaction is the most common bias. On the other hand, individuals also over-react to information in various settings. For example, De Bondt and Thaler (1985) find evidence of over-reaction in financial markets (in particular, when news is unexpected); more recently, Thaler (2021) finds evidence of over-reaction to weak signals and under-reaction to strong signals.

When information is "ego-relevant," subjects may respond asymmetrically to information. Eil and Rao (2011), for example, find that when information concerns personal attributes such as attractiveness, individuals under-react to negative signals but are approximately Bayesian when processing positive signals; see also Sharot and Garrett (2016) for a survey of related studies.

To represent such behavior in the Coarse Bayesian framework, I follow the literature by considering two-state settings; this way, the probability simplex Δ can be represented by the unit interval. Figures 3a and 3b illustrate under- and over-reaction. In 3a, the agent never over-reacts but typically under-reacts: his posterior belief (solid dot) is as close as possible

[6]See Phillips and Edwards (1966) and Edwards (1968) for early experiments on conservative updating.



(a) Under-reaction (b) Over-reaction

(c) Upward Bias (d) A "Typical" Coarse Bayesian

Figure 3: Four Coarse Bayesian Representations on $\Delta = [0, 1]$.

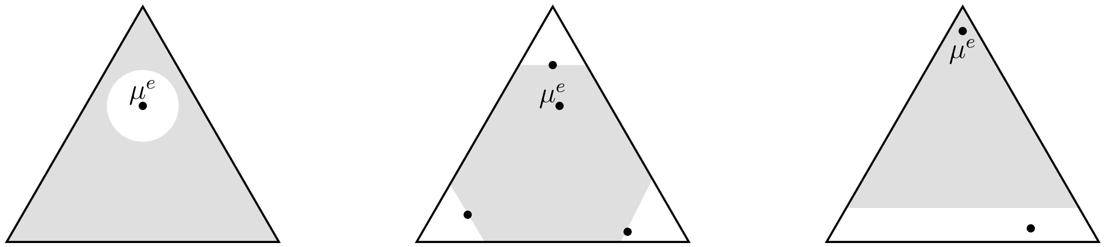
to μ^e given the partition of Δ into sub-intervals (light/dark gray regions representing different cells). In 3b, the agent never under-reacts but typically over-reacts: his posterior is farthest away from μ^e given the partition. Figure 3c exhibits a biased agent who favors one state: posteriors typically assign higher probability to state 1 than the Bayesian posterior, but never less. Thus, it is relatively easy for this agent to revise beliefs upward, but relatively difficult to revise downward. Finally, Figure 3d depicts a "typical" Coarse Bayesian: representative points do not necessarily sit on the boundaries of cells, and therefore both over- and under-reaction occur, depending on the signal.

2. Limited Perception, Extreme-Belief Aversion, and Reactions to Unexpected News. The model can also capture limited perception or attention. For example, consider Figure 4a. In this representation, the agent retains prior μ^e unless the Bayesian posterior is sufficiently far away from μ^e, in which case he applies Bayes' rule. An interpretation is that the agent only notices signals that are sufficiently strong or provocative to yield a large shift in the Bayesian posterior.[7]

Figure 4b exhibits rather the opposite behavior: the agent is Bayesian unless posterior beliefs are too "extreme"—that is, close to degenerate distributions representing certainty about the state. Ducharme (1970) argues that such behavior may explain some of the experimental evidence for under-reaction (see also Benjamin et al. (2016), who introduce the term "extreme-belief aversion"). Indeed, a Coarse Bayesian employing the representation in Figure 4b would effectively under-react to signals that strongly support a particular state.

Figure 4c illustrates an updating rule that coincides with Bayes' rule unless the observed signal is sufficiently "surprising." In this case, the prior strongly supports a particular state, and the agent exhibits non-Bayesian behavior only if the signal has a low probability of

[7]This makes the most sense in the signal-distortion interpretation of the model, where the agent transforms signals before applying Bayes' rule. The underlying signal distortion function d represents the agent's attention, avoiding the "circularity" of having the agent compute the Bayesian posterior of ignored signals—a common criticism of rational inattention models.



(a) Limited Perception (b) Extreme-Belief Aversion (c) Unexpected News

Figure 4: Limited Perception, Extreme-Belief Aversion, and Reactions to Unexpected News. Each point in the shaded regions represents a singleton cell.

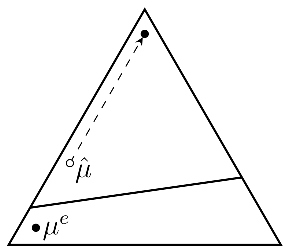


Figure 5: A straw man fallacy.

occurrence in that state. Several studies, such as De Bondt and Thaler (1985), find that updating behavior at such unexpected signals may be inconsistent with Bayes' rule. See also Ortoleva (2012), who develops a model (discussed in the next section) to accommodate this, and related, evidence.

3. Logical Fallacies. Coarse Bayesian updating can give rise to—or make agents susceptible to—various logical or rhetorical fallacies that are common in real life. By restricting attention to a subset of competing theories, agents may succumb to faulty generalizations and false extrapolations. Consequently, they may become susceptible to slippery-slope arguments, leading them to conclusions that are more extreme than what the evidence suggests.

For example, a Coarse Bayesian may be susceptible to "straw man" arguments: by providing evidence that refutes some particular theory, the agent may conclude that other theories are refuted as well, even if they do not conflict with the data. To illustrate, consider Figure 5. Here, the agent's prior places high probability on state 1 (bottom-left point), and a signal casts doubt on state 2 (bottom-right point) but not state 1: the Bayesian posterior $\hat{\mu}$ is near the edge of the simplex where only states 1 and 3 have positive probability. However, the realized posterior effectively eliminates state 1. Thus, by refuting the "straw man" theory (state 2), the signal makes the agent abandon a theory that does not conflict with the evidence.

3.2 Paradigm Shifts and Maximum-Likelihood Updating

In the competing-theories interpretation of the model, the agent employs subject thresholds (the partition) for switching among candidate beliefs. It is natural to wonder if such behavior can be reformulated in terms of second-order beliefs. If an agent assigns a prior degree of confidence to each feasible theory, can Coarse Bayesian updating be reconciled with Bayesian updating of such second-order beliefs?

To answer this question, I take an approach similar to that of Ortoleva (2012), who proposes the Hypothesis-Testing (HT) model of belief updating. Under HT, an agent applies Bayes' rule for signals of sufficiently high prior likelihood (that is, above some threshold $\varepsilon \geq 0$, an individual parameter). For unexpected signals (likelihood less than ε), the agent experiences a "paradigm shift" and updates beliefs by applying a maximum-likelihood criterion to a *second-order prior*, or "prior over priors." Specifically, the agent updates the second-order prior via Bayes' rule, then adopts as posterior a belief of maximal probability under the revised second-order distribution. In this section, I consider a similar maximum-likelihood procedure, adapted to the domain S of noisy signals.[8]

Definition 2. A Homogeneous, Convex updating rule μ has a **Maximum-Likelihood (ML) Representation** if there exists a probability distribution Γ over Δ (with density γ) such that, for all $s \in S$,

$$\mu^s \in \operatorname*{argmax}_{\hat{\mu} \in \Delta} \gamma(\hat{\mu})\hat{\mu} \cdot s.$$

The function $L : \Delta \times S \to \mathbb{R}$ given by $L(\hat{\mu}|s) = \gamma(\hat{\mu})\hat{\mu} \cdot s$ is the **likelihood function**.

In a Maximum-Likelihood Representation, the agent has a second-order prior Γ that he updates, via Bayes' rule, upon arrival of signal s. Then, he selects a belief of maximal probability under the new second-order distribution. This procedure selects among beliefs $\hat{\mu}$ that maximize the likelihood function at s.[9] Intuitively, ML updating captures the behavior of an agent who assigns prior degrees of confidence to competing theories, updates these values in a Bayesian fashion, and selects the most-likely theory given available information.[10]

[8]As Weinstein (2017) explains, the HT model allows essentially any updating to occur for unexpected news (ie, likelihood less than ε). As we shall see, extending maximum-likelihood updating procedures to the domain of noisy signals does rule out some updating behavior.

[9]Notice that L is homogeneous (of degree 0) and convex in s. The restriction to Homogeneous, Convex updating rules, therefore, only takes effect when there are ties—multiple candidate beliefs that maximize L.

[10]There are other ways of reducing a second-order belief to a first-order belief. For example, one might use the second-order distribution to compute an average belief. However, such a procedure is continuous in s while Coarse Bayesian updating, in general, exhibits discontinuities in s.

Proposition 2.

(i) Not every Maximum-Likelihood rule can be expressed as a Coarse Bayesian rule.

(ii) Not every Coarse Bayesian rule can be expressed as a Maximum-Likelihood rule.

(iii) If $N = 2$, then every Coarse Bayesian rule is a Maximum-Likelihood rule.

(iv) Bayesian updating is a special case of both Coarse Bayesian and Maximum-Likelihood updating. To express Bayesian updating as a Maximum-Likelihood rule, take

$$\gamma(\hat{\mu}) \propto \left\| \frac{\hat{\mu}}{\sqrt{\mu^e}} \right\|^{-1} \tag{2}$$

where $\sqrt{\mu^e} := (\sqrt{\mu^e_\omega})_{\omega \in \Omega}$.

Proposition 2 establishes that neither updating procedure subsumes the other—there exist updating rules that have Coarse Bayesian Representations but not ML Representations, and there exist updating rules that have ML Representations but not Coarse Bayesian Representations. These claims are demonstrated by Examples 1 and 2 below. Part (iii) establishes an important special case: if there are only two states, then every Coarse Bayesian rule can be expressed as a ML rule. Part (iv) asserts that Bayesian updating is a special case of both models and provides an explicit formula for a second-order prior generating Bayesian updating in the ML procedure. For proof of claims (iii) and (iv), see the appendix.

Example 1. Not every ML rule can be expressed as a Coarse Bayesian rule. Suppose $N = 2$ and consider the distribution γ such that $\gamma(\mu^1) = 3/4$ and $\gamma(\mu^2) = 1/4$, where $\mu^1 = (1/3, 2/3)$ and $\mu^2 = (3/4, 1/4)$. Observe that $L(\mu^1|e) = \gamma(\mu^1)\mu^1 \cdot e = \gamma(\mu^1) > \gamma(\mu^2) = \gamma(\mu^2)\mu^2 \cdot e = L(\mu^2|e)$; thus, $\mu^e = \mu^1$. It is easy to verify that $B(\mu^e|s) = \mu^2$ if and only if $s_1/s_2 = 6$. Therefore, to be consistent with a Coarse Bayesian updating rule, we must have $L(\mu^2|s) \geq L(\mu^1|s)$ whenever $s_1/s_2 = 6$. Take $s = (1, 1/6)$. Then $L(\mu^2|s) = 19/96 < 19/72 = L(\mu^1|s)$, so that the ML rule selects μ^1 at s. This means the rule is not Confirmatory, and therefore is inconsistent with Coarse Bayesian updating. ♦

Example 2. Not every Coarse Bayesian rule can be expressed as a ML rule. Suppose $N = 3$ and consider a Coarse Bayesian Representation where $\mathcal{P}$ has two cells, P and P', with $\mu^P = \mu^e$ and $\mu^{P'} = \mu' \neq \mu^e$. The boundary between P and P' corresponds to a hyperplane, H, in S. We will choose H (hence, $\mathcal{P}$) in such a way that no distribution γ on Δ (with support $\{\mu^e, \mu'\}$) generates the same updating behavior as $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ under the ML procedure.

Observe that if γ generates the same updating behavior, then $L(\mu^e|s) = L(\mu'|s)$ for all $s \in H$. In particular, $[\gamma(\mu^e)\mu^e - \gamma(\mu')\mu'] \cdot s = 0$ for all $s \in H$. Therefore, to be consistent with ML updating, the normal vector for H must be (a scalar multiple of) a member of the set $\{\lambda\mu^e - (1-\lambda)\mu' : 0 < \lambda < 1\}$; the span of this set is a 2-dimensional subset of $\mathbb{R}^3$. Thus, we may perturb the hyperplane H so that its normal does not belong to the required set. ♦

As demonstrated by Example 1 above, ML updating rules may be incompatible with Coarse Bayesian updating due to violations of Confirmation: ML rules are measurable with respect to some partition of Δ into convex cells, but cells need not contain their representative elements. I show in appendix B that the categorical-thinking model of Mullainathan (2002) also violates Confirmation in some cases, and for a similar reason. Rather than employing a second-order prior to compute likelihoods and select posteriors, Mullainathan's model uses a particular formula to calculate "base rates" for candidate beliefs. Thus, the categorical-thinking model is similar in spirit to a ML procedure, and the particular functional form employed can produce violations of Confirmation.

3.3 Dynamics

This section examines some basic dynamic properties of the model. Suppose an agent observes a sequence of signals $\vec{s} = (s^1, \ldots, s^n)$, where s^t is the signal generated in period t. How do properties of $\vec{s}$ affect the agent's final belief? Must beliefs converge to the truth?

For standard Bayesians, terminal beliefs do not depend on how signals are pooled or ordered. For example, consider a sequence $\vec{s} = (s^1, s^2, s^3)$. The terminal Bayesian belief is $B(\mu^e|s^1s^2s^3)$ regardless of whether the sequence is rearranged (eg. (s^2, s^1, s^3)), pooled differently (eg. (s^1, s^2s^3)), or both.[11] Another feature of Bayesian updating is that, for sufficiently informative structures σ, repeated draws of signals from σ make beliefs converge to the truth (a point mass δ_ω on the true state ω). More precisely, suppose the true state is ω and that for every n, s^n is an independent draw from σ (if $t \in \sigma$, then $s^n = t$ with probability t_ω). For a Bayesian, the sequence $(s^n)_{n=1}^\infty$ induces a sequence $(B^n)_{n=1}^\infty$ of beliefs $B^n = B(\mu^e|s^1s^2\ldots s^n)$ such that $B^n \to \delta_\omega$ almost surely, provided σ is sufficiently informative.[12]

Under non-Bayesian updating, including Coarse Bayesian updating, dynamics are more nuanced. For example, the terminal belief of an agent who incorporates the full history of signal realizations typically differs from that of one who performs signal-by-signal updating.

[11] See Cripps (2018) for a general analysis of updating rules that are invariant to how an agent partitions histories of signals.

[12] For example, the uninformative structure $\sigma = e$ yields $B^n = \mu^e$ for all n, so that beliefs converge to μ^e instead of δ_ω. For beliefs to converge to the truth, the distribution over signals $s \in \sigma$ for state ω (that is, the row of matrix σ corresponding to state ω) must differ from that of other states.

Similarly, matters of belief convergence depend not only on σ, but on how the (static) non-Bayesian updating rule is extended to a dynamic updating rule. Fortunately, Coarse Bayesian updating yields fairly simple results.

Some additional terminology and notation is needed to proceed. A signal s is **interior** if $s_\omega > 0$ for all $\omega \in \Omega$; let S^0 denote the set of interior signals. A **dynamic updating rule** associates a belief $\mu^{(s^1,\dots,s^n)}$ to every finite **history** $\vec{s} = (s^1, \dots, s^n)$ of interior signals. Interpreting a signal s as a history of length 1, a dynamic updating rule gives rise to an updating rule with prior μ^e.

Definition 3. A dynamic updating rule μ is:

(i) **Invariant to signal ordering** if $\mu^{\vec{s}} = \mu^{\pi(\vec{s})}$ for all histories $\vec{s}$ and permutations $\pi(\vec{s})$ of $\vec{s}$.

(ii) **Invariant to signal pooling** if, for all histories $\vec{s} = (s^1, \dots, s^n)$ of length $n \geq 2$ and all $k < n$, $\mu^{\vec{s}} = \mu^{(s^1,\dots,s^{k-1},s^k s^{k+1},s^{k+2},\dots,s^n)}$.

Definition 3 formalizes two different notions of history independence. Under invariance to signal ordering, any history $\vec{s}$ can be reordered without affecting the final belief.[13] Invariance to signal pooling, by contrast, requires that any signal in a history can be pooled with its successor without affecting the final belief. Clearly, invariance to signal pooling implies invariance to signal ordering.

Consider first the following dynamic extension of a Coarse Bayesian updating rule:

Definition 4. A dynamic updating rule μ is a **Pooled Coarse Bayesian** updating rule if either of the following equivalent conditions hold:

(i) There is a Coarse Bayesian Representation $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ such that, for all histories $(s^1, \dots, s^n)$, $\mu^{(s^1,\dots,s^n)} = \mu^P$ where $B(\mu^e | s^1 s^2 \dots s^n) \in P \in \mathcal{P}$.

(ii) There is a Signal Distortion Representation $d : S \to S$ such that, for all histories $(s^1, \dots, s^n)$, $\mu^s = B(\mu^e | d(s^1 s^2 \dots s^n))$.

A Pooled Coarse Bayesian updating rule works by applying, at every n, the full history of signals up to that point. The pooled signal $s^1 s^2 \dots s^n$ represents the joint likelihood of having

[13] Rabin and Schrag (1999) analyze a model of history-dependent updating where, at each time period, information is distorted to support the agent's current belief. Such a procedure would not be invariant to signal ordering.

observed the sequence, and these likelihoods are applied either to the Coarse Bayesian Representation $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ or its associated Signal Distortion Representation d. Naturally, Pooled Coarse Bayesian updating rules are invariant to signal pooling and, hence, signal ordering.

To study belief convergence, an additional definition is needed. A Coarse Bayesian Representation $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ is **stable at** ω if there exists $P \in \mathcal{P}$ and $\varepsilon > 0$ such that the ε-ball $\{\hat{\mu} \in \Delta : \|\hat{\mu} - \delta_\omega\| < \varepsilon\}$ around δ_ω is contained in P. The next result summarizes the dynamic properties of Pooled Coarse Bayesian updating rules.

Proposition 3. *Pooled Coarse Bayesian updating rules are invariant to signal ordering and pooling. If $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ is stable at ω, $(s^n)_{n=1}^{\infty}$ is the stochastic sequence generated by σ in state ω (that is, $s^n = t \in \sigma$ with probability t_ω), and $B(\mu^e | s^1 \ldots s^n) \overset{a.s.}{\to} \delta_\omega$, then $\mu^{(s^1, \ldots, s^n)} \overset{a.s.}{\to} \mu^P$, where $\delta_\omega \in P \in \mathcal{P}$.*

Proposition 3 states that if $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ is stable at ω and σ is sufficiently informative for Bayesian beliefs to converge to δ_ω, then Pooled Coarse Bayesian beliefs converge to the representative μ^P of the cell P containing δ_ω. Thus, Pooled Coarse Bayesian beliefs converge whenever Bayesian beliefs do, but not necessarily to the point δ_ω.

Next, consider the following two types of signal-by-signal updating:

Definition 5. A dynamic updating rule μ is:

(i) A **Sequential Coarse Bayesian** updating rule if there is a Coarse Bayesian Representation $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ for histories of length 1 such that, for every history $(s^1, \ldots, s^n)$ of length $n \geq 2$, $\mu^{(s^1, \ldots, s^n)} = \mu^P$ where $B(\mu^{(s^1, \ldots, s^{n-1})} | s^n) \in P \in \mathcal{P}$.

(ii) A **Sequential Signal Distortion** updating rule if there is a Signal Distortion Representation $d : S^0 \to S^0$ for histories of length 1 such that, for every history $(s^1, \ldots, s^n)$ of length $n \geq 2$, $\mu^{(s^1, \ldots, s^n)} = B(\mu^{(s^1, \ldots, s^{n-1})} | d(s^n))$.[14]

A Sequential Coarse Bayesian updating rule employs a fixed Coarse Bayesian Representation to perform signal-by-signal updating. Starting at prior μ^e, the agent applies $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ to reach posterior μ^{s^1} after observing s^1. Then, treating μ^{s^1} as the prior, the agent applies the same representation $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ to reach posterior $\mu^{(s^1, s^2)}$ after observing s^2, and so on. A Sequential Signal Distortion rule follows a similar procedure, substituting d for $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$. Thus, sequential rules apply to agents who have imperfect memory and rely on current beliefs as summary statistics of the history.

[14] Restricting d to take values in S^0 ensures that $B(\mu^{(s^1, \ldots, s^{n-1})} | d(s^n))$ is well defined at all possible histories.

Proposition 4. *Let μ^e have full support. Then:*

(i) Sequential Signal Distortion rules are invariant to signal ordering but not necessarily to signal pooling.

(ii) Sequential Coarse Bayesian updating rules need not be invariant to signal ordering nor to signal pooling. If there are full-support representatives $\mu^P \neq \mu^{P'}$ and a signal s^* such that both $B(\mu^P|s^*) \in P$ and $B(\mu^{P'}|s^*) \in P$, then the updating rule is not invariant to signal ordering.

Proposition 4 establishes that the path-dependence properties of sequential updating rules depend on whether the rule is implemented as a Coarse Bayesian or Signal Distortion rule: Sequential Signal Distortion rules are invariant to signal ordering, but Sequential Coarse Bayesian rules need not satisfy either type of path-independence. The requirements specified by the second part of (ii) are satisfied by many Coarse Bayesian rules; such rules fail to be invariant to signal ordering and, therefore, fail to be invariant to signal pooling as well.

The distinction between Sequential Coarse Bayesian and Signal Distortion rules also has implications for belief convergence. In general, sequences of beliefs induced by Sequential Coarse Bayesian rules need not converge to the true state, or even to converge at all. Sequential Signal Distortion rules, however, do induce belief convergence, though not necessarily to the true state.

Proposition 5. *Suppose μ is a Sequential Signal Distortion rule with distortion function d. Fix $\sigma = [t^1, \ldots, t^J]$ and $\omega \in \Omega$. Let $(s^n)_{n=1}^{\infty}$ denote a sequence of random vectors $s^n \in \sigma$ independently and identically distributed by σ in state ω (for all n, $s^n = t^j \in \sigma$ with probability t^j_ω). Let*

$$t^* = d(t^1)^{t^1_\omega} d(t^2)^{t^2_\omega} \ldots d(t^J)^{t^J_\omega} := \left(d(t^1)^{t^1_\omega}_{\omega'} d(t^2)^{t^2_\omega}_{\omega'} \ldots d(t^J)^{t^J_\omega}_{\omega'} \right)_{\omega' \in \Omega}. \tag{3}$$

Then $\mu^{(s^1,\ldots,s^n)} \to B(\mu^e|t_{E^*})$ almost surely, where $E^* = \{\omega' \in \Omega : t^*_{\omega'} \geq t^*_{\omega''} \ \forall \omega'' \in \Omega\}$ and $t_{E^*} = 1_{[\omega' \in E^*]} \in S$ is the indicator vector for E^*.

Proposition 5 states that, in the limit, Sequential Signal Distortion narrows the set of possible states down to $E^* = \text{argmax}_{\omega'} t^*_{\omega'}$, where t^* is the "average" distorted signal generated by σ in state ω. For standard Bayesians, $E^* = \{\omega\}$ provided σ is sufficiently informative. As the next example illustrates, however, E^* need not contain the true state; thus, although beliefs converge, they need not converge to the true state.

Example 3. Consider a two-state setting. Let

$$d(s) = \begin{cases} e & \text{if } \frac{s_2}{s_1} \leq 3 \\ (\frac{1}{4}, 1) & \text{otherwise} \end{cases}$$

and $\sigma = [s, t]$ where $s = (\frac{1}{5}, \frac{4}{5})$ and $t = (\frac{4}{5}, \frac{1}{5})$. Then $d(s) = (\frac{1}{4}, 1)$ and $d(t) = e$, so that in state 1 we have $t^* := d(s)^{s_1} d(t)^{t_1} = ((\frac{1}{4})^{1/5}, 1)$. Since $t^*_1 < t^*_2$, beliefs converge to state 2. ♦

4 Application: The Value of Information

Assessing the value of information is a fundamental part of decision making in many economic models. In this section, I study the Coarse Bayesian value of information, including its relationship to the Bayesian value of information, the Blackwell (1951) ordering, and notions of cognitive sophistication and bias.

Throughout this section, μ denotes an updating rule with Coarse Bayesian Representation $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$. Let $\mathcal{A}$ denote the set of all nonempty, compact subsets of $\mathbb{R}^\Omega$. Each $A \in \mathcal{A}$ is a **menu**, and elements $x = (x_\omega)_{\omega \in \Omega} \in A$ represent feasible **actions** the agent may take. Action $x \in A$ yields payoff x_ω in state ω. For each $A \in \mathcal{A}$ and $s \in S$, let $c^s(A) := \text{argmax}_{x \in A}\, x \cdot \mu^s$ denote the actions in A that maximize expected utility at beliefs μ^s.

Definition 6. Let $A \in \mathcal{A}$.

(i) The **value of information** at A is given by the function $V^A : \mathcal{E} \to \mathbb{R}$ where

$$V^A(\sigma) := \max \sum_{\omega} \mu^e_\omega \sum_{s \in \sigma} s_\omega x^s_\omega \;\text{ subject to }\; x^s \in c^s(A). \tag{4}$$

(ii) The **Bayesian value of information** at A is given by the function $\overline{V}^A : \mathcal{E} \to \mathbb{R}$ where

$$\overline{V}^A(\sigma) := \max \sum_{\omega} \mu^e_\omega \sum_{s \in \sigma} s_\omega x^s_\omega \;\text{ subject to }\; x^s \in \underset{x \in A}{\text{argmax}}\, x \cdot \frac{s\mu^e}{s \cdot \mu^e}. \tag{5}$$

Equation (4) expresses ex-ante expected utility for a Coarse Bayesian agent. Faced with a menu A and experiment σ, the agent calculates expected utility by applying weight μ^e_ω to the average payoff in state ω given that signals are generated by σ. Consistent with the Cognizance axiom, the agent correctly forecasts his own signal-contingent beliefs and, hence, signal-contingent choices. Equation (5) expresses a similar formula for an agent with the same prior μ^e but who applies Bayes' rule: signal-contingent choices maximize expected utility at beliefs $B(\mu^e|s)$ instead of beliefs μ^s.

It will be convenient to express V^A in a slightly different form. For any $\hat{\mu} \in \Delta$ and $A \in \mathcal{A}$, let

$$c^{\hat{\mu}}(A) := \operatorname*{argmax}_{x \in A} x \cdot \mu^P \text{ subject to } \hat{\mu} \in P \qquad \text{and} \qquad v^A(\hat{\mu}) := \max_{x \in c^{\hat{\mu}}(A)} x \cdot \hat{\mu}.$$

That is, $c^{\hat{\mu}}(A)$ consists of the actions in A that maximize expected utility for the Coarse Bayesian if the *Bayesian* posterior is $\hat{\mu}$. Similarly, $v^A(\hat{\mu})$ represents expected utility at A conditional on the Bayesian posterior being $\hat{\mu}$. These mappings are well-defined because $\mathcal{P}$ partitions Δ and each cell $P \in \mathcal{P}$ has a unique representative μ^P. For a standard Bayesian, analogous mappings are given by

$$\overline{c}^{\hat{\mu}}(A) := \operatorname*{argmax}_{x \in A} x \cdot \hat{\mu} \qquad \text{and} \qquad \overline{v}^A(\hat{\mu}) := \max_{x \in \overline{c}^{\hat{\mu}}(A)} x \cdot \hat{\mu}.$$

If $\sigma \in \mathcal{E}$ and $\hat{\mu} \in \Delta$, let $\tau^\sigma(\hat{\mu}) := \sum_{s \in \sigma : B(\mu^e|s) = \hat{\mu}} s \cdot \mu^e$; this is the total probability of generating Bayesian posterior $\hat{\mu}$ under information σ and prior μ^e. That is, given μ^e, σ generates a distribution of Bayesian posteriors where $\tau^\sigma(\hat{\mu})$ is the probability of posterior $\hat{\mu}$.

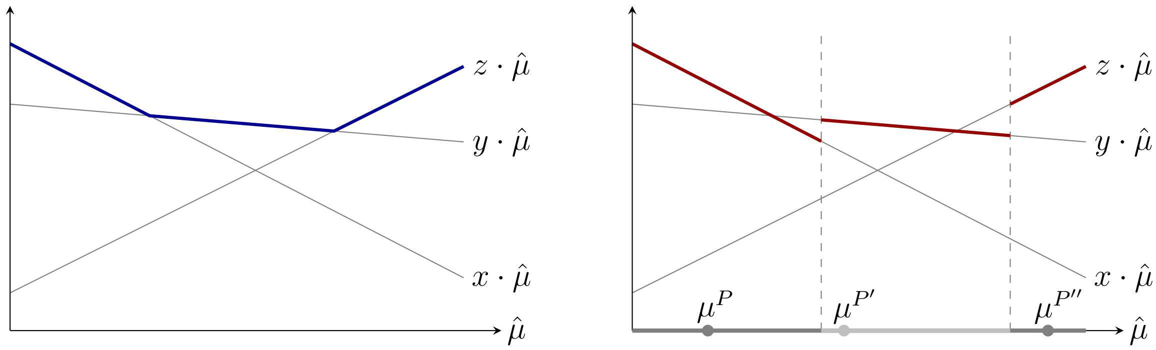
Proposition 6. *For all $A \in \mathcal{A}$ and $\sigma \in \mathcal{E}$, $V^A(\sigma) = \sum_{\hat{\mu} \in \Delta} \tau^\sigma(\hat{\mu}) v^A(\hat{\mu})$.*

Proposition 6 states that V^A can be written in posterior-separable form. In particular, it is as if the agent associates value $v^A(\hat{\mu})$ to Bayesian posterior $\hat{\mu}$, so that the distribution of Bayesian posteriors can be used to calculate expected utility. This also facilitates comparisons between Bayesian and Coarse Bayesian payoffs (see Figure 6); clearly, $v^A(\hat{\mu}) \leq \overline{v}^A(\hat{\mu})$ for all $\hat{\mu}$ and, hence, $V^A(\sigma) \leq \overline{V}^A(\sigma)$ for all σ—the Bayesian always does better. Intuitively, Proposition 6 holds because a Coarse Bayesian updating rule is Homogeneous and, hence, a function of the Bayesian posterior;[15] I omit the straightforward proof.

4.1 The Blackwell Ordering

This section examines whether and when Coarse Bayesians benefit from improvements to information. For experiments σ, σ', the relation $\sigma \sqsupseteq \sigma'$ indicates that σ is more informative than σ' in the sense of Blackwell (1951). An experiment σ' is a **garbling** of σ if there is a

[15] This is the fundamental assumption of de Clippel and Zhang (2022), who study persuasion with non-Bayesian agents. A similar result appears in Galperti (2019). In general, posterior separability means the model is well-suited for the analysis of persuasion models, as in Kamenica and Gentzkow (2011).



(a) $\overline{v}^A(\hat{\mu})$ (b) $v^A(\hat{\mu})$

Figure 6: Bayesian vs. Coarse Bayesian value of information for $A = \{x, y, z\}$.

matrix M with entries in $[0, 1]$ such that every row is a probability distribution and $\sigma' = \sigma M$. For the purposes of this paper, the ordering $\sqsupseteq$ is defined by: $\sigma \sqsupseteq \sigma'$ if and only if σ' is a garbling of σ.

The function V^A **satisfies the Blackwell ordering** if $\sigma \sqsupseteq \sigma'$ implies $V^A(\sigma) \geq V^A(\sigma')$; if there exists $\sigma \sqsupseteq \sigma'$ such that $V^A(\sigma) < V^A(\sigma')$, then V^A **violates the Blackwell ordering**. An important part of Blackwell's characterization is that a Bayesian's value of information satisfies the Blackwell ordering in all menus A—in fact, $\sigma \sqsupseteq \sigma'$ *if and only if* $\overline{V}^A(\sigma) \geq \overline{V}^A(\sigma')$ for all $A \in \mathcal{A}$. For Coarse Bayesians, this need not be the case.

For every menu A and signal s, let $b^s(A) \subseteq A$ denote the Bayesian-optimal actions in A conditional on s. Formally, $b^s(A) := \{x \in A : x \cdot \frac{s\mu^e}{s \cdot \mu^e} \geq y \cdot \frac{s\mu^e}{s \cdot \mu^e} \ \forall y \in X\}$. Let $c(A) = \bigcup_{s \in S} c^s(A)$ and $b(A) = \bigcup_{s \in S} b^s(A)$. That is, $c(A)$ is the set of actions in A that are chosen by the Coarse Bayesian—and $b(A)$ the set of actions chosen by the Bayesian—for at least one s. Observe that, by Confirmation, $c(A) \subseteq b(A)$.

Proposition 7. *Let $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ be a regular Coarse Bayesian Representation and $A \in \mathcal{A}$. The following are equivalent:*

(i) V^A satisfies the Blackwell ordering.

(ii) v^A is convex.

(iii) $c^s(A) \cap b^s(c(A)) \neq \emptyset$ for all s.

Proposition 7 characterizes, for regular Coarse Bayesians, the class of menus A such that V^A satisfies the Blackwell ordering.[16] The key property is (iii), asserting that Coarse

[16] The regularity requirement only serves to establish (i) $\Rightarrow$ (iii). In particular, the implication (iii) $\Rightarrow$ (i) holds for all Coarse Bayesian Representations, as does the equivalence of (i) and (ii). The implication

Bayesian choices from A agree with Bayesian choices from the menu $c(A) \subseteq A$ (that is, the submenu of actions that are actually chosen at some signal realization). When (iii) is satisfied, Coarse Bayesian behavior at A coincides with Bayesian behavior at $c(A)$, making $v^A = \overline{v}^{c(A)}$ convex and $V^A = \overline{V}^{c(A)}$ satisfy the Blackwell ordering. Since (iii) is a rather strong requirement, this means Blackwell violations are a fairly common occurrence.

Example 4. Some non-Bayesians satisfy the Blackwell ordering in all menus. Suppose $N = 2$, so that Δ is represented by the interval $[0, 1]$ of values $\hat{\mu}_1$. First, consider $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ where $\mathcal{P}$ contains two cells: $P = \{0\}$ and $P' = (0, 1]$. Assume $\mu^{P'} < 1$. Then, for every A, v^A is convex; this implies V^A satisfies the Blackwell ordering, even though choices generated by $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ violate condition (iii) of Proposition 7 in some menus. Next, let $\langle \mathcal{Q}, \mu^{\mathcal{Q}} \rangle$ consist of a cell $Q = [0, \mu^*]$ where $0 < \mu^* < 1$ and, for every $\hat{\mu} > \mu^*$, a singleton cell $\{\hat{\mu}\}$. Let $\mu^Q = \mu^*$. Choices generated by $\langle \mathcal{Q}, \mu^{\mathcal{Q}} \rangle$ satisfy condition (iii) of Proposition 7 for all A; this implies the corresponding value of information function satisfies the Blackwell ordering in all menus, even though $\langle \mathcal{Q}, \mu^{\mathcal{Q}} \rangle$ violates the regularity assumption (see footnote 16). ♦

Example 4 shows that it is possible for non-Bayesian representations to generate functions V^A satisfying the Blackwell ordering for all A with or without condition (iii) of Proposition 7. Such representations are quite rare, however, in that small perturbations of the cells or representative points guarantee that V^A violates both the Blackwell ordering and condition (iii) for some A. Intuitively, violations of the Blackwell ordering arise through discontinuities in v^A because such discontinuities, except possibly on the boundary of Δ, make v^A non-convex. Most non-Bayesian representations have the property that any violation of (iii) introduces a non-convexity in v^A for some A because the gap between Bayesian and non-Bayesian choices creates points of discontinuity. For regular representations, violations of (iii) are both necessary and sufficient for the existence of such discontinuities.

While it is perhaps not too surprising that non-Bayesian updating behavior can generate violations of the Blackwell ordering, it turns out that, for Coarse Bayesians, the connection to the Blackwell ordering runs much deeper:

Proposition 8. *Suppose $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ and $\langle \mathcal{Q}, \dot{\mu}^{\mathcal{Q}} \rangle$ are regular Coarse Bayesian Representations of μ and $\dot{\mu}$, respectively, such that $\mu^e = \dot{\mu}^e$. The following are equivalent:*

(i) $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle = \langle \mathcal{Q}, \dot{\mu}^{\mathcal{Q}} \rangle$.

(ii) For all $\sigma \sqsupseteq \sigma'$ and $A \in \mathcal{A}$, $V^A(\sigma) \geq V^A(\sigma') \Leftrightarrow \dot{V}^A(\sigma) \geq \dot{V}^A(\sigma')$.

(ii) ⇒ (i) is part of Blackwell's characterization, but the converse implication is not, and relies on the assumption that μ^e has full support (see Lemma 1 in the appendix).

Proposition 8 states that, for a regular Coarse Bayesian, the parameters $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ are pinned down by the agent's ranking of Blackwell-comparable experiments. Thus, by observing when the agent benefits (or expects to benefit) from a Blackwell improvement, one can uniquely identify the parameters of the representation. A key takeaway, then, is not just that Coarse Bayesians exhibit violations of the Blackwell ordering, but that they do so in a way that fully reveals their updating behavior.

4.2 Measures of Sophistication and Bias

In this section, I explore different notions of cognitive ability and how they relate to a Coarse Bayesian's value of information. In addition to providing basic comparative static results for the model, the findings are potentially relevant for endogenizing non-Bayesian updating rules and, hence, developing theories of where they "come from" (see the discussion at the end of the section). For any updating rule μ and signal $s \in S$, let

$$D_\mu(s) := \left\| \frac{s\mu^e}{\|s\mu^e\|} - \frac{\mu^s}{\|\mu^s\|} \right\|.$$

This is the Euclidean distance between μ^s and the Bayesian posterior $\frac{s\mu^e}{s \cdot \mu^e}$ after normalizing each vector to length 1. Thus, $D_\mu(s)$ provides a measure of how distorted the agent's beliefs are at signal s.

Definition 7. Suppose μ and $\dot{\mu}$ have full-support priors $\mu^e = \dot{\mu}^e$ and Coarse Bayesian Representations $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ and $\langle \mathcal{Q}, \dot{\mu}^{\mathcal{Q}} \rangle$, respectively. Then:

(i) $\dot{\mu}$ is **more sophisticated** than μ if every $P \in \mathcal{P}$ is a union of cells in $\mathcal{Q}$.

(ii) $\dot{\mu}$ is **less biased** than μ if $D_{\dot{\mu}}(s) \leq D_\mu(s)$ for all $s \in S$.

Definition 7 provides two comparative notions of cognitive ability. Part (i) states that a Coarse Bayesian is more sophisticated if he employs a finer partition, while part (ii) states the agent is less biased if, for every signal, posterior beliefs are closer to the Bayesian posterior. Each ordering captures some aspect of what it means to be "more Bayesian," but the two concepts are quite different: higher sophistication entails higher responsiveness to information, while lower bias entails less skewness in the updating rule (see Figure 7).

The goal of this section is to characterize these orderings in terms of the welfare of the agent. A natural conjecture, for example, is that a more sophisticated agent always enjoys a higher expected utility than a less sophisticated agent, or benefits from more information whenever a less sophisticated agent does. As the next example shows, this conjecture is false.

(a) More biased (b) Less biased

Figure 7: An illustration of the bias ordering. The two updating rules employ the same pair feasible beliefs, but rule (b) is less biased than rule (a) because it exhibits smaller distortions away from Bayesian posteriors; this makes the cutoff between cells more "centered."

Example 5. Consider a two-state setting, so that $\Delta = [0,1]$. Let $\mathcal{P} = \{P, P'\}$ where $P = \{0\}$ and $P' = (0,1]$ and $\mathcal{Q} = \{Q, Q', Q''\}$ where $Q = \{0\}$, $Q' = [\frac{3}{4}, 1]$, and $Q'' = (0, \frac{3}{4})$. Finally, let $\mu^P = \dot{\mu}^Q = 0$, $\mu^e = \mu^{P'} = \frac{4}{5} = \dot{\mu}^{Q'} = \dot{\mu}^e$, and $\dot{\mu}^{Q''} = \frac{1}{3}$. Clearly, $\dot{\mu}$ is more sophisticated than μ. Let $A = \{x, y\}$ where $x = (1, 0)$ and $y = (0, 1)$. Then

$$v^A(\hat{\mu}_1) = \begin{cases} 1 & \text{if } \hat{\mu}_1 = 0 \\ \hat{\mu}_1 & \text{otherwise} \end{cases} \qquad \text{and} \qquad \dot{v}^A(\hat{\mu}_1) = \begin{cases} 1 - \hat{\mu}_1 & \text{if } \hat{\mu}_1 < \frac{3}{4} \\ \hat{\mu}_1 & \text{otherwise} \end{cases},$$

so that $\dot{v}^A(\hat{\mu}_1) < v^A(\hat{\mu}_1)$ for $\frac{1}{2} < \hat{\mu}_1 < \frac{3}{4}$. Thus, $\dot{V}^A(\sigma) < V^A(\sigma)$ for some σ (for example, any σ such that $\tau^\sigma(\frac{2}{3}) = \frac{3}{5}$ and $\tau^\sigma(1) = \frac{2}{5}$). Moreover, v^A is convex but $\dot{v}^A$ is not; thus, V^A satisfies the Blackwell ordering but $\dot{V}^A$ does not. ◆

In general, greater sophistication does not guarantee greater welfare because, on its own, greater sophistication does not rule out the possibility of wider gaps between Bayesian and Coarse Bayesian choices at some menu-signal pairs. Similarly, lower bias need not imply welfare improvements. At the end of this section, I return to this question and examine the conditions under which one Coarse Bayesian is better off than another in all decision problems (Proposition 11).

To characterize the sophistication ordering, an additional definition is required. Given $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$, a pair (A, σ) is **$\mu^{\mathcal{P}}$-decisive** if $c^s(A)$ is a singleton for all $s \in \sigma$; that is, if no posterior μ^P induced by σ makes the agent indifferent between two or more options in A. For any $\sigma, \sigma' \in \mathcal{E}$, $V(\sigma) = V(\sigma')$ **$\mu^{\mathcal{P}}$-decisively** if $V^A(\sigma) = V^A(\sigma')$ for all A such that (A, σ) and (A, σ') are $\mu^{\mathcal{P}}$-decisive.

Proposition 9. *Suppose $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ and $\langle \mathcal{Q}, \dot{\mu}^{\mathcal{Q}} \rangle$ are regular Coarse Bayesian Representations of μ and $\dot{\mu}$, respectively, and that $\mu^e = \dot{\mu}^e$. The following are equivalent:*

(i) $\dot{\mu}$ is more sophisticated than μ.

(ii) If $\sigma, \sigma' \in \mathcal{E}$ and $\dot{V}(\sigma) = \dot{V}(\sigma')$ $\dot{\mu}^{\mathcal{Q}}$-decisively, then $V(\sigma) = V(\sigma')$ $\mu^{\mathcal{P}}$-decisively.

This result states that for regular Coarse Bayesians, greater sophistication means welfare is more responsive to information: as sophistication increases, fewer pairs σ, σ' yield identical ex-ante expected utility for (almost) all menus A. The proof of Proposition 9 shows that the characterization holds even if one restricts attention to experiments σ, σ' that are Blackwell comparable. Thus, higher sophistication means greater responsiveness to *improvements* to information. More-responsive welfare, of course, does not imply greater welfare.

The characterization of the bias ordering does not involve the responsiveness of welfare, but rather a comparison to that of a Bayesian. For each $s \in S$ and $A \in \mathcal{A}$, let $\overline{V}^A(s) := \overline{v}^A(B(\mu^e|s))$ and $V^A(s) := v^A(B(\mu^e|s))$ denote the Bayesian and Coarse Bayesian payoffs at menu A conditional on signal s. Let

$$L_\mu(s) := \sup_{A \in \mathcal{A}^*} \overline{V}^A(s) - V^A(s)$$

where $\mathcal{A}^*$ denotes the set of menus A such that $\|x\| \leq 1$ for all $x \in A$. Intuitively, $L_\mu(s)$ is the maximum loss, relative to a Bayesian, that the Coarse Bayesian can incur under any decision problem A.[17] Alternatively, $L_\mu(s)$ may be interpreted as the maximum rate at which a Bayesian agent can "money pump" the Coarse Bayesian agent under public information s. So, if actions x represent bets or gambles, and a Bayesian agent is free to specify a set $A \in \mathcal{A}^*$ after both agents have observed s, then $L_\mu(s)$ is the amount of money the Bayesian can extract from the Coarse Bayesian.[18]

Proposition 10. *Suppose μ and $\dot{\mu}$ are Coarse Bayesian and $\mu^e = \dot{\mu}^e$. Then $L_{\dot{\mu}}(s) \leq L_\mu(s)$ if and only if $D_{\dot{\mu}}(s) \leq D_\mu(s)$. Thus, $\dot{\mu}$ is less biased than μ if and only if $L_{\dot{\mu}}(s) \leq L_\mu(s)$ for all $s \in S$.*

Proposition 10 establishes that $\dot{\mu}$ is less biased than μ if and only if $\dot{\mu}$ is less exploitable than μ in that worst-case losses for $\dot{\mu}$, relative to a Bayesian, are smaller than those for μ.

As indicated above, neither greater sophistication nor lower bias guarantee higher payoffs in all decision problems. The next result establishes that, under mild regularity conditions, a particular refinement of these orderings is needed to improve payoffs in all decision problems.

Proposition 11. *Suppose $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ and $\langle \mathcal{Q}, \dot{\mu}^{\mathcal{Q}} \rangle$ are Coarse Bayesian Representations of μ and $\dot{\mu}$ such that $\mu^e = \dot{\mu}^e$ and non-singleton cells are regular. The following are equivalent:*

[17] The restriction to normalized menus $A \in \mathcal{A}^*$ is needed because $V^{\lambda A} = \lambda V^A$ for all $\lambda > 0$.

[18] Indeed, as shown in the appendix, one may restrict attention to menus of the form $A = \{0, x\}$ where, conditional on s, the Bayesian prefers the safe option 0 but the Coarse Bayesian strictly prefers x. On average, the Bayesian profits by $|x \cdot B(\mu^e|s)|$.

(i) $\dot{v}^A(\hat{\mu}) \geq v^A(\hat{\mu})$ *for all* $A \in \mathcal{A}$ *and* $\hat{\mu} \in \Delta$.

(ii) $\dot{\mu}$ *is less biased, more sophisticated and, for every* $\dot{\mu}^Q \in \dot{\mu}^{\mathcal{Q}} \backslash \mu^{\mathcal{P}}$, *the cell* Q *is a singleton.*

Proposition 11 states that payoffs increase at all menu-signal pairs if and only if the agent becomes more sophisticated and all "new" feasible posteriors $\dot{\mu}^Q$ are contained in singleton cells Q. This means the agent becomes perfectly Bayesian in some subset of Δ, preventing the introduction of new or different distortions that can yield lower payoffs in some menu-signal pair. It follows immediately that the agent is less biased and that $\dot{V}^A(\sigma) \geq V^A(\sigma)$ for all $A \in \mathcal{A}$ and $\sigma \in \mathcal{E}$.

I conclude this section with a brief discussion of how my results might enable various approaches for selecting, endogenizing, or rationalizing different Coarse Bayesian updating rules. One approach is to solve for an optimal updating rule in a given *environment*—a menu and signaling structure—under some constraint (for example, a fixed number of cells or a cost per additional cell). Pioneered by Wilson (2014) and Brunnermeier and Parker (2005), versions of this approach can provide a theory of where the updating rule "comes from." A drawback is that an updating rule adapted to one environment may be ill-suited for another. Only the robust ordering given by statement (ii) of Proposition 11 ensures weakly greater payoffs at all menu-signal pairs. So, rather than considering updating rules adapted to specific environments, one might instead endogenize them by selecting rules that are unimprovable (given constraints or costs) under the robust ordering. Alternatively, one might consider the weaker objective of minimizing worst-case losses (Proposition 10). These approaches are suitable if agents are unable to form probabilistic beliefs about their environment and, consequently, seek rules or heuristics robust to such uncertainty. Naturally, different criteria yield different predictions about updating rules; minimization of worst-case losses, for example, leads to representations exhibiting less skewness. Analysis of endogenous updating rules is beyond the scope of this paper, but—as illustrated by the characterizations in this section—the general framework of Coarse Bayesian updating provides a natural and tractable setting in which to carry it out.

5 Conclusion

Real people often revise their beliefs in a way that conflicts with the standard Bayesian model. In this paper, I have proposed a simple generalization of Bayes' rule, *Coarse Bayesian updating*, that can account for a variety of biases and individual heterogeneity in updating behavior. Three axioms—*Homogeneity*, *Cognizance*, and *Confirmation*—fully characterize the model and have the property that strengthening any of them to an if-and-only-if form makes the agent fully Bayesian. Thus, Coarse Bayesian updating may be viewed as a "small" departure from Bayes' rule.

An advantage of my framework is that it employs standard primitives that frequently appear in applications. The use of noisy signals over a state space, for example, allows one to directly import Coarse Bayesian updating into familiar settings in economics and game theory. I illustrate this by embedding the model in a standard setting of decision under risk, leading to a close relationship with the Blackwell information ordering and comparative notions of cognitive sophistication and bias. Hopefully, the framework and results developed in this paper will prove useful for further analysis of economic and game-theoretic applications.

A Proofs

A.1 Proof of Theorem 2

First, suppose μ has a Coarse Bayesian Representation. Note that for every $s \in S$ the signal $\frac{\mu^s/\mu^e}{\|\mu^s/\mu^e\|}$ is well-defined because μ^e has full support. Define $d : S \to S$ by

$$d(s) = \begin{cases} s & \text{if } \mu^s = B(\mu^e|s) \\ \frac{\mu^s/\mu^e}{\|\mu^s/\mu^e\|} & \text{otherwise} \end{cases}.$$

It is straightforward to verify that $\mu^s = B(\mu^e|d(s))$ for all s and that d satisfies properties (i)–(iii) of Theorem 2.

Conversely, suppose μ has a Signal Distortion Representation d. Define a binary relation $\sim$ on S by $s \sim t$ if and only if $d(s) \approx d(t)$. Clearly, $\sim$ is an equivalence relation; thus, its equivalence classes partition S. By (i) and (ii), each equivalence class is a convex cone. Thus, as in the proof of Theorem 1, each equivalence class is associated with a convex subset of Δ, and these subsets form a partition $\mathcal{P}$ of Δ. For each cell $P \in \mathcal{P}$, let $\mu^P := B(\mu^e|d(s))$ such that s belongs to the equivalence class associated with P. By (iii), $\mu^P \in P$.

A.2 Proof of Proposition 1

It is straightforward to verify that Bayesian updating satisfies properties (i)–(iii). So, suppose μ has a Coarse Bayesian Representation $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$. We show that each of properties (i)–(iii) forces each cell of $\mathcal{P}$ to be a singleton, making the agent Bayesian.

For (i), suppose $\mu^s = \mu^t$ implies $s \approx t$. Let $P \in \mathcal{P}$ and $\hat{\mu}, \hat{\mu}' \in P$. Choose signals s, t such that $B(\mu^e|s) = \hat{\mu}$ and $B(\mu^e|t) = \hat{\mu}'$. Then $\mu^s = \mu^t = \mu^P$, so that $s \approx t$ and, hence,

$\hat{\mu} = B(\mu^e|s) = B(\mu^e|t) = \hat{\mu}'$. Thus, every cell $P \in \mathcal{P}$ is a singleton.

For (ii), suppose $\mu^{s+t} = \mu^s$ implies $\mu^s = \mu^t$. Suppose toward a contradiction that $\mathcal{P}$ contains a non-singleton cell P. Since μ is non-constant, there exists $P' \in \mathcal{P}$ such that $\mu^P \neq \mu^{P'}$. Since μ^e has full support, there exist signals $\hat{s}, \hat{t}$ such that $B(\mu^e|\hat{s}) = \mu^P$ and $B(\mu^e|\hat{t}) = \mu^{P'}$; thus, $\mu^{\alpha\hat{s}} = \mu^P$ and $\mu^{\beta\hat{t}} = \mu^{P'}$ for all $\alpha, \beta \in (0,1)$. By equation (1) in the main text, it follows that if $\alpha\hat{s} + \beta\hat{t} \in S$, then $B(\mu^e|\alpha\hat{s} + \beta\hat{t}) = \frac{\alpha\hat{s}\cdot\mu^e}{(\alpha\hat{s}+\beta\hat{t})\cdot\mu^e}\mu^P + \frac{\beta\hat{t}\cdot\mu^e}{(\alpha\hat{s}+\beta\hat{t})\cdot\mu^e}\mu^{P'}$, which converges to μ^P as $\beta \to 0$. By regularity of P, there is an ε-ball $B^\varepsilon \subseteq P$ around μ^P. Thus, for sufficiently small $\alpha, \beta \in (0,1)$, we have $\mu^{\alpha\hat{s}+\beta\hat{t}} \in S$ and $B(\mu^e|\alpha\hat{s} + \beta\hat{t}) \in B^\varepsilon$; but then $\mu^{\alpha\hat{s}+\beta\hat{t}} = \mu^P = \mu^{\alpha\hat{s}}$ while $\mu^{\beta\hat{t}} = \mu^{P'} \neq \mu^P$, contradicting property (ii).

For (iii), suppose $\mu^t = \mu^s$ implies $t \approx \mu^s/\mu^e$. Consider the case $t = s$. Then $\mu^t = \mu^s$, so $s = t \approx \mu^s/\mu^e$. This implies $\mu^s \approx s\mu^e$, so that $\mu^s = B(\mu^e|s)$.

A.3 Proof of Proposition 2

Proof of part (iii). If every cell of $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ is a singleton, then the agent is Bayesian and the ML representation is established independently by the proof of part (iv) below. So, let $P^* \in \mathcal{P}$ be a non-singleton cell. Let I denote the set of all Coarse Bayesian Representations $i = \langle \mathcal{Q}(i), \dot{\mu}^{\mathcal{Q}(i)} \rangle$ such that $\mathcal{Q}(i)$ is finite, $\mathcal{P}$ is finer than $\mathcal{Q}(i)$, $\dot{\mu}^{\mathcal{Q}(i)} \subseteq \mu^{\mathcal{P}}$, and $P^* \in \mathcal{Q}(i)$. Define a partial order $\geq_I$ on I by $i \geq_I i'$ if and only if $\mathcal{Q}(i)$ is finer than $\mathcal{Q}(i')$ and $\dot{\mu}^{\mathcal{Q}(i)} \supseteq \dot{\mu}^{\mathcal{Q}(i')}$. It is straightforward to verify that $\geq_I$ is a partial order and that for all $i, i' \in I$, there exists $i^* \in I$ such that $i^* \geq_I i$ and $i^* \geq_I i'$. Thus, $(I, \geq_I)$ is a directed set.

For each $\langle \mathcal{Q}, \dot{\mu}^{\mathcal{Q}} \rangle \in I$, define a function $\gamma : \Delta \to [0, \infty)$ as follows. Since $N = 2$, the (finite) set $\dot{\mu}^{\mathcal{Q}}$ can be arranged in decreasing order of state 1: $\dot{\mu}^{\mathcal{Q}} = \{\dot{\mu}^{Q_1}, \ldots, \dot{\mu}^{Q_M}\}$, where $\dot{\mu}_1^{Q_1} > \dot{\mu}_1^{Q_2} > \ldots > \dot{\mu}_1^{Q_M}$. Since $P^* \in \mathcal{Q}$, there exists m^* such that $\dot{\mu}^{Q_{m^*}} = \mu^{P^*}$. For $1 \leq m < M$, let $\dot{\mu}^m$ denote the (unique) belief belonging to $\partial Q_m \cap \partial Q_{m+1}$ (the boundaries of Q_m and Q_{m+1}) and choose a signal s^m such that $B(\mu^e|s^m) = \dot{\mu}^m$. Now choose scalars $\alpha_m > 0$ such that, for all $1 \leq m < M$, $\alpha_m \dot{\mu}^{Q_m} \cdot s^m = \alpha_{m+1} \dot{\mu}^{Q_{m+1}} \cdot s^m$; taking $\alpha_{m^*} = 1$ pins down the α_m uniquely. Now define γ by

$$\gamma(\hat{\mu}) = \begin{cases} \alpha_m & \text{if } \hat{\mu} = \dot{\mu}^{Q_m} \\ 0 & \text{otherwise} \end{cases}.$$

By construction, $\dot{\mu}^{Q_m} \in \operatorname{argmax}_{\hat{\mu}} \gamma(\hat{\mu})\hat{\mu} \cdot s$ (that is, $\dot{\mu}^{Q_m}$ maximizes the likelihood function associated with γ) if and only if $B(\mu^e|s) \in Q_m$. Moreover, every point $\gamma(\hat{\mu})\hat{\mu}$, viewed as a point in $\mathbb{R}^2$, is contained in the half-space bounded above by the line with normal s^* passing through μ^{P^*}, where s^* is any signal such that $B(\mu^e|s^*) = \mu^{P^*}$. Thus, there exists a scalar $\overline{\gamma} > 0$ such that $\gamma(\hat{\mu}) \in [0, \overline{\gamma}]$ for all $\hat{\mu}$. Observe that the bound $\overline{\gamma}$ is independent of i.

Having defined a function $\gamma^i : \Delta \to [0, \overline{\gamma}]$ for every $i \in I$, the family $\{\gamma^i\}_{i \in I}$ forms a net. Each γ^i is an element of the (compact) product set $[0, \overline{\gamma}]^\Delta$, so that $\{\gamma^i\}_{i \in I}$ has a convergent subnet. This means there is a directed set $(J, \geq_J)$ and a function $\iota : J \to I$ such that (a) $j \geq_J j'$ implies $\iota(j) \geq_I \iota(j')$, (b) for every $i \in I$, there exists $j \in J$ such that $\iota(j') \geq_I i$ for all $j' \geq_J j$, and (c) the net $\{\gamma^{\iota(j)}\}_{j \in J}$ converges to some γ^*. Thus, for every $\hat{\mu} \in \Delta$, $\gamma^{\iota(j)}(\hat{\mu})$ converges to a point $\gamma^*(\hat{\mu})$.

Let $P \in \mathcal{P}$. By definition of $(I, \geq_I)$ and properties (a) and (b) of $(J, \geq_J)$, there exists $j^P \in J$ such that $P \in \mathcal{Q}(\iota(j))$ and $\mu^P \in \dot{\mu}^{\mathcal{P}(\iota(j))}$ for all $j \geq_J j^P$. Suppose s satisfies $B(\mu^e|s) \in P$. By construction, μ^P maximizes the likelihood function associated with $\gamma^{\iota(j)}$ at s if $j \geq_J j^P$: for every $\hat{\mu} \in \Delta$, $\gamma^{\iota(j)}(\mu^P)\mu^P \cdot s \geq \gamma^{\iota(j)}(\hat{\mu})\hat{\mu} \cdot s$. Taking the limit of both sides with respect to j yields $\gamma^*(\mu^P)\mu^P \cdot s \geq \gamma^*(\hat{\mu})\hat{\mu} \cdot s$; thus, μ^P maximizes the likelihood function associated with γ^* at s. □

Proof of part (iv). Notice that $B(\mu^e|s) = \mu'$ if and only if $s \approx \mu'/\mu^e := (\mu'_\omega/\mu^e_\omega)_{\omega \in \Omega}$. Thus, it will suffice to verify that $L(\cdot|s)$ is maximized at μ' for such signals s. This is done as follows. Let $s \in S$. Then, for any $\hat{\mu} \in \Delta$, we have

$$\begin{aligned}
L(\hat{\mu}|s) &= \gamma(\hat{\mu})\hat{\mu} \cdot s \\
&= \frac{\hat{\mu}}{\|\hat{\mu}/\sqrt{\mu^e}\|} \cdot s \\
&= \frac{\hat{\mu}/\sqrt{\mu^e}}{\|\hat{\mu}/\sqrt{\mu^e}\|} \cdot s\sqrt{\mu^e} \\
&= \left\| \frac{\hat{\mu}/\sqrt{\mu^e}}{\|\hat{\mu}/\sqrt{\mu^e}\|} \right\| \|s\sqrt{\mu^e}\| \cos\theta \\
&= \|s\sqrt{\mu^e}\| \cos\theta
\end{aligned}$$

where θ is the angle (in radians) between $\hat{\mu}/\sqrt{\mu^e}$ and $s\sqrt{\mu^e}$. Thus, $L(\cdot|s)$ is maximized at $\hat{\mu}$ where $\hat{\mu}/\sqrt{\mu^e} \approx s\sqrt{\mu^e}$ (because then $\theta = 0$), implying $\hat{\mu} \approx s\mu^e \approx \frac{\mu'}{\mu^e}\mu^e = \mu'$. □

A.4 Proof of Proposition 3

Suppose μ is a Pooled Coarse Bayesian updating rule induced by a representation $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ that is stable at ω. Let $\{s^n\}_{n=1}^\infty$ denote a sequence of signal realizations from σ such that $B(\mu^e|s^1 \dots s^n) \to \delta_\omega$, and $\{B^n\}_{n=1}^\infty$ the associated sequence of Pooled Coarse Bayesian beliefs; formally, $B^n := \mu^P$ such that $B(\mu^e|s^1 \dots s^n) \in P \in \mathcal{P}$. By stability at ω, there is an $\varepsilon > 0$ and a cell $P^* \in \mathcal{P}$ such that the ε-ball in Δ around δ_ω is a subset of P^*. Thus, for all n sufficiently large, $B(\mu^e|s^1 \dots s^n) \in P^*$ and, hence, $B^n = \mu^{P^*}$. So, if $B(\mu^e|s^1 \dots s^n) \to \delta_\omega$ almost surely, then $\mu^{(s^1, \dots, s^n)} = B^n \to \mu^{P^*}$ almost surely.

A.5 Proof of Proposition 4

For (i), let μ be a Sequential Signal Distortion rule. Observe that for every signal r, $\mu^r = B(\mu^e|d(r)) \approx d(r)\mu^e$. It follows immediately that $\mu^{(s,t)} \approx d(t)d(s)\mu^e \approx \mu^{(t,s)}$, so that μ is invariant to signal ordering. However, μ need not be invariant to signal ordering. For example, consider a model with two states and distortion function

$$d(s) = \begin{cases} (1/5, 4/5) & \text{if } \frac{s_2}{s_1} \geq 2 \\ e & \text{else} \end{cases}.$$

Let $s = (1/5, 4/5)$ and $t = (3/4, 1/4)$. Then $st = (3/20, 4/20)$, $d(st) = e$, $d(s) = (1/5, 4/5)$, and $d(t) = e$; thus, $d(s)d(t) = (1/5, 4/5) \neq e = d(st)$, so that $\mu^{(s,t)} \neq \mu^{st}$.

For (ii), consider a Sequential Coarse Bayesian updating rule satisfying all requirements in the second part of the statement. Since μ^e has full support, there is a signal r such that $B(\mu^e|r) = \mu^P$. Similarly, there is a signal t such that $B(\mu^P|t) = \mu^{P'}$ because μ^P has full support. It follows that $\mu^{(r,t,s^*)} = \mu^P \neq \mu^{P'} = \mu^{(r,s^*,t)}$.

A.6 Proof of Proposition 5

Given a finite sequence $s^1, \ldots, s^n \in \sigma = [t^1, \ldots, t^J]$ and $1 \leq j \leq J$, let n_j denote the number of signals s^i such that $s^i = t^j$. Then $\mu^{(s^1,\ldots,s^n)} = B(\mu^e|r^n)$ where $r^n := d(s^1)d(s^2)\ldots d(s^n) = d(t^1)^{n_1}d(t^2)^{n_2}\ldots d(t^J)^{n_J}$. Observe that, in state ω, $\frac{n_j}{n} \to t^j_\omega$ almost surely. Thus,

$$(r^n)^{1/n} := d(t^1)^{n_1/n}d(t^2)^{n_2/n}\ldots d(t^J)^{n_J/n} \to d(t^1)^{t^1_\omega}d(t^2)^{t^2_\omega}\ldots d(t^J)^{t^J_\omega} := t^*$$

almost surely. Consider the likelihood ratio $\ell^n_{\omega',\omega''} := \frac{r^n_{\omega'}}{r^n_{\omega''}}$. If $\frac{t^*_{\omega'}}{t^*_{\omega''}} < 1$, then $\ell^n_{\omega',\omega''} \to 0$ almost surely because $\ell^n_{\omega',\omega''} = \left(\frac{(r^n_{\omega'})^{1/n}}{(r^n_{\omega''})^{1/n}}\right)^n$ and $\frac{(r^n_{\omega'})^{1/n}}{(r^n_{\omega''})^{1/n}} \to \frac{t^*_{\omega'}}{t^*_{\omega''}} \in [0,1)$ almost surely. So, take any $\omega^* \in E^*$. Then, as $n \to \infty$, we have

$$B(\mu^e|r^n) = B\left(\mu^e \middle| \frac{1}{t^*_{\omega^*}} r^n\right) = \frac{\mu^e \frac{r^n}{t^*_{\omega^*}}}{\mu^e \cdot \frac{r^n}{t^*_{\omega^*}}} \overset{a.s.}{\to} \frac{\mu^e 1_{[\omega' \in E^*]}}{\mu^e \cdot 1_{[\omega' \in E^*]}} = B(\mu^e|t_{E^*}).$$

A.7 Proof of Proposition 7

Lemma 1. *Let $\varphi : \Delta \to \mathbb{R}$ and $\Phi : \mathcal{E} \to \mathbb{R}$ such that $\Phi(\sigma) = \sum_{\hat{\mu}} \varphi(\hat{\mu})\tau^\sigma(\hat{\mu})$. Suppose Φ satisfies the Blackwell ordering: $\sigma \sqsupseteq \sigma'$ implies $\Phi(\sigma) \geq \Phi(\sigma')$. Then φ is convex.*

Proof. Let $\hat{\mu}, \hat{\mu}' \in \Delta$, $\alpha \in (0,1)$, and $\hat{\mu}^\alpha := \alpha\hat{\mu} + (1-\alpha)\hat{\mu}'$. Since μ^e has full support, there exists $\hat{\mu}^* \in \Delta$ and $\lambda \in (0,1]$ such that $\lambda\hat{\mu}^* + (1-\lambda)\hat{\mu}^\alpha = \mu^e$. Let $\sigma = [s^*, s, s']$ and $\sigma' = [s^*, s+s']$ where $s^* = \lambda\frac{\hat{\mu}^*}{\mu^e}$, $s = (1-\lambda)\alpha\frac{\hat{\mu}}{\mu^e}$, and $s' = (1-\lambda)(1-\alpha)\frac{\hat{\mu}'}{\mu^e}$. Clearly, $\sigma \sqsupseteq \sigma'$, so that $\Phi(\sigma) \geq \Phi(\sigma')$. Moreover, $\mu^e \cdot s^* = \lambda$, $\mu^e \cdot s = (1-\lambda)\alpha$, $\mu^e \cdot s' = (1-\lambda)(1-\alpha)$, and $\mu^e \cdot (s+s') = 1-\lambda$, while $B(\mu^e|s^*) = \hat{\mu}^*$, $B(\mu^e|s) = \hat{\mu}$, $B(\mu^e|s') = \hat{\mu}'$, and $B(\mu^e|s+s') = \hat{\mu}^\alpha$. Thus, $\Phi(\sigma) = \varphi(\hat{\mu}^*)\lambda + \varphi(\hat{\mu})(1-\lambda)\alpha + \varphi(\hat{\mu}')(1-\lambda)(1-\alpha)$ and $\Phi(\sigma') = \varphi(\hat{\mu}^*)\lambda + \varphi(\hat{\mu}^\alpha)(1-\lambda)$, so that $\Phi(\sigma) \geq \Phi(\sigma')$ yields $\alpha\varphi(\hat{\mu}) + (1-\alpha)\varphi(\hat{\mu}') \geq \varphi(\hat{\mu}^\alpha)$, as desired. □

To prove Proposition 7, let $A \in \mathcal{A}$ and observe that (i) ⇒ (ii) by Lemma 1 (taking $\varphi = v^A$). The converse implication, (ii) ⇒ (i), follows from Blackwell's theorem. To see that (iii) ⇒ (i), observe that if $c^s(A) \cap b^s(c(A)) \neq \emptyset$ for all s, then every Coarse Bayesian choice from A is Bayesian-optimal in the menu $A' = c(A)$. Since Coarse Bayesian choices from A are identical to those from A', it follows that $V^A(\sigma) = V^{A'}(\sigma) = \overline{V}^{A'}(\sigma)$ for all σ. That is, V^A coincides with the Bayesian value of information in some menu, and therefore satisfies the Blackwell ordering.

Finally, we prove that (i) ⇒ (iii). Suppose (iii) is violated; that is, there exists $s \in S$ such that $c^s(A) \cap b^s(c(A)) = \emptyset$. Let $\hat{\mu} = B(\mu^e|s)$. Then there exists $x \in c(A)$ such that $v^A(\hat{\mu}) = x \cdot \hat{\mu} < y \cdot \hat{\mu}$ for all $y \in b^s(c(A))$. Choose any $y \in b^s(c(A))$ and $P \in \mathcal{P}$ such that $y \in c^{\mu^P}(A)$. Let $t \in S$ such that $B(\mu^e|t) = \mu^P$. By regularity, P has full dimension in Δ and μ^P belongs to the interior of P; therefore, we may assume $B(\mu^e|s+t) \in P$ (if necessary, scale s and t down by some $\lambda > 0$ sufficiently small). Observe that

$$B(\mu^e|s+t) = \frac{s \cdot \mu^e}{(s+t)\cdot\mu^e}\hat{\mu} + \frac{t\cdot\mu^e}{(s+t)\cdot\mu^e}\mu^P := \hat{\mu}',$$

and that there exists $y' \in c^{\mu^P}(A)$ such that

$$v^A(\hat{\mu}') = y' \cdot \hat{\mu}' = \frac{s\cdot\mu^e}{(s+t)\cdot\mu^e}y'\cdot\hat{\mu} + \frac{t\cdot\mu^e}{(s+t)\cdot\mu^e}y'\cdot\mu^P.$$

In particular, y' maximizes the above expression, so we have $y' \cdot \hat{\mu} \geq y \cdot \hat{\mu}$ and $y' \cdot \mu^P = y \cdot \mu^P$ because $y \in c^{\mu^P}(A)$. Now let $\sigma = [s, t, e-s-t]$ and $\sigma' = [s+t, e-s-t]$. Clearly, $\sigma \sqsupseteq \sigma'$. Let $V^A(e-s-t) := v^A(B(\mu^e|e-s-t))[(e-s-t)\cdot\mu^e]$. Then

$$\begin{aligned}
V^A(\sigma') &= v^A(\hat{\mu}')[(s+t)\cdot\mu^e] + V^A(e-s-t) \\
&= (y'\cdot\hat{\mu})(s\cdot\mu^e) + (y'\cdot\mu^P)(t\cdot\mu^e) + V^A(e-s-t) \\
&\geq (y\cdot\hat{\mu})(s\cdot\mu^e) + (y\cdot\mu^P)(t\cdot\mu^e) + V^A(e-s-t) \\
&> (x\cdot\hat{\mu})(s\cdot\mu^e) + (y\cdot\mu^P)(t\cdot\mu^e) + V^A(e-s-t) \\
&= V^A(\sigma),
\end{aligned}$$

so that V^A violates the Blackwell ordering.

A.8 Proof of Proposition 8

The implication (i) ⇒ (ii) is clear; the converse is an immediate consequence of the following lemma.

Lemma 2. *Suppose $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ and $\langle \mathcal{Q}, \dot{\mu}^{\mathcal{Q}} \rangle$ are regular representations of μ and $\dot{\mu}$, respectively, such that $\mu^e = \dot{\mu}^e$. Furthermore, suppose that for all $\sigma \sqsupseteq \sigma'$ and $A \in \mathcal{A}$, $\dot{V}^A(\sigma) \geq \dot{V}^A(\sigma') \Rightarrow V^A(\sigma) \geq V^A(\sigma')$. Then $\mathcal{Q}$ is finer than $\mathcal{P}$ and $\mu^{\mathcal{P}} \subseteq \dot{\mu}^{\mathcal{Q}}$.*

Proof of Lemma 2. The proof is divided into three steps.

Step 1: for every $Q \in \mathcal{Q}$, there is a unique $P \in \mathcal{P}$ such that int$(Q) \subseteq$ int(P).

First, observe that for every $Q \in \mathcal{Q}$ there is at least one $P \in \mathcal{P}$ such that $\text{int}(Q) \cap \text{int}(P) \neq \emptyset$; this holds because at least one P intersects the (nonempty, by regularity) set int(Q), which implies $\text{int}(Q) \cap \text{int}(P) \neq \emptyset$ by regularity of Q and P.

So, suppose toward a contradiction that there exist $Q \in \mathcal{Q}$ and distinct $P, P' \in \mathcal{P}$ such that $\text{int}(Q) \cap \text{int}(P) \neq \emptyset$ and $\text{int}(Q) \cap \text{int}(P') \neq \emptyset$. Then there exist $\hat{\mu}, \hat{\mu}' \in \text{int}(Q)$ such that $\hat{\mu} \in \text{int}(P)$ and $\hat{\mu}' \in \text{int}(P')$. Note that $\hat{\mu} \neq \hat{\mu}'$ since $P \cap P' = \emptyset$. Moreover, we may assume $\mu^P \notin \text{co}\{\hat{\mu}, \hat{\mu}'\}$ since, by regularity, we can replace $\hat{\mu}$ with a point in the interior of $\text{co}\{\mu^P, \hat{\mu}'\} \cap P$ if $\mu^P \in \text{co}\{\hat{\mu}, \hat{\mu}'\}$. Similarly, we may assume $\mu^{P'} \notin \text{co}\{\hat{\mu}, \hat{\mu}'\}$.

Next, we argue that it is without loss to assume that either $\mu^P \notin \text{co}\{\hat{\mu}, \hat{\mu}', \mu^{P'}\}$ or $\mu^{P'} \notin \text{co}\{\hat{\mu}, \hat{\mu}', \mu^P\}$. First, consider the case $N = 2$ (2 states). Since $\mu^P \notin \text{co}\{\hat{\mu}, \hat{\mu}'\}$ and $\mu^{P'} \notin \text{co}\{\hat{\mu}, \hat{\mu}'\}$, it follows immediately that $\mu^P \notin \text{co}\{\hat{\mu}, \hat{\mu}', \mu^{P'}\}$ because otherwise $\mu^P \in \text{co}\{\mu^{P'}, \hat{\mu}'\} \subseteq P'$. Similarly, $\mu^{P'} \notin \text{co}\{\hat{\mu}, \hat{\mu}', \mu^P\}$. Now consider the case $N \geq 3$. By regularity, we may assume that the points $\hat{\mu}$, $\hat{\mu}'$, μ^P, and $\mu^{P'}$ are distinct and not collinear (regularity allows us to perturb the points if necessary). It follows immediately that $\mu^P \notin \text{co}\{\hat{\mu}, \hat{\mu}', \mu^{P'}\}$ or $\mu^{P'} \notin \text{co}\{\hat{\mu}, \hat{\mu}', \mu^P\}$.

Suppose $\mu^P \notin \text{co}\{\hat{\mu}, \hat{\mu}', \mu^{P'}\}$ (the argument for the other case is similar). Then we may strictly separate μ^P and $\text{co}\{\hat{\mu}, \hat{\mu}', \mu^{P'}\}$; in particular, there exists x such that $x \cdot \mu^P < 0$ and $x \cdot \tilde{\mu} > 0$ for $\tilde{\mu} \in \{\hat{\mu}, \hat{\mu}', \mu^{P'}\}$. If necessary, perturb x so that $x \cdot \mu^Q \neq 0$. Let $A = \{x, 0\}$ and $s, t \in S$ such that $B(\mu^e|s) = \hat{\mu}$ and $B(\mu^e|t) = \hat{\mu}'$. For sufficiently small $\alpha, \beta > 0$, we have $\alpha s + \beta t \in S$; moreover, by equation (1) in the main text, $B(\mu^e|\alpha s + \beta t) \to \hat{\mu}'$ as $\alpha \to 0$. Thus, we assume without loss of generality (replacing s and t with appropriate αs and βt) that $B(\mu^e|s+t) \in \text{int}(P')$. It follows that $c^s(A) = c^{\mu^P}(A) = 0$ while $c^t(A) = c^{s+t}(A) = c^{\mu^{P'}}(A) = x$. Finally, let $\sigma = [s, t, e - s - t]$ and $\sigma' = [s + t, e - s - t]$. Clearly, $\sigma \sqsupseteq \sigma'$ and $\dot{V}^A(\sigma) = \dot{V}^A(\sigma')$ since $\hat{\mu}$, $\hat{\mu}'$, and $B(\mu^e|s+t)$ belong to the same cell $Q \in \mathcal{Q}$. However,

$V^A(\sigma') > V^A(\sigma)$ because $V^A(s+t) > V^A(s) + V^A(t)$, where $V^A(\tilde{s}) := v^A(B(\mu^e|\tilde{s}))(\tilde{s} \cdot \mu^e)$. This contradicts the second assumption of the lemma.

We have shown that for every $Q \in \mathcal{Q}$, there is a unique $P \in \mathcal{P}$ such that $\text{int}(Q) \cap \text{int}(P) \neq \emptyset$. Since $\mathcal{P}$ partitions Δ and cells are regular, it follows that, in fact, $\text{int}(Q) \subseteq \text{int}(P)$.

Step 2: $\mu^{\mathcal{P}} \subseteq \dot{\mu}^{\mathcal{Q}}$.

Suppose toward a contradiction that there is a cell $P \in \mathcal{P}$ such that $\mu^P \neq \dot{\mu}^Q$ for all $Q \in \mathcal{Q}$. Let Q denote the (unique) cell in $\mathcal{Q}$ such that $\mu^P \in Q$. By regularity, there is a neighborhood of μ^P contained in $\text{int}(P)$; since $\mu^P \in Q$, such a neighborhood intersects $\text{int}(Q)$. Thus, by Step 1, $\text{int}(Q) \subseteq \text{int}(P)$. Moreover, since $\mu^e = \dot{\mu}^e$, we have $P \neq P^e$ and $Q \neq Q^e$, where $\mu^e \in P^e \in \mathcal{P}$, $\dot{\mu}^e \in Q^e \in \mathcal{Q}$, and $\text{int}(Q^e) \subseteq \text{int}(P^e)$. There are two cases: either $\mu^P \notin \text{co}\{\dot{\mu}^Q, \mu^e\}$ or $\mu^P \in \text{co}\{\dot{\mu}^Q, \mu^e\}$.

If $\mu^P \notin \text{co}\{\dot{\mu}^Q, \mu^e\}$, there exists x such that $x \cdot \mu^P < 0$ and $x \cdot \tilde{\mu} > 0$ for $\tilde{\mu} \in \text{co}\{\dot{\mu}^Q, \mu^e\}$. Let $A = \{x, 0\}$ and $s, t \in S$ such that $B(\mu^e|s) = \dot{\mu}^Q$ and $B(\mu^e|t) = \mu^e$. As in Step 1, we may choose s and t so that $s + t \in S$ and $B(\mu^e|s+t) \in \text{int}(Q^e) \subseteq \text{int}(P^e)$. Thus, $c^s(A) = c^{\mu^P}(A) = 0$ and $c^t(A) = c^{s+t}(A) = c^{\mu^e}(A) = x$. Letting $\sigma = [s, t, e-s-t]$ and $\sigma' = [s+t, e-s-t]$, it follows that $\sigma \sqsupseteq \sigma'$, $\dot{V}^A(\sigma) = \dot{V}^A(\sigma')$, and $V^A(\sigma') > V^A(\sigma)$, contradicting the second assumption of the lemma.

If instead $\mu^P \in \text{co}\{\dot{\mu}^Q, \mu^e\}$, we may strictly separate μ^e from $\text{co}\{\dot{\mu}^Q, \mu^P\}$: there exists x such that $x \cdot \mu^e < 0$ and $x \cdot \tilde{\mu} > 0$ for $\tilde{\mu} \in \text{co}\{\dot{\mu}^Q, \mu^P\}$. Moreover, we may choose x so that the line $x \cdot \hat{\mu}' = 0$ passes through $\text{int}(P)$ and, therefore, so that there exists $\hat{\mu} \in P$ so that $x \cdot \hat{\mu} < 0$. Let $s, t \in S$ so that $s+t \in S$, $B(\mu^e|s) = \hat{\mu}$, $B(\mu^e|t)$, and $B(\mu^e|s+t) \in \text{int}(Q^e)$. Let $A = \{x, 0\}$. Then $c^s(A) = c^{\mu^P}(A) = x$ and $c^t(A) = c^{s+t}(A) = 0$. Letting $\sigma = [s, t, e-s-t]$ and $\sigma' = [s+t, e-s-t]$, it follows that $\sigma \sqsupseteq \sigma'$, $\dot{V}^A(\sigma) = \dot{V}^A(\sigma')$, and $V^A(\sigma') > V^A(\sigma)$, contradicting the second assumption of the lemma.

Step 3: for every $Q \in \mathcal{Q}$, there exists $P \in \mathcal{P}$ such that $Q \subseteq P$.

Let $Q \in \mathcal{Q}$. By Step 1, there is a unique $P \in \mathcal{P}$ such that $\text{int}(Q) \subseteq \text{int}(P)$. Suppose toward a contradiction that there exists $\hat{\mu} \in Q$ such that $\hat{\mu} \notin P$; such a $\hat{\mu}$ must be on the boundary of Q, so $\hat{\mu} \neq \dot{\mu}^Q$ by regularity. Since $\text{int}(Q) \subseteq \text{int}(P)$, we also have that $\hat{\mu}$ is on the boundary of P (otherwise there is a neighborhood of $\hat{\mu}$ contained in the complement of P; but every such neighborhood intersects $\text{int}(Q)$, contradicting $\text{int}(Q) \subseteq \text{int}(P)$).

Since $\mathcal{P}$ partitions Δ and $\hat{\mu} \notin P$, there is a cell $P' \in \mathcal{P}$ ($P' \neq P$) such that $\hat{\mu} \in P'$. By regularity, $\mu^{P'} \in \text{int}(P')$. Moreover, since $\hat{\mu}$ is on the boundary of P, $\hat{\mu}$ is also on the boundary of P'. Thus, we may strictly separate $\mu^{P'}$ from the closure of P; in particular, there exists x such that $x \cdot \mu^{P'} < 0$ and $x \cdot \tilde{\mu} > 0$ for $\tilde{\mu} \in \text{co}\{\hat{\mu}, \dot{\mu}^Q\}$. Choose $s, t \in S$ so that $B(\mu^e|s) = \hat{\mu}$, $B(\mu^e|t) = \dot{\mu}^Q$, and $B(\mu^e|s+t) \in \text{int}(Q)$. Letting $\sigma = [s, t, e-s-t]$ and $\sigma' = [s+t, e-s-t]$, it follows that $\sigma \sqsupseteq \sigma'$, $\dot{V}^A(\sigma) = \dot{V}^A(\sigma')$, and $V^A(\sigma') > V^A(\sigma)$, contradicting the second assumption of the lemma. $\square$

A.9 Proof of Proposition 9

For any $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ and $P \in \mathcal{P}$, let $S^P := \{s \in S : B(\mu^e|s) \in P\}$. For any σ, let $s^{P,\sigma} := \sum_{s \in \sigma \cap S^P} s$. Experiments σ and σ' are **$\mathcal{P}$-equivalent** if $s^{P,\sigma} = s^{P,\sigma'}$ for all $P \in \mathcal{P}$.

Lemma 3. *Suppose $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ is regular and let $\sigma, \sigma' \in \mathcal{E}$. Then σ and σ' are $\mathcal{P}$-equivalent if and only if $V^A(\sigma) = V^A(\sigma')$ for every A such that (A, σ) and (A, σ') are $\mu^{\mathcal{P}}$-decisive.*

Proof. Suppose σ and σ' are $\mathcal{P}$-equivalent. Observe that for every $\mu^{\mathcal{P}}$-decisive pair $(A, \hat{\sigma})$, $V^A(\hat{\sigma}) = \sum_{P \in \mathcal{P}} (\mu^e s^{P,\hat{\sigma}}) \cdot c^{\mu^P}(A)$ because decisiveness implies $c^{\mu^P}(A)$ is a singleton for all $P \in \mathcal{P}$ where $s^{P,\hat{\sigma}} \neq 0$. Thus, $V^A(\sigma) = V^A(\sigma')$ because $s^{P,\sigma} = s^{P,\sigma'}$ for all $P \in \mathcal{P}$.

For the converse, suppose σ and σ' are not $\mathcal{P}$-equivalent. We construct a menu A such that (A, σ) and (A, σ') are $\mu^{\mathcal{P}}$-decisive but $V^A(\sigma) \neq V^A(\sigma')$. For each $P \in \mathcal{P}$, let $\delta^P := s^{P,\sigma} - s^{P,\sigma'}$. Since experiments consist of finitely many signals, there are finitely many (but at least two) cells P such that $\delta^P \neq 0$. Let $\mu^\delta := \{\mu^P : \delta^P \neq 0\}$ and let μ^{P^*} be an extreme point of the convex hull of μ^δ. Since μ^δ is finite, μ^{P^*} can be strictly separated from the convex hull of $\mu^\delta \backslash \{\mu^{P^*}\}$; that is, there exists x such that $x \cdot \mu^{P^*} > 0 > x \cdot \mu^{P'}$ for all $\mu^{P'} \in \mu^\delta \backslash \{\mu^{P^*}\}$. By regularity, we may assume that x is such that the menu $A = \{x, 0\}$ makes (A, σ) and (A, σ') $\mu^{\mathcal{P}}$-decisive (if necessary, perturb x so that $c^s(A)$ is a singleton for all $s \in \sigma \cup \sigma'$). Then $V^A(\sigma) - V^A(\sigma') = \sum_{P \in \mathcal{P}} (\mu^s \delta^P) \cdot c^{\mu^P}(A) = (\mu^e \delta^{P^*}) \cdot x$ because $c^{\mu^P}(A) = 0$ for all $\mu^P \in \mu^\delta \backslash \{\mu^{P^*}\}$. Thus, $V^A(\sigma) \neq V^A(\sigma')$ provided $(\mu^e \delta^{P^*}) \cdot x \neq 0$. Since the separation is strict and $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ is regular, we may perturb x if necessary to ensure $(\mu^e \delta^{P^*}) \cdot x \neq 0$. $\square$

Proof that (i) implies (ii). Let $\sigma, \sigma' \in \mathcal{E}$ and suppose $\dot{V}^A(\sigma) = \dot{V}^A(\sigma')$ for all A such that (A, σ) and (A, σ') are $\dot{\mu}^{\mathcal{Q}}$-decisive. By Lemma 3, σ and σ' are $\mathcal{Q}$-equivalent. Since $\mathcal{Q}$ is finer than $\mathcal{P}$, it follows that σ and σ' are $\mathcal{P}$-equivalent. Thus, by Lemma 3, $V^A(\sigma) = V^A(\sigma')$ for all $\dot{\mu}^{\mathcal{Q}}$-decisive A.

Proof that (ii) implies (i). Let $Q \in \mathcal{Q}$ and suppose $s, t \in S^Q$. Let $\sigma = [s, t, e-s-t]$ (if necessary, scale s and t down by a factor $\lambda > 0$ to make σ well-defined), and let $\sigma' = [s+t, e-s-t]$. By Convexity, $s+t \in S^Q$ and, thus, σ and σ' are $\mathcal{Q}$-equivalent. By Lemma 3 and the hypothesis of (ii), this implies σ and σ' are $\mu^{\mathcal{P}}$-equivalent. Thus, there exists $P \in \mathcal{P}$ such that $s, t \in S^P$ (otherwise, there are distinct cells $P', P'' \in \mathcal{P}$ such that $s \in P'$ and $t \in P''$; but then σ and σ' are not $\mathcal{P}$-equivalent, as $s+t$ belongs to a single cell). We have shown that any two signals belonging to a common S^Q ($Q \in \mathcal{Q}$) belong to a common S^P ($P \in \mathcal{P}$). Thus, $\mathcal{Q}$ is finer than $\mathcal{P}$.

A.10 Proof of Proposition 10

Fix $s \in S$ and let $\mu^* = B(\mu^e|s)$ and $\mu^P = \mu^s$ where $\mu^s \in P \in \mathcal{P}$. If $A \in \mathcal{A}^*$, then there exist $x^*, y^* \in A$ such that $\overline{V}^A(s) = x^* \cdot \mu^*$ and $V^A(s) = y^* \cdot \mu^P$. In particular, $x^* \cdot \mu^* \geq x \cdot \mu^*$ and $y^* \cdot \mu^P \geq y \cdot \mu^P$ for all $x, y \in A$. Let $A^* = \{x^* - x^*, y^* - x^*\} = \{0, y^* - x^*\}$. Then $\overline{V}^A(s) - V^A(s) = \overline{V}^{A^*}(s) - V^{A^*}(s)$. Hence, to compute $L_\mu(s)$, it is without loss of generality to consider menus of the form $\{0, y\}$ where $\|y\| \leq 1$. We therefore rewrite the $L_\mu(s)$ as

$$\begin{aligned} L_\mu(s) &= \sup_{\|y\|\leq 1} 0 \cdot \mu^* - y \cdot \mu^* \text{ subject to: } 0 \cdot \mu^* \geq y \cdot \mu^* \text{ and } y \cdot \mu^P > 0 \cdot \mu^P \\ &= \inf_{\|y\|\leq 1} y \cdot \mu^* \text{ subject to: } 0 \geq y \cdot \mu^* \text{ and } y \cdot \mu^P > 0. \end{aligned}$$

The first constraint ensures the Bayesian prefers action 0 over y at signal s while the second ensures the Coarse Bayesian prefers y over 0 at s. Hence, we seek the infimum of $y \cdot \mu^*$ over all y on the unit (hyper)sphere bounded by the planes $y \cdot \mu^* \leq 0$ and $y \cdot \mu^P > 0$. Clearly, the infimum is characterized by a point y^* on the plane $y \cdot \mu^P = 0$. Thus, we seek a point on the disc $\{y : y \cdot \mu^P = 0 \text{ and } \|y\| \leq 1\}$ tangent to a plane $y \cdot \mu^* = c$ with normal μ^*. There are two such points; one maximizes $y \cdot \mu^*$, the other minimizes it.

Restricting attention to the case $\mu^* \neq \mu^P$, the first constraint does not bind. Thus, the Lagrangian is

$$\mathscr{L} = -y \cdot \mu^* + \lambda_1(y \cdot \mu^P) + \lambda_2(y \cdot y - 1).$$

Setting $\frac{\partial \mathscr{L}}{\partial y_\omega} = 0$ gives $2\lambda_2 y = \mu^* - \lambda_1 \mu^P$. Then $y \cdot \mu^P = 0$ implies $0 = \mu^* \cdot \mu^P - \lambda_1 \|\mu^P\|^2$ and $y \cdot y = 1$ implies $2\lambda_2 = \mu^* \cdot y - \lambda_1 \mu^P \cdot y = \mu^* \cdot y$. Thus, $\lambda_1 = \frac{\mu^* \cdot \mu^P}{\|\mu^P\|^2}$, so that

$$2\lambda_2 y = \mu^* - \left(\frac{\mu^* \cdot \mu^P}{\|\mu^P\|^2}\right) \mu^P.$$

Since $2\lambda_2 = \mu^* \cdot y$, this implies $(\mu^* \cdot y) y = \mu^* - \left(\frac{\mu^* \cdot \mu^P}{\|\mu^P\|^2}\right) \mu^P$. Thus, any solution y satisfies

$$\begin{aligned} (\mu^* \cdot y)^2 &= \|\mu^*\|^2 - \frac{(\mu^* \cdot \mu^P)^2}{\|\mu^P\|^2} \\ &= \|\mu^*\|^2 - \frac{\|\mu^*\|^2 \|\mu^P\|^2 \cos^2 \theta}{\|\mu^P\|^2} \\ &= \|\mu^*\|^2 \sin^2 \theta \end{aligned}$$

where $\theta \in (0, \frac{\pi}{2}]$ is the angle (in radians) between μ^* and μ^P. Thus, $L_\mu(s) = |y \cdot \mu^*| = \|\mu^*\| \sin\theta$, which is increasing in θ. Observe that $D_\mu(s) = \left\| \frac{\mu^*}{\|\mu^*\|} - \frac{\mu^P}{\|\mu^P\|} \right\|$ is the length of the chord connecting the points $\frac{\mu^*}{\|\mu^*\|}$ and $\frac{\mu^P}{\|\mu^P\|}$ on the unit circle. The length of a chord with central angle θ is $2\sin\left(\frac{\theta}{2}\right)$, which is strictly increasing on $[0, \frac{\pi}{2}]$. Thus, $D_\mu(s) = 2\sin\left(\frac{\theta}{2}\right)$ increases if and only if θ increases, so that $D_\mu(s)$ increases if and only if $L_\mu(s)$ increases.

A.11 Proof of Proposition 11

To see that (ii) implies (i), observe that $\dot{v}^A(\hat{\mu}) \neq v^A(\hat{\mu})$ only if $\hat{\mu}$ belongs to a cell Q such that $\dot{\mu}^Q \notin \mu^{\mathcal{P}}$. Every such Q is a singleton because $\dot{\mu}$ is less biased than μ, which implies $\mu^{\mathcal{P}} \subseteq \dot{\mu}^{\mathcal{Q}}$ and, hence, that Q is a "new" cell. Thus, $\dot{v}^A(\hat{\mu}) = \overline{v}^A(\hat{\mu}) \geq v^A(\hat{\mu})$.

To prove that (i) implies (ii), first apply Proposition 10 to get that $\dot{\mu}$ is less biased than μ. Therefore, $\mu^{\mathcal{P}} \subseteq \dot{\mu}^{\mathcal{Q}}$. We need to show that $\mathcal{Q}$ is finer than $\mathcal{P}$ and that every cell Q such that $\dot{\mu}^Q \in \dot{\mu}^{\mathcal{Q}} \backslash \mu^{\mathcal{P}}$ is a singleton.

First, we verify that $\mathcal{Q}$ is finer than $\mathcal{P}$. Suppose toward a contradication that there is a cell $Q \in \mathcal{Q}$ that intersects two or more distinct cells of $\mathcal{P}$. There is a unique $P \in \mathcal{P}$ such that $\dot{\mu}^Q \in P$. Let $P' \neq P$ be another cell of $\mathcal{P}$ such that $Q \cap P' \neq \emptyset$. Clearly, $\dot{\mu}^Q \notin P'$. Let $\partial P'$ denote the boundary of P'. There are two cases.

Case 1: $\dot{\mu}^Q \notin \partial P'$. Then, since P' is convex, there exists $x \in \mathbb{R}^N$ that strictly separates $\dot{\mu}^Q$ and P': $x \cdot \dot{\mu}^Q > 0 > x \cdot \hat{\mu}$ for all $\hat{\mu} \in P'$. Let $A = \{x, 0\}$. Then $v^A(\hat{\mu}') = 0$ for all $\hat{\mu}' \in Q \cap P'$ because $0 > x \cdot \mu^{P'}$. However, $\dot{v}^A(\hat{\mu'}) = x \cdot \hat{\mu}'$ for all $\hat{\mu}' \in Q \cap P'$ because $x \cdot \dot{\mu}^Q > 0$. Since $0 > x \cdot \hat{\mu}'$ for all $\hat{\mu}' \in Q \cap P'$, it follows that $\dot{v}^A(\hat{\mu}') < v^A(\hat{\mu}')$ for such $\hat{\mu}'$, a contradiction.

Case 2: $\dot{\mu}^Q \in \partial P'$. Then P' is not a singleton (otherwise $\mu^{P'} = \dot{\mu}^Q \notin P'$), forcing P' to be regular. Moreover, Q is regular because it intersects the (disjoint) sets P and P'. Thus, there are disjoint open neighborhoods $N_Q \subseteq Q$ and $N_{P'} \subseteq P'$ of $\dot{\mu}^Q$ and $\mu^{P'}$. Since N_Q and $N_{P'}$ are convex, there exists $x \in \mathbb{R}^N$ that strictly separates them: $x \cdot \hat{\mu} > 0 > x \cdot \hat{\mu}'$ for all $\hat{\mu} \in N_Q$ and $\hat{\mu}' \in N_{P'}$. Moreover, $\dot{\mu}^Q \in Q \cap \partial P'$ implies $N_Q \cap P' \neq \emptyset$, where $\partial P'$; by regularity, $N_Q \cap P'$ is a full-dimensional subset of $Q \cap P'$. Perturb x so that the plane $x \cdot \hat{\mu} = 0$ passes through the interior of $N_Q \cap P'$ (but not the point $\dot{\mu}^Q$); this can be done by shifting the plane toward the point $\dot{\mu}^Q$. Then x no longer separates N_Q and $N_{P'}$, but the set $C := \{\hat{\mu} \in N_Q \cap P' : 0 > x \cdot \hat{\mu}\}$ is nonempty, and we still have $x \cdot \dot{\mu}^Q > 0$ and $0 > x \cdot \hat{\mu}'$ for all $\hat{\mu}' \in N_{P'}$. Letting $A = \{x, 0\}$, it follows that $v^A(\hat{\mu}) = 0 > x \cdot \hat{\mu} = \dot{v}^A(\hat{\mu})$ for all $\hat{\mu} \in C$, a contradiction.

Next, we verify that every cell Q such that $\dot{\mu}^Q \in \dot{\mu}^{\mathcal{Q}} \backslash \mu^{\mathcal{P}}$ is a singleton. Suppose toward a contradiction that there exists $\dot{\mu}^Q \in \dot{\mu}^{\mathcal{Q}} \backslash \mu^{\mathcal{P}}$ such that Q is not a singleton. Since $\dot{\mu}$ is more sophisticated than μ, there is a unique $P \in \mathcal{P}$ such that $Q \subseteq P$. Note that $\mu^P = \dot{\mu}^P \in \dot{\mu}^{\mathcal{P}}$. Since μ^Q belongs to the relative interior of Q, there exists $\mu^* \in Q$ such that $\dot{\mu}^Q \notin \{\alpha\mu^* + (1-\alpha)\mu^P : \alpha \in [0,1]\} := L$. The set L is closed and convex, and therefore can be strictly separated from $\dot{\mu}^Q$: there exists $x \in \mathbb{R}^N$ such that $x \cdot \dot{\mu}^Q > 0 > x \cdot \hat{\mu}$ for all $\hat{\mu} \in L$. In particular, both $x \cdot \mu^* < 0$ and $x \cdot \mu^P < 0$. Let $A = \{0, x\}$. Then, at (Bayesian) posterior $\mu^* \in Q \subseteq P$, the $\langle \mathcal{P}, \mu^{\mathcal{P}} \rangle$ representation selects 0 from A: $v^A(\mu^*) = 0$. Under representation $\langle \mathcal{Q}, \mu^{\mathcal{Q}} \rangle$, however, x is selected from A at posterior μ^* because $\mu^* \in Q$ and $x \cdot \dot{\mu}^Q > 0$. Thus, $\dot{v}^A(\mu^*) = x \cdot \mu^* < 0$, so that $\dot{v}^A(\mu^*) < v^A(\mu^*)$.

B Relationship to Mullainathan (2002)

In a working paper, Mullainathan (2002) develops a model of categorical thinking sharing several features of Coarse Bayesian updating. In this appendix, I show that the categorical thinking model (adapted to my framework of states and signals) satisfies Homogeneity and Cognizance but not necessarily Confirmation.

Mullainathan works with a type space T and prior p where $p(t)$ is the prior probability of type $t \in T$. The analogous components in my model are the state space Ω and prior μ^e. Data d in Mullainathan's model is expressed as conditional probabilities $p(d|t)$ indicating the probability of observing the data given type t; in my model, data corresponds to a signal realization s, and s_ω plays the role of $p(d|t)$.

In Mullainathan's model, a set C of probability distributions over T constitutes a set of "categories"; these are feasible beliefs the agent can hold. Thus, the set C is analogous to the set $\{\mu^P : P \in \mathcal{P}\}$ in my model. For a category c and data d, $p(d|c)$ is the probability of generating data d in category c; this is analogous to $s \cdot \mu^P$, which is the probability of observing signal s if μ^P is the true probability law. Finally, Mullainathan defines $p(c) := \int_t p(t)c(t)$ to be the "base rate" of category c.[19] In my model, the analogous rate is $\mu^e \cdot \mu^P$.

Like Coarse Bayesians, agents in Mullainathan's model partition the probability simplex and assign posterior beliefs as a function of the cell containing the Bayesian posterior. Any set of C of categories is permitted; however, the partition is derived from C using an optimality criterion resembling that of Maximum-Likelihood rules in section 3.2. In particular, let

[19] I have modified the notation slightly; Mullainathan writes $q_c(\cdot)$ instead of $c(\cdot)$ to indicate the probability distribution over T associated with category $c \in C$.

$c^*(d) \in C$ denote the agent's posterior after observing data d. Mullainathan requires that

$$c^*(d) \in \operatorname*{argmax}_{c \in C} p(d|c)p(c). \tag{6}$$

In my framework, the analogous condition is

$$\mu^s \in \operatorname*{argmax}_{\hat{\mu} \in \hat{C}} (s \cdot \hat{\mu})(\mu^e \cdot \hat{\mu}), \tag{7}$$

where $\hat{C} \subseteq \Delta$ is some set of feasible posteriors. This is very similar to maximization of the likelihood function specified in section 3.2; the main difference is that my likelihood functions use a second-order belief γ instead of the base rate $p(c)$ proposed by Mullainathan.

Thus, Mullainathan's model works by specifying a set C of categories (feasible posteriors) from which the criterion (6) selects posteriors after observing data d. Because of the functional forms employed, it is as if there is a partition of the probability simplex such that the agent's selected posterior only depends on which cell contains the Bayesian posterior.

Unlike Coarse Bayesians, categorical thinkers need not satisfy Confirmation because condition (6) does not guarantee that beliefs $c^*(d)$ belong to the cell containing the Bayesian posterior associated with data d.[20] Below, I prove these claims in my framework (in particular, employing condition (7)).

First, let $\hat{C}$ be a nonempty set of feasible posteriors. Suppose that some $\mu^* \in \hat{C}$ is a solution to the maximization problem in (7) for both s and t. That is, μ^* solves both

$$\max_{\hat{\mu} \in \hat{C}} (s \cdot \hat{\mu})(\mu^e \cdot \hat{\mu}) \qquad \text{and} \qquad \max_{\hat{\mu} \in \hat{C}} (t \cdot \hat{\mu})(\mu^e \cdot \hat{\mu}).$$

Then, if $\alpha, \beta \geq 0$, it follows at μ^* solves

$$\max_{\hat{\mu} \in \hat{C}} ((\alpha s + \beta t) \cdot \hat{\mu})(\mu^e \cdot \hat{\mu}).$$

It follows that the map $s \mapsto \operatorname{argmax}_{\hat{\mu} \in \hat{C}} (s \cdot \hat{\mu})(\mu^e \cdot \hat{\mu})$ is measurable with respect to a partition of S into convex cones. As demonstrated in the proof of Theorem 1, such convex cones can be associated with convex subsets of Δ by mapping signals s to Bayesian posteriors $B(\mu^e|s)$.

Thus, any updating rule satisfying (7) satisfies Homogeneity and Cognizance if one restricts attention to signals that yield unique solutions to the optimization problem. For signals that involve ties, Homogeneity and/or Cognizance may be violated if the agent's

[20]Note that the partition in Mullainathan's model typically has convex cells. Convexity fails only if the maximization problem in (6) has more than one solution and the agent's tie-breaking criterion is not convex.

tie-breaking selection is not Homogeneous or Convex.

A more substantive difference between Mullainathan's model and Coarse Bayesian updating is that condition (7) does not guarantee that the updating rule satisfies Confirmation. To see this, suppose $|\Omega| = 2$ and let $\mu^e = (\frac{1}{3}, \frac{2}{3})$. Suppose $\hat{\mu}, \hat{\mu}' \in \hat{C}$ where $\hat{\mu} = (\frac{1}{4}, \frac{3}{4})$ and $\hat{\mu}' = (\frac{1}{5}, \frac{4}{5})$. Let $s = (\frac{3}{8}, \frac{9}{16})$. It follows that $B(\mu^e|s) = \hat{\mu}$; so, Confirmation requires $\hat{\mu}$ to solve

$$\max_{\tilde{\mu} \in \hat{C}} (s \cdot \tilde{\mu})(\mu^e \cdot \tilde{\mu}).$$

However,

$$(s \cdot \hat{\mu})(\mu^e \cdot \hat{\mu}) = \frac{77}{256} < \frac{63}{200} = (s \cdot \hat{\mu}')(\mu^e \cdot \hat{\mu}').$$

Thus, $\hat{\mu}$ is not selected at s, violating Confirmation.

References

Barberis, N., A. Shleifer, and R. Vishny (1998). A model of investor sentiment. *Journal of financial economics 49*(3), 307–343.

Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. In *Handbook of Behavioral Economics: Applications and Foundations 1*, Volume 2, pp. 69–186. Elsevier.

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

Blackwell, D. (1951). Comparison of experiments. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, Volume 1, pp. 93–102.

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

Camerer, C. (1995). Individual decision making. *Handbook of experimental economics*.

Cripps, M. W. (2018). Divisible updating. *Working Paper*.

De Bondt, W. F. and R. Thaler (1985). Does the stock market overreact? *The Journal of finance 40*(3), 793–805.

de Clippel, G. and X. Zhang (2022). Non-bayesian persuasion. *Journal of Political Economy (forthcoming)*.

Ducharme, W. (1970). Response bias explanation of conservative human inference. *Journal of Experimental Psychology 85*, 66–74.

Edwards, W. (1968). Conservatism in human information processing. *Formal representation of human judgment*, 17–52.

Eil, D. and J. M. Rao (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics 3*(2), 114–38.

Epstein, L. G. (2006). An axiomatic model of non-Bayesian updating. *The Review of Economic Studies 73*(2), 413–436.

Epstein, L. G., J. Noor, and A. Sandroni (2008). Non-Bayesian updating: A theoretical framework. *Theoretical Economics 3*(2), 193–229.

Fryer, R. G., P. Harms, and M. O. Jackson (2019). Updating beliefs when evidence is open to interpretation: Implications for bias and polarization. *Journal of the European Economic Association 17*(5), 1470–1501.

Galperti, S. (2019). Persuasion: The art of changing worldviews. *American Economic Review 109*(3), 996–1031.

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

Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *The Quarterly journal of economics 95*(3), 537–557.

Jakobsen, A. M. (2021). An axiomatic model of persuasion. *Econometrica 89*(5), 2081–2116.

Kahneman, D. and A. Tversky (1972). Subjective probability: A judgment of representativeness. *Cognitive psychology 3*(3), 430–454.

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *American Economic Review 101*(6), 2590–2615.

Kovach, M. (2020). Conservative updating. *Working Paper*.

Mullainathan, S. (2002). Thinking through categories. *Working Paper*.

Mullainathan, S., J. Schwartzstein, and A. Shleifer (2008). Coarse thinking and persuasion. *The Quarterly Journal of Economics 123*(2), 577–619.

Ortoleva, P. (2012). Modeling the change of paradigm: Non-Bayesian reactions to unexpected news. *American Economic Review 102*(6), 2410–36.

Phillips, L. D. and W. Edwards (1966). Conservatism in a simple probability inference task. *Journal of experimental psychology 72*(3), 346.

Rabin, M. (1998). Psychology and economics. *Journal of economic literature 36*(1), 11–46.

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

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.

Sharot, T. and N. Garrett (2016). Forming beliefs: Why valence matters. *Trends in cognitive sciences 20*(1), 25–33.

Thaler, M. (2021). Overinference from weak signals, underinference from strong signals, and the psychophysics of interpreting information. *arXiv preprint arXiv:2109.09871*.

Tversky, A. and D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *science 185*(4157), 1124–1131.

Tversky, A. and D. Kahneman (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological review 90*(4), 293.

Weinstein, J. (2017). Bayesian inference tempered by classical hypothesis testing. *Working paper*.

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

Zhao, C. (2022). Pseudo-bayesian updating. *Theoretical Economics 17*(1), 253–289.