PIER
PENN INSTITUTE *for* ECONOMIC RESEARCH
UNIVERSITY *of* PENNSYLVANIA

The Ronald O. Perelman Center for Political Science and Economics (PCPSE)
133 South 36th Street
Philadelphia, PA 19104-6297

pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper
19-018

Learning under Diverse World Views: Model-Based Inference

GEORGE J. MAILATH
University of Pennsylvania

LARRY SAMUELSON
Yale University

September 30, 2019

https://ssrn.com/abstract=3463205

Learning under Diverse World Views: Model-Based Inference*

George J. Mailath[†] Larry Samuelson[‡]

September 30, 2019

Abstract

People reason about uncertainty with deliberately incomplete models, including only the most relevant variables. How do people hampered by different, incomplete views of the world learn from each other? We introduce a model of "model-based inference." Model-based reasoners partition an otherwise hopelessly complex state space into a manageable model. We find that unless the differences in agents' models are trivial, interactions will often *not* lead agents to have common beliefs, and indeed the correct-model belief will typically lie outside the convex hull of the agents' beliefs. However, if the agents' models have enough in common, then interacting will lead agents to similar beliefs, even if their models also exhibit some bizarre idiosyncrasies and their information is widely dispersed.

*We thank many seminar audiences, the coeditor and four referees for helpful comments and discussions. Previous versions of this paper were circulated under the title, "The Wisdom of a Confused Crowd: Model Based Inference." We thank the National Science Foundation (SES-1459158 and SES-1559369) for financial support.

[†]Department of Economics, University of Pennsylvania, and Research School of Economics, Australian National University; gmailath@econ.upenn.edu

[‡]Department of Economics, Yale University, New Haven, CT 06525, larry.samuelson@yale.edu

Contents

1 Introduction **1**

2 The Setting **2**
2.1 The Environment . . . 2
2.2 Model-Based Reasoning . . . 3
2.3 Beliefs . . . 6
2.3.1 Full-Information Beliefs . . . 6
2.3.2 Prior Beliefs . . . 7
2.3.3 Interim Beliefs . . . 7
2.3.4 Why Don't Agents Choose the Right Model? . . . 8

3 Learning From Others **9**
3.1 Unknown Sense or Known Nonsense? . . . 9
3.2 The Details of Updating Beliefs . . . 11
3.3 Example . . . 13
3.4 How Revealing are Beliefs? . . . 15
3.5 The Bliss of Others' Ignorance . . . 16
3.6 Properties of the Belief Updating Process . . . 18

4 Agents and Oracles **21**
4.1 Do Agents Agree? . . . 21
4.2 The Wisdom of the Crowd? . . . 23

5 Information Aggregation **26**
5.1 Correlated Model Predictions . . . 26
5.2 Models with a Common Component . . . 27
5.3 Dispersed Information . . . 28

6 Discussion and Related Literature **32**

A Appendices **34**
A.1 Proof of Proposition 6 . . . 34
A.2 Proof of Proposition 7 . . . 37
A.3 Proof of Lemma 2 . . . 38
A.4 Proof of Lemma 4 . . . 40

B Online Appendices **42**
B.1 What's Wrong with Different Priors? . . . 42
B.2 Subsequent Updating in Example 1 . . . 45
B.3 An Example with Infinite Iterations . . . 46
B.4 Common Knowledge . . . 48
B.5 Implications of Necessary Agreement . . . 50
B.6 An Example Illustrating Redundancy and Correlation . . . 51

References **52**

Learning under Diverse World Views:
Model-Based Inference

1 Introduction

Economists' theories typically model people as having a complete and perfect understanding of their world. When people lack information, such uncertainty is captured by a (common) state space, whose elements "resolve all uncertainty." Physicists may not be sure which is the right reconciliation of quantum theory and general relativity, but they understand every detail of every possibility.

In practice, people work with models that are deliberately incomplete, including the most salient variables and excluding others. At best, the states in such models resolve all *relevant* uncertainty. Moreover, different people work with different models. Civil engineers building bridges and electrical engineers designing quantum computers persist with models that are incomplete.

People routinely interact, exchanging information and beliefs. These exchanges seldom lead to complete agreement, but people do learn from each other. How do they do this when hampered by different incomplete views of the world? We address this question by developing and analyzing a model of "model-based inference."

We conduct our analysis in a particularly stylized interaction. As in Geanakoplos and Polemarchakis (1982), our agents observe information, form beliefs about an event, exchange these beliefs, update in response to the other agents' beliefs, announce their new beliefs, and so on. We are obviously ignoring much of the complexity of actual interactions, but retain precisely what is needed to examine how people learn from others with diverse world views.

What do we mean by "diverse world views?" One obvious source of different views is different information. However, Aumann's (1976) agreeing-to-disagree theorem (as adapted by Geanakoplos and Polemarchakis, 1982) tells us that asymmetric information alone cannot be a source of persistently different views of the world.

We assume people follow Savage's (1972) recommendation that they escape the hopeless complexity of states that resolve all uncertainty by partitioning the state space into elements designed to capture the most important factors and ignore less important ones. The elements of these partitions become the states in people's models. But we expect different people to construct this partition differently, leading to different world views. Once they have formed their models, our agents apply Bayes' rule.

Since different information *can* give rise to persistently different world views when people have different prior beliefs, why not simply assume agents hold different priors? The important advantage of working directly with different models is that we can then reasonably insist that agents have a common prior on the *common* elements of their models. This restores much of the discipline whose absence typically pushes research away from models with heterogeneous priors. Appendix B.1 illustrates the lack of discipline that arises with heterogeneous priors.

Sections 2 and 3 introduce model-based reasoning and interactions between model-based reasoners. In Section 4, we show that unless the differences in agents' models are trivial, interactions will not lead agents to common beliefs. More problematically, any conventional aggregate of the agents' beliefs will often be off the mark, in the sense that the correct-model belief will lie outside the convex hull of the agents' beliefs. In general, we cannot expect people with different models to effectively aggregate their information.

Section 5 shows that if the agents' models have enough in common, then interacting will lead agents to similar beliefs, even if their models also exhibit some bizarre idiosyncracies. Perhaps more importantly, we identify conditions under which agents who collectively have sufficient information will have average belief close to the correct-model belief, even if their information is widely dispersed. The key to effective information aggregation is thus *not* that people have common information or a common model, but that the different models people use imply a sufficiently common interpretation of whatever information they have.

2 The Setting

2.1 The Environment

We begin with a familiar model of uncertainty. A state of the world is an element of the set $\Omega = X^N$, where $X \subseteq \mathbb{R}$ and $N \subseteq \mathbb{N}$ is possibly infinite.[1] Nature draws a state ω from Ω according to the probability measure ρ on Ω. Agents form beliefs about the occurrence of an event F. It is convenient

[1] With the exception of Section 4 and Proposition 8 in Section 5.3, our results hold for a more general model in which states of the world are given by a complete, separable metric space Ω, endowed with the associated Borel σ-algebra. See the preceding working paper (Mailath and Samuelson, 2019) for details.

to describe this event in terms of its indicator function $f : \Omega \to \{0,1\}$, so that $f(\omega) = 1$ if and only if $\omega \in F$.

For illustration, we often assume that X is finite (at which point nothing is lost by taking X to be the set $\{0,1\}$) and N is finite, which we refer to as the finite case.

2.2 Model-Based Reasoning

It is standard in economic analyses to equip agent i with the state space Ω, nature's distribution ρ as prior belief, and description f of the event. We refer to such a reasoner as an *(agent) oracle*.

In contrast, we are concerned with *model-based reasoners*. A model-based reasoner is a faithful adherent of Savage's (1972) *Foundations of Statistics*. Savage explains that it is a hopeless undertaking to work with a state space that resolves all uncertainty, i.e., that specifies "[t]he exact and entire past, present, and future history of the universe, understood in any sense, however wide" (Savage, 1972, p. 8).[2] Savage argues on the next page that "the use of modest little worlds, tailored to particular contexts, is often a simplification, the advantage of which is justified." A model-based reasoner's "modest little world" effectively partitions the state space into equivalence classes that he or she believes capture relevant information about F while ignoring irrelevant information. These equivalence classes then become "states" in the reasoner's model.[3]

We capture this reliance on models by assuming that each agent i explains the occurrence of the event F by a function (her *theory*)

$$f^i : X^{M^i} \to [0,1]$$

that depends only on the realizations of the variables contained in a subset $M^i \subseteq N$.[4] We refer to the set M^i as agent i's *model*. Once the agent has a model and attendant theory, we can talk about reasoning, Bayesian or otherwise.

[2] Savage (1972, p. 16) describes this logical extreme of "look before you leap" as "utterly ridiculous."

[3] This is not the only possible interpretation of Savage (1972). An alternative is that Bayesian decision making in the spirit of Savage (1972) is only justified in situations where a decision maker can legitimately specify a state space that resolves *all* uncertainty. The distinction between these interpretations is particularly relevant when discussing the origin of the prior, an issue we avoid by assuming (1) and (2) below.

[4] The corresponding formulation in the general case is that i's theory f^i is measurable with respect to a sub-σ-algebra $\mathcal{M}^i$ of the Borel σ-algebra.

If the variables in M^i suffice to determine whether F has occurred, then agent i correctly understands the forces determining F. We are most interested in cases in which agent i works with a set M^i that omits some of the variables required to determine whether F has occurred. Agent i realizes that such a model cannot be expected to perfectly explain the event F, reflected in f^i mapping into $[0, 1]$, giving the probability that the event F has occurred, rather than $\{0, 1\}$.

Denote by ω_{M^i} an element of the set X^{M^i}. An element ω_{M^i} specifies the realizations drawn from X of each of the variables that appear in i's model. In terms of Savage's procedure for creating a "modest little world", the model M^i identifies the equivalence classes, of the form $\{\omega_{M^i}\} \times X^{-M^i}$, that convert X^N into agent i's view of the world.

For example, suppose the event F corresponds to an increase in the price of a financial asset. Even upon restricting attention to professionals, we encounter a variety of approaches. A fundamentalist will typically seek information on the cash reserves, debt load, volume of sales, profit margin (and so on) of the underlying firm; these are the variables that would appear in her model M^i. A chartist's M^i will include variables corresponding to recent share trading volumes, price trends, reversals in price movements, the existence of apparent price ceilings, and so on. An efficient marketer will ask for a coin to flip. And even among professionals, there are forecasters whose models focus on astrological data. The fundamentalist is likely to exclude much of the asset-price history from her model, while the chartist may neglect various aspects of the firm's current financial position. Both will typically exclude information about zodiac signs. All of the agents are likely to miss factors whose relevance has not yet been imagined, as well as factors they are convinced are irrelevant, while possibly including irrelevant factors.

We *assume* that agent i's theory f^i is consistent with the event F's indicator function f. The probability agent i attaches to the event F given ω_{M^i} matches the probability that the prior probability measure ρ attaches to the event F, conditional on ω_{M^i}. In the finite case, this is the requirement that for every ω_{M^i} (for which $\rho(\omega_{M^i}) = \rho(\{\omega_{M^i}\} \times X^{-M^i})$ is positive),

$$f^i(\omega_{M^i}) = \sum_{\omega \in \Omega} f(\omega)\rho(\omega | \{\omega_{M^i}\} \times X^{-M^i}). \tag{1}$$

To formulate the infinite version of this requirement, let $\mathcal{M}^i$ be the σ-algebra on X^N induced by the model M^i.[5] Agent i's theory $f^i : X^{M^i} \to [0, 1]$ can

[5] That is, $\mathcal{M}^i$ is generated by the equivalence classes $\{\omega_{M^i}\} \times X^{-M^i}$.

also be viewed as a function from X^N to $[0,1]$ that is measurable with respect to $\mathcal{M}^i$, i.e., that takes on the value $f^i(\omega_{M^i})$ for every $\omega \in \{\omega_{M^i}\} \times X^{-M^i}$. We adopt this view throughout our formal development, while thinking of f^i as a function on X^{M^i} for interpretation. To keep notation uncluttered, we refer to both formulations as f^i, with the context making clear which is appropriate. Our assumption for the infinite case is

$$f^i(\omega) = \mathbb{E}[f|\mathcal{M}^i](\omega). \tag{2}$$

As in (1), the probability agent i attaches to event F having observed the variables in $\mathcal{M}^i$ is the probability attached to the event F by the prior measure ρ, conditional on the event.[6]

One interpretation for the correct beliefs assumed in (2) is that, as in Spiegler (2016), agent i builds her theory from her model M^i and her access to a record of an unlimited number of independent draws from the prior distribution ρ.[7] Recall that the agent restricts attention to the variables in her model. For each of the possible realizations ω_{M^i} of such variables (focusing for interpretation on the finite case), the agent identifies the draws in the record whose realizations match ω_{M^i} for the variables in M^i and calculates the frequency with which the event F has occurred among these realizations, giving rise to the probabilities $f^i(\omega_{M^i})$ in (1).

Of course we do not expect agent i to literally have access to an infinite number of independent draws from the distribution ρ. In practice, the agent must construct f^i using a finite number of observations. Hence, even while using the device of an infinite record to interpret the beliefs in (1)–(2), we still feel free to appeal to the case of finite data when developing intuition for our analysis. Our goal in imposing (1)–(2) is to isolate the implications of agent i's model-based inference from the statistical problems that invariably arise with finite sets of data, much as econometricians prefer to separate questions of estimation and identification. We thus endow each agent with correct beliefs and examine the implication of these beliefs. In particular, (2) imposes a natural consistency requirement across agents.

Remark 1 (Clinging to Models) Our agents are dogmatic in their models, in the sense that they never entertain the possibility of adopting a different model. People indeed go to great lengths to defend models to which they resolutely cling. Einstein is reputed to have argued that "God does

[6]Because Ω in the general case is complete, separable, and metric (i.e., Polish), we can assume conditional beliefs exist for *all* ω (Stroock, 2011, Theorem 9.2.1).

[7]Eyster and Piccione (2013, p. 1492) analogously motivate a condition similar to (1)–(2) as the outcome of a statistical learning process.

not play dice with the universe" and Dirac to have argued that "God used beautiful mathematics in creating the world." Both are examples of defending particular (types of) model. We do not examine the process by which agent i might come to focus on the model M^i or might consider alternatives to the model M^i. Importantly, one should *not* expect such processes to eliminate differences in models, since different agents may follow different model selection processes.[8]

Section 2.3.4 returns to this issue. ♦

2.3 Beliefs

Forming beliefs about uncertain events, whether by an oracle or a model-based reasoner, requires two steps. First, the agent identifies their complete-information beliefs. Then the agent takes expectations of these complete-information beliefs with respect to an appropriately updated probability measure.

2.3.1 Full-Information Beliefs

If agent i observed all of the information she deemed relevant, i.e., if agent i observed the realization ω_{M^i} of the variables in her model M^i, then she would regard herself as having full information and would attach to the event F the probability

$$f^i(\omega_{M^i}) = f^i(\omega),$$

whose value is given by (2). We write the notation-abusing equality as a reminder that f^i is viewed as a function on X^N throughout our formal development, measurable with respect to $\mathcal{M}^i$, but often interpreted as a function on X^{M^i}. We refer to this as a *full information belief.* Agent i *ignores* any variables outside of M^i, but she *correctly* uses the implications of the information she *does* think relevant, namely ω_{M^i}. It follows immediately from (2) that the full-information beliefs of a model-based reasoner agree with those of an agent oracle having information ω_{M^i}.

[8] Hong, Stein, and Yu (2007) examine a model in which agents restrict attention to a class of models simpler than that of an (in our terms) agent oracle, but update their beliefs about which model in the simple class is appropriate. The extension of our model to such a setting would involve agents who restrict attention to *different* classes of models, or follow different model-updating rules. The difficulties agents face in learning from other agents would only be exacerbated in such a setting.

2.3.2 Prior Beliefs

If agent i has no information about the state, then she attaches to the event F the expectation of her full-information belief, given by

$$\mathbb{E}[f^i(\omega)] = \mathbb{E}[\mathbb{E}[f|\mathcal{M}^i](\omega)] = \mathbb{E}[f(\omega)], \tag{3}$$

where the first equality is from (2) and the second is an application of the law of iterated expectations.

This indicates that agent i's prior belief matches that of an agent oracle. This is no surprise. Recall we interpret i's belief as having arisen from a consultation of the record. Without any information, no observations from the record are eliminated as possible candidates, and the empirical frequency calculated by i matches that calculated by an oracle.

2.3.3 Interim Beliefs

We now consider the case where agent i observes the realizations ω_{I^i} for some subset $I^i \subseteq N$. Agent i then forms her interim beliefs, which we denote by β^i.

Analogously to our treatment of agent i's theory f^i, for purposes of interpretation, we treat β^i as a function $X^{I^i} \to [0,1]$, identifying for each realization $\omega_{I^i} \in X^{I^i}$ the updated probability agent i attaches to the event F. For the formal development, we view β^i as an equivalent (and identically named) function on X^N that is measurable with respect to the information contained in I^i.[9]

We think of agent i as observing her information ω_{I^i} and then consulting the record. She identifies all those realizations that match her observation ω_{I^i}, and calculates the frequency of the various values of ω_{M^i} in these observations. This gives her an updated distribution of the realizations of the variables ω_{M^i}. Each realization of ω_{M^i} gives rise to a full-information belief, and she takes the expectation of these full-information beliefs with respect to this updated distribution.

Importantly, the essence of our model-based inference "model" of the agent is that agent i does *not* simply look at the empirical frequency of the occurrence of F under I^i. This reflects our assumption that i acts as if only her model variables are relevant for predicting F, and the only value in information is to help her in inferring the variables in M^i. Indeed, calculating β^i as the empirical frequency of F under I^i is equivalent to taking

[9] This latter viewpoint is particularly useful when we include announcements by other agents in i's information.

$M^i = N$. Given the stark simplicity of our model of model-based reasoners, agent i appears to be throwing away useful information. Recall, though, that our model is a representation of a vastly more complex reality, in which (as we explain in Section 2.3.4) the agent adopts a model, uses whatever means she has available for formulating full-information beliefs (for convenience assumed to be correct in our analysis), and then uses information to update her beliefs about the variables in her model.

Denote by $\mathcal{I}^i$ the σ-algebra induced by I^i.[10] Then agent i's interim belief, denoted by $\beta^i(\omega)$, is

$$\beta^i(\omega) = \mathbb{E}[f^i \mid \mathcal{I}^i](\omega). \tag{4}$$

An oracle can similarly be viewed as examining those observations in the record that match ω_{I^i}, but then taking the expectation of the description $f(\omega)$ over this set. If $I^i \subseteq M^i$, then these two updating procedures are equivalent.[11] We have

$$\begin{aligned} \beta^i(\omega) &= \mathbb{E}[f^i \mid \mathcal{I}^i](\omega) \\ &= \mathbb{E}[\mathbb{E}[f|\mathcal{M}^i] \mid \mathcal{I}^i](\omega) \\ &= \mathbb{E}[f \mid \mathcal{I}^i](\omega), \end{aligned}$$

where the first equality repeats the definition (4) of the interim belief β^i, the next line follows from the definition of the full-information belief from (2), and the last line follows from the law of iterated expectations.

We hereafter assume $I^i \subseteq M^i$,allowing us to focus on interactions between agents.[12]

2.3.4 Why Don't Agents Choose the Right Model?

Given an unlimited record of previous draws from the distribution ρ, why doesn't the agent use the record to identify the correct model? Equivalently, why doesn't the agent choose the largest possible model, N, ensuring that she never omits anything relevant?

In practice, agents are confronted with finite data and a state space of potentially infinite complexity. Even big data cannot "slip the surly bonds" of finiteness, while Arrow and Hurwicz (1972, p. 2) note that variables

[10] Similar to footnote 5, $\mathcal{I}^i$ is generated by the equivalence classes $\{\omega_{I^i}\} \times X^{-M^i}$.

[11] If $I^I \not\subseteq M^I$, the agent's interim beliefs and the agent oracular beliefs need not coincide, since we cannot apply the law of iterated expectations.

[12] We view the restriction of I^i to a subset of M^i as reasonable in many circumstances, on the grounds that people are likely to not process information they deem irrelevant.

are unlimited: "How we describe the world is a matter of language, not of fact. Any description of the world can be made finer by introducing more elements to be described." An agent will never encounter data that unambiguously contradicts whatever model she holds or unambiguously identifies a correct model. Instead, anomalous observations can be explained away by unobserved factors, and for every event and every set of data, there will an infinite collection of models that explain the data perfectly, making it impossible to use the data to find the "right" model. And for every event, there will be an infinite list of variables about which the agent could collect information, making it impossible to be a pure empiricist. Al-Najjar (2009) and Gilboa and Samuelson (2012) elaborate on the futility of interpreting data without models. If she is to make any meaningful use of the data, agent i must then appeal to *some* model M^i and proceed as if only the variables captured by M^i matter. In order to focus on how and whether agents can "learn" from each other, we make the extreme assumption captured in (1)–(2) that the agent makes perfect use of whatever model she has.

Giacomini, Skreta, and Turén (2017) describe the behavior of 75 professional forecasters in terms we recognize as model-based reasoning. The object of each participant was to predict the US inflation rate, for each of the years 2007–2014. Forecasting typically began at the beginning of July of the preceding year (with slightly later initial forecasts for 2007 and 2008), with individual forecasters updating their predictions at any time until the end of the year in question. Giacomini, Skreta, and Turén (2017) argue that the forecasters in their sample appear to be Bayesians (albeit much more so in non-crisis years), but with different models that lead them to different forecasts. In response to this disagreement, the agents persevere in their belief in their models (again, more so in non-crisis years) and in their disagreement. Such agents would find themselves well at home in our setting.

3 Learning From Others

We now assume there are K agents. Each agent $i = 1, \ldots, K$ has a model M^i exhibiting the properties outlined in Section 2.2, and has access to information $I^i \subseteq M^i$.

3.1 Unknown Sense or Known Nonsense?

How does agent i extract information from agent j's beliefs, given that i and j may have different models? What, if anything, must agent i know about

j's model to draw such inferences?

Our leading interpretation is that agent i need know *nothing* about j's model. We refer to this as the case of *unknown sense*—agent i may allow for the possibility that there is some sense behind j's reasoning, but i might have no idea how this reasoning proceeds. The fundamentalist may comment that "I have no idea how the chartist comes up with these conclusions."

Our model is also consistent with the assumption that each agent knows the models of other agents. In this case, agent i may think that agent j is ill-advised in her choice of model—this is presumably why agent i sticks with her own model rather than adopting j's model—and hence we refer to this as the case of *known nonsense*. The fundamentalist might remark that "I can read the charts and see how the astrologer comes up with these predictions, but I do not believe it an improvement to incorporate the zodiac into my model."

We follow Geanakoplos and Polemarchakis (1982) in examining the following information-exchange protocol:

(a) First, each agent i observes her information ω_{I^i} and forms her interim belief.

(b) Agents (by assumption truthfully) simultaneously announce their interim beliefs.

(c) Agents update their beliefs in response to these announcements.

(d) Agents then announce their revised beliefs, update, and announce, and so on.

Formally, this process continues indefinitely; we say that the process terminates if a stage is reached at which beliefs are not subsequently revised.[13]

In the known nonsense case, one can readily imagine how i draws inferences from j's beliefs, since i can invert j's reasoning. But how does i draw inferences under our leading interpretation, in which i may have no understanding of how j reasons?

We interpret the updating that occurs in steps (c) and (d) of this process as follows. Consistent with our discussion in Section 2, we think of the record

[13]Geanakoplos and Polemarchakis (1982) assume the agents have finite information partitions, ensuring that the belief revision process terminates in a finite number of steps. Sethi and Yildiz (2012) apply Geanakoplos and Polemarchakis (1982) to a model of deliberation with different priors. There is no belief updating in their model after the first round because the first round signals are a sufficient statistic of the relevant private information.

as including not only the realizations of the state, but also the sequences of announcements made by the agents. Agent i forms her interim belief by identifying those states in the record corresponding to her observation ω_{I^i}, then identifying the relative frequencies of ω_{M^i} in these states, and then taking the expectation of her full-information belief over this set. She forms her next update by restricting attention to the subset of such states in which the other agents' first announcements match those she has observed, and again identifying the relative frequencies of ω_{M^i} in these states and taking the expectation of her full-information belief over this set. She continues similarly in subsequent rounds.[14]

3.2 The Details of Updating Beliefs

Fix ω and suppose that agent i has observed her information ω_{I^i} and the other agents have announced the vector $b_0^{-i} = (b_0^1, \ldots, b_0^{i-1}, b_0^{i+1}, \ldots, b_0^K)$, where the j^{th}-element of the vector b_0^{-i} corresponds to agent j's announced interim belief $b_0^j = \beta^j(\omega)$ of the event F (determined by j's observation of his information). Agent i then forms a belief denoted by $\beta^i(\omega_{I^i}, b_0^{-i})$. Letting $\mathcal{B}_0^{-i}$ be the σ-algebra generated by the announcement b^{-i}, agent i forms her *model-based* belief about the event F as[15]

$$\beta^i(\omega, b_0^{-i}) = \mathbb{E}[f^i|\mathcal{I}^i, \mathcal{B}_0^{-i}](\omega). \tag{5}$$

Denote the interim belief announced by agent i by b_0^i, the second posterior by b_1^i, and so on; the vector of announced posteriors is similarly denoted by $b_0 = (b_0^i, b_0^{-i})$, $b_1 = (b_1^i, b_1^{-i})$, and so on. The beliefs we have examined to this point are

$$b_0^i = \beta^i(\omega) \text{ and } b_1^i = \beta^i(\omega, b_0^{-i}).$$

Let $\mathcal{B}_1^{-i}$ denote the σ-algebra induced by the announcements (b_0^{-i}, b_1^{-i}). Then given the beliefs $b_0^i = \beta^i(\omega)$ and $b_1^i = \beta^i(\omega, b_0^{-i})$, an announcement by the remaining agents of their updated posteriors $b_1^j = \beta^j(\omega, b_0^{-j})$ results in agent

[14] Each agent's beliefs in any round of the protocol depend only on previous-round beliefs of the other agents. There is then no circularity of the type that arises in Spiegler (2016) when considering decisions.

[15] Throughout the formal analysis, agent i's beliefs at each stage of the updating process are given by a function on X^N that is measurable with respect to the σ algebra generated by the relevant information. We conserve on notation by using $\beta^i(\cdot)$ to denote these beliefs, such as $\beta^i(\omega)$, $\beta^i(\omega, b_0^{-i})$, $\beta^i(\omega, b_0^{-i}, b_1^{-i})$, and so on, where $\cdot$ identifies the relevant information. (Recall footnote 9.) In examples, we replace ω by ω_{I^i} in this notation to serve as a reminder of agent i's initial information.

i updating her beliefs to

$$b_2^i = \beta^i(\omega, b_0^{-i}, b_1^{-i}) = \mathbb{E}[f^i|\mathcal{I}^i, \mathcal{B}_1^{-i}](\omega).$$

Letting $\mathcal{B}_n^{-i}$ and $\mathcal{B}_n$ denote the σ-algebras induced by $(b_0^{-i}, \dots, b_n^{-i})$ and $(b_0, \dots, b_n)$, we have, for all n,

$$b_{n+1}^i = \beta^i(\omega, b_0^{-i}, \dots, b_n^{-i}) = \mathbb{E}[f^i|\mathcal{I}^i, \mathcal{B}_n^{-i}](\omega) = \mathbb{E}[f^i|\mathcal{I}^i, \mathcal{B}_n](\omega),$$

where the last equality follows from $\sigma(\mathcal{I}^i, \mathcal{B}_n^{-i}) = \sigma(\mathcal{I}^i, \mathcal{B}_n)$ (and $\sigma(\mathcal{A}, \mathcal{B})$ is the σ-algebra generated by the σ-algebras $\mathcal{A}$ and $\mathcal{B}$). Denote by $\mathbf{b} := (b_0, b_1, \dots)$ the infinite sequence of announcements and by $\mathcal{B}_\infty$ the σ-algebra induced by $\mathbf{b}$. It will also be useful to keep track of the beliefs of the *public oracle*,

$$\mathbb{E}[f|\mathcal{B}_n] \text{ and } \mathbb{E}[f|\mathcal{B}_\infty].$$

Intuitively, a public oracle is an agent whose model is given by N and whose theory is thus given by f, and who observes the announcements of all players, but no other information.

Since each agent i and the public oracle follow Bayesian updating on the sequence of increasingly informative announcements (filtrations) $(\mathcal{B}_n^{-i})$ and $(\mathcal{B}_n)$, the resulting sequence of updates are martingales and so converge (with probability one under ρ) to limits which are measurable with respect to the limit σ-algebras. Summarizing this discussion:

Lemma 1 *The updated beliefs*

$$(\mathbb{E}[f^i|\mathcal{I}^i, \mathcal{B}_n^{-i}])_{n=1}^\infty \text{ and } (\mathbb{E}[f|\mathcal{B}_n])_{n=1}^\infty$$

are martingales, with ρ-almost-sure limits

$$\mathbb{E}[f^i|\mathcal{I}^i, \mathcal{B}_\infty^{-i}] \text{ and } \mathbb{E}[f|\mathcal{B}_\infty].$$

Remark 2 (Different Models or Different Events?) We interpret our analysis as that of agents forming beliefs about a single event F, but with different models. The challenge is then to examine how agents infer information relevant to their own models from other agents who have different models. Returning to our example, the fundamentalist may recognize that there is information to be gleaned about fundamentals from another agent who is primarily concerned with charts.

Much of the analysis of this section could be recast as one in which every agent is an oracle, but the agents are forming beliefs about different events.

The challenge is then to examine how agents infer information about their own events from other agents who are concerned with other events. One fundamentalist may be concerned with industrial stocks, while recognizing that there is useful information to be gleaned from the beliefs of an analyst who specializes in agricultural futures. For the purposes of much of this section, it is a taste question which interpretation is most congenial. However, most of the results in subsequent sections are consistent only with our preferred interpretation of agents using different models to reason about a common event. ♦

3.3 Example

Agent i may find j's beliefs relevant for two reasons. First, j may observe a variable that appears in i's model but i does not observe. A fundamentalist may be convinced that the outcome of a firm's recent drug trial is important, but may not be privy to that outcome, and so may glean inferences from the beliefs of an insider. Second, there may be correlations between the variables. A fundamentalist may take note when the CEO of a firm opens a secret bank account in the Cayman Islands, not because such accounts appear directly in the list of fundamentals, but because they are correlated with other variables that do. If the president of the country believes in astrology, then government policy may cause firm fundamentals to be correlated with astrological phenomena, inducing the fundamentalist to glean information from the beliefs of the astrologer.

Example 1 Suppose the state space is given by $\{0,1\}^4$. The pair (ω_1, ω_2) is drawn from the distribution $\Pr\{(\omega_1, \omega_2) = (0,0)\} = \Pr\{(\omega_1, \omega_2) = (1,1)\} = \frac{3}{8}$, $\Pr\{(\omega_1, \omega_2) = (0,1)\} = \Pr\{(\omega_1, \omega_2) = (1,0)\} = \frac{1}{8}$, and the pair (ω_3, ω_4) is independently drawn from a distribution with an identical correlation structure. The event is

$$F = \left\{ \omega : \sum\nolimits_{k=1}^{4} \omega_k \geq 2 \right\}.$$

There are two agents. Agent 1's model and information are given by

$$M^1 = \{1,2,3,4\} \quad \text{and} \quad I^1 = \{2,3\}$$

while agent 2's are given by

$$M^2 = \{3,4\} \quad \text{and} \quad I^2 = \{4\}.$$

State	Prior		2's theory	Interim beliefs		First-round updates		Second round
$(\omega_1, \omega_2, \omega_3, \omega_4)$	ρ	$f(\omega)$	$f^2(\omega_{M^2})$	$\beta^1(w_{I^1})$	$\beta^2(w_{I^2})$	$\beta^1(\omega_{I^1}, b_0^2)$	$\beta^2(\omega_{I^2}, b_0^1)$	$\beta^2(\omega_{I^2}, b_0^1, b_1^1, b_0^2)$
$(0,0,0,0)$	9/64	0	3/8	1/16	14/32	0	3/8	3/8
$(0,0,0,1)$	3/64	0	5/8	1/16	29/32	1/4	5/8	5/8
$(0,0,1,0)$	3/64	0	5/8	13/16	14/32	1/4	14/32	5/8
$(0,1,0,0)$	3/64	0	3/8	13/16	14/32	3/4	14/32	3/8
$(1,0,0,0)$	3/64	0	3/8	1/16	14/32	0	3/8	3/8
$(0,0,1,1)$	9/64	1	1	13/16	29/32	1	29/32	29/32
$(0,1,0,1)$	1/64	1	5/8	13/16	29/32	1	29/32	29/32
$(1,0,0,1)$	1/64	1	5/8	1/16	29/32	1/4	5/8	5/8
$(0,1,1,0)$	1/64	1	5/8	1	14/32	1	5/8	5/8
$(1,0,1,0)$	1/64	1	5/8	13/16	14/32	1/4	14/32	5/8
$(1,1,0,0)$	9/64	1	3/8	13/16	14/32	3/4	14/32	3/8
$(1,1,1,0)$	3/64	1	5/8	1	14/32	1	5/8	5/8
$(1,1,0,1)$	3/64	1	5/8	13/16	29/32	1	29/32	29/32
$(1,0,1,1)$	3/64	1	1	13/16	29/32	1	29/32	29/32
$(0,1,1,1)$	3/64	1	1	1	29/32	1	1	1
$(1,1,1,1)$	9/64	1	1	1	29/32	1	1	1

$$\begin{aligned} M^1 &= \{1,2,3,4\}, & M^2 &= \{3,4\}, \\ I^1 &= \{2,3\}, & I^2 &= \{4\}. \end{aligned}$$

Figure 1: The beliefs for Example 1. Because agent 1 is an oracle, her theory agrees with the indicator f and so is not listed separately.

This information is summarized in Figure 1, together with the interim beliefs $\beta^1(\omega_{I^1})$ and $\beta^2(\omega_{I^2})$.

Now we turn to updating in response to others' beliefs. First, consider agent 1, who is an oracle. Agent 2 observes only one piece of information, namely ω_4, and different realizations of ω_4 cause agent 2 to announce different interim beliefs. Agent 1's first-round update, given by (5), is then identical to the interim belief agent 1 would have if 1 observed $\{\omega_2, \omega_3, \omega_4\}$. We report these beliefs in Figure 1, in the column labeled $\beta^1(\omega_{I^1}, b_0^2)$. There is nothing more agent 1 can learn, and hence agent 1 does no further updating.

Turning to agent 2, suppose, first, $b_0^1 = 1$. Agent 2 observes ω_4, and infers that agent 1 has observed $\omega_3 = 1$. Agent 2 then has (from her point of view) full information. Any information about ω_2 in agent 1's belief agent

2 considers irrelevant. Agent 2's updated beliefs $\beta^2(\omega_{I^2}, b_0^1)$ about the event F are then given by

$$\beta^2(0,1) = 5/8 \quad \text{and} \quad \beta^2(1,1) = 1.$$

We see here the difference between model-based and oracular updating. An agent-2 oracle who observed $\omega_4 = 0$ and $b_0^1 = 1$ would infer that the state is $(0,1,1,0)$ with probability 1/2 and $(1,1,1,0)$ with probability 1/2. Both states give rise to the event F, and so the agent-2 oracle would attach posterior probability 1 to the event. In contrast, the model-based updater who has observed $\omega_4 = 0$ and $b_0^1 = 1$ draws the inference that the state (in her model) (ω_3, ω_4) equals $(1,0)$. The agent then calculates her full information probability of F, given $(\omega_3, \omega_4) = (1,0)$, which is 5/8.

The case of $b_0^1 = 1/16$ is similar.

Finally, suppose $b_0^1 = \frac{13}{16}$. Unlike the previous two cases, this observation does not unambiguously identify player 1's observation, instead pooling the realizations $(0,1)$ and $(1,0)$ of (ω_2, ω_3). Let $\rho^2(\omega_{M^2} | \omega_{I^2}, b_0^1)$ identify the probabilities agent 2 attaches to the values of the states (in her model) $\omega_{M^2} = (\omega_3, \omega_4)$ given the information ω_{I^2} and the announcement b_0^1. Then

$$\begin{aligned} &\rho(0,0 \mid 0, 13/16) = 3/4, & & \rho(1,0 \mid 0, 13/16) = 1/4, \\ &\rho(0,1 \mid 1, 13/16) = 1/4, & \text{and} \quad & \rho(1,1 \mid 1, 13/16) = 3/4. \end{aligned}$$

Agent 2's updated beliefs $\beta^2(\omega_{I^2}, b_0^1)$ about the event F are then given by

$$\beta^2(0, 13/16) = 14/32 \quad \text{and} \quad \beta^2(1, 13/16) = 29/32.$$

Again, these beliefs differ from those of an agent-2 oracle, who attaches probabilities 5/8 (after observing $(\omega_4, b_0^1) = (0, 13/16)$) and 1 (after observing $(\omega_4, b_0^1) = (1, 13/16)$) to event F. The results of agent 2's updating are reported in the column $\beta^2(\omega_{I^2}, b_0^1)$. This concludes the first round of updating.

The subsequent round of updating is described in Appendix B.2. ♦

3.4 How Revealing are Beliefs?

Why not have agents simply announce their *information* rather than their beliefs? We are comfortable in abstracting from the details of agents' interactions by using the exchange of beliefs as a convenient proxy for the workings of such interactions, but we are not comfortable simply assuming the interaction will reveal *all* of the agents' information.

This difference matters. As illustrated by Geanakoplos and Polemarchakis (1982, Proposition 3) and Figure 4 below, an agent oracle need not hold the same beliefs as someone who can observe the information contained in $\cup_{k=1}^{K} I^k$.[16] Instead, some player k's belief announcements may pool together some of the information contained in I^k. Constructing such examples is straightforward, even when the agents are all oracles.

One might counter that the pooling encountered in these examples is nongeneric (Geanakoplos and Polemarchakis, 1982, Proposition 4). Indeed, one might argue that for a generic specification of prior beliefs, each agent's *first* announcement reveals that agent's information, and hence we need not worry about multiple rounds of announced beliefs.

We first note that if the state space is a (multi-dimensional) continuum with agents receiving continuously distributed signals, and if an agent observed several signals, then a *one*-dimensional announcement will typically (and generically) not reveal all the agent's information. We find it convenient in the examples to strip away complications by working with discrete signals, but are then unwilling to appeal to genericity arguments. Second, even within a discrete framework, the space of prior beliefs may not be the appropriate space to seek genericity. For example, the factors determining which state has occurred may be summarized by a tree, with random moves at decision nodes and terminal nodes corresponding to states. We would then apply genericity arguments to the mixtures appearing in the tree. If this tree has a nontrivial structure, then generic specifications of the probabilities appearing in the tree will induce probability distributions over states that appear nongeneric, but that we nonetheless view as robust.

We believe that the repeated announcement of beliefs gives us information transmission similar to that allowed by (for example) the common knowledge that agents are willing to trade, sufficiently so that we are willing to avoid modeling the fine details of market microstructure by working directly with sequences of belief announcements. However, we are not convinced that market or other interactions will necessarily reveal every detail of every agent's information, and so would be skeptical of a model that precluded pooling.

3.5 The Bliss of Others' Ignorance

Our next example illustrates a phenomenon that can only arise with agents having different models: increasing the information of one agent (even when

[16] The limit beliefs held by an agent oracle are Geanakoplos and Polemarchakis's (1982) indirect communication equilibrium beliefs.

State	Prior		Interim beliefs		First-round update
$(\omega_1, \omega_2, \omega_3)$	ρ	$f^*(\omega)$	$\beta^1(\omega_{I^1})$	$\beta^2(\omega_{I^2})$	$\beta^1(\omega_{I^1}, b_0^2)$
$(0,0,0)$	$1/10$	0	$(2x+y)/5$	$(x+y)/4$	0
$(1,0,0)$	$1/10$	x	$(x+y)/5$	$(x+y)/4$	$(x+y)/2$
$(0,0,1)$	$1/10$	0	$(2x+y)/5$	$(x+y)/4$	0
$(1,0,1)$	$1/10$	y	$(x+y)/5$	$(x+y)/4$	$(x+y)/2$
$(0,1,0)$	$2/10$	x	$(2x+y)/5$	$(2x+y)/6$	$(2x+y)/3$
$(1,1,0)$	$2/10$	0	$(x+y)/5$	$(2x+y)/6$	0
$(0,1,1)$	$1/10$	y	$(2x+y)/5$	$(2x+y)/6$	$(2x+y)/3$
$(1,1,1)$	$1/10$	0	$(x+y)/5$	$(2x+y)/6$	0

$$\Omega = \{0,1\}^3, \quad \begin{array}{rcl} M^1 &=& \{1,2,3\}, \\ I^1 &=& \{1\}, \end{array} \quad \begin{array}{rcl} M^2 &=& \{2,3\}, \\ I^2 &=& \{2\}. \end{array}$$

Figure 2: The beliefs for Example 2.

another agent thinks the information is valuable) can result in a deterioration of inferences.

Example 2 We jump immediately to the tabular presentation of this example, which includes all the relevant information, presented in Figure 2. In contrast to the presentation of our earlier examples, we replace the column specifying the indicator function, f, with f^*, its expected value conditional on *all* the agents' model variables, i.e., $f^*(\omega) := \mathbb{E}[f(\omega) \mid \omega_1, \omega_2, \omega_3]$. In Example 1, $N = \cup_i M^i$, which is to say that the variables contained in $\{0,1\}^{\cup_i M^i}$ suffice to determine the value of f. In the current example, there are additional variables in the state space that we have not presented. These variables lie outside all agents' models, and play a role in the analysis only to the extent that they shape the values of f^* and so we omit them from the table.

Since ω_1 is independent of (ω_2, ω_3), agent 2 learns nothing from agent 1 and does no updating. Agent 1 learns the realization of ω_2 from agent 2, and so does one round of updating. In four of the states, agent 1 learns the probability of F, namely 0. Agent 1 overestimates the value of F in two of the remaining four states and underestimates it in the remaining two states.

Now suppose we give agent 2 more information, as displayed in Figure 3. Agent 2 again does not update, while agent 1 does one round of updating. As a result of the additional information, agent 2 now pools her states. Agent 1 does *not* estimate the probability of F correctly in any state. ♦

State	Prior		Interim beliefs		First-round update
$(\omega_1, \omega_2, \omega_3)$	ρ	$f^*(\omega)$	$\beta^1(\omega_{I^1})$	$\beta^2(\omega_{I^2})$	$\beta^1(\omega_{I^1}, b_0^2)$
$(0,0,0)$	$1/10$	0	$(2x+y)/5$	$x/2$	$2x/3$
$(1,0,0)$	$1/10$	x	$(x+y)/5$	$x/2$	$x/3$
$(0,0,1)$	$1/10$	0	$(2x+y)/5$	$y/2$	$y/2$
$(1,0,1)$	$1/10$	y	$(x+y)/5$	$y/2$	$y/2$
$(0,1,0)$	$2/10$	x	$(2x+y)/5$	$x/2$	$2x/3$
$(1,1,0)$	$2/10$	0	$(x+y)/5$	$x/2$	$x/3$
$(0,1,1)$	$1/10$	y	$(2x+y)/5$	$y/2$	$y/2$
$(1,1,1)$	$1/10$	0	$(x+y)/5$	$y/2$	$y/2$

$$\Omega = \{0,1\}^3, \quad \begin{array}{rcl} M^1 & = & \{1,2,3\}, \\ I^1 & = & \{1\}, \end{array} \quad \begin{array}{rcl} M^2 & = & \{2,3\}, \\ I^2 & = & \{2,3\}. \end{array}$$

Figure 3: The result of giving agent 2 in Figure 2 increased information.

3.6 Properties of the Belief Updating Process

The following proposition gathers some information about the belief-updating process. Recall that throughout, we maintain the assumption that $I^i \subseteq M^i$ for all i, and that $\mathbf{b} = (b_0, b_1, \dots)$ denotes the complete sequence of publicly announced beliefs with associated σ-algebra $\mathcal{B}_\infty$. We introduce the *omniscient* oracle who, in addition to having the model N, knows the realization of the state.

Proposition 1

1. *If $\mathcal{M}^i$ is finite for all $i \in \{1, \dots, K\}$, then $\mathbf{b}$ is eventually constant, i.e., the updating process terminates. If the M^i are infinite, the updating process need not terminate.*

2. *The limiting beliefs of different agents need not be equal.*

3. *With ρ-probability one, once an agent's belief equals 0 or 1, that agent's beliefs agree with those of the omniscient oracle, and so are never subsequently revised.*[17] *Thus, two agents cannot simultaneously assign a belief of* 0 *and* 1 *to the event F.*

[17] So model-based reasoners cannot match the common description of being "often wrong but never in doubt."

4. *Agent i's private information is only pooled in the limit if it does not matter to agent i, that is,*

$$\mathbb{E}[f^i \mid \mathcal{I}^i, \mathcal{B}_\infty] = \mathbb{E}[f^i \mid \mathcal{B}_\infty].$$

5. *If f depends only on the variables in M^i, then $f^i = f$ and agent i's limit belief equals the agent oracular and public oracular belief, that is,*

$$\begin{aligned}\mathbb{E}[f \mid \mathcal{I}^i, \mathcal{B}_\infty] &= \mathbb{E}[f^i \mid \mathcal{I}^i, \mathcal{B}_\infty] \\ &= \mathbb{E}[f^i \mid \mathcal{B}_\infty] \\ &= \mathbb{E}[f \mid \mathcal{B}_\infty].\end{aligned}$$

6. *If $\cup_j I^j \subseteq M^i$, then agent i's limit belief equals the agent oracular and public oracular belief,*

$$\begin{aligned}\mathbb{E}[f \mid \mathcal{I}^i, \mathcal{B}_\infty] &= \mathbb{E}[f^i \mid \mathcal{I}^i, \mathcal{B}_\infty] \\ &= \mathbb{E}[f^i \mid \mathcal{B}_\infty] \\ &= \mathbb{E}[f \mid \mathcal{B}_\infty].\end{aligned}$$

Proof.

1. At each round n of the updating process, agent i's belief about the event F is the expectation of i's full-information belief conditioning on the σ-algebra reflecting i's information and the information revealed by the collective announcements of the agents, $\sigma(\mathcal{I}^i, \mathcal{B}_n)$. The sequence $(\sigma(\mathcal{I}^i, \mathcal{B}_n))_{n=0}^{\infty}$ is a filtration, with each σ-algebra being coarser than $\sigma(\mathcal{I}^1, \ldots, \mathcal{I}^K)$. If all $\mathcal{M}^j$ are finite, then each $\sigma(\mathcal{I}^i, \mathcal{B}_n)$ and $\sigma(\mathcal{I}^1, \ldots, \mathcal{I}^K)$ are generated by finite partitions, and so the filtration must eventually be constant, ensuring that the updating process terminates. Appendix B.3 describes an example with infinite M^1 and M^2 in which updating proceeds for an infinite number of rounds.

2. Example 1 shows that the limit beliefs need not agree. With positive probability, the limiting beliefs in Appendix B.3 are not equal.

3. A belief b_n^i for agent i can equal an extreme value (0 or 1) at some round n if and only if the full-information belief $f^i(\omega)$ takes the same extreme value on a full ρ-measure event in $\sigma(\mathcal{I}^i, \mathcal{B}_n)$, which implies the omniscient oracle has the same beliefs on a full ρ-measure event in $\sigma(\mathcal{I}^i, \mathcal{B}_n)$, and so on every subsequent subevent in the sequence.

Since the omniscient oracle cannot have two distinct beliefs, it is then immediate that two agents cannot simultaneously assign a beliefs of 0 and 1 to the event F.

4. Proof is by contradiction. Suppose that

$$\mathbb{E}[f^i \mid \mathcal{I}^i, \mathcal{B}_\infty] \neq \mathbb{E}[f^i \mid \mathcal{B}_\infty].$$

   Then, $\mathcal{B}_\infty$ must pool together some states that agent i does not pool together, and on which f^i is not constant.[18] But if this were the case, then there would be an announcement from agent i not contained in $\mathcal{B}_\infty$, a contradiction.

5. Immediately follows from the definitions and item 4.

6. We verify the first equality. Since $\sigma(\mathcal{I}^i, \mathcal{B}_\infty) \subseteq \sigma(\cup_j \mathcal{I}^j)$, if $\cup_j \mathcal{I}^j \subseteq \mathcal{M}^i$, then $\sigma(\mathcal{I}^i, \mathcal{B}_\infty) \subseteq \sigma(\mathcal{M}^i)$, and so (using (2) and the law of iterated expectations)

$$\begin{aligned}\mathbb{E}[f^i \mid \mathcal{I}^i, \mathcal{B}_\infty] &= \mathbb{E}[\mathbb{E}[f \mid \mathcal{M}^i] \mid \mathcal{I}^i, \mathcal{B}_\infty] \\ &= \mathbb{E}[f \mid \mathcal{I}^i, \mathcal{B}_\infty].\end{aligned}$$

   The second equality is just item 5 above, while the third equality is established by an identical argument to that which verified the first equality.

■

Remark 3 (Common Knowledge) If we adopt the interpretation that the agents know each others' models, then their limit beliefs are common knowledge. Appendix B.4 provides details. ⧫

[18] More precisely, there exist two positive probability events E and E' in $\sigma(\mathcal{I}^i, \mathcal{B}_\infty)$ not separated by $\mathcal{B}_\infty$ (i.e., for all events $B \in \mathcal{B}_\infty$, we have either $E, E' \subseteq B$ or $(E \cup E') \cap B = \varnothing$) for which $\mathbb{E}[f^i \mid E] \neq \mathbb{E}[f^i \mid E']$.

4 Agents and Oracles

4.1 Do Agents Agree?

In one sense, it is obvious (and our earlier discussion confirms) that when agents have substantively different models, their limit beliefs may not agree. We now explore the sources and implications of agreement.

To ease notation, we consider the case of two agents. We say that limit beliefs *necessarily agree* if, for all $\omega \in \Omega$, the limit beliefs of agents 1 and 2 are equal, i.e.,

$$\mathbb{E}[f^1|\mathcal{I}^1, \mathcal{B}_\infty](\omega) = \mathbb{E}[f^2|\mathcal{I}^2, \mathcal{B}_\infty](\omega). \tag{6}$$

The left side is agent 1's model-based belief, giving 1's observation of ω_{I^1} and the announced sequence of beliefs, and the right side is agent 2's corresponding belief. From Proposition 1.4, equation (6) can be rewritten as

$$\mathbb{E}[f^1|\mathcal{B}_\infty](\omega) = \mathbb{E}[f^2|\mathcal{B}_\infty](\omega).$$

Given I^1 and I^2, we say that the variable $k \in M^i$ is *redundant* in agent i's model if $\mathbb{E}[f^i \mid \mathcal{I}^1, \mathcal{I}^2]$ is constant in ω_k.

When the variables in $I^i \setminus M^j$ are *not* redundant for agent i, it seems difficult to achieve necessary agreement. Suppose $k \in I^i \setminus M^j$ is not redundant for agent i. Then we would expect agent i's belief to vary as the value of ω_k varies, while agent j's *theory* is *not* responsive to variations in the values of ω_k for $k \notin M^j$. It may still be that agent j's *belief* varies with ω_k (reflecting changes in j's beliefs over ω_{M^j}), but agent j's theory averages f over ω_k and so it seems difficult to obtain agreement on the more confident beliefs, precluding necessary agreement. Appendix B.5 illustrates this intuition.

While suggestive, this intuition fails when there is no variation in beliefs. The example in Figure 4 has the feature that both agents pool their information (agent 1 because she is not fully informed). Consequently, both agents' beliefs agree with the prior.

Under independence, necessary agreement implies either that the variables in $I^i \setminus M^j$ are redundant, or the information is not being revealed (as in Figure 4):

Proposition 2 *Suppose the variables ω_k are drawn independently and $K = 2$. Suppose the beliefs of agents 1 and 2 necessarily agree. If for some $i \in \{1, 2\}$, the variables in $I^i \setminus M^j$ are not redundant for agent i, then for some $j \in \{1, 2\}$ (which may but need not equal i)*

$$\mathbb{E}[f^j \mid \mathcal{B}_\infty] \neq \mathbb{E}[f^j \mid \mathcal{I}^1, \mathcal{I}^2],$$

State	Prior		Theories		Beliefs	
(ω_1, ω_2)	ρ	$f(\omega)$	$f^1(\omega_{M^1})$	$f^2(\omega_{M^2})$	$\beta^1(\omega_{I^1})$	$\beta^2(\omega_{I^2})$
$(0,0)$	$1/4$	0	0	$1/2$	$1/2$	$1/2$
$(0,1)$	$1/4$	1	1	$1/2$	$1/2$	$1/2$
$(1,0)$	$1/4$	1	1	$1/2$	$1/2$	$1/2$
$(1,1)$	$1/4$	0	0	$1/2$	$1/2$	$1/2$

$$X = \{0,1\}, \quad \begin{array}{rcl} M^1 &=& \{1,2\}, \quad M^2 = \{2\}, \\ I^1 &=& \{1\}, \quad\;\; I^2 = \{2\}. \end{array}$$

Figure 4: An example with agreement, even though variable 1 is not redundant for agent 1.

that is, not all the agents' information is revealed.

Note that if agents' information is not all revealed, then even if the different variables in N are independent, agent's beliefs conditional on $\mathcal{B}_\infty$ need not be independent.

Proof. Suppose the beliefs of agents 1 and 2 necessarily agree and that $I^1 \setminus M^2$ is not redundant for agent 1, Suppose, moreover, that for both $i = 1, 2$,

$$\mathbb{E}[f^i \mid \mathcal{B}_\infty] = \mathbb{E}[f^i \mid \mathcal{I}^1, \mathcal{I}^2], \tag{7}$$

so that all agents are effectively conditioning on all the information. Since the variables are independent, $\mathbb{E}[f^2 \mid \mathcal{I}^1, \mathcal{I}^2]$ does not depend on ω_k for $k \in I^1 \setminus M^2$. But necessary agreement implies $\mathbb{E}[f^1 \mid \mathcal{B}_\infty] = \mathbb{E}[f^2 \mid \mathcal{B}_\infty]$ while nonredundancy of $I^1 \setminus M^2$ implies $\mathbb{E}[f^1 \mid \mathcal{I}^1, \mathcal{I}^2]$ must depend on ω_k for some $k \in I^1 \setminus M^2$, contradicting (7). ∎

Hence, when variables are independent, agents necessarily agree only if either there are effectively no differences in information (i.e., information is either redundant or common) or not all information is revealed.

Correlation in variables may allow necessary agreement even when there are nonredundant variables and all information is revealed. If there is correlation, then agent 1 may observe information that is useful to agent 2, not because it appears in 2's model but because it is correlated with the values of other variables in 2's model (that 2 does not observe). The example in Appendix B.6 illustrates that this can occur. Moreover, Section 5 shows that strong correlation implies limit beliefs will be close.

Conversely, redundancy in general leads to agreement, with or without independence. In the following, we must strengthen redundancy since $\mathbb{E}[f^1 \mid \mathcal{I}^1, \mathcal{I}^2]$ constant in ω_k for all $k \in I^1 \setminus M^2$ need not imply that $\mathbb{E}[f^1 \mid \mathcal{I}^1]$ is constant in ω_k for all $k \in I^1 \setminus M^2$ (for example, ω_k may be correlated with a variable in $I^2 \cap M^1$).

Proposition 3 *Suppose $K = 2$, $\mathbb{E}[f^1 \mid \mathcal{I}^1, \mathcal{G}^2]$ is constant in ω_k for all $k \in I^1 \setminus M^2$ and all sub-σ-algebras $\mathcal{G}^2$ of $\mathcal{I}^2$, and $\mathbb{E}[f^2 \mid \mathcal{I}^2, \mathcal{G}^1]$ is constant in ω_k for all $k \in I^2 \setminus M^1$ and all sub-σ-algebras $\mathcal{G}^1$ of $\mathcal{I}^1$. Then, the limit beliefs of agents 1 and 2 necessarily agree, and agree with the public oracular belief.*

Proof. The strengthened redundancy assumptions imply that no matter what agent 1, for example, learns from agent 2's announcements, 1's prediction is independent of ω_k for all $k \in I^1 \setminus M^2$. This implies that, without loss of generality, we may assume $I^i \setminus M^j = \varnothing$ for all i. This implies that for each agent i, $I^1 \cup I^2 \subseteq M^i$, and so by Proposition 1.6, agent i's limit belief necessarily agrees with the public oracular belief, and so with agent j's limit belief. ∎

4.2 The Wisdom of the Crowd?

The idea of the "wisdom of the crowd" (e.g., Surowiecki (2004), Wolfers and Zitzewitz (2004)) is that groups or "crowds" of people effectively aggregate information, even if their members disagree. We can thus reasonably assert that information is aggregated, even though various agents disagree, as long as the crowd forms beliefs that are "correct on average."

We have introduced agent oracles, the public oracle and the omniscient oracle. We now introduce the *universal oracle*, who has access to all of the agents' information and hence has beliefs $\mathbb{E}[f \mid \mathcal{I}^1, \ldots, \mathcal{I}^K]$.

All oracular beliefs are based on the true indicator function f. The difference between the different oracles is the information on which they condition. In order of increasing information, the public oracle has the least information (namely, $\mathcal{B}_\infty$), followed by an agent's oracle (who has both $\mathcal{B}_\infty$ and that agent's information $\mathcal{I}^i$), then the universal oracle, and finally the omniscient oracle.

One interpretation of "correct on average" is that there is some statistic φ of the crowd's limit beliefs that necessarily agrees with the universal oracle. There are many possible candidates for such a statistic (mean, median, etc).

It is easy to see that if there is *any* such statistic, then the public oracle effectively aggregates all the agents' information.

Proposition 4 *Suppose there exists a function $\varphi : [0,1]^K \to [0,1]$ that necessarily agrees with the universal oracle, i.e., for all $\omega \in \Omega$,*

$$\varphi\left(\mathbb{E}[f^1 \mid \mathcal{I}^1, \mathcal{B}_\infty](\omega), \ldots, \mathbb{E}[f^K \mid \mathcal{I}^1, \mathcal{B}_\infty](\omega)\right) = \mathbb{E}[f \mid \mathcal{I}^1, \ldots, \mathcal{I}^K](\omega). \quad (8)$$

Then, for all $\omega \in \Omega$, the public and universal oracular beliefs coincide:

$$\mathbb{E}[f \mid \mathcal{I}^1, \ldots, \mathcal{I}^K](\omega) = \mathbb{E}[f \mid \mathcal{B}_\infty](\omega).$$

Proof. From Proposition 1.4,

$$\varphi(\mathbb{E}[f^1 \mid \mathcal{I}^1, \mathcal{B}_\infty](\omega), \ldots, \mathbb{E}[f^K \mid \mathcal{I}^1, \mathcal{B}_\infty](\omega)) = \varphi(\mathbb{E}[f^1 \mid \mathcal{B}_\infty](\omega), \ldots, \mathbb{E}[f^K \mid \mathcal{B}_\infty](\omega)),$$

and so the statistic (as a function of limit beliefs) must be measurable with respect to the sequence of public announcements $\mathcal{B}_\infty$, and by (8) then so must be the belief of the universal oracle. But then the universal oracle must agree with the public oracle. ■

The least demanding standard for beliefs being correct on average is that the universal oracular belief lies in the convex hull of the agents' updated beliefs. Unfortunately, even this mild requirement is not guaranteed.

Example 3 We examine a case in which $M^1 \cup M^2 = \Omega = I^1 \cup I^2 = \Omega$, so every variable appears in the model of at least one agent and is also observed by at least one agent. This presents conditions most favorable to information aggregation. Consider the environment in Figure 5. Both agents observe the information they deem relevant, neither updates, and their limiting beliefs are given by their theories. In every state, the universal oracular belief (given by $f^*(\omega)$) lies outside the convex hull of the agents' limit beliefs.

★

Our next proposition shows it is a pervasive result that the universal oracular belief lies outside the convex hull of the model-based beliefs. A subset $\tilde{N} \subset N$ is *sufficient* if the variables in $\tilde{N}$ suffice to determine whether F has occurred. Obviously N is always a sufficient set. There is always at least one minimal sufficient set, and there may be multiple minimal sufficient sets (e.g., if the realizations of some variables are perfectly correlated).

State (ω_1, ω_2)	Prior ρ	$f^*(\omega)$	Theories $f^1(\omega_{M^1})$	$f^2(\omega_{M^2})$
$(0,0)$	$1/4$	$7/8$	$1/2$	$9/16$
$(0,1)$	$1/4$	$1/8$	$1/2$	$1/2$
$(1,0)$	$1/4$	$2/8$	$9/16$	$9/16$
$(1,1)$	$1/4$	$7/8$	$9/16$	$1/2$

$$X = \{0,1\}, \quad M^1 = \{1\}, \quad M^2 = \{2\},$$
$$I^1 = \{1\}, \quad I^2 = \{2\}.$$

Figure 5: The universal oracular beliefs (given by $f^*(\omega)$) are not in the convex hull of agent beliefs.

Proposition 5 *Suppose X and N are finite and let $\cup_{k=1}^K M^k = \cup_{k=1}^K I^k = \widetilde{N}$ for some minimal sufficient set $\widetilde{N}$, with the collection $\{M^k\}_k$ pairwise disjoint and $M^k \subsetneq \widetilde{N}$ for each k. Suppose ρ has full support. Then there exist states for which the universal oracular belief lies outside the convex hull of the model-based beliefs.*

Proof Because no model is sufficient, f is not constant. Suppose first that beliefs reveal the agent's information, i.e., $\mathcal{B}_\infty = \sigma(\mathcal{I}^1, \ldots, I^K)$. Then the agents' limit model-based beliefs will be their full-information beliefs. However, because f is not measurable with respect to any $\mathcal{M}^j$, for every agent there is a state at which her beliefs do not equal 0 or 1.

We now note that there is a state at which every agent's belief is strictly between 0 and 1. Otherwise it could not be the case that for every state, there is at least one agent whose belief is either 0 or 1.[19]

Because $\widetilde{N}$ is sufficient, the universal oracular belief will always be 0 or 1, and hence must sometimes lie outside the convex hull of the agents beliefs.

Finally, if beliefs are not revealing, then the agents have less information, and so again there cannot be an agent whose beliefs are always either 0 or 1. ■

This result does not require that announcements pool information, and so is not simply a statement that the universal oracle has more information

[19] Since ρ has full support, if $f^i(\omega_{M^i}) = 1$ for some ω_{M^i}, then $f(\omega_{M^i}, \omega_{-M^i}) = 1$ for all ω_{-M^i}. Suppose for some ω'_{M^j}, $f^j(\omega'_{M^j}) = 0$. Then for all ω_{M^i} and all $\omega_{-M^i-M^j}$, $f(\omega_{M^i}, \omega'_{M^j}, \omega_{-M^i-M^j}) = 0$. But this is impossible, and so $f(\omega) = 1$ for all ω, a contradiction (because f is not constant).

than does any single agent—in the limit they will often have identical information. However, the universal oracle has a more encompassing model than any of the individuals, and hence makes use of more information, leading to more extreme beliefs.

5 Information Aggregation

This section presents three variations on the idea that groups of agents will effectively aggregate information if they have a sufficiently common understanding. The agents need not have similar information, and indeed each individual agent may have very little information. The protocol will aggregate their information, as long as their models by which they interpret this information are not too different.

5.1 Correlated Model Predictions

If the realizations of the variables in the agents' different models are sufficiently correlated across models, then their limit beliefs will be close. We view this correlation as an indication that the agents' models are, for practical purposes, nearly the same. The extreme case involves agents whose models are disjoint, but whose realizations are perfectly correlated, so that the agents effectively have the same model described in different languages.

Proposition 6 *Fix agent i's theory f^i. For any $\varepsilon > 0$, there is an $\eta < 1$ such that if the coefficient of correlation between agent i's theory f^i and agent j's theory f^j is at least $1 - \eta$, then agents i's and j's limiting beliefs are within ε of one another with probability $1 - \varepsilon$.*

This proposition imposes the correlation requirement on f^i and f^j, rather than imposing the stronger requirement on the correlation between ω_{M^i} and ω_{M^j}, because correlation is relevant only for those variables that play a role in affecting beliefs about F.

The proof first shows that if two agents' theories are perfectly correlated ex ante, then their updated beliefs must be identical. In this case, the agents effectively have identical models with different descriptions. We then show that if two agents' theories are *close* ex ante, then their limit beliefs must, with high probability, be close. The delicateness in establishing this seemingly intuitive result arises in showing that it holds *irrespective* of the nature of the agents' information. Appendix A.1 contains the proof.

5.2 Models with a Common Component

The next result shows that if the agents' models share a large enough common component, then their beliefs cannot be too different from one another. We fix a group of K agents and examine a sequence $(M_n^1, \ldots, M_n^K)_{n=1}^\infty$, with each element $(M_n^1, \ldots, M_n^K)$ specifying a model for each agent, and with these models growing (at least weakly) larger along the sequence.

Proposition 7 *Consider a sequence of $(M_n^1, \ldots, M_n^K)_{n=1}^\infty$, with $M_n^k \subset M_{n+1}^k$. Suppose that for all i and j, $\cup_{n=0}^\infty M_n^i = \cup_{n=0}^\infty M_n^j$. Then for every $\delta > 0$, there exists N_δ such that for any accompanying sequence of information $(I_n^1, \ldots, I_n^K)_{n=1}^\infty$, for all $n > N_\delta$, with probability at least $1 - \delta$, the limit beliefs of all agents are within δ of each other.*

The requirement that $\cup_n M^i = \cup_n M^j$ ensures that any variable that eventually appears in i's model also eventually appears in j's model (and that if the models are constant, then they agree). Notice that we do *not* require that F is completely determined by the variables in $\cup_n M_n^i$. It may be that no agent i ever acquires a complete understanding of the event F.

The conditions of the proposition are consistent with the agents having an arbitrarily small, even zero, proportion of their models in common, for every term in the sequence.[20]

The proposition indicates that as the agents' models come to share an increasing common component, the agents come to share common beliefs. If the agents have access to little information, these beliefs will be rather uninformative, while if the agents have access to ample information, these beliefs will be close to those of an omniscient oracle.

Appendix A.2 contains the proof. The idea behind the proof is that even though f may depend on an infinite number of variables (in the sense that there is no finite set of variables L such that f is measurable with respect to X^L, there are a limited number of variables that can be "important" in determining whether F occurs. Eventually, the important variables are either included in everyone's model, ensuring that each agent makes use of the information that is important to the public oracle, or such variables never appear in the agents' models, in which case the public oracle also lacks access to such information. In either case, the agents' beliefs must grow close to those of the public oracle.

[20]For example, agent 1's n^{th} model may be $\{1, 2, 3, \ldots, n\} \cup \{1, 3, 5, 7, \ldots\}$ and agent 2's n^{th} model may be $\{1, 2, 3, \ldots, n\} \cup \{2, 4, 6, 8, \ldots\}$.

5.3 Dispersed Information

We now investigate the sense in which the limiting beliefs of agents with diverse models and dispersed information approximate the belief of the omniscient oracle as the number of agents grows.

Our first step toward an information-aggregation result is to show that if an agent's model is sufficiently sophisticated, than that agent's full-information belief cannot be too far from the omniscient belief. For an arbitrary subset $Z \subseteq N$, let $\mathcal{G}^Z$ denote the σ-algebra induces by Z.[21] Appendix A.3 contains the proof of the following:

Lemma 2 *For all $\zeta > 0$, there exists a finite set $Z_\zeta \subseteq \mathbb{N}$ such for all σ-algebras $\mathcal{H}$,*

$$\rho\left\{\omega : \left|\mathbb{E}[f|\mathcal{G}^{Z_\zeta}, \mathcal{H}](\omega) - f(\omega)\right| < \zeta\right\} \geq 1 - \zeta.$$

Lemma 2 implies that if an agent's model includes the variables Z_ζ, not only is the agent's full-information belief close to the omniscient belief, but there is no additional information that can change the agent's belief significantly. It is not surprising that if we put enough of the right variables into an agent's model, then their full-information belief will be close to the omniscient oracular belief. The more delicate part of Lemma 2 lies in showing that nothing else the agent could possibly include in her model can drive the agent outside the ζ-margin of error.

Remark 4 (Proposition 7 Redux) Lemma 2 allows a variation on Proposition 7. Any sequence whose agents' models eventually include Z_ζ must eventually have beliefs that are close to those of an public oracle (and hence each other). Lemma 2 plays the role of Lemma A.1 in the proof of Proposition 7, mutatis mutandis. ♦

We need to replace the full-information beliefs of Lemma 2 with limiting beliefs. It is clear that doing so will require some conditions, even if all agents were agent oracles—Geanakoplos and Polemarchakis (1982) and Figure 4 above present examples in which agents' announcements convey no information, despite all agents being agent oracles. The culprit behind the agents' discombobulation in Figure 4 is that their initial announcements pool information, preventing even oracular agents from making use of this information. If we are to even get off the ground, we must ensure some information transmission. Section 3.4 explains why we do not rule out pooling

[21] For example, $\mathcal{G}^{M^i} = \mathcal{M}^i$.

by making the stronger assumption that ρ is generic. It is also worth noting that genericity would not simplify the argument.

Definition 1 *The indicator f is* discernible *if, for any sets $I^1, I^2 \subseteq N$, we have*

$$\mathbb{E}[f|\mathcal{I}^1, \mathcal{I}^2] = \mathbb{E}[f|\sigma(\mathbb{E}[f|\mathcal{I}^1], \mathbb{E}[f|\mathcal{I}^2])], \tag{9}$$

where $\sigma(\mathbb{E}[f|\mathcal{I}^1], \mathbb{E}[f|\mathcal{I}^2])$ is the σ-algebra induced by the announcements of the beliefs $\mathbb{E}[f|\mathcal{I}^1]$ and $\mathbb{E}[f|\mathcal{I}^2]$.

Discernibility requires that the announcements $\mathbb{E}[f|\mathcal{I}^1]$ and $\mathbb{E}[f|\mathcal{I}^2]$ allow an agent oracle to infer the information contained in $I^1 \cup I^2$ that is relevant for determining f (but may not allow the agent oracle to identify $\omega_{I^1 \cup I^2}$, and hence allows the pooling of information that is irrelevant for determining F). For any f, discernibility will approximately hold if I_1 and I_2 are sufficiently large. Discernibility fails in Figure 4.

We note that discernibility extends to finite numbers of sets. The following result is immediate:

Lemma 3 *Suppose f is discernible. Then*

$$\begin{aligned}\mathbb{E}[f|\mathcal{I}^1, \mathcal{I}^2, \mathcal{I}^3] &= \mathbb{E}[f|\sigma(\mathbb{E}[f|\mathcal{I}_1, \mathcal{I}_2], \mathbb{E}[f|\mathcal{I}^3])] \\ &= \mathbb{E}[f|\sigma(\mathbb{E}[f|\mathcal{I}^1], \mathbb{E}[f|\mathcal{I}^2], \mathbb{E}[f|\mathcal{I}^3])].\end{aligned}$$

The first equality is the statement of discernibility, and the second again applies discernibility.[22]

We can hope to replace the full-information beliefs of Lemma 2 with limiting beliefs only if the sequence of announced beliefs conveys sufficient information. Our route to ensuring this is to examine a large group of agents. To make the notion "large" precise, we construct a sequence of groups of agents whose models and information are randomly determined, designed to capture our interest in agents with diverse models and dispersed information.

Definition 2 *A sequence of groups of agents (with the number of agents going to infinity) is* canonical *if there is a probability measure λ and a family of measures $\{\mu_M : M \subset \mathbb{N}\}$ satisfying*

[22]The discernibility condition (9) implies that the expectation $\mathbb{E}[f|\mathcal{I}^1, \mathcal{I}^2]$ of f conditioning on the σ-algebra $\sigma(\mathcal{I}^1, \mathcal{I}^2)$ is measurable with respect to the coarser σ-algebra $\sigma(\mathbb{E}[f|\mathcal{I}^1], \mathbb{E}[f|\mathcal{I}^2])$, and hence the random variable $\mathbb{E}[f|\mathcal{I}^1, \mathcal{I}^2]$ is measurable with respect to $\sigma(\mathbb{E}[f|\mathcal{I}^1], \mathbb{E}[f|\mathcal{I}^2])$, or $\sigma(\mathbb{E}[f|\mathcal{I}^1, \mathcal{I}^2]) = \sigma(\mathbb{E}[f|\mathcal{I}^1], \mathbb{E}[f|\mathcal{I}^2])$.

1. *λ is a full support measure on the space of all finite models,*

2. *there is a $\Gamma \in \mathbb{N}$ such that for each finite model $M \subset \mathbb{N}$, μ_M is a full support probability distribution over the subsets of M with at most Γ elements, and*

3. *in each group, each agent's model and information are determined by independent draws from λ and μ_M.*

The agents in a canonical sequence will exhibit a wide variety of models, reflecting the full-support measure λ. Importantly, sophisticated models arise with positive probability, and so in large groups, with very high probability there will be very sophisticated agents. At the same time, the information of each individual may be paltry (Γ may be small), and hence individuals may lack the requisite information to form interim beliefs that are close to full information beliefs.

In a sufficiently large group drawn from a canonical sequence, any agent whose model contains the variables Z_ζ (from Lemma 2) will, with high probability, effectively learn F from the sequence of announced beliefs. Appendix A.4 proves:

Lemma 4 *Suppose f is discernible and fix a canonical sequence of groups of agents. For all $\varepsilon > 0$, there exists an N_ε such that for all $n > N_\varepsilon$, every agent with a finite model M containing $Z_{\varepsilon^2/8}$ has, with probability at least $1 - \varepsilon$, a limiting belief within ε of the omniscient oracle.*

The argument first identifies a set of agents $K^{\mathbb{I}}$ whose interim beliefs, given discernibility, allow any agent whose model is precisely $Z_{\varepsilon^2/8}$ to form full-information beliefs. In a sufficiently large set of agents drawn from a canonical sequence, such a set $K^{\mathbb{I}}$ will exist with high probability. We then argue that the limiting beliefs of an agent whose model *contains* $Z_{\varepsilon^2/8}$ must be close to that agent's full information beliefs, and use Lemma 2 to argue that the latter must be close to the beliefs of the omniscient oracle. The triangle inequality (with the help of our choice of the parameter $\varepsilon^2/8$) then ensures that the agents' limiting beliefs are close to those of the omniscient oracle.

The key to bringing the average of the agents' limiting beliefs close to those of the limiting oracle is to have the measure λ insert enough sufficiently sophisticated agents into the groups in the canonical sequence. Let $\Lambda(\zeta)$ be the probability that a model drawn according to λ does *not* contain Z_ζ.

Proposition 8 *Suppose f is discernible and fix a canonical sequence of groups of agents. For all $\varepsilon > 0$, there exists N^*_ε such that for all $n > N^*_\varepsilon$,* *the probability the average of beliefs is within* $3\varepsilon + \Lambda(\varepsilon^2/8)$ *of the omniscient belief is at least* $1-\varepsilon$*.*

Proof. We can choose $N^*_\varepsilon \geq N_\varepsilon$ so that for all $n > N^*_\varepsilon$, with probability at least $1-\varepsilon$ the proportion of agents whose models include $Z_{\varepsilon^2/8}$ will be at least $1 - \Lambda(\varepsilon^2/8) - \varepsilon$ and the $K^{\mathbb{I}}$-agents are present. Conditional on this event, the difference between the average belief and the belief of an omniscient oracle is then at most

$$(1 - \Lambda(\varepsilon^2/8) - \varepsilon)((1-\varepsilon)\varepsilon + \varepsilon) + \Lambda(\varepsilon^2/8) + \varepsilon \leq 3\varepsilon + \Lambda(\varepsilon^2/8).$$

■

The discernibility requirement in Lemma 4 gets the information-revelation process off the ground in a particularly brutal way, ensuring that for any information set I, the first announcement by an agent i with $M^i = I^i = I$ reveals all the relevant information contained in I. It would suffice that *limiting* beliefs reveal such information, and we could formulate discernibility in terms of limiting beliefs, at the costs of greater complexity and pushing the assumption further away from the fundamentals of the problem. Alternatively, discernibility is more demanding when applied to small information sets. We could work with a version of γ-discernibility that applies the discernibility requirement only to sets with at least γ elements, but then would need to place stronger requirements on the presence of agents with larger information sets.

Proposition 8 establishes that if *enough* agents have large enough (i.e., containing $Z_{\varepsilon^2/8}$) models, then it is very likely that the average belief will be "close" to that of an omniscient oracle. How close? This depends on the characteristics of the canonical sequence. If the measure λ puts sufficient probability on large models, then $\Lambda(\varepsilon^2/8)$ will be small. This is the case of a sophisticated crowd, whose members entertain sufficiently nuanced models of the forces determining the event F that the revelation of information can bring their beliefs close to those of the omniscient oracle. If λ concentrates its probability of small models, $\Lambda(\varepsilon^2/8)$ will be large, giving us a dogmatic crowd whose beliefs are impervious to the onslaught of overwhelming information.

Proposition 8 may appear to be nothing more than the statement that if enough people get it right, then the average will be about right. The more delicate part of the argument involves showing the beliefs of those who would otherwise "get it right" are not disrupted by the presence of some agents with bizarre models. This requires a uniformity condition across σ-algebras (in Lemma 2). The average belief is then driven toward the omniscient belief, not by having those who get it right convincing or converting those who are confused, but by having the former swamp the latter. Notice, however, that for this to happen there must be sufficiently many agents with sufficiently large and common models. There is no similar requirement on the commonality of information. Interactions can indeed effectively aggregate dispersed information, *if* the agents have a sufficiently common understanding of the meaning of that information.

6 Discussion and Related Literature

People cannot help but reason via models, different people use different models, and yet people learn from one another. We provide a tractable description of model-based reasoning that describes how people can learn from each other in a disciplined manner. Our analysis confirms that we should not expect people to agree after exchanging opinions, nor should we expect the average opinion of a group to be particularly accurate. But our analysis elucidates the sense in which interactions can effectively aggregate information and generate approximate consensus, no matter how dispersed is information and no matter what idiosyncracies various models contain, as long as the models share a sufficient common core.

A key element of our description is the notion of a model-based reasoner, who forms beliefs about the occurrence of an event using only the variables in her model. Such a reasoner uses information not in her model only to infer those variables in her model that she does not know. In this sense, a model-based reasoner uses a particularly simple Bayesian network. A Bayesian network is a directed acyclic graph, with the arrows capturing conditional dependencies. A model-based reasoner effectively reasons using a Bayesian network with two maximal cliques, one capturing the dependence of the event F on the variables in agent i's model M^i, and another capturing the possible dependence of variables in M^i (but *not* F) on variables not in M^i. See Pearl (2009) for an exposition of Bayesian networks, Spiegler (2016) for an application to decision making, and Spiegler (2019) for a review of recent work. That work focuses the behavioral implications of misinterpreted correlations and causations, while our focus is on inference and information aggregation.

Our work shares with Jehiel (2005) the idea that agents will simplify the description of the world by aggregating states. In particular, we can view an element of agent i's model partition $\{\omega_{M^i}\} \times X^{-M^i}$ as an analogy class.

Jehiel (2005) introduced the idea of an analogy class and used it to define an equilibrium notion for games of perfect information, and Jehiel and Koessler (2008) discuss an extension to games of incomplete information. Eyster and Piccione (2013) examine financial traders who partition the state space into analogy classes, and then form expectations on these analogy classes as in our (1)–(2).

The idea of the "wisdom of the crowd" has attracted considerable attention (e.g., Surowiecki (2004), Wolfers and Zitzewitz (2004), Page (2017)). One can view our model of information exchange as a stylized model of the process by which the crowd grows wise. Arieli, Babichenko, and Smorodinsky (2018) take a different perspective in a similar setting, examining a model in which the members of a crowd of agents receive signals, update their beliefs, and then (once) report their beliefs. The question is when an observer can infer the identity of the underlying state, despite knowing nothing about the agents' signal structures. The (rough) answer is that even if the crowd is arbitrarily large, no inferences can be drawn unless a signal drives a posterior belief to either 0 or 1. The flavor of this result is reminiscent of our observations that (Proposition 1) beliefs of 0 or 1 must match those of an omniscient oracle and that (from Appendix B.1) when beliefs are interior, Bayes' rule places very little discipline on models in the absence of a common prior.

Analyses of information exchange inevitably takes place in the shadow of Aumann's agreeing-to-disagree theorem (Aumann (1976)) and Milgrom and Stokey's no-trade theorem (Milgrom and Stokey (1982)). An extensive literature has arisen motivated by a desire to break these results by relaxing various assumptions underlying them. Perhaps the most obvious approach is to allow heterogenous priors (e.g., Morris (1994)).[23] We believe that agents will often hold different prior beliefs. For much the same reasons, we believe that agents will hold different models. We explain in Appendix B.1 that model-based reasoning imposes more discipline than simply allowing priors to differ.[24]

[23]Disagreement can also arise, reflecting considerations similar to those that arise with different prior beliefs, when agents have a common prior but nonpartitional information structures (e.g., Geanakoplos (1989) and Brandenburger, Dekel, and Geanakoplos (1992)), while the possibility of noise traders (e.g., Kyle (1985) and Ostrovsky (2012) figures prominently among the many other approaches to disagreement and trade.

[24]Sethi and Yildiz (2016) examine a setting in which agents with different and unknown prior beliefs learn, but often incompletely, about others' prior beliefs. Acemoglu, Chernozhukov, and Yildiz (2016) in which uncertainty about the signal-generating process causes posterior beliefs to be sensitive to prior beliefs even after receiving an infinite sequence of signals, and so allows limiting beliefs of agents with different prior beliefs to

A Appendices

A.1 Proof of Proposition 6

We first prove a preliminary result that looks obvious, but for the fact that we require uniformity over σ-algebras $\mathcal{H}$.

Lemma A.1 *Suppose $(f_n)_n$ is a sequence of $\mathcal{F}$-measurable functions converging almost surely to the $\mathcal{F}$-measurable function $f^\dagger$. For all $\delta > 0$, there exists a set Ω_δ with $\rho(\Omega_\delta) > 1-\delta$ and an integer N_δ such that for all $n > N_\delta$ and for all σ-algebras $\mathcal{H} \subseteq \mathcal{F}$,*

$$\Big|\mathbb{E}[f^\dagger|\mathcal{H}](\omega) - \mathbb{E}[f_n|\mathcal{H}](\omega)\Big| < \delta \qquad \forall \omega \in \Omega_\delta.$$

Proof. Fix a value $\delta > 0$. Choose λ and ε such that

$$\begin{aligned} \lambda &> 1/\delta \\ \varepsilon(1+\lambda) &< \delta. \end{aligned}$$

By Egorov's theorem, there exists a value N_δ and a set Ω_δ of measure at least $1-\varepsilon$ (which is at least $1-\delta$) with the property that for all $n > N_\delta$, we have

$$|f_n(\omega) - f^\dagger(\omega)| < \varepsilon \qquad \forall \omega \in \Omega_\delta.$$

We now argue that with probability at least $1-\delta$, we have

$$\Big|\mathbb{E}[f_n \mid \mathcal{H}] - \mathbb{E}[f^\dagger \mid \mathcal{H}]\Big| < \delta.$$

Define

$$h(\omega) := \begin{cases} \varepsilon, & \omega \in \Omega_\delta, \\ 1, & \omega \notin \Omega_\delta. \end{cases}$$

Then,

$$|f_n(\omega) - f^\dagger(\omega)| \le h(\omega),$$

and

$$\begin{aligned} \mathbb{E}[h \mid \mathcal{H}] &= \varepsilon \Pr(\Omega_\delta \mid \mathcal{H}) + \Pr(\Omega \setminus \Omega_\delta \mid \mathcal{H}) \\ &\le \varepsilon + \Pr(\Omega \setminus \Omega_\delta \mid \mathcal{H}). \end{aligned}$$

---

differ.

Let $A := \{\omega : \Pr(\Omega \setminus \Omega_\delta \mid \mathcal{H})(\omega) > \lambda\varepsilon\}$. Then $A \in \mathcal{H}$ and so

$$\begin{aligned}\lambda\varepsilon \Pr A &< \int_A \mathbb{E}[\chi_{\Omega\setminus\Omega_\delta} \mid \mathcal{H}] d\rho \\ &= \int_A \chi_{\Omega\setminus\Omega_\delta} d\rho \\ &\leq \varepsilon,\end{aligned}$$

and so

$$\Pr\{\omega : \Pr(\chi_{\Omega\setminus\Omega_\delta} \mid \mathcal{H})(\omega) > \lambda\varepsilon\} \leq 1/\lambda,$$

and hence we have

$$\Pr\{\omega : \varepsilon + \Pr(\Omega \setminus \Omega_\delta \mid \mathcal{H})(\omega) < (1+\lambda)\varepsilon\} > 1 - 1/\lambda.$$

Invoking our conditions on λ and ε yields

$$\Pr\{\omega : \mathbb{E}[h \mid \mathcal{H}](\omega) < \delta\} > 1 - \delta,$$

and since

$$\left|\mathbb{E}[f_n \mid \mathcal{H}] - \mathbb{E}[f^\dagger \mid \mathcal{H}]\right| \leq \mathbb{E}\left[|f_n - f^\dagger| \mid \mathcal{H}\right],$$

we have the desired inequality.

∎

Proof of Proposition 6. Suppose first that the coefficient of correlation between f^i and f^j equals 1. Then $f^j - \mathbb{E}f = \alpha(f^i - \mathbb{E}f)$ ρ-almost surely for some constant $\alpha > 0$ (recall (3)). Suppose f^i is not constant (if it were, the result is trivial), so that for some $x > 0$, $\mathbb{E}f + x$ is in the support of f^i.

We now argue that $\alpha = 1$. En route to a contradiction, suppose $\alpha > 1$ (a similar argument rules out $\alpha < 1$). Fix $\varepsilon > 0$ so that $\alpha(x - \varepsilon) > x$ and set $B(x) := \{\omega : x - \varepsilon \leq f^i(\omega) - \mathbb{E}f \leq x\}$. We may assume $\rho(B(x)) > 0$ (if not, marginally increasing the value of x yields a positive measure set). Then, for $y = \alpha x$ and $B'(y) := \{\omega : y - \alpha\epsilon \leq f^j(\omega) - \mathbb{E}f \leq y\}$, we have $\rho(B(x)\Delta B'(y)) = 0$.[25] From (2), since $B(x) \in \mathcal{M}^i$ and $B'(y) \in \mathcal{M}^j$, we

[25] The notation $A\Delta B := (A \setminus B) \cup (B \setminus A)$ is the symmetric difference of the two sets A and B.

then have

$$\begin{aligned} x\rho(B(x)) &\geq \int_{B(x)} (f^i(\omega) - \mathbb{E}f)\ d\rho \\ &= \int_{B(x)} (f(\omega) - \mathbb{E}f) d\rho \\ &= \int_{B'(y)} (f(\omega) - \mathbb{E}f) d\rho \\ &= \int_{B'(y)} (f^j(\omega) - \mathbb{E}f) d\rho \geq (y - \alpha\varepsilon)\rho(B'(y)), \end{aligned}$$

and so $x \geq \alpha(x - \varepsilon)$, a contradiction.

From Proposition 1.4, we have that agent i's limiting belief $\mathbb{E}[f^i|\mathcal{I}^i, \mathcal{B}_\infty]$ equals $\mathbb{E}[f^i|\mathcal{B}_\infty]$ which equals $\mathbb{E}[f^j|\mathcal{B}_\infty]$, and so agent i and j's limiting beliefs agree on any sequence of announced posteriors.

Turning to the approximation, it is enough to prove that we can make $\mathbb{E}|f^i - f^j|$ arbitrarily small by choosing η sufficiently small, where $1 - \eta$ is the lower bound on the correlation. We prove the latter by contradiction. If not, then there exists $\varepsilon > 0$ such that for all $n > 0$ there exists f^j_n such that the correlation between f^i and f^j_n is at least $1 - 1/n$ and yet $\mathbb{E}|f^i - f^j_n| > \varepsilon$.

Define $X := f^i - \mathbb{E}f$ and $Y_n := f^j_n - \mathbb{E}f$. Then,

$$\begin{aligned} \mathbb{E}[Y_n\mathbb{E}(X^2) - X\mathbb{E}(XY_n)]^2 &= \mathbb{E}(X^2)[\mathbb{E}(X^2)\mathbb{E}(Y_n^2) - \mathbb{E}(XY_n)^2] \\ &\leq \mathbb{E}(X^2)[\mathbb{E}(X^2)\mathbb{E}(Y_n^2) - (1 - 1/n)^2\mathbb{E}(X^2)\mathbb{E}(Y_n^2)] \\ &= (\mathbb{E}X^2)^2\mathbb{E}(Y_n^2)[1 - 1 + 2/n - 1/n^2], \end{aligned}$$

and so $Y_n\mathbb{E}(X^2) - X\mathbb{E}(XY_n)$ converges in mean square to 0 as $n \to \infty$ (since $\mathbb{E}(Y_n^2)$ is bounded above by $\frac{1}{4}$). If $(Y_n)_n$ (or any subsequence) has a limit in mean square (and so a limit in mean), then that limit must equal X (for the reasons above). We will show that every subsequence has a convergent subsubsequence, which implies that the original sequence converges to X.

We use n to index an arbitrary subsequence and let $\alpha_n := \mathbb{E}XY_n/\mathbb{E}(X^2)$, so that $Y_n - \alpha_n X$ converges to 0 in mean square. We claim that (α_n) has a convergent subsequence. For, if not, then $|\alpha_n| \to \infty$, which implies $\mathbb{E}(Y_n^2) \to \infty$, which is impossible.

Suppose (α_{n_k}) converges to some α. Then,

$$0 \leq [\mathbb{E}(Y_{n_k} - \alpha X)^2]^{\frac{1}{2}} \leq [\mathbb{E}(Y_{n_k} - \alpha_{n_k}X)^2]^{\frac{1}{2}} + [\mathbb{E}(\alpha - \alpha_{n_k})^2X^2]^{\frac{1}{2}} \to 0,$$

and so Y_n converges in mean square to αX, and so $\alpha = 1$.

It remains to argue that for n sufficiently large, with probability at least $1-\varepsilon$,

$$\left|\mathbb{E}[Y_n|\mathcal{I}^j, \mathcal{B}_\infty] - \mathbb{E}[X|\mathcal{I}^i, \mathcal{B}_\infty]\right| < \varepsilon.$$

By Proposition 1.5, this inequality can be rewritten as

$$|\mathbb{E}[Y_n|\mathcal{B}_\infty] - \mathbb{E}[X|\mathcal{B}_\infty]| < \varepsilon.$$

Since every subsequence of $(Y_n)_n$ has a sub-subsequence almost surely converging to X, the desired result is implied by Lemma A.1. ∎

A.2 Proof of Proposition 7

Proof. Fix a value $\delta > 0$. Define

$$\mathcal{M}_\infty := \sigma(M_1^i, M_2^i, \dots)$$

(which is, by assumption, independent of i) and set $\hat{f} := \mathbb{E}[f \mid \mathcal{M}_\infty]$.

Agent i's theory under her n^{th} model is given by

$$f_n^i = \mathbb{E}[f \mid \mathcal{M}_n^i] = \mathbb{E}[\hat{f} \mid \mathcal{M}_n^i]$$

(where the second equality follows from $\mathcal{M}_n^i$ being coarser than $\mathcal{M}_\infty$ and the law of iterated expectations). Since $(\mathcal{M}_n^i)_n$ is a filtration, with limit σ-algebra $\mathcal{M}_\infty$,

$$f_n^i \to \hat{f} \qquad \rho\text{-a.s.}$$

Our goal is to show that with probability at least $1-\delta$, we have

$$\left|\mathbb{E}[f_n^i \mid \mathcal{I}_n^i, \mathcal{B}_\infty(n)] - \mathbb{E}[f \mid \mathcal{B}_\infty(n)]\right| < \delta,$$

where $\mathcal{B}_\infty(n)$ is the σ-algebra induced by the sequence of publicly announced beliefs for the n^{th} term in the sequence.

By Lemma A.1, with probability at least $1-\delta$, we have

$$\left|\mathbb{E}[f_n^i \mid \mathcal{B}_\infty(n)] - \mathbb{E}[\hat{f} \mid \mathcal{B}_\infty(n)]\right| < \delta.$$

By Proposition 1.5,

$$\mathbb{E}[f_n^i \mid \mathcal{I}_n^i, \mathcal{B}_\infty(n)] = \mathbb{E}[f_n^i \mid \mathcal{B}_\infty(n)],$$

and so with probability at least $1-\delta$, we have

$$\Big|\mathbb{E}[f_n^i \mid \mathcal{I}_n^i, \mathcal{B}_\infty(n)] - \mathbb{E}[\hat{f} \mid \mathcal{B}_\infty(n)]\Big| < \delta.$$

Finally, since $\mathcal{B}_\infty(n)$ is coarser than $\mathcal{M}_\infty$ (since $\cup_j \mathcal{I}_n^j \subseteq \cup_j \mathcal{M}_n^j \subseteq \mathcal{M}_\infty$) and $\hat{f} := \mathbb{E}[f \mid \mathcal{M}_\infty]$, we have that with probability at least $1-\delta$,

$$\big|\mathbb{E}[f_n^i \mid \mathcal{I}_n^i, \mathcal{B}_\infty(n)] - \mathbb{E}[f \mid \mathcal{B}_\infty(n)]\big| < \delta.$$

■

A.3 Proof of Lemma 2

Recall that $\mathcal{G}^t$ is the σ-algebra generated by the t coordinate of $\Omega = \{0,1\}^N$, and set $\mathcal{F}^t := \sigma(\mathcal{G}^1, \dots, \mathcal{G}^t)$. Since f is measurable with respect to $\sigma(\mathcal{G}^1, \mathcal{G}^2, \dots)$, we have

$$\mathbb{E}[f \mid \mathcal{F}^t] \to f \qquad \text{a.s. } [\rho].$$

Egorov's theorem implies that for all $\zeta > 0$, there exists T_ζ and an event Ω_ζ, with $\rho(\Omega_\zeta) \geq 1 - \zeta/4$, for which

$$\big|\mathbb{E}[f \mid \mathcal{F}^t](\omega) - f(\omega)\big| < \zeta^2/4 \qquad \forall t \geq T_\zeta^*, \forall \omega \in \Omega_\zeta. \tag{A.1}$$

Set $Z_\zeta := \{1, \dots, T_\zeta\}$, so that $\mathcal{G}^{Z_\zeta} = \mathcal{F}^{T_\zeta}$.

Claim A.1 *On a full probability subset of* $\Omega_\zeta \cap F$,

$$\Pr\{\mathbb{E}[f \mid \mathcal{G}^{Z_\zeta}, \mathcal{H}] \leq 1 - \zeta \mid \mathcal{G}^{Z_\zeta}\} < \zeta/4 \tag{A.2}$$

and on a full probability subset of $\Omega_\zeta \setminus F$,

$$\Pr\{\mathbb{E}[f \mid \mathcal{G}^{Z_\zeta}, \mathcal{H}] \geq \zeta \mid \mathcal{G}^{Z_\zeta}\} < \zeta/4. \tag{A.3}$$

Proof. We prove (A.2); the proof of (A.3) follows similar lines. Define $g^\dagger(\omega) := \mathbb{E}[f \mid \mathcal{G}^{Z_\zeta}, \mathcal{H}](\omega)$, and $g(\omega) := \Pr\{g^\dagger \leq 1 - \zeta \mid \mathcal{G}^{Z_\zeta}\}(\omega)$. Note that g is only measurable with respect to $\mathcal{G}^{Z_\zeta}$ (so in particular, the inequality in (A.2) is measurable with respect to $\mathcal{G}^{Z_\zeta}$), while $g^\dagger$ is measurable with respect to the finer $\sigma(\mathcal{G}^{Z_\zeta}, \mathcal{H})$.

Recalling that f is the indicator function of the event F, for $\omega \in \Omega_\zeta \cap F$, (A.1) is

$$1 - \zeta^2/4 < \mathbb{E}[f \mid \mathcal{G}^{Z_\zeta}](\omega),$$

and so (A.2) is implied by for ρ-almost all $\omega \in \Omega_\zeta \cap F$,

$$\mathbb{E}[f \mid \mathcal{G}^{Z_\zeta}](\omega)] \leq 1 - \zeta g(\omega).$$

Since the left and right sides of the above inequality are measurable with respect to $\mathcal{G}^{Z_\zeta}$, if the inequality does not hold, there is a positive ρ-probability event $B \in \mathcal{G}^{Z_\zeta}$ such that,

$$\mathbb{E}[f \mid \mathcal{G}^{Z_\zeta}](\omega) > 1 - \zeta g(\omega) \qquad \forall \omega \in B. \tag{A.4}$$

Since $B \in \mathcal{G}^{Z_\zeta}$, where χ_A is the indicator function of the event A, and the first and last (respectively, third) equalities hold because the integrating events are measurable with respect to $\mathcal{G}^{Z_\zeta}$ (resp., $\sigma(\mathcal{G}^{Z_\zeta}, \mathcal{H})$),

$$\begin{aligned}
\int_B \mathbb{E}[f \mid \mathcal{G}^{Z_\zeta}] d\rho &= \int_B f d\rho \\
&= \int_{B \cap \{g^\dagger \leq 1-\zeta\}} f d\rho + \int_{B \cap \{g^\dagger > 1-\zeta\}} f d\rho \\
&= \int_{B \cap \{g^\dagger \leq 1-\zeta\}} g^\dagger d\rho + \int_{B \cap \{g^\dagger > 1-\zeta\}} g^\dagger d\rho \\
&\leq (1-\zeta) \int_B \chi_{\{g^\dagger \leq 1-\zeta\}} d\rho + \int_B 1 - \chi_{\{g^\dagger \leq 1-\zeta\}} d\rho \\
&= \int_B 1 - \zeta \chi_{\{g^\dagger \leq 1-\zeta\}} d\rho \\
&= \int_B 1 - \zeta g d\rho,
\end{aligned}$$

contradicting (A.4). □

Defining

$$B' := \left\{\omega : \left|\mathbb{E}[f|\mathcal{G}^{Z_\zeta}, \mathcal{H}](\omega) - f(\omega)\right| \geq \zeta\right\}$$

and

$$F' := \{\omega : \Pr\{\mathbb{E}[f \mid \mathcal{G}^{Z_\zeta}, \mathcal{H}] \leq 1 - \zeta \mid \mathcal{G}^{Z_\zeta}\}(\omega) < \zeta/4\}$$

we have (since, up to a zero probability event, $\Omega_\zeta \cap F \subseteq F'$ and $F' \in \mathcal{G}^{Z_\zeta}$)

$$\begin{aligned}
\Pr(B' \cap F) &= \Pr\left\{\{\mathbb{E}[f|\mathcal{G}^{Z_\zeta},\mathcal{H}](\omega) \le 1-\zeta\} \cap F\right\} \\
&\le \Pr\left\{\{\mathbb{E}[f|\mathcal{G}^{Z_\zeta},\mathcal{H}](\omega) \le 1-\zeta\} \cap F'\right\} \\
&= \mathbb{E}\big[\mathbb{E}[\chi_{\{\mathbb{E}[f|\mathcal{G}^{Z_\zeta},\mathcal{H}](\omega)\le 1-\zeta\}\cap F'} \mid \mathcal{G}^{Z_\zeta}]\big] \\
&= \mathbb{E}\big[\mathbb{E}[\chi_{\{\mathbb{E}[f|\mathcal{G}^{Z_\zeta},\mathcal{H}](\omega)\le 1-\zeta\}} \mid \mathcal{G}^{Z_\zeta}]\chi_{F'}\big] \\
&\le \int_{\Omega_\zeta} \Pr\{\mathbb{E}[f|\mathcal{G}^{Z_\zeta},\mathcal{H}](\omega) \le 1-\zeta\} \mid \mathcal{G}^{Z_\zeta}]\chi_{F'} d\rho + \rho(\Omega \setminus \Omega_\zeta) \\
&\le \zeta/4 + \zeta/4.
\end{aligned}$$

Applying a similar argument to $B' \setminus F$, we obtain

$$\Pr(B' \setminus F) \le \zeta/2,$$

so that $\rho(B') \le \zeta$.

A.4 Proof of Lemma 4

Proof. Denote by $\mathbb{I}$ a collection of subsets of $\mathbb{N}$ satisfying $Z_\zeta = \cup\mathbb{I}$ such that no set in $\mathbb{I}$ has more than Γ elements. Choose N_ε sufficiently large that for $N > N_\varepsilon$ with probability at least $\sqrt{1-\varepsilon}$, for every set $I^i \in \mathbb{I}$, there is an agent whose model and information set consist precisely of that set. Denote this set of agents by $K^{\mathbb{I}}$, and suppose this set is present.

Consider now an agent j whose model contains Z_ζ. We now argue that

$$\rho\left\{\omega : \left|\mathbb{E}[f^j \mid \mathcal{B}_\infty](\omega) - f(\omega)\right| < \varepsilon\right\} \ge \sqrt{1-\varepsilon}. \tag{A.5}$$

Observe first that since for $i \in K^{\mathbb{I}}$, agent i's interim belief is $\mathbb{E}[f|\mathcal{G}^{I^i}]$, we have

$$\begin{aligned}
\mathbb{E}[f \mid \mathcal{B}_\infty] &= \mathbb{E}\left[f \mid \mathbb{E}[f \mid \mathcal{B}_\infty], (\mathbb{E}[f \mid \mathcal{G}^{I^i}])_{i\in K^{\mathbb{I}}}\right] \\
&= \mathbb{E}\left[f \mid \mathbb{E}[f \mid \mathcal{B}_\infty], \mathbb{E}[f \mid \mathcal{G}^{Z_\zeta}]\right] \\
&= \mathbb{E}[f \mid \mathcal{B}_\infty, \mathcal{G}^{Z_\zeta}],
\end{aligned}$$

where the second equality following from the discernability of f.

From Lemma 2, with probability at least $1-\zeta$,

$$\left|\mathbb{E}[f \mid \mathcal{B}_\infty, \mathcal{G}^{Z_\zeta}] - f\right| < \zeta < \frac{\varepsilon}{2}. \tag{A.6}$$

We now bound the first term in the triangle inequality,

$$\left|\mathbb{E}[f^j \mid \mathcal{B}_\infty] - \mathbb{E}[f \mid \mathcal{B}_\infty]\right| + \left|\mathbb{E}[f \mid \mathcal{B}_\infty, \mathcal{G}^{Z_\zeta}] - f\right| \ge \left|\mathbb{E}[f^j \mid \mathcal{B}_\infty] - f\right|. \tag{A.7}$$

Claim A.2 *With probability at least* $1 - 4\zeta/\varepsilon$,

$$\left|\mathbb{E}[f^j \mid \mathcal{B}_\infty] - \mathbb{E}[f \mid \mathcal{B}_\infty]\right| < \frac{\varepsilon}{2}.$$

Proof. From Lemma 2, there is a set Ω_ζ, $\rho(\Omega_\zeta) > 1 - \zeta$ such that on Ω_ζ,

$$\left|f^j - f\right| < \zeta.$$

Let

$$B := \left\{\left|f^j \mid \mathcal{B}_\infty] - \mathbb{E}[f \mid \mathcal{B}_\infty]\right| \geq \varepsilon/2\right\}.$$

Then,

$$\int_B \left|\mathbb{E}[f^j \mid \mathcal{B}_\infty] - \mathbb{E}[f \mid \mathcal{B}_\infty]\right| \geq \frac{\varepsilon\rho(B)}{2},$$

But since $B \in \mathcal{B}_\infty$,

$$\begin{aligned}\int_B \left|\mathbb{E}[f^j \mid \mathcal{B}_\infty] - \mathbb{E}[f \mid \mathcal{B}_\infty]\right| &= \int_B \left|f^j - f\right| \\ &< \zeta\rho(B \cap \Omega_\zeta) + \rho(B \setminus \Omega_\zeta) \\ &\leq 2\zeta,\end{aligned}$$

and so

$$\rho(B) < \frac{4\zeta}{\varepsilon}.$$

□

From (A.6) and the claim, we have that the right side of (A.7) is bounded above by ε with probability at least

$$1 - \zeta - 4\zeta/\varepsilon > \sqrt{1 - \varepsilon},$$

where the inequality follows from $\zeta = \varepsilon^2/8$.

Since the realization of uncertainty under ρ is independent of the determination of agents' information and models, the probability that the set $K^{\mathbb{I}}$ of agents is present *and* (A.5) holds is $\sqrt{1-\varepsilon}\sqrt{1-\varepsilon} = 1 - \varepsilon$. ■

B Online Appendices

B.1 What's Wrong with Different Priors?

Could it be that our analysis of model-based reasoning is simply a repackaged version of allowing agents to hold different priors?

The starkest difference is that models with different prior beliefs impose virtually no discipline on the relationship of the beliefs of different agents, and hence on the "collective" beliefs of the agents.[26] In contrast, model-based reasoning ensures that agents' beliefs about the events they deem relevant are anchored to the data. This imposes restrictions on the beliefs of individual agents as well as restrictions on how the beliefs of various agents can differ.

It is a common characterization of Bayesian updating that (under natural conditions) at least eventually "the data swamps the prior." This suggests that the discordance allowed by differing priors should be only temporary, with the data eventually imposing as much discipline on a group of agents with different priors as it does on a group of model-based reasoners. To investigate this, we examine a sequence in which agents receive increasing amounts of information. In order to focus clearly in the discipline imposed on beliefs by this information, we assume the agents have common information. In particular, let $(I_n)_{n=0}^{\infty}$ be an increasing sequence of subsets of $\mathbb{N}$. We consider a sequence in which every agent's information set I_n^i in the n^{th} term is given by I_n.

We begin with a model of different priors, holding fixed the other aspects of agents' models. Suppose each agent has the correct state space and description f (i.e., is an oracle), but we place no restrictions on the priors ρ^i, and in particular no restrictions on how these priors may differ across agents.

Given the sequence, let $(\beta_{\infty}^{i,n})_{i=1,n=0}^{K\ \ \infty}$ be the sequence of induced limiting beliefs, for each agent, about the event F. We now argue that once we allow priors to differ, there are few restrictions placed on the sequence of limit posteriors $(\beta_{\infty}^{i,n})_{i=1,n=1}^{K,\ \infty}$, even though the agents are oracles.

Of course, the agents' limit posteriors are not completely arbitrary, as the mere fact that they are derived from Bayes' rule imposes some restrictions. Say that the sequence $(\beta_{\infty}^{i,n})_{i=1,n=0}^{K\ \ \infty}$ has the *martingale property* if, for any agent i and ω_{I_n}, there exists $\omega_{I_{n+1}}$ consistent with ω_{I_n} with

$$\beta_{\infty}^{i,n+1}(\omega_{I_{n+1}}) < \beta_{\infty}^{i,n}(\omega_{I_n}), \tag{B.1}$$

[26] Di Tillio, Lehrer, Polemarchakis, and Samet (2019) come to a similar conclusion.

if and only if there also exists $\omega'_{I_{n+1}}$ consistent with ω_{I_n} with

$$\beta_\infty^{i,n+1}(\omega'_{I_{n+1}}) > \beta_\infty^{i,n}(\omega_{I_n}). \tag{B.2}$$

Intuitively, an agent can receive encouraging news if and only if it is also possible for the agent to receive discouraging news. Note that this implies that zero and unitary beliefs are absorbing.

We also impose minimal consistency with f. The consistency requirement is the following, where the antecedents should be interpreted as the joint hypothesis that the limit exists and has the indicated sign, and $[\omega_{I_n}]$ is the cylinder set given by $\{\omega_{I_n}, \omega_{-I_n}\}$,

$$\lim_n \beta_\infty^{i,n}(\omega_{I_n}) > 0 \quad \implies \quad \exists \omega \in [\omega_{\cup_n I_n}] \text{ } s.t. \text{ } f(\omega) = 1 \tag{B.3}$$

$$\text{and } \lim_n \beta_\infty^{i,n}(\omega_{I_n}) < 1 \quad \implies \quad \exists \omega \in [\omega_{\cup_n I_n}] \text{ } s.t. \text{ } f(\omega) = 0. \tag{B.4}$$

Requirements (B.3) and (B.4) are the only ones that connect the event F with agent beliefs. Without them, there is nothing precluding an agent from, for example, assigning positive probability to F on the basis of some information $\omega_{\cup_n I_n}$ when F is inconsistent with that information. If that were to happen, there is clearly no hope for $\beta_\infty^{i,n}(\omega_{I_n}) = \mathbb{E}_{\rho^i}[f \mid \omega_{I_n}]$.

Proposition B.1 *Consider a sequence of groups of agent oracles indexed by $n = 0, \ldots,$ with each agent's information set in group n given by I_n, where the sequence $(I_n)_{n=0}^\infty$ is increasing. Suppose the sequences $(\beta_\infty^{i,n})_{i=1,n=0}^{K,\ \infty}$ satisfy the martingale property and* (B.3) *and* (B.4)*. Then there exists a vector of prior beliefs $(\rho^1, \ldots, \rho^K)$ generating the limiting posterior beliefs $(\beta_\infty^{i,n})_{i=1,n=0}^{K,\ \infty}$, i.e., $\beta_\infty^{i,n}(\omega_{I_n}) = \mathbb{E}_{\rho^i}[f \mid \omega_{I_n}]$.*

Before proving this result, we make three observations. First, if $\cup_{n=0}^\infty I_n = \Omega$, then since beliefs are a martingale, $\beta_\infty^{i,n} \to f$ ρ^i-almost surely. For states with positive probability under ρ^i and ρ, the data then swamps the prior—agent i attaches probability one to the event that her beliefs about F converge to those of an omniscient oracle. However, the convergence in the previous observation is pointwise, not uniform. That is, for any finite sequence $(\beta_\infty^{i,n})$ satisfying the martingale property given in (B.1)–(B.2), there is a prior rationalizing $(\beta_\infty^{i,n})$. Notice that there need be no connection between such a sequence and the event F. Hence, Bayesian updating from different priors places no restrictions on finite sequences of agents' beliefs, no matter how long. Moreover, if $\cup_{n=0}^\infty I_n \subsetneq \Omega$, then beliefs over states conditional on $\cup_{n=0}^\infty I_n$ are essentially arbitrary, needing only to satisfy the property that the conditional probability of F equals the limit of $\beta_\infty^{i,n}$. Hence, unless we are dealing with a case in which the agents will eventually resolve every vestige of uncertainty, updating places few restrictions on beliefs. If agents with different priors are also sufficiently romantic as to think the world will always contain some mystery, then we cannot expect their beliefs to be coherent.

Proof. We fix an agent i and construct the prior belief ρ^i, proceeding by induction. Note that $\beta_\infty^{i,0}$ is the agent's prior probability of F. If this prior is either 0 or 1, then so must be all subsequent updates, and then any prior belief with support contained either on the event F^c or on the event F (respectively, with the requisite set nonempty, by the martingale property) suffices.

Suppose $\beta_\infty^{i,0} \in (0,1)$. By assumption, the measure $\beta_\infty^{i,1}$ attaches conditional probabilities to a collection of cylinder sets of the form $[\omega_{I_1}]$, with some of these values larger than $\beta_\infty^{i,0}$ and some smaller. Assign probabilities $\rho^i([\omega_{I_1}])$ to these sets so that the average of the conditional probabilities is $\beta_\infty^{i,0}$. Continuing in this fashion, we attach a probability to every cylinder set $[\omega_{I_n}]$. It follows from Kolmogorov's theorem (Billingsley, 2012, p. 517) that this measure extends to a probability measure ψ^i over $X^{\cup_{n=0}^\infty I_n}$. By construction, $(\beta_\infty^{i,n})$ is a martingale with respect to ψ^i, and so converges ψ^i-almost surely to some $\beta_\infty^{i,\infty}$ (which is measurable with respect to $\cup_{n=0}^\infty I_n$).

Suppose f is measurable with respect to $\cup_{n=0}^\infty I_n$. Then (B.3) and (B.4) imply that $\beta_\infty^{i,\infty} = f$ almost surely: If $\cup_{n=0}^\infty I_n = N$, set $\rho^i = \psi^i$ and we have $\beta_\infty^{i,n}(\omega_{I_n}) = \mathbb{E}_{\rho^i(\cdot|\omega_{I_n})}[f(\omega)]$. If $\cup_{n=0}^\infty I_n$ is a strict subset of N, then let ρ^i be any probability measure whose marginal on $X^{\cup_n I_n}$ agrees with ψ^i and we again have $\beta_\infty^{i,n}(\omega_{I_n}) = \mathbb{E}_{\rho^i(\cdot|\omega_{I_n})}[f(\omega)]$.

Suppose f is not measurable with respect to $\cup_{n=0}^\infty I_n$. This implies that $\cup_{n=0}^\infty I_n$ is a strict subset of N. Requirements (B.3) and (B.4) imply that we can choose $\rho^i \in \Delta(\Omega)$ so that its marginal on $X^{\cup I_n}$ agrees with ψ^i and $\beta_\infty^{i,\infty}(\omega_{\cup I_n}) = \mathbb{E}_{\rho^i(\cdot|\omega_{\cup I_n})}[f(\omega)]$. This then implies $\beta_\infty^{i,n}(\omega_{I_n}) = \mathbb{E}_{\rho^i(\cdot|\omega_{I_n})}[f(\omega)]$.

■

We now contrast this result with a group of model-based reasoners. We again consider a sequence that receives increasing amounts of information (I_n) and assume the agents have common information. We maintain our running assumption that agents observe information contained in their models.

Proposition 1 immediately implies the following.

Corollary 1 *Consider a sequence $n = 1, \ldots,$ of groups of model-based reasoners, with agent i's model given by M^i, and each agent i's information* *set I^i_n in group n given by I_n. For each n and each agent i, $I_n \subseteq M^i$. Then every agent's limit belief equals the public oracular belief.*

Model-based updating thus places considerably more structure on agents' beliefs. Even when removing all other obstacles to disagreement, including making information common, agents with different priors face virtually unlimited possibilities for disagreement. In contrast to the case of different priors, the only sources of disagreement among agents with different models arise out of the different ways agents interpret information they think irrelevant.

B.2 Subsequent Updating in Example 1

We complete the discussion of updating in Example 1. The second round calls for the agents to announce their updated beliefs to one another. Agent 1 learns nothing new from this new announcement. Agent 2's original announcement revealed all of 2's information to 1, namely the value of ω_4, and so agent 1 draws no further inferences (and the table contains no further column for agent 1).

Agent 2 does update in response to agent 1's announcement, giving rise to the column $\beta_2(\omega_{I^2}, b^1_0, b^1_1, b^2_0)$. First, suppose agent 1 announces the belief 1/16 on the first round. This announcement reveals to agent 2 that $\omega_3 = 0$ (and also that $\omega_2 = 0$, though 2 considers this information irrelevant), and there is nothing more for 2 to learn from 1's subsequent announcement of either 0 or 1/4. Agent 2's beliefs are unchanged in this case. A similar argument applies if agent 1 announces a belief of 1.

Suppose that 1's initial announcement was 13/16, and 2's observation is $\omega_4 = 1$ (and hence 2's report was 29/32). Agent 1's updated belief is always 1 in this case, and hence there is no new information for agent 2 to process on the second round. In this case, 2's beliefs remain unchanged. Suppose, however, that 2's initial observation was $\omega_4 = 0$ (and hence 2's report was 14/32). Now suppose 2 observes that 1 has revised her belief to 1/4. This reveals to 2 that $\omega_3 = 1$. (It also potentially reveals that $\omega_2 = 0$, but 2 considers this information irrelevant.) Agent 2 then notes that when $(\omega_3, \omega_4) = (1, 0)$, the full-information belief of the event F is 5/8, and this becomes 2's new belief. Analogously, suppose that 2's initial observation was $\omega_4 = 0$ (and hence 2's report was 14/32). Now 2 observes that 1 has revised her belief to 3/4. This reveals to 2 that $\omega_3 = 0$. Agent 2 then notes that when $(\omega_3, \omega_4) = (0, 1)$, the full information belief of the event F is 3/8, and this becomes 2's new belief. We report these beliefs in column

$\beta_2(\omega_{I^2}, b_0^1, b_1^1, b_0^2)$.

It is straightforward to check that subsequent rounds of announcements have no further effect on beliefs.

B.3 An Example with Infinite Iterations

Let $N = \mathbb{N}$ and $\Omega = \{0,1\}^\infty$. There are two agents, with $M^1 = \mathbb{N} \setminus \{1\}$ and $M^2 = \mathbb{N} \setminus \{2\}$. The data generating process ρ independently chooses each variable to be 0 or 1 with probability 1/2. Agents 1 and 2 observe

$$I^1 = \{1, 3, 4, 6, 8, 10, \ldots\} \text{ and } I^2 = \{2, 3, 5, 7, 9, 11, \ldots\}.$$

We first define two events, G and H, which are constituents of the event F.

The event G occurs if and only if $(\omega_1, \omega_2) = (1, 0)$.

The event H occurs if at least one of the following statements holds:

$$\begin{aligned}
\omega_3 = \omega_4 &= \omega_5, \\
(\omega_3 + \omega_5)_{\text{mod } 2} = \omega_6 &= (\omega_8 + \omega_9)_{\text{mod } 2} = (\omega_{10} + \omega_{11})_{\text{mod } 2}, \\
(\omega_3 + \omega_4)_{\text{mod } 2} = \omega_7 &= (\omega_8 + \omega_9)_{\text{mod } 2} = (\omega_{10} + \omega_{11})_{\text{mod } 2}, \\
(\omega_3 + \omega_7)_{\text{mod } 2} = \omega_8 &= (\omega_{10} + \omega_{11})_{\text{mod } 2} = (\omega_{12} + \omega_{13})_{\text{mod } 2} \\
&= (\omega_{14} + \omega_{15})_{\text{mod } 2}, \\
(\omega_3 + \omega_6)_{\text{mod } 2} = \omega_9 &= (\omega_{10} + \omega_{11})_{\text{mod } 2} = (\omega_{12} + \omega_{13})_{\text{mod } 2} \\
&= (\omega_{14} + \omega_{15})_{\text{mod } 2}, \\
(\omega_3 + \omega_9)_{\text{mod } 2} = \omega_{10} &= (\omega_{12} + \omega_{13})_{\text{mod } 2} = (\omega_{14} + \omega_{15})_{\text{mod } 2} \\
&= (\omega_{16} + \omega_{17})_{\text{mod } 2} = (\omega_{18} + \omega_{17})_{\text{mod } 2}, \\
(\omega_3 + \omega_8)_{\text{mod } 2} = \omega_{11} &= (\omega_{12} + \omega_{13})_{\text{mod } 2} = (\omega_{14} + \omega_{15})_{\text{mod } 2} \\
&= (\omega_{16} + \omega_{17})_{\text{mod } 2} = (\omega_{18} + \omega_{19})_{\text{mod } 2}, \\
&\vdots
\end{aligned}$$

The probability of event H lies between 1/4 (the probability that $\omega_3 = \omega_4 = \omega_5$) and 3/4 (the sum of the probabilities of each of the statements on the list).

Now consider beliefs about the event $F := G \cup H$.

Upon observing ω_{I^1}, agent 1's posterior belief about every statement in the definition of H other than the first is unchanged. However, 1 updates positively the posterior probability that H holds if $\omega_3 = \omega_4$, and updates negatively if this equality fails. Agent 1's first announcement of the probability of F thus reveals the realization of ω_4 to agent 2, but reveals no additional information. Similarly, agent 2's first announcement of the probability of F reveals the realization of ω_5 (but no additional information) to agent 1.

The first round of announcements may reveal that the event H occurs, but with positive probability this is not the case. In the latter case, the agents now update their posteriors about the second and third statements in the definition of H (and no others), depending on their realizations of ω_6 and ω_7, and their next announcements of the probability of F reveal these values. This in turn allows them to update their beliefs about the fourth and fifth statements (and no others), and so on.

With positive probability, the event H has indeed occurred, in which case the belief updating about the event H terminates after a finite number of iterations, with probability 1 attached to H. However, with positive probability H has not occurred, in which case beliefs about H are revised forever.

We then have the following possibilities concerning the event $F = G \cup H$ (in all cases, after the initial exchange, subsequent exchanges of beliefs have no effect on the probability they attach to event G, and cause them to update the probability that H as described above):

- $(\omega_1, \omega_2) = (0, 1)$. Both agents attach interim probability 0 to event G, and each agent attaches the same probability to event F as they do to event H. Beliefs about H converge to a common limit.

- $(\omega_1, \omega_2) = (1, 0)$. Both agents attach interim probability 1/2 to the event that G has occurred. Beliefs about F converge to either 1/2 (if H has not occurred) or 1 (if H has occurred). In either case, beliefs converge to a common limit.

- $(\omega_1, \omega_2) = (0, 0)$. Agent 1 attaches interim probability 0 and agent 2 attaches interim probability 1/2 to event G. If H has occurred, the beliefs of both agents will eventually place probability 1 on event F. However, if H has not occurred, it will take an infinite number of exchanges for beliefs about event F to converge to 0 for agent 1 and 1/2 for agent 2.

- $(\omega_1, \omega_2) = (1, 1)$. This duplicates the previous case, with the roles of agents 1 and 2 reversed.

Remark B.1 A simplification of this example shows that Geanakoplos and Polemarchakis's (1982) protocol on an infinite space with a common prior and model also need not terminate in a finite number of steps. Take the event to be H, the common model to be $N \setminus \{1,2\}$, and let agent 1 observe $\{3,4,6,8,\dots\}$, and agent 2 observe $\{3,5,7,9,\dots\}$. ♦

B.4 Common Knowledge

We now explore the sense in which, once beliefs in the belief revision process have converged, the resulting beliefs, though different, are common knowledge. Here, we find it most natural to adopt the interpretation that the agents understand each others' models.

We first discuss the case where each agent's model $\mathcal{M}^i$ is finite. We can think of agent i's model as described by a finite partition of Ω and, since $\mathcal{I}^i \subseteq \mathcal{M}^i$, agent i's information as a coarser partition of Ω. The announcement of a belief b^i implies that the event that led agent i to having that belief is common knowledge, and so all agent's information partitions are refined. After round n announcements, all agents have new partitions, and the intersection of the events leading to the round n announcements is common knowledge (though beliefs conditional on the intersection need not be common knowledge).

We say that a vector of beliefs $(b^1,\dots,b^K)$ is *common knowledge at state* ω if these beliefs prevail at every state in that element of the meet of the agents' partitions containing ω, and their announcement does not lead to further revision of the partitions.

Intuitively, if the true state was not contained in a common knowledge event containing the final posteriors to be announced, then there would be further revision. This leads to:

Proposition B.2 *If $\mathcal{M}^i$ is finite for all i, then once the updating process terminates, the resulting beliefs are common knowledge.*

Proof. Each agent's interim belief, and each subsequent announcement by that agent, must be measurable with respect to the agent's partition. Each announcement thus gives rise to a common knowledge event. Moreover, for each player, these common knowledge events are descending, and hence form a sequence that is eventually constant. By Proposition 1.4, the limit beliefs are constant on this limit set, and so their announcement does not change agents' partitions. Moreover, since the $\mathcal{M}^i$ are finite, all players know the finite time by which the updating process terminates, and so at that time the beliefs are common knowledge. ■

The common knowledge of limit beliefs implies an agreement theorem.

Proposition B.3 *If all agents have the same (finite) model $\mathcal{M}$, then all agents have the same limit beliefs, for all possible information structures.*

Proof. In each round, all agents are updating their beliefs on the same partition $\mathcal{M}$, and since beliefs are common knowledge, they must agree (Aumann, 1976). ∎

When the models are infinite, as in Example B.3, the belief revision process may continue without end. At no stage during the belief-revision process in Appendix B.3 are the beliefs common knowledge. Despite this difficulty, there is an appropriate notion of common knowledge when the models are infinite.

Since we now must deal with conditioning on potentially zero probability events, we follow Brandenburger and Dekel (1987) in defining knowledge as probability one belief, and requiring conditional probabilities to be regular and proper.[27] Recall that the state space has prior ρ, and suppose that each player's information is described by a σ–algebra $\mathcal{G}^i$. For each agent i, there is a mapping $\rho^i : \mathcal{F} \times \Omega \to [0,1]$, where $\rho^i(\cdot \mid \omega)$ is a probability measure on $\mathcal{F}$ for all $\omega \in \Omega$; for each $G \in \mathcal{F}$, $\rho^i(G \mid \cdot)$ is a version of $\rho(G \mid \mathcal{G}^i)$; and $\rho^i(G \mid \omega) = \chi_G(\omega)$ for all $G \in \mathcal{G}^i$ (in other words, ρ^i is a regular and proper conditional probability). These are the beliefs used to define what it means for agent i to know (assign probability 1 to) an event. By Brandenburger and Dekel (1987, Lemma 2.1), an event G is common knowledge at some ω (in the sense that every agent assigns probability one to the event, every agent assigns probability one to every agent assigning probability one to the event, and so on) if there is a set G' in the meet $\wedge\mathcal{G}^i$ such that $\omega \in G'$ and $\rho^i(\{\omega' \in G' : \omega' \notin G\} \mid \omega'') = 0$ for all $\omega'' \in \Omega$.[28] The last requirement is simply that G' is a subset of G, up to a zero measure set, under each agent's beliefs ρ^j.

We will say that limit beliefs are common knowledge if they are common knowledge given the information provided to the agents by the entire infinite sequence of belief announcements.

Proposition B.4 *Limit beliefs are common knowledge.*

[27] Bogachev (2007, Corollary 10.4.10) ensures the existence of such conditional probabilities.

[28] This is a sufficient condition for common knowledge. The characterization requires a little more (Brandenburger and Dekel, 1987, Lemma 2.3 and Proposition 2.1), which we do not need.

Proof. Recall that $\mathcal{B}_n$ denotes the round n σ-algebra generated by the announcements from the first n rounds. For each $\omega \in \Omega$, all events G satisfying $\omega \in G \in \mathcal{B}_n$ are common knowledge at ω. Recall also that $(\mathcal{B}_n)_n$ is a filtration with limit $\mathcal{B}_\infty$, so that the beliefs $\beta^i_{n+1} = \mathbb{E}[f \mid \mathcal{I}^i, \mathcal{B}_n]$ are a martingale and converge almost surely to $\mathbb{E}[f \mid \mathcal{I}^i, \mathcal{B}_\infty] =: \beta^i_\infty$. Moreover, $\beta^i_\infty = \int f d\rho^i_\infty$.

Fix b^i in the range of β^i_∞ and let $A := (\beta^i_\infty)^{-1}(b^i)$. We now prove that for all $\omega \in A$ there is a subset A' in the meet $\wedge\sigma(\mathcal{I}^i, \mathcal{B}_\infty)$ containing ω. Fix $\omega \in A$, and define $A_n := \cap_j (\beta^j_n)^{-1}(b^j)$ where $b^j = \beta^j_n(\omega)$. Since $A_n \in \wedge\sigma(\mathcal{I}^i, \mathcal{B}_\infty)$, we have $\cap_n A_n \in \wedge\sigma(\mathcal{I}^i, \mathcal{B}_\infty)$. Suppose $\cap_n A_n \not\subseteq A$, so that there exists $\tilde{\omega} \in \cap_n A_n \setminus A$. But then $\beta^i_n(\omega) = \beta^i_n(\tilde{\omega})$ for all n, and since the beliefs converge,[29] $\beta^i_\infty(\tilde{\omega}) = b^i$, a contradiction. ■

Green (2012) presents an agreeing-to-disagree result for infinite models that would allow us to extend Proposition B.3 to this case.

B.5 Implications of Necessary Agreement

Suppose X and N are finite, $M^2 \subsetneq M^1$, $I^1 = M^1$, and $I^2 = \varnothing$, that is, player 1 has full information, and player 2 observes nothing, but thinks only a subset of the variables in 1's model are relevant. We can, without loss of generality, write $X^{M^1} = X_1 \times X_2$, $X^{M^2} = X_2$, and suppose all states $(\omega_1, \omega_2) \in X_1 \times X_2$ have positive probability. Since player 2 has no information, there is no updating after 2 has updated from the announcement of 1's full information beliefs. Necessary agreement in this context means that for all $(\omega_1, \omega_2) \in X_1 \times X_2$, if $f^1(\omega_1, \omega_2) = \alpha$, then

$$\begin{aligned}\alpha &= \mathbb{E}[f^2(\omega_2) \mid b^1 = \alpha] \\ &= \mathbb{E}[\mathbb{E}[f^1(\omega_1, \omega_2) \mid \omega_2] \mid b^1 = \alpha].\end{aligned}$$

We now argue that necessary agreement implies that the first coordinate of 1's model is redundant for 1, that is, f^1 is independent of ω_1.

Let $\bar{\alpha} := \max f^1(\omega_1, \omega_2)$, and let $(\bar{\omega}_1, \bar{\omega}_2)$ be values that achieve $\bar{\alpha}$, i.e., $f^1(\bar{\omega}_1, \bar{\omega}_2) = \bar{\alpha}$. Conditional on the announcement $\bar{\alpha}$, necessary agreement implies

$$\bar{\alpha} = \mathbb{E}[f^2(\omega_2) \mid b^1 = \bar{\alpha}],$$

which implies $\mathbb{E}[f^1(\omega_1, \omega_2) \mid \omega_2] = \bar{\alpha}$ for all ω_2 in the support of the conditional beliefs $\rho_2(\cdot \mid b^1 = \bar{\alpha}) \in \Delta(X_2)$, But this implies that for all $\omega_1 \in X_1$

[29] The sentence previously footnoted implies we can assume beliefs converge on $\cap_n A_n$.

State	Prior		Theories		Interim beliefs		First-round update
(ω_1, ω_2)	ρ	$f^*(\omega)$	$f^1(\omega_{M^1})$	$f^2(\omega_{M^2})$	$\beta^1(\omega_{I^1})$	$\beta^2(\omega_{I^2})$	$\beta^2(\omega_{I^2}, b_0^1)$
$(0,0)$	a	x	$\frac{ax+by}{a+b}$	$\frac{ax+cz}{a+c}$	$\frac{ax+by}{a+b}$	$ax+by+cz+dw$	$\frac{a}{a+b}\frac{ax+cz}{a+c} + \frac{b}{a+b}\frac{by+dw}{b+d}$
$(0,1)$	b	y	$\frac{ax+by}{a+b}$	$\frac{by+dw}{b+d}$	$\frac{ax+by}{a+b}$	$ax+by+cz+dw$	$\frac{a}{a+b}\frac{ax+cz}{a+c} + \frac{b}{a+b}\frac{by+dw}{b+d}$
$(1,0)$	c	z	$\frac{cz+dw}{c+d}$	$\frac{ax+cz}{a+c}$	$\frac{cz+dw}{c+d}$	$ax+by+cz+dw$	$\frac{c}{c+d}\frac{ax+cz}{a+c} + \frac{d}{c+d}\frac{by+dw}{b+d}$
$(1,1)$	d	w	$\frac{cz+dw}{c+d}$	$\frac{by+dw}{b+d}$	$\frac{cz+dw}{c+d}$	$ax+by+cz+dw$	$\frac{c}{c+d}\frac{ax+cz}{a+c} + \frac{d}{c+d}\frac{by+dw}{b+d}$

$$X = \{0,1\}, \quad M^1 = \{1\}, \quad M^2 = \{2\},$$
$$I^1 = \{1\}, \quad I^2 = \varnothing.$$

Figure B.1: Agreement need not imply redundancy in the presence of correlation.

and for all ω_2 in the support of the conditional beliefs $\rho_2(\cdot \mid b^1 = \bar{\alpha}) \in \Delta(X_2)$, $f^1(\omega_1, \omega_2) = \bar{\alpha}$; in particular, for such ω_2, f^1 is independent of ω_1.

Since $X_1 \times X_2$ is finite, we can now argue inductively. Suppose $\alpha' := \max\{f^1(\omega_1, \omega_2) < \bar{\alpha}\}$ and let (ω_1', ω_2') be values that achieve α', i.e., $f^1(\omega_1', \omega_2') = \alpha'$. After agent 1's announcement of α', agent 2 assigns positive probability to ω_2'. Moreover, from the previous paragraph, for all ω_1, $f^1(\omega_1, \omega_2') \leq \alpha'$ (if $f^1(\omega_1, \omega_2') = \bar{\alpha}$ for some ω_1, then ω_2' is in the support of $\rho_2(\cdot \mid b^1 = \bar{\alpha})$ and so $f^1(\omega_1', \omega_2') = \bar{\alpha} \neq \alpha'$, a contradiction). But then, for all $\omega_1 \in X_1$ and for all ω_2 in the support of the conditional beliefs $\rho_2(\cdot \mid b^1 = \alpha')$, $f^1(\omega_1, \omega_2) = \alpha'$; in particular, for such ω_2, f^1 is independent of ω_1. Proceeding in this way for progressively lower values of beliefs of agent 1, we conclude that f^1 is independent of ω_1 for all ω_2.

B.6 An Example Illustrating Redundancy and Correlation

We start with the general specification given in Figure B.1. Agent 1 observes every variable in 1's model, and so never does any updating past the interim belief. Agent 2, who observes nothing, ceases updating after the first round. If the values of ω_1 and ω_2 are independently drawn, then it follows immediately from Proposition 2 that beliefs can necessarily agree only if ω_1 is redundant for agent 1.

We now seek values of the parameters for which ω_1 is not redundant for player 1, i.e.,

$$\frac{ax + by}{a + b} \neq \frac{cz + dw}{c + d} \tag{B.5}$$

and for which there is necessary agreement, i.e. (after simplification),

$$(a+c)by = ac(z-x) + b\frac{a+c}{b+d}(by+dw) \tag{B.6}$$

and

$$(b+d)cz = bd(y-w) + c\frac{b+d}{a+c}(ax+cz). \tag{B.7}$$

Setting $b = c = 0$ gives the case where the two variables are perfectly correlated (ω_2 is simply a relabeling of ω_1), and we trivially have necessary agreement without redundancy.

It is straightforward that there are many parameters with the desired characteristics. If we set $z = x$ and $y = w$, then *any* specification of a, b, c, d satisfies these equations, including values that also satisfy (B.5). In this case, ω_1 plays no role in the determination of F, and agent 1's observation of ω_1 is informative only to the extent that it is correlated with ω_2. In addition, agent 2 receives no information of her own, and so must similarly rely on gleaning information from the correlation of ω_1 with ω_2, leading the two agents to agree. In the case of independence, or $a = b = c = d$, agent 1 learns nothing about the state, and the two agents necessarily agree on the uninformative posterior of 1/2.

When at least one of $z = x$ and $y = w$ fails, then ω_1 plays a role in determining the event F. There then exist particular values of a, b, c, d satisfying the equations (B.6)–(B.7) for necessary agreement.

References

Acemoglu, D., V. Chernozhukov, and M. Yildiz (2016): "Fragility of Asymptotic Agreement under Bayesian Learning," *Theoretical Economics*, 11(1), 187–225.

Al-Najjar, N. I. (2009): "Decision Makers as Statisticians: Diversity, Ambiguity, and Learning," *Econometrica*, 77(5), 1371–1401.

Arieli, I., Y. Babichenko, and R. Smorodinsky (2018): "When is the Crowd Wise?," Faculty of Industrial Engineering and Management, Technion.

Arrow, K. J., and L. Hurwicz (1972): "An Optimality Criterion for Decision-Making under Ignorance," in *Uncertainty and Expectations in*

Economics, ed. by C. Carter, and J. Ford, pp. 1–11. Basil Blackwell, Oxford.

Aumann, R. J. (1976): "Agreeing to Disagree," *Annals of Statistics*, 4(6), 1236–1239.

Billingsley, P. (2012): *Probability and Measure*. John Wiley and Sons, New York, anniversary edn.

Bogachev, V. I. (2007): *Measure Theory, Volume 2*. Springer, Berlin.

Brandenburger, A., and E. Dekel (1987): "Common Knowledge with Probability 1," *Journal of Mathematical Economics*, 16(3), 237–245.

Brandenburger, A., E. Dekel, and J. Geanakoplos (1992): "Correlated Equilibrium with General Information Structures," *Games and Economic Behavior*, 4(2), 182–201.

Di Tillio, A., E. Lehrer, H. Polemarchakis, and D. Samet (2019): "Bayesian Monolgoues and Dialogues," Bocconi University, University of Tel Aviv, and Unversity of Warwick.

Eyster, E., and M. Piccione (2013): "An Approach to Asset Pricing Under Incomplete and Diverse Perceptions," *Econometrica*, 81(4), 1483–1506.

Geanakoplos, J. (1989): "Game Theory Without Partitions, and Applications to Speculation and Concensus," Cowles Foundation discussion paper 914, Yale University.

Geanakoplos, J., and H. Polemarchakis (1982): "We Can't Disagree Forever," *Journal of Economic Theory*, 27(2), 192–200.

Giacomini, R., V. Skreta, and J. Turén (2017): "Models, Inattention and Bayesian Updates," *American Economic Journal: Macroeconomics*, Forthcoming.

Gilboa, I., and L. Samuelson (2012): "Subjectivity in Inductive Inference," *Theoretical Economics*, 7(2), 183–216.

Green, E. J. (2012): "Events Concerning Knowledge," Working paper, Penn State University.

Hong, H., J. C. Stein, and J. Yu (2007): "Simple Forecasts and Paradigm Shifts," *Journal of Finance*, 52(3), 1207–1242.

Jehiel, P. (2005): "Analogy-Based Expectation Equilibrium," *Journal of Economic Theory*, 123(2), 81–104.

Jehiel, P., and F. Koessler (2008): "Revisiting Games of Incomplete Information with Analogy-Based Expectations," *Games and Economic Behavior*, 62(2), 533–557.

Kyle, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53(6), 1315–1335.

Mailath, G. J., and L. Samuelson (2019): "The Wisdom of a Confused Crowd: Model Based Inference," Cowles Foundation Discussion Paper 2161, Yale University.

Milgrom, P. R., and N. Stokey (1982): "Information, Trade, and Common Knowledge," *Journal of Economic Theory*, 26(1), 17–27.

Morris, S. (1994): "Trade with Heterogeneous Prior Beliefs and Asymmetric Information," *Econometrica*, 62(6), 1327–1347.

Ostrovsky, M. (2012): "Information Aggregation in Dynamic Markets with Strategic Traders," *Econometrica*, 80(2), 2595–2647.

Page, S. E. (2017): *The Diversity Bonus*. Princeton University Press, Princeton.

Pearl, J. (2009): *Causality*. Cambridge University Press.

Savage, L. J. (1972): *The Foundations of Statistics*. Dover Publications, New York, Orignally 1954.

Sethi, R., and M. Yildiz (2012): "Public Disagreement," *American Economic Journal: Microeconomics*, 4(3), 57–95.

Sethi, R., and M. Yildiz (2016): "Communication with Unknown Perspectives," *Econometrica*, 84(6), 2029–2069.

Spiegler, R. (2016): "Bayesian Networks and Boundedly Rational Expectations," *Quarterly Journal of Economics*, 131, 1243–1290.

——— (2019): "Behavioral Implications of Causal Misperceptions," Discussion paper, Tel Aviv University and University College London.

Stroock, D. W. (2011): *Probability Theory: An Analytic View*. Cambridge University Press, second edn.

Surowiecki, J. (2004): *The Wisdom of Crowds*. Anchor Books, New York.

Wolfers, J., and E. Zitzewitz (2004): "Prediction Markets," *Journal of Economic Perspectives*, 18(2), 107–126.