ABSTRACT

Title of Dissertation:	DELIBERATION, LEARNING CONDITIONAL PROBABILITY, AND RISK: THREE ESSAYS
	Boning Yu, Doctor of Philosophy, 2026
Dissertation directed by:	Professor, Eric Pacuit, Department of Philosophy

This dissertation investigates how rational inquiry and decision-making depend on the structure of information, evidence, and risk. Across three chapters, I examine how individuals or groups update their beliefs, communicate what they learn, and choose under uncertainty.

Chapter 2, "Deliberation Under Dynamic Discovery of Evidence," develops a model of group deliberation in which evidence discovery is dynamic: agents are more likely to uncover high-quality evidence when their current beliefs are closer to the truth. Using computer simulations, I compare different practices of communicating evidence and conclusions. The results show that communication about evidence is most valuable when agents can access good evidence, since it enables groups to pool and accumulate information. By contrast, when high-quality evidence is difficult to obtain, communication about conclusions can be more effective, allowing agents to benefit from others' relatively successful inquiries. More generally, extensive communication

of conclusions usually improves collective performance, whereas extensive communication of evidence is not always beneficial. The optimal communication strategy depends on how difficult the scenario is.

Chapter 3, "A Bayesian Reflection on the Judy Benjamin Problem," addresses the Judy Benjamin problem and argues that it is under-specified. Bayesian conditioning on learned conditional probabilities requires an extended probability space, but the original story does not determine a unique such space. Models with different extended spaces, therefore, yield different posterior beliefs. I argue that these competing models reflect different structural representations of Judy's situation, and that accuracy considerations can choose among them only when the case is supplemented with further facts about the actual structure of Judy's situation.

Chapter 4, "A Calibration Theorem for Risk-Weighted Expected Utility Theory" establishes a calibration theorem for risk-weighted expected utility theory (REU). REU is intuitively appealing as it seems to resolve some of the classical challenges to EU, such as the Allais paradox and Rabin's calibration result. I calibrate REU using Allais preferences. Thus, REU resolves the Allais paradox only at the cost of becoming susceptible to a Rabin-style challenge (being calibratable). Compared to existing calibration results against REU, my approach relies on less demanding and more realistic assumptions.

DELIBERATION, LEARNING CONDITIONAL PROBABILITY, AND RISK: THREE ESSAYS

by

Boning Yu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2026

Advisory Committee:
Professor Eric Pacuit, Chair
Professor Fabrizio Cariani
Professor Paolo Santorio
Associate Professor Harjit Bhogal
Professor Wedad J. Elmaghraby, Dean's Representative

Acknowledgements

I am excited and deeply grateful to have reached the completion of this dissertation. Finishing a PhD is an accomplishment that I will always appreciate. This project represents far more than the three papers collected here. It reflects the guidance, patience, encouragement, and generosity of many people who helped me become the philosopher and scholar I am today.

First and foremost, I thank my advisor, Professor Eric Pacuit. Since I began my graduate studies at the University of Maryland, College Park, Eric has met with me regularly to discuss my research. His patience and guidance have been invaluable in helping me develop and explore my philosophical interests. Eric gave me the freedom to pursue questions that genuinely excited me while also helping me shape those questions into serious research projects. I have learned a great deal from his philosophical insight. His passion for philosophy has shown me what it is like to approach philosophy as a living and continually rewarding activity. I am deeply grateful for his support throughout my time in the program.

I am also very grateful to Professor Fabrizio Cariani. Fabrizio first introduced me to the Judy Benjamin problem, which eventually became one of the papers in this dissertation. His influence on my work goes far beyond that initial introduction. He has helped me see how technical results can be presented not merely as formal observations, but as philosophically significant insights. Fabrizio's comments on my papers have repeatedly pushed me to clarify my arguments and better explain why the formal details matter. Many of his suggestions inspired substantial revisions. I am also grateful for his generous career advice, which has helped me think more carefully about my future as a philosopher.

I thank Professor Paolo Santorio for his detailed and thoughtful feedback on my work. Paolo has consistently provided comments that helped me improve my papers. I am especially grateful for his patience, not only with my research, but with every aspect of my life as a graduate student. His support, advice, and understanding have meant a great deal to me throughout the program.

I thank Professor Harjit Bhogal for serving on my dissertation committee. Harjit's feedback on the dissertation and questions during my defense were thoughtful. I also thank Professor Wedad J. Elmaghraby for serving as the dean's representative at my defense.

I am grateful to the Department of Philosophy at the University of Maryland for supporting me over the past six years. The department's generous financial support made it possible for me to focus on my studies and research. I also appreciate the department's flexibility and openness in encouraging me to take courses outside philosophy. Courses in areas such as information theory and probability theory equipped me with the technical skills necessary to complete this dissertation. That intellectual freedom was crucial for the development of my research.

I also owe thanks to two committee members from my master's program at Texas Tech University. I thank Professor Joel Velasco, my M.A. advisor, for helping me develop my interest in Bayesian epistemology. That interest has continued to shape my philosophical work and is directly reflected in this dissertation. I also thank Professor Jeremy Schwartz, whose class first introduced me to decision theory. That early exposure opened up a field of questions that eventually became central to one of the chapters of this dissertation.

I am very grateful to the logic group at the University of Maryland. Over the years, members of the group listened to presentations of my work at different stages and gave many helpful comments. Their questions and suggestions helped me refine my ideas, clarify my

arguments, and see problems from new perspectives. I especially thank Professor Steven Kuhn, Professor Ilaria Canavotto, Dr. Masayuki Tashiro, and Justin Helms for their support and feedback.

I also thank my peers at the University of Maryland. Graduate school would have been much harder without their encouragement and companionship. I am grateful for the conversations, shared frustrations, advice, and friendship that carried me through the many stages of this process. Their presence made the department not only a place to work, but also a community in which I could grow.

Finally, I thank my parents. They have supported me emotionally throughout my studies and created the worry-free conditions that allowed me to focus on my work. This support did not begin with my PhD; it has continued throughout more than twenty years of school. Their love, patience, and sacrifices made this accomplishment possible. I cannot adequately express how grateful I am for everything they have done for me. This dissertation is, in a very real sense, also the result of their lifelong support.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

This dissertation consists of three self-standing papers in Bayesian epistemology. The central lesson of the dissertation is that formal models do not provide normative conclusions in isolation. Their implications depend on background assumptions about epistemic structure. In social inquiry, the effect of communication depends not only on how agents talk, but also on what kind of information they exchange and how that information is generated. In belief dynamics, the result of conditioning depends not only on the prior and the learned information, but also on the probability space in which that information is represented. In decision theory, the success of a risk-sensitive alternative to expected utility theory depends not only on its ability to accommodate intuitive preferences, but also on the further patterns of choice to which that accommodation commits the agent. Across these cases, the philosophical work lies in identifying which structural assumptions are doing the explanatory and normative work.

Chapter 2, "Deliberation Under Dynamic Discovery of Evidence," belongs to formal social epistemology. It investigates how groups of agents can improve their chances of discovering the truth by communication. Much of the literature on deliberation and epistemic networks asks whether more communication improves or harms collective accuracy. One influential lesson from this literature is that more communication is not always epistemically beneficial, i.e., frequent communication can reduce diversity, lead groups to converge too quickly, or cause agents to become collectively trapped by misleading information (Zollman, 2007; Zollman, 2010). My chapter contributes to this debate by modeling evidence discovery as a dynamic process. In my model, agents' current beliefs affect the quality of evidence they are likely to discover. Agents whose beliefs are closer to the truth are more likely to encounter truth-conducive evidence, while agents farther from the truth are less likely to do so.

This feature is meant to capture the phenomenon that one's belief affects how one investigates. As a result, agents' epistemic positions partly shape the quality of the evidence they encounter. This dynamic feature changes how we should think about deliberation. Communication does not merely distribute a fixed stock of evidence; it also affects where agents move in epistemic space, which in turn affects the future evidence they are likely to discover.

The chapter also distinguishes between two kinds of communicable information: evidence and conclusions. Evidence is information that agents acquire during inquiry and can share with others. Conclusions are the results agents assert after they have become sufficiently stable in their beliefs. This distinction matters because evidence and conclusions play different epistemic roles. Evidence is timely and continues to evolve as agents move through epistemic space. Conclusions appear later and reflect a period of successful inquiry, but they can also be false. The simulations show that the value of communicating evidence and conclusions depends on the epistemic environment. When agents have access to high-quality evidence, communication about evidence helps by allowing agents to accumulate a larger body of information. When agents lack access to high-quality evidence, communication about conclusions can be more useful, because it allows agents to benefit from the results of those who happen to discover the truth.

The chapter's broader conclusion is that there is no single communication policy that is best across all deliberative contexts. In difficult settings, where agents are unlikely to find truth-conducive evidence on their own, the best strategy may be to learn all asserted conclusions while avoiding the communication of low-quality evidence. In more favorable settings, where agents have decent access to high-quality evidence, the best strategy combines extensive communication of conclusions with some communication of evidence. Thus, the chapter refines the idea that "less communication can be more."

Chapter 3, "A Bayesian Reflection on the Judy Benjamin Problem," turns from group inquiry to individual belief updating. Bayesian epistemology is often presented as a clear rule for learning. When an agent learns new information, she should update by conditioning on that information. However, conditioning requires that the learned information be representable as an event or a constraint of an appropriate probability space. The Judy Benjamin problem shows that this requirement is not trivial. Judy receives information in the form of a conditional probability, e.g., if she is in enemy territory, the probability that she is in one particular region is three-fourths. The challenge is that this learned conditional probability is not itself an event in the naïve probability space consisting of the four possible regions.

To apply Bayesian conditioning, an extended probability space must be introduced. But the original Judy Benjamin story does not determine a unique extended space. Different extensions represent Judy's situation differently and yield different posterior probabilities. Grove and Halpern's (1997) model represents Judy's uncertainty as uncertainty over possible probability distributions on the four regions and yields an intuitive answer to Judy's inquiry. Vasudevan's (2020) model represents Judy's uncertainty using sequences of region-outcomes and recovers the minimum-relative-entropy answer.

This disagreement reveals that the original problem is under-specified. The story tells us Judy's prior probabilities and the conditional probability she learns, but it does not tell us what structure gives rise to those probabilities. Thus, it does not determine which extended space is the correct one for Bayesian conditioning. I also challenge the proposed qualitative contrast between the intuitive answer and the minimum-relative-entropy answer. Vasudevan suggests that his model reflects epistemic charity toward Judy's informant, while Grove and Halpern's model reflects conservativity. I argue that this contrast is not well supported. Both models incorporate structurally

similar assumptions once their extended spaces have been chosen. The real difference lies not in one model being charitable and the other conservative, but in the different representational structures they impose.

To make this point clearer, I introduce a third, two-stage model. The model represents Judy's uncertainty as first concerning the division between red and blue territory and then concerning the division between camps within each color. Like Grove and Halpern's model, the two-stage model vindicates the intuitive answer to the original Judy Benjamin question. At the same time, it differs from Grove and Halpern's model in how it represents Judy's higher-order uncertainty. This shows that even models agreeing on the original posterior can encode different structural assumptions. The broader moral is that a choice of an extended probability space is a choice of representation. Accuracy considerations can select among such representations only when the case is supplemented with further facts about the actual structure of the situation.

Chapter 4, "A Calibration Theorem for Risk-Weighted Expected Utility Theory," shifts from belief updating to decision-making under risk. Expected utility theory has long served as the standard theory of rational choice under uncertainty, but it faces two famous challenges. The Allais paradox suggests that expected utility theory cannot accommodate certain intuitive preferences over lotteries (Allais, 1953). Rabin's calibration theorem suggests that expected utility theory can accommodate seemingly mild small-stakes risk aversion only at the cost of generating implausibly extreme risk aversion over larger stakes (Rabin, 2000). Risk-weighted expected utility theory (REU), developed by Buchak (2013) and closely related to rank-dependent expected utility theory (Quiggin, 1982), appears to offer a promising response. By adding a risk function to the utility function, REU allows agents to weight potential improvements according to their risk attitudes.

This enables REU to accommodate Allais-style preferences without forcing all risk aversion into the curvature of the utility function.

I argue that this advantage is unstable. I establish a calibration theorem for REU showing that, under weak and plausible assumptions, accommodating Allais preferences commits REU agents to extreme risk-aversion in other situations. In the linear-utility case, the argument shows that if an REU agent has Allais preferences, then the agent must severely undervalue small-probability gains relative to small-probability losses. The result is that the agent will reject gambles in which the possible gain is many times larger than the possible loss, even when the probabilities of gain and loss are equal and small. I then extend the argument to the more realistic case of concave utility by assuming constant relative risk aversion. The general result is that no combination of a suitably concave utility function and a risk function can both accommodate Allais preferences and avoid calibration.

The significance is that the very mathematical structure that allows REU to accommodate the Allais paradox also makes it vulnerable to a Rabin-style challenge. REU appears to improve on expected utility theory by separating attitudes toward outcomes from attitudes toward risk. However, once the risk function fits Allais preferences, it imposes strong constraints on how the agent evaluates other gambles. In this way, REU resolves one classical problem only by recreating the structure of another. Compared with existing calibration results against rank-dependent or risk-weighted theories, my argument relies on less demanding assumptions and applies even when the agent begins with fixed wealth rather than uncertainty about initial wealth.

The dissertation proceeds as follows. Chapter 1 develops and analyzes an agent-based model of group deliberation under dynamic evidence discovery. Chapter 2 examines the Judy Benjamin problem and argues that the original case underdetermines the appropriate Bayesian

extended space. Chapter 3 establishes a calibration theorem for risk-weighted expected utility theory and argues that REU does not fully escape the challenges it was designed to address.

Chapter 2: Deliberation Under Dynamic Discovery of Evidence

2.1 Introduction

Because human preferences are often complex and heterogeneous, aggregating them is difficult. Classic impossibility theorems (e.g., Arrow, Gibbard–Satterthwaite) show that no single rule can meet all reasonable fairness and rationality conditions simultaneously. Dryzek and List (2003) argues that appropriately structured deliberation can mitigate these limits by inducing more truthful revelation, narrowing the live agenda, and structuring preferences so that preferences can be aggregated in a meaningful way. From an epistemic perspective, however, "more talk" is not automatically "more truth." Zollman (2007; 2010), building on Bala and Goyal's (1998) bandit framework, reveals that frequent communication can lock groups prematurely onto the wrong hypothesis (hereafter termed the Zollman effect). Related work on communicating electorates confirms and extends this point. Hahn, von Sydow, and Merdes (2019) recognize that communication within deliberation can raise individual accuracy yet lower the accuracy of the collective decision, because dependence among voters erodes the diversity that underwrites the wisdom of crowds. Hahn (2022) generalizes the point: collective performance depends jointly on individual accuracy and on the diversity of errors; indeed, a group is most accurate when its members are reasonably right and their mistakes do not all point the same way.

Against this backdrop, the core question is not aggregation versus deliberation, but, in a specific situation, how to organize communication so that groups learn the truth. In this paper, I introduce a model of deliberation that has two novel features. First, my model treats evidence discovery as a dynamic process, where agents' chance of discovering truth-conducive evidence (a piece of evidence that, once learned, brings an agent closer to the truth) depends on how closely their current beliefs align with the truth. This concept builds on the idea that one's beliefs shape

one's actions (for my interest, how one gathers evidence).[1] Namely, one's beliefs influence how one engages with the external world and, consequently, the type of evidence one encounters.[2] People with beliefs close to the truth are more likely to adopt methodologies that tend to yield truth-conducive evidence. Conversely, those with beliefs further from the truth may not engage with the external world in ways that lead to such quality evidence. Second, my model incorporates communication about both evidence and conclusions. These are two kinds of information being communicated in real deliberations, and omitting either aspect leaves any model of deliberation incomplete. In long-term deliberative processes, agents may conclude that they have reached the truth and cease further investigation. Ideally, conclusions are rationally derived from a body of evidence, which, in turn, makes them appear to be more credible. However, even when assuming that such conclusions are offered in good faith, there is no guarantee that a conclusion is indeed true. It remains unclear how others should respond to these announcements. Therefore, it is crucial to evaluate the effects of both the communication about evidence and conclusions as well as how they interact.

With two primary goals in mind, I develop a model of deliberation that extends the Hegselmann–Krause (H-K) model (Hegselmann & Krause, 2002). In the H-K model, agents synchronously update their opinions by communicating with others whose views are not too far from their own. Building on this framework, I also consider an alternative mode of random mixing, in which agents update by interacting with some random others, regardless of similarity.[3] My first

[1] As a general discussion, Putnam (1981) argues that conceptual schemes, including beliefs and language, shape human interactions with reality. Regarding scientific discovery, Hanson's (1979) concept of the Theory-Ladenness of Observation suggests that scientific observation is influenced by scientists' beliefs, such as the theoretical background, prior knowledge, and conceptual framework under which they operate.

[2] My work here is concerned with how one's beliefs affect the discovery of evidence. Although the idea that one's beliefs affect the discovery of evidence is not new, modeling and examining this idea in the context of deliberation is innovative. In Zollman's (2007) bandit game model, an agent's belief determines their decision about which arm of the bandit machine to pull. However, the feedback from the machine does not depend on the strength of the agent's belief. So, Zollman's model captures the fact that beliefs affect one's actions but does not model the fact that people tend to receive different evidence depending on their beliefs.

[3] This contrasts with the Deffuant–Weisbuch (D-W) model, where agents meet in random pairs, and an update occurs only if their

goal is to provide an overview of how different communicative practices influence a group's ability to track the truth, and my second goal is to identify, under various scenarios, the exact practice of deliberation that maximizes the majority of a group's chance of arriving at the truth. I define a group's communicative practice in terms of three key components: (1) the kinds of information being communicated (evidence, conclusions, or both); (2) the mode of communication (while communicating, whether agents randomly interact with other agents or only with agents that are epistemically close to themselves); (3) the extent of communication (in the random mode, the extent represents the frequency of communication, or in the neighborhood mode, the extent represents the epistemic distance within which this communication occurs). When analyzing the effects of practices in various scenarios, I focus on two aspects of deliberative scenarios: (1) accessibility—whether agents have access to truth-conducive evidence, and (2) ambiguity—due to its nature, whether the truth tends to give clear or ambiguous information.

From my simulations, I draw three major conclusions. First, when agents have access to high-quality evidence (given the specific situation, agents are reasonably likely to obtain truth-conducive evidence), communication about evidence improves informational clarity by allowing agents to accumulate a larger body of evidence; however, when agents lack such access, communication about conclusions improves accessibility to high-quality information. Second, for any given practice of communication about evidence, extensive communication about conclusions almost always improves a group's ability to converge on the truth; nevertheless, when agents are communicating their conclusions, extensive communication about evidence, despite promoting diversified information, is not necessarily beneficial. Together, these findings lead to a third conclusion about optimal communicative practices. In challenging scenarios (i.e., scenarios in

opinions are sufficiently close to each other.

which agents lack access to high-quality evidence), the best strategy is to learn all asserted conclusions without communicating evidence; however, in scenarios where agents have decent access to high-quality evidence, combining full communication of conclusions with some degree of communication about evidence is the best practice. Overall, extensively communicating both evidence and conclusions is often not the optimal strategy; in other words, the so-called Zollman effect occurs more often than not.

In the rest of the paper, I use the term *truth-conducive* to describe individual pieces of information, such as single items of evidence or a conclusion. To describe a body of accessible information, I use the terms *high-quality* or *low-quality*. For evidence, *high-quality* means that accessible evidence is, on average, truth-conducive and not ambiguous—sufficiently indicative of the truth. Accordingly, in each inquiry, an agent with access to high-quality evidence is likely to get a piece of truth-conducive item. In some cases, specifically for conclusions, the terms *truth-conducive* and *high-quality* can coincide when the context makes this clear.

In Section 2.2, I go over related literature on simulating deliberation. In Section 2.3, I introduce my model and present a sample computer simulation. In Section 2.4, I analyze the effects of various communicative practices and their interactions. In Section 2.5, I identify the optimal communicative practice across various scenarios. Section 2.6 concludes the paper and points to potential avenues for future work.

2.2 Related Works

The H-K (Hegselmann & Krause, 2002) and the D-W (Deffuant, Neau, Amblard, & Weisbuch, 2000) models are among the first widely recognized agent-based models (ABMs) of continuous belief dynamics. Both models assume that agents update their beliefs only when their opinions are sufficiently close, and simulations show that depending on the size of this confidence

threshold, the group either converges to consensus or fragments into polarized clusters. In the H-K model, each agent averages the opinions of all its epistemic neighbors, whereas in the D-W model, agents meet in random pairs and update only if the partner qualifies as an epistemic neighbor. Building on the 2002 framework, the H-K model has been extended to include a fixed truth attractor (Hegselmann & Krause, 2006) and, more recently, a new type of agents, campaigners, that actively broadcast some misleading information (Douven & Hegselmann, 2021). In addition, Douven and Riegler (2010a; 2010b) provide a systematic investigation of various possible extensions of the H-K model.

A major subsequent strand (Betz, 2013; Dupuis de Tarlé, Michelini, Šešelja, & Straßer, 2025), argumentative ABMs (AABMs), represents reasons and their relations (support/attack) so as to model argumentative communication in deliberation. Other approaches, including my study, do not model full argumentative structure but focus on specific aspects of deliberation. For example, Hartmann and Rad (2018) model deliberation as a procedure by which participants share beliefs at a stage in time and get to know whether each other is reliable. In their model, with better knowledge of other agents' reliability, participants can strategically assess and integrate their beliefs, increasing the group's likelihood of discovering the truth. Zollman (2007; 2010) instead studies truth-tracking in communicative networks using a bandit model and highlight that greater connectivity increases the speed of consensus formation but reduces the accuracy of groups' opinions. Not all models of deliberation rely heavily on ABM simulations: some derive results analytically. Ding and Pivato (2019) analyze agents' motivation to share evidence and argue that through deliberation, without agents sharing and learning all evidence, a group's opinion will still converge on the same outcome as if agents had fully shared and learned all evidence. Dietrich and Spiekermann (2025) argue that deliberation helps mitigate various failures in evidence processing,

ultimately leading to better group opinions. While these models primarily focus on the exchange of information, they do not address the generation of new evidence during deliberation. These works nevertheless provide important insights into modeling essential components of deliberation.

Beyond accuracy and speed, two further phenomena are especially relevant here: polarization and anchoring. Research on polarization analyzes how communication prevents convergence and sustains opposing camps in epistemic communities. O'Connor and Weatherall, building on Zollman's (2007), demonstrate that polarization can arise even when one hypothesis has a payoff advantage. Mäs and Flache (2013) propose an AABM in which bi-polarization can emerge from homophilous exchange of arguments without negative influence. Anchoring, by contrast, is a structural property of sequential deliberation, whereby early speakers can disproportionately affect the outcomes. Hartmann & Rad (2018) show that such order-of-speech bias persists even with (boundedly) rational agents. Braun, Rad, and Roy (2024) extend anchoring to preference aggregation and find that, on average, anchoring is stronger than relative social influence, popularity, or group homogeneity.

2.3 The Model and a Sample Simulation

In this section, I introduce each component of my model and discuss what occurs in a typical simulation round. After outlining the components, I present a sample simulation to show how agents in my model discover evidence, communicate with each other, and update their beliefs.

2.3.1 Modeling Dynamic Discovery of Evidence and Deliberation

My model builds on the one-dimensional space setup from the H-K and the D-W models. In my model, agents are deliberating about two alternatives (one of which is true), each of which is represented by a point on the real line, and an agent's *doxastic state* is represented by their

position on this line. At any given time, the distance between an agent's position and the truth (the true alternative) indicates how far their beliefs deviate from the truth. Similarly, the distance between two agents in the space represents the divergence between their doxastic states. However, in my model, an agent's doxastic state alone does not fully determine their *epistemic position* in seeking the truth.

I follow Dietrich and Spiekermann's (2025) setup of generating evidence. Each piece of evidence is represented by a real number. To capture the observation that people who are close to the truth tend to discover higher-quality evidence, I adopt a two-step procedure: (1) a sample is drawn from a probability distribution (for ease of reference, I call it the *truth's distribution*) whose mean aligns with the position of the truth, and if the sample falls within a certain threshold distance (referred to as the agent's *scope* which describes the ability to discover evidence) to the distance, then the agent learns it as evidence; (2) if the sample falls outside the agent's scope, the agent instead learns a random value within their scope as evidence. Together, for each agent, these two steps create a unique distribution (termed the agent's *evidential distribution*) whose shape (i.e., skewness away from the agent's position/towards the truth) depends on the agent's position. The truth is the underlying source of all information, but the evidence learned by an agent is a distortion (due to the agent's scope and their doxastic divergence from the truth) of the information from the truth. Thus, an agent's scope and position (in the space) together define that agent's epistemic position in evidential discovery. I model evidence discovery stochastically, and an agent's evidential distribution is a representation of the agent's epistemic position in terms of their overall access to evidence[4] (i.e., their tendencies in discovering evidence).[5] To illustrate, suppose the truth

[4] An agent having access to high-quality evidence essentially means that the agent's evidential distribution is sufficiently skewed toward the truth.

[5] Despite different motivations, my model shares a similar mechanism with Fryer et al.'s (2019) model of agents' bias in interpreting evidence and Baccini et al.'s (2023) model of myside bias. In all three models, external information is distorted to some extent.

is located at 0 and generates samples from a normal distribution with a mean of 0 and a standard deviation of 5. Below are two figures showing the distribution of evidence for four agents at points 5 or 20, with scope ranges of 5 and 10, respectively.

For agents with the same scope, those closer to the truth have a distribution more skewed towards the truth (compare Figure 2.1 and 2.3.2). For agents in the same position, those with a larger scope have a more skewed distribution. A more skewed evidential distribution away from the agent's position (towards the truth) indicates being in an epistemic position that is more apt to discover truth-conducive evidence. Notice in Figure 2.2 that the agent whose scope is 5 has a nearly flat evidential distribution, making it unlikely for them to move towards the truth quickly. Overall, a person's evidence is expected to become increasingly truth-conducive as they move closer to the truth.

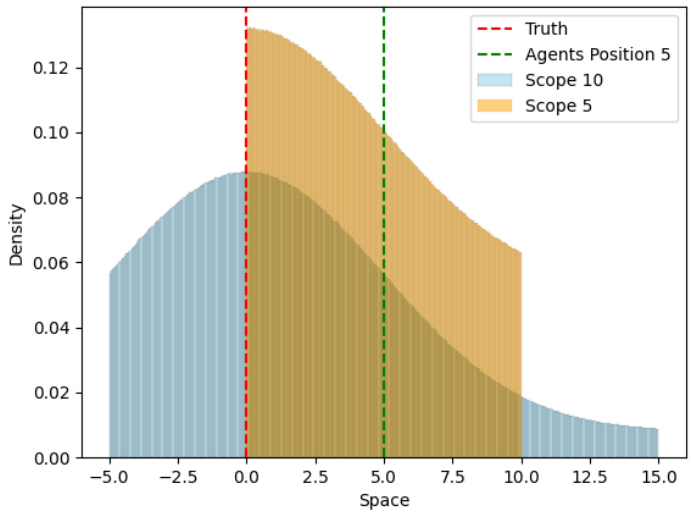

Figure 2.1 Two agents' (located at 5, with a scope of 5 and 10) evidential distributions when the truth is at 0 (red dashed line)

Fryer et al.'s model distorts information to a greater extent than mine, such that, in the long run, agents are likely to converge on their priors. In contrast, in my model, although information is distorted, it still indicates the truth (in some cases, in a minimal sense). Additionally, in Baccini et al.'s model, agents' beliefs are represented by their credence in a binary issue. Learned information is distorted such that if it is close to the agent's belief, it has a stronger impact on the agent's posterior than the raw information would. However, my motivation for including a component that "distorts" information differs from the motivations in the other two models.

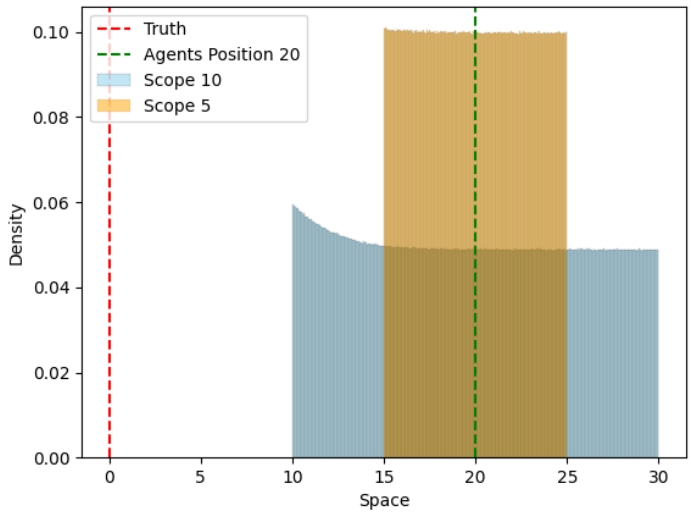

Figure 2.2 Two agents' (located at 20, with a scope of 5 and 10) evidential distributions when the truth is at 0 (red dashed line)

Another key parameter is the standard deviation of the truth's distribution (for simplicity, I refer to it as the standard deviation). The standard deviation describes informational clarity and reflects a tension between accessibility and ambiguity. The standard deviation and the agents' scope together define a learning scenario. The scope and agents' distance to the truth have straightforward effects on evidential distributions. However, the two extremes of the truth distribution's standard deviation are associated with different kinds of challenges. When the standard deviation is low, the truth's distribution is narrow, providing clear but narrowly accessible evidence. In this case, agents with a large scope or that are close enough to the truth will quickly move towards the truth in response to their gathered evidence. In contrast, when the standard deviation is high, the information is spread out in a large range (i.e., ambiguous). Even if agents have large scope or are close to the truth, the evidence they acquire is not likely to push them towards the truth. In this case, accumulating a large amount of evidence is a more efficient way of converging on the truth. Finally, standard deviations in the middle range contribute to an environment that is more conducive to agents' learning. In real-life scenarios, a truth distribution

with a low standard deviation may represent problems that require specialized skills or knowledge, while truth distributions with a high standard deviation correspond to problems that, by their nature, lack clear evidence or are filled with noise, regardless of agents' abilities or positions. Thus, my model captures two distinct aspects of learning: epistemic accessibility, represented by agents' scope and distance from the truth, and informational clarity, represented by the standard deviation of the truth distribution.

In my model, there are two kinds of information that can be communicated: evidence and conclusions. Communication about evidence takes place in one of two modes; within each mode, the extent of communication is characterized by an independent parameter. In the first mode, which I call *random communication of evidence* (RE), agents learn evidence from random other agents in the group. In many cases, communication incurs costs in terms of time, energy, and other resources. Even if agents wish to engage with diversified opinions by interacting with random others, communicating with the entire group may not be practical. So, I characterize the extent of RE by the probability that an agent learns from any given other agent.[6] For example, an extent of $\frac{1}{2}$ implies that, on average, each agent learns evidence from about half of the group members in each round. The second mode, which I call *neighborhood communication of evidence* (NE), restricts communication to agents within a fixed epistemic distance from each other. In many contexts, agents view those with significantly different opinions as unreliable and are less inclined to consider their evidence seriously. So, I define the extent of NE as the maximum distance between agents allowed for learning evidence. The specific parameter I use in neighborhood communication is proportional to the distance between the two alternatives. In my simulation, the

[6] In my simulation, this probability is uniform across all agents in the group.

numbers I adopt are 0.65 and 1.5.[7] For communication about evidence, there are two extreme cases: communicating with everyone else and communicating with no one else. For ease of reference, I call these situations *full communication of evidence* (FE) and *private evidence* (PE), respectively.

Asserting a conclusion essentially acts as a convergence check in my model, where agents stop gathering or updating their beliefs after a certain period of stability. When an agent is close to the truth, newly acquired evidence is unlikely to cause significant changes in their beliefs. So, once an agent's belief remains close to an alternative for several rounds, the agent concludes that the alternative is likely the truth and ceases further investigation.[8] Asserting a conclusion terminates an agent's journey through the space. This process of gathering evidence and then asserting a conclusion can be viewed as analogous to performing repetitive tests before drawing a conclusion. Still, it is possible for an agent to assert a false conclusion. When an agent is far from the truth but relatively close to the false alternative, they may remain close to the false alternative. As illustrated in Figure 2.2, an agent far from the truth may draw evidence from a nearly flat distribution, making it difficult for them to move closer to the truth.

Communication about conclusions is also modeled in two modes. Agents can either randomly learn other agents' asserted conclusions based on a fixed probability (*random communication of conclusions*, RC) or only learn conclusions within a specified distance (*neighborhood communication of conclusions*, NC). The two extremes are learning all the available conclusions (*full communication of conclusions*, FC) and communicating no conclusions (*private conclusions*, PC). An agent's conclusion remains fixed once asserted. I propose and model

[7] The numbers are determined by reference to agents' prior positions. Later, I explain how agents' priors are collectively distributed according to a normal distribution whose standard deviation equals the distance between the two alternatives. In this case, setting the distance at 0.65 times the distance between alternatives means that, on average, an agent's prior is within this range for about one-fourth of the other agents' priors. The number 1.5 indicates that roughly half of the other agents' priors fall within this range for each agent. Details of determining agents' priors are given in Section 2.3.2.

[8] Admittedly, continuously discovering new evidence is beneficial in the long run. However, my model reflects the tendency for agents to tire and lose motivation to pursue the same issue indefinitely, especially when they anticipate little change in their beliefs.

a heuristic: agents disregard conclusions that have been circulating for a long time. The aim is to abandon false conclusions in a timely manner while leaving a true conclusion sufficient time to affect the agents. The heuristic avoids the complexity of evaluating belief changes in response to old conclusions.[9] Together, the mode and extent of communicating evidence and conclusions define a communicative practice. Table 2.1 presents the sixteen practices that my simulation investigates. Each cell shows an abbreviation for the corresponding practice. A blank cell indicates that the combination is not included in the simulation,[10] and a bolded abbreviation indicates that the practice's performance is extensively discussed in this paper.[11]

Finally, for agents in the model, updating involves moving to new positions in the space. Regardless of a particular practice, in each simulation round, an agent gathers all of the evidence and conclusions (if there are any, given the practice) they have acquired into a piece of integrated information. This includes both the evidence the agent has personally acquired and the evidence acquired from other agents (if any) during that round of the simulation. The function that integrates agents' information represents their attitudes to information from various sources. In the long run, agents are generally expected to move towards the truth, indicating that recent information is generally more valuable than old information. So, to update their beliefs, agents refer to information gathered from a certain number of recent rounds (i.e., discount the older ones). Details on these steps are given in the next subsection when I later present a sample simulation.

[9] If a conclusion is true, given sufficient time, it should have already led some agents to the truth. In this case, even if ignoring the conclusion, other agents can learn the *same* content by learning the successors of the early asserted conclusion. The group is not missing any valuable information. If a conclusion is false, ignoring it is beneficial for deliberation.

[10] The major reason for not including the blank cells is that those combinations are not well motivated as plausible deliberative practices. For example, it is hard to imagine why a group wants to fully communicate evidence but only learn conclusions from their neighbors. The practices on the diagonal are motivated by the thought that agents treat evidence and conclusions in the same manner.

[11] I only discuss practices whose performance is representative and helpful in establishing qualitative claims.

Table 2.1 Sixteen Communicative practices investigated in this paper

Evidence \ Conclusions	Private	1/4 Random	1/2 Random	0.65 Neighborh-ood	1.5 Neighborh-ood	Full
Private	**PE-PC**					**PE-FC**
1/4 Random	1/4RE-PC	1/4RE-1/4RC				1/4RE-FC
1/2 Random	1/2RE-PC		1/2RE-1/2RC			1/2RE-FC
0.65 Neighborhood	.65NE-PC			**.65NE-.65NC**		**.65NE-FC**
1.5 Neighborhood	1.5NE-PC				1.5NE-1.5NC	1.5NE-FC
Full	**FE-PC**					**FE-FC**

Numbers before RC, RE, NC, or NE indicate the extent parameter. Abbreviations: FC, full communication of conclusions; FE, full communication of evidence; NC, neighborhood communication of conclusions; NE, neighborhood communication of evidence; PC, private conclusions; PE, private evidence; RC, random communication of conclusions; RE, random communication of evidence. Accordingly, **PE-PC**, private evidence - private conclusions; **PE-FC**, private evidence - full communication of conclusions; **.65NE - .65NC**, .65 neighborhood communication of evidence - .65 neighborhood communication of conclusions; **.65NE - FC**, .65 neighborhood communication of evidence - full communication of conclusions; **FE-PC**, full communication of evidence - private conclusions; **FE-FC**, full communication of evidence - full communication of conclusions.

Here is a summary of what happens to each agent in a typical simulation round, using the practice .65NE-FC to illustrate:

1) Discovering Evidence

A sample is drawn from the truth's distribution. If the sample is within the agent's scope, the agent learns it as evidence. Or, if the sample lies outside of the agent's scope, the agent draws a sample from a uniform distribution centered at the agent's position. [The piece of evidence learned is the agent's personal evidence in this round.]

2) Sharing

- The agent learns evidence from neighbors who are within the specified distance (0.65 times the distance between the two alternatives).
- The agent learns all conclusions that have been asserted.

3) Updating

- The agent integrates their personal evidence and the evidence and conclusions acquired from others into an information pack.
- The agent refers to a certain number of recent information packs and updates their beliefs, thereby moving to a new position in the space.
- The agent checks whether they have been at a small distance from any alternative for a sufficiently long time. If so, the agent announces that alternative as a conclusion and ceases further discovery or updates. If not, the agent repeats steps 1–3 in the next round of the simulation.

My model resembles later extensions (Hegselmann & Krause, 2006; Douven & Riegler, 2010) of the H-K model in including a truth component that actively guides belief updates. In

contrast, my model also has a false alternative conclusion that may trap agents.[12] Within my model, agents in my model are motivated to converge on the truth, but this tendency may nevertheless be disrupted by the presence of a false alternative.

2.3.2 A Sample Simulation

To illustrate my simulations, I present a set of sample "paths" of agents' motions in Figure 2.3. The parameters of the sample simulations apply to my later simulations unless otherwise specified. The two alternatives are randomly placed in the interval of $[-50, 50]$, with the red dashed line indicating the truth and the blue dashed line indicating the false alternative. The truth's distribution is a normal distribution whose mean aligns with the truth, and its standard deviation is set to 10. For simplicity, to integrate available information in each round, agents take an average of them; to determine a movement, agents take an average of the 50 most recent information packs. To assert a particular conclusion, an agent needs to remain within a distance of 1/50 of the distance between the two alternatives for 5 consecutive rounds. Agents only keep track of conclusions made within the last 100 rounds. The difficulty varies non-monotonically with the distance between the two alternatives: both extremely small and extremely large separations pose challenges for agents.[13] To determine prior positions, I divide the agents into two groups of random sizes, with each group scattered according to a normal distribution whose mean aligns with one of the two

[12] Douven and Hegselmann (2021) introduce campaigners, a type of agent that actively broadcasts misleading information. Both campaigners and the false alternative in my model can disrupt agents' convergence to the truth. In my model, however, the false alternative functions only as a reference point for asserting a conclusion; it does not generate any information.

[13] The false alternative influences agents' trajectories as they move toward the truth. The effect depends on the distance between the two alternatives relative to the agents' epistemic position (determined jointly by agents' scope and the truth's standard deviation). The condition for asserting a conclusion (i.e., an agent needs to remain within a distance of 1/50 of the distance between the two alternatives) scales with the distance between the two alternatives. If the alternatives are too far apart relative to the standard deviation or the agents' scope, agents may fail to obtain sufficiently truth-conducive evidence to move toward the truth. In such cases, their movements remain limited, and agents who happen to be near the false alternative are likely to mistakenly treat it as their conclusion. Conversely, if the distance is small relative to the standard deviation, agents near the truth may not receive evidence strong enough to remain anchored near it. These dynamics will be examined in detail in the simulation results presented in Section 2.4.

alternatives. Thus, at the outset, each alternative has a cluster of agents positioned around it. The standard deviation of each agents-positioning distribution is equal to the distance between the two alternatives.[14] I examine a .65NE-FC group with 9 agents, with green crosses showing each agent's priors and red dots representing their final positions. Agents' scopes are uniformly set to 10. The simulation runs until 1,500 rounds or until all agents have reached a conclusion.

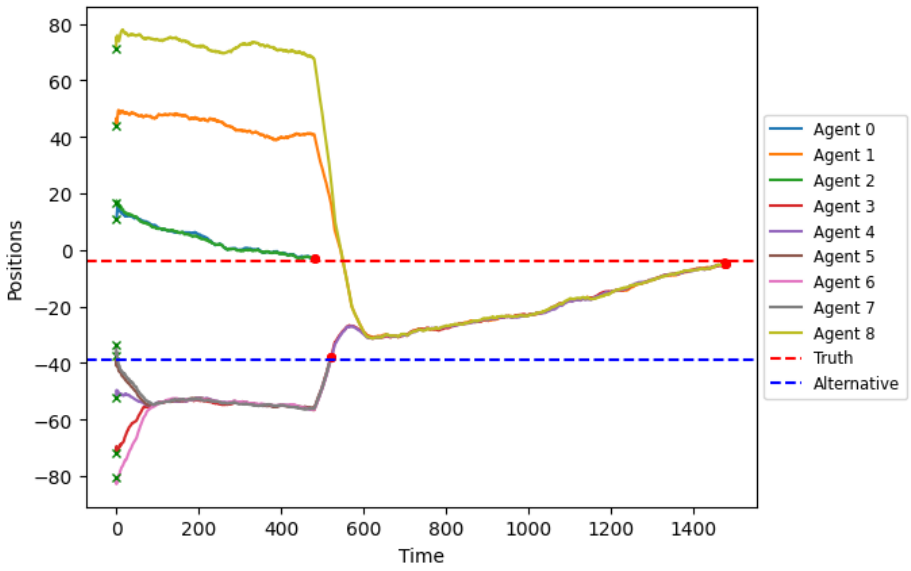

Figure 2.3 Sample paths for a group of nine agents in the neighborhood communication of evidence and full communication of conclusions, .65NE-FC.

After approximately 1,400 rounds, all agents have reached a conclusion. Under NE, agents polarize into subgroups. Once a subgroup forms, agents in the subgroup *move together*.[15] This sample also demonstrates the impact of communication about conclusions. Agent 2 and Agent 0, whose priors are closest to the truth, find the truth first, and their conclusion effectively pulls everyone else towards the truth. However, the majority of the subgroup at the bottom of the figure, unfortunately, is trapped by the false alternative, where they assert a false conclusion. One

[14] This ensures that agents' priors are moderately biased towards either alternative.
[15] Polarization in my model takes place in a manner largely consistent with O'Connor and Weatherall's (2018) analysis. However, my model does not consider anti-updating.

exception in that subgroup is Agent 4, who joins the subgroup with Agent 1 and Agent 8. With five agents asserting the false alternative and only two recognizing the truth, the false conclusion dominates within the group and draws the middle subgroup towards the false alternative (around steps 550–600). Eventually, the middle subgroup converges on the truth as they discard old conclusions while learning from current evidence.

2.4 Evaluating Practices of Communication

In this section, I investigate the general effects of various practices of communication about evidence, conclusions and their respective interactions. To achieve this goal, I examine the performance of groups of agents under various combinations of the agents' scope and the standard deviation of the truth's distribution. In each simulation, the two alternatives are randomly positioned to ensure that the results do not depend on any specific placement of the alternatives. I adopt a 50×50 parameter space, in which each parameter can take an integer value from 1 to 50. A group consists of an odd number of agents from 7 to 21. To evaluate a group's performance, I observe whether the majority of the group's members succeed in finding the truth within a certain period of time.[16] I record each groups' performance after 750, 1,500, and 3,000 steps,[17] which is taken as an exogenous variable. I use data from 1,500 steps to support my main claims and data from simulations with shorter (750 steps) and longer (3,000 steps) stopping times to analyze the effect of different deliberation durations. At each of the 2,500 parameter settings, I run 1,200 simulations for each of the 16 practices of communication. I place the two alternatives randomly

[16] Although agents in my model are always expected to find the truth in a finite amount of time, in real-life scenarios, attaining a truth more quickly is important for saving resources. Such concerns are affirmed by Kummerfeld and Zollman (2016). My paper has different interests from their paper. However, the idea of evaluating performance under a certain time frame is solid (Zollman, 2010).

[17] The groups' performance stabilizes around 1,500 steps, at the highest value of scope, for the majority (except for PE-PC and .65NE-PC) of practices.

and generate agents' initial positions relative to those alternatives (see Section 2.3.2 for details), so that the results are not driven by any particular arrangement of agents' initial positions. Admittedly, this setup abstracts away from potentially interesting differences that might arise from specific configurations of alternatives or initial agent positions.[18]

Through comparing the performance of two relevant practices, I can trace the difference in performance to the differences in the two practices. For example, by comparing PE-FC to FE-PC, I attribute performance differences to the effects of communicating the two different kinds of information. The rest of this section focuses on two primary aspects: the differing effects of communicating evidence versus conclusions and the interactions between communicating these two kinds of information. Additional findings are provided in Appendices A and B. To avoid redundancy, I only present the most representative data, though the following claims are robust across all the data and apply broadly to the relevant kinds of communication.

2.4.1 Communication About Evidence vs. Communication About Conclusions

Evidence represents timely discoveries made by agents and is expected to become more truth-conducive over time. In contrast, conclusions differ from evidence in three important aspects. First, an agent can assert a conclusion only if they have been sufficiently close to an alternative for five rounds. Hence, asserting conclusions, unlike sharing evidence, reflects agents' confidence and requires a period of consideration (below, I refer to this as the prudential feature of conclusions). Second, a conclusion does not evolve once agents announce it. Because of this static nature, if an asserted conclusion happens to be false, its negative effect on group dynamics is relatively long-

[18] Future work could explore alternative ways of positioning agents, such as assigning initial positions independently of the alternatives (e.g., uniformly across the space), or using skewed distributions that favor or disfavor particular alternatives.

lasting. Third, communication about evidence takes place throughout the simulation round, but communication about conclusions takes place only after agents have asserted their conclusions.

To demonstrate the differing effects of communication about evidence versus conclusions, Figure 2.4 compares the practice PE-FC against FE-PC. The black curves (solid and dashed) are iso-value contour lines of the performance difference at 10, 0, and −10. Areas to the left and below the contour labeled 10 (dark gray) indicate regions (roughly, Regions I and IV) where PE-FC outperforms FE-PC by more than 10%. The light gray area between the contours labeled 10 and 0 (roughly Region V) marks cases where PE-FC still performs better than FE-PC, but by less than 10%. By contrast, the light blue area between the contours labeled 0 and −10 (roughly Region III) corresponds to cases where PE-FC underperforms FE-PC by less than 10%, while the dark blue area to the upper right of the contour labeled −10 (roughly Region II) indicates regions where PE-FC underperforms FE-PC by more than 10%. Using the contour lines, I draw the colored dash-dotted boundaries to delineate five interpretive regions that guide the discussion of the results.

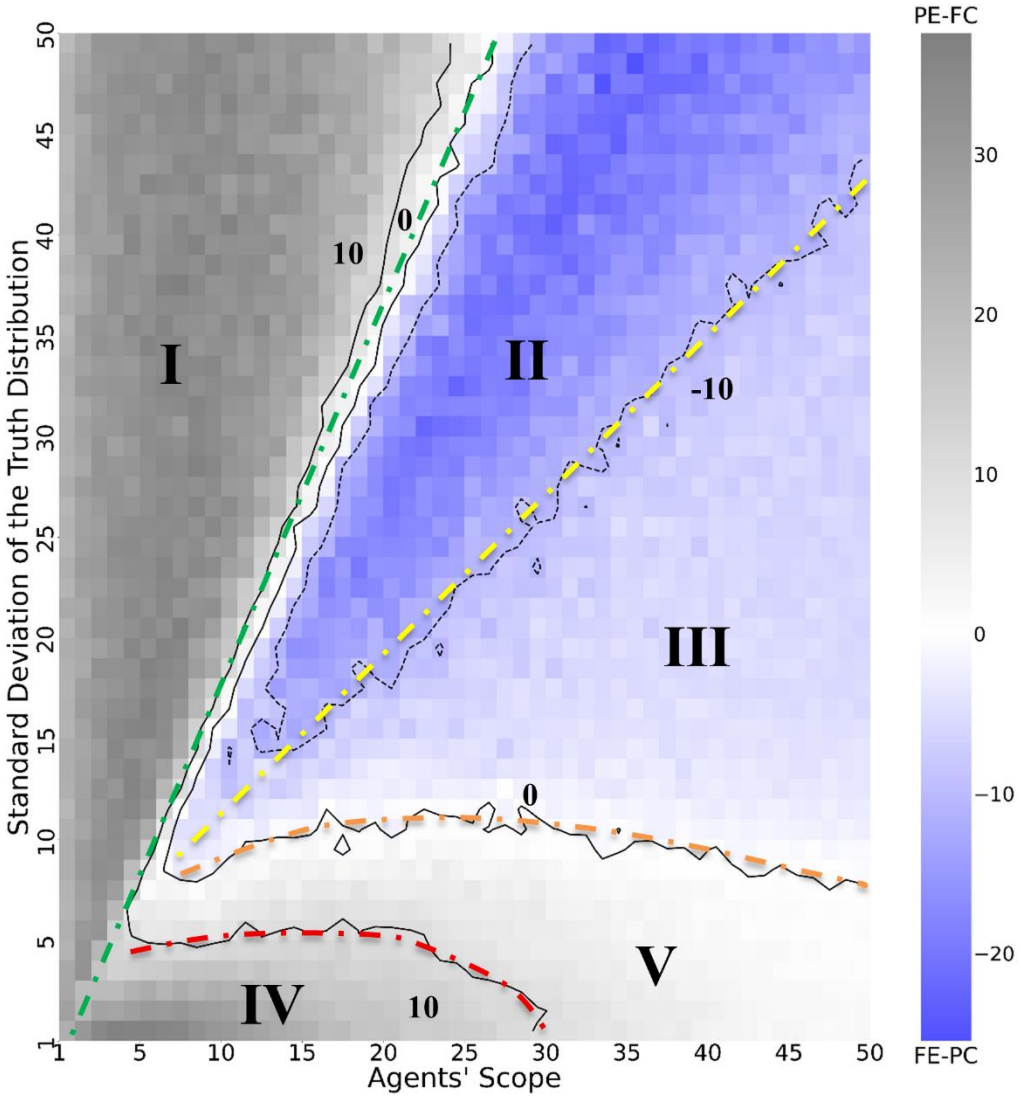

Figure 2.4 Number of groups in which the majority finds the truth: PE-FC compared to FE-PC at 1,500 steps. Gray indicates that PE-FC has more groups whose majority finds the truth than FE-PC, and blue indicates that FE-PC has more groups whose majority finds the truth than PE-FC. Lightness indicates the magnitude of performance difference. Abbreviations: FC, full communication of conclusions; FE, full communication of evidence; PC, private conclusions; PE, private evidence. The black curves (solid and dashed) show iso-value contours of the performance difference at 10, 0, and -10. The colored dash-dotted lines use these contours to delineate five interpretive regions.

In Figure 2.4, a gray cell means that PE-FC outperforms FE-PC, and a blue cell means that FE-PC outperforms PE-FC.[19] For each practice, I calculate the percentage (out of 1,200

[19] For a heatmap comparing two practices, a gray cell means that the one coming in the first place outperforms the second; a blue cell means the opposite.

simulations) of groups in which the majority finds the truth. The difference in these percentages between the two practices determines the color lightness in each cell. In general, because the comparison is between FE-PC and PE-FC, the color also reflects the comparative quality of evidence versus conclusions; for example, blue cells indicate parameter combinations in which evidence is, on average, more truth-conducive than conclusions.

Region I marks a triangular gray area in which agents' scopes are "small" relative to the corresponding standard deviations. The slope of the green boundary illustrates this comparative relationship rather than a fixed threshold. The solid gray color suggests that agents are in epistemic positions with limited access to high-quality evidence,[20] because they lack sufficient scope to "see" evidence drawn from a sufficiently wide range. By contrast, the prudential feature of conclusions makes, on average, communicated conclusions better indicators of the truth than the evidence exchanged under FE-PC.

Except for extremely low standard deviations (roughly below 8),[21] as agents' scopes increase relative to the standard deviation (moving horizontally in Figure 2.4), the cells turn blue (Region II and Region III), which suggests a relative advantage of FE-PC over PE-FC. In both Regions II and III, agents have sufficiently broad epistemic access to obtain truth-conducive evidence on average. Thus, unlike in Region I, frequent exchange of evidence does not harm overall accuracy. Furthermore, FE has two features that are potentially beneficial in Regions II and III. First, by accumulating a larger body of evidence, frequent exchange of evidence smooths

[20] As in Section 2.1, I use the terms *low-quality* and *high-quality* evidence to describe the body of accessible evidence relative to an agent's epistemic position. Although it is possible to quantify evidence quality under each parameter setting, such calculations add little to the qualitative analysis of simulation outcomes. Instead, I make retrospective qualitative judgments, labeling the evidence as low- or high-quality on the basis of comparative results. For instance, in Figure 2.4, the gray area (where PE-FC outperforms FE-PC) illustrates scenarios in which agents lack access to high-quality evidence.

[21] Because this model is hypothetical, the precise parameter values have no physical interpretation; they are meaningful only in relation to one another. Whenever I cite specific numbers, it is solely for ease of reference to a specific Roman-numbered region in a figure, not to highlight their intrinsic significance. Moreover, when describing an area, I will specify whether the relevant scope and standard deviation are high or low, and whether the epistemic position provides access to high- or low-quality evidence.

out ambiguity in evidence and randomness in evidence gathering, thereby bringing the average of the actually learned evidence closer to the expected value of the agents' evidential distributions[22] (an instance of the law of large numbers). Second, unlike static conclusions, the evidence available under FE continues to improve as agents approach the truth. The superiority of FE-PC in these regions indicates that dynamically acquiring increasingly truth-conducive evidence can be more beneficial than exclusively waiting for conclusions.[23] The green boundary suggests that the more the truth's distribution spreads out (in the vertical direction of Figure 2.4), agents require proportionally larger scopes to benefit from accumulating additional evidence, i.e., from (scope 7, standard deviation 10) to (scope 25, standard deviation 50).

In contrast to Region III, the V-shaped Region II marks a set of situations in which scopes are moderately large (not too large) relative to the corresponding standard deviation. The agents' evidential distributions are only modestly skewed toward the truth and retain substantial variance. As a result, any individual piece of evidence may still be misleading with a relatively high probability. For example, with a standard deviation of 30 and a scope of 20, in a single inquiry, an agent located 25 units from the truth faces a 0.4872 probability of obtaining misleading evidence.[24] Given the setup, moving farther away from the truth places an agent in an even less favorable epistemic position, making recovery increasingly unlikely; conversely, moving closer makes

[22] Recall from Section 2.3.1 that an agent's evidential distribution refers to the distribution of evidence they tend to acquire given their epistemic position.

[23] This comparison focuses on conclusions versus the group's evidence around the same time. In this context, only two possible conclusions exist: the truth and the false alternative. Theoretically, no evidence is more truth-conducive than a true conclusion. However, agents are fallible and may sometimes arrive at a false conclusion. With only a small number of active conclusions at any given time, a single false conclusion can significantly lower the overall quality of conclusions. Moreover, a false conclusion will not be corrected, while evidence is generally expected to evolve. Hence, when agents' scope is reasonably high but not excessive, the overall quality of evidence may surpass that of conclusions at the moment when certain conclusions have been asserted.

[24] The standard deviation of 30 and a scope of 20 place this example in one of the darkest blue areas of Figure 2.4. A distance of 25 from the truth is not particularly large, given that the two alternatives are randomly positioned in the interval $[-50, 50]$. The probability quoted here is estimated using 10^{10} samples. This probability increases as the distance between the agent and the truth grows.

further progress easier. In such cases, FE's ability to reduce the chance of moving away from the truth at each step becomes especially important. Lastly, the lightness of blue fades in Region III where agents' scopes are extremely large relative to the corresponding standard deviations. In these cases, the evidence accessible to agents is already of very high quality, which also contributes to accurate conclusions. Thus, overall performance remains strong regardless of the kind of information being exchanged, and the difference between FE-PC and PE-FC naturally narrows. The fact that FE-PC still slightly outperforms PE-FC reflects the persistent, though diminished, benefits of accumulating evidence. The yellow boundary conveys a similar relationship to the green one: as the standard deviation increases, proportionally larger scopes are required for the performance difference between the two modes to shrink.

Lastly, when the standard deviation is extremely low (Regions IV and V), increasing scope does not produce a transition from gray to blue. This may seem puzzling at first glance, but it follows directly from the behavior of the truth's distribution. A distribution with an extremely low standard deviation generates evidence that is narrowly concentrated and therefore less accessible to agents with limited scope.[25] Meanwhile, the information from the truth tends to be highly precise and unambiguous. Consequently, there leaves little room for frequent exchange to smooth out ambiguity or provide any substantial advantage. As a result, PE-FC consistently dominates FE-PC throughout Regions IV and V. Furthermore, when scope becomes very large relative to this small standard deviation, the difference between PE-FC and FE-PC gradually decreases. This is

[25] Although Region IV includes scopes of up to 30, which may not appear small relative to the overall scale of 50, agents are not in a favorable epistemic position given the extremely low standard deviation. For example, with a standard deviation of 1 and a scope of 30, an agent located 25 units from the truth faces approximately a 0.5000 probability of obtaining misleading evidence in a single inquiry (5000036127 non-misleading samples out of 10^{10}). Moreover, an agent's scope is defined as the length of the entire interval within which information can be received from the truth. In this one-dimensional setup, this means that an agent can only learn evidence within 15 units (half the scope) of their position. From this perspective, a scope of 30 is not as large as it might initially appear.

consistent with the overall analysis: (1) both evidence and conclusions are already of high quality, and (2) there is minimal ambiguity for FE to smooth out.

In addition, as the duration of deliberation increases, the blue areas expand. This is consistent with the assumption that agents are more likely to encounter truth-conducive evidence as they move closer to the truth. With more time, agents have more opportunities to improve their epistemic positions, so they tend to accumulate evidence of higher quality. Under these conditions, FE-PC gains a stronger advantage, because the benefits of learning a larger body of evidence become more significant.

> **Claim 2.1** In general, learning all asserted conclusions (PE-FC) improves groups' performance when agents lack access to high-quality evidence (due to either relatively small agents' scope or the truth's low standard deviation, shown in Regions I and IV). When agents have access to high-quality evidence, communicating evidence results in better performance.

If a group, on average, lacks access to truth-conducive evidence, communication about conclusions improves the quality of accessible information. After all, a conclusion is something that an agent feels reasonably confident about. If that agent is lucky (e.g., happens to be close to the truth or learns something truth-conducive in a small probability event), learning a conclusion from the lucky pioneer who discovered the truth essentially improves the quality of information that the group can access. Correspondingly, Claim 2.1 can be phrased more precisely with respect to the quality of available information.

> **Claim 2.1′** When evidence is of low quality, communication about conclusions improves the group's overall access to high-quality information (as shown in Regions I and IV). In contrast, communicating evidence accumulates a larger body of evidence, which is

especially beneficial when the available evidence is expected to be truth-conducive but ambiguous (as shown in Region II).

2.4.2 Interactions Between Communication about Evidence and Conclusions

To illustrate the interaction between communication about the conclusions versus evidence, I first examine the effects of adding communication about one kind on top of full communication about the other kind. Here, I present two comparisons: FE-PC vs. FE-FC and PE-FC vs. FE-FC.

I begin by examining the effects of adding communication about conclusions on top of FE by comparing FE-PC to FE-FC (see Figure 2.5). Communication about evidence is a more extensive form of communication than that of communication about conclusions. Furthermore, the evidence shared by an agent shortly before they announce a conclusion should be close to the conclusion to be announced. So, extensively exchanging evidence still takes account of the asserted conclusions but in a way that limits their impact. However, when agents lack access to high-quality evidence (i.e., where limited scope combines with comparatively large standard deviations, shown by the blue area in Figure 2.5), the drawbacks of such a limitation become significant. In such cases, FE-FC outperforms FE-PC, though the performance difference is relatively small compared to how much PE-FC outperforms FE-PC (in Figure 2.4).[26]

Claim 2.2 In most cases, groups that include communication about conclusions (FE-FC) do not perform significantly better than groups that already have FE (FE-PC), except when agents lack access to high-quality evidence; in that case, additional communication about conclusions improves groups' performance.

[26] The gray area in Figure 2.4 indicates that the performance difference can exceed 30%, whereas the difference shown in the blue area of Figure 2.5 is consistently below 10%.

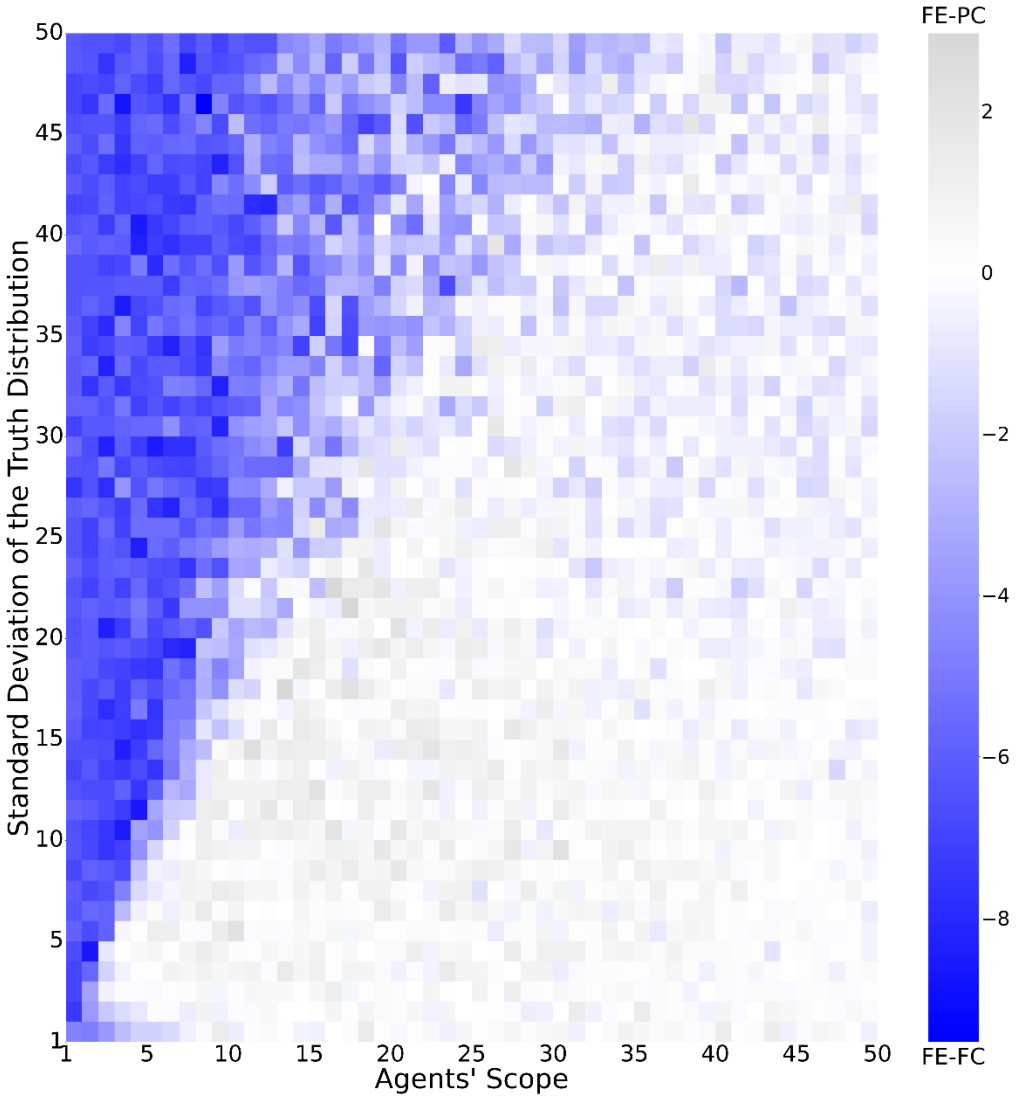

Figure 2.5 Number of groups in which the majority finds the truth: FE-PC compared to FE-FC at 1,500 steps. Gray indicates that FE-PC has more groups whose majority finds the truth than FE-FC, and blue indicates the opposite. Lightness indicates the magnitude of performance difference. Abbreviations: FC, full communication of conclusions; FE, full communication of evidence; PC, private conclusions.

Next, the relative performance of PE-FC and FE-FC shown in Figure 2.6 looks very similar to that of PE-FC and FE-PC shown in Figure 2.4. According to Claim 2.2, the difference between FE-PC and FE-FC is minimal, except in scenarios where agents lack access to high quality evidence. Thus, the overall pattern in Figure 2.6 mirrors that in Figure 2.4, and the explanations for Figure 2.4 also apply to the pattern in Figure 2.6. I derive two claims from the data.

Claim 2.3 When agents do not have access to high-quality evidence (due to relatively small agents' scopes or the truth's low standard deviation, shown by the gray area in Figure 2.6), fully communicating evidence dilutes the benefits of learning all the asserted conclusions. Furthermore, I generalize the comparison to a tension between two kinds of information of different quality.

Claim 2.3′ In general, adding information of low-quality, e.g., evidence that is not truth-conducive, undermines groups' performance.

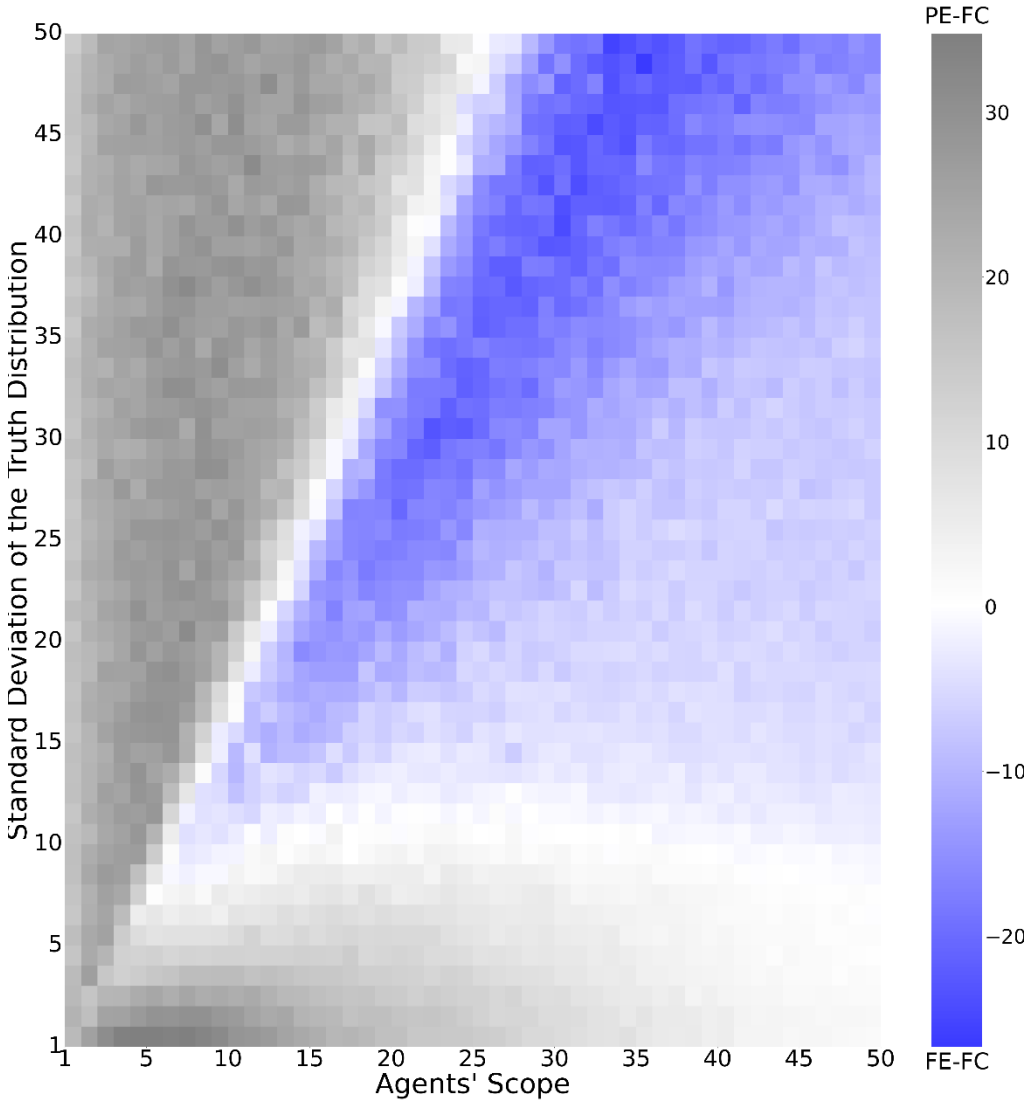

Figure 2.6 Number of groups in which the majority finds the truth: PE-FC compared to FE-FC at 1,500 steps. Gray indicates that PE-FC has more groups whose majority finds the truth than FE-

FC, and blue indicates the opposite. Lightness indicates the magnitude of performance difference. Abbreviations: FC, full communication of conclusions; FE, full communication of evidence; PE, private evidence.

2.4.3 Restricting Communication about Evidence

For further investigation of the two modes of evidence communication, I examine the effects of varying their extent. The impact of restricting RE is straightforward: except at extremely low levels, different extents of RE have no significant effect on group performance.[27] By contrast, I demonstrate the effect of restricting NE by comparing .65NE-FC to FE-FC in Figure 2.7. Overall, when agents learn all asserted conclusions (FC), restricting communication about evidence to one's neighborhood (.65NE) yields better performance than extensive evidence exchange (FE) in most situations, unless the available information carries substantial ambiguity.

[27] See Appendix A for more details.

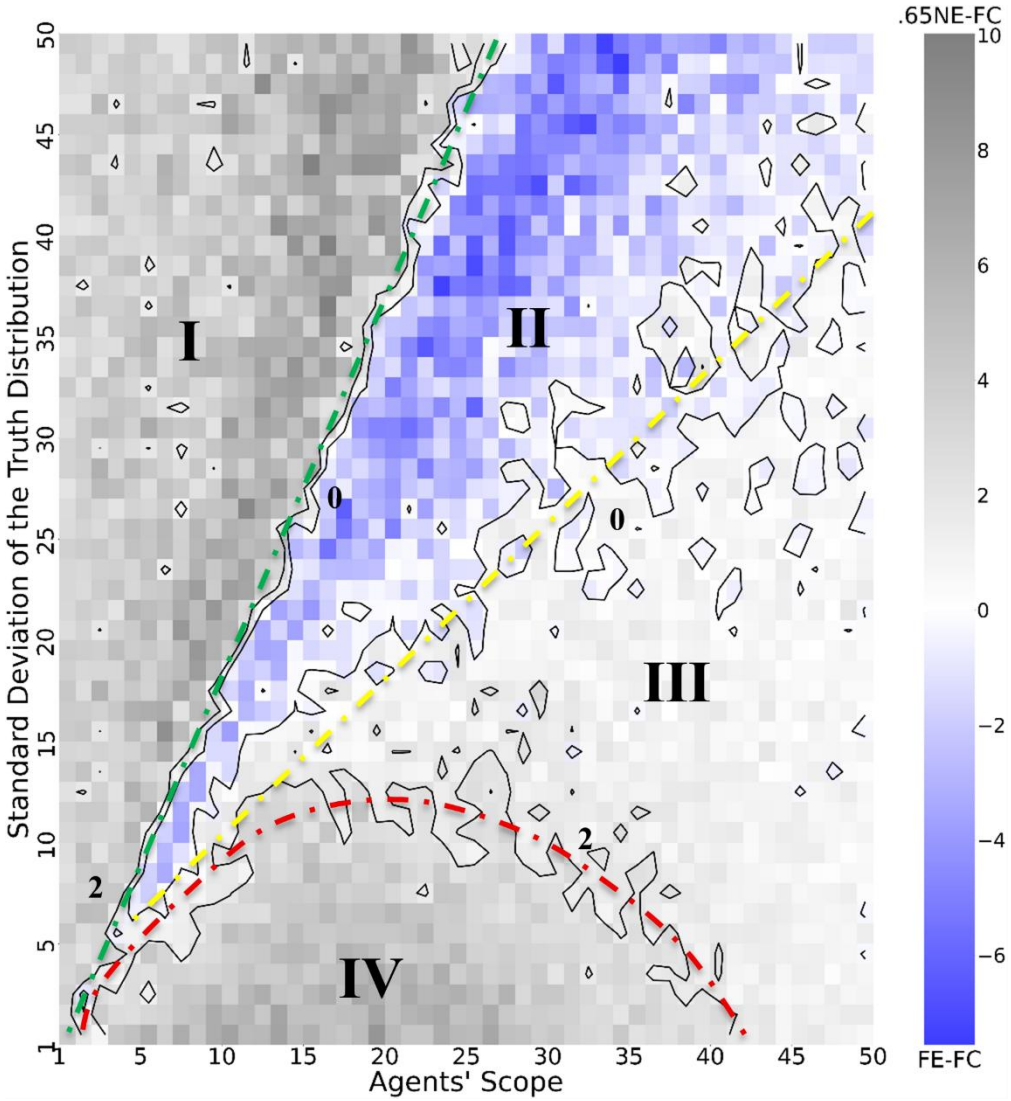

Figure 2.7 Number of groups in which the majority finds the truth: .65NE-FC compared to FE-FC at 1,500 steps. Gray indicates that .65NE-FC has more groups whose majority finds the truth than FE-PC, and blue indicates the opposite. Lightness indicates the magnitude of performance difference. Abbreviations: FC, full communication of conclusions; FE, full communication of evidence; NE, neighborhood communication of evidence. The black curves (solid and dashed) show iso-value contours of the performance difference at 0 and 2. The colored dash-dotted lines use these contours to delineate four interpretive regions.

In Regions I and IV, where agents lack access to evidence of high quality,[28] restricting communication about evidence to a small epistemic neighborhood performs better than extensively

[28] Similar to Region IV in Figure 2.4 (see Footnote 25), an agent's scope of 40 is not sufficient to place agents in a favorable epistemic position given the extremely low standard deviation of 1. For illustration, with a standard deviation of 1 and a scope of 30, an agent located 25 units from the truth faces approximately a 0.5000 probability of obtaining misleading evidence in a single

exchanging evidence. Unsurprisingly, limiting the transmission of low-quality evidence mitigates the potential negative effects of learning from such information. More specifically, under communication about conclusions, it is important for the first asserted conclusion to be correct and for the first correct conclusion to emerge early because the first asserted conclusion has a significant impact on the group's deliberation.[29] With NE, agents polarize into subgroups with their neighbors. Agents whose prior positions are close to the truth may rapidly find the truth and share their findings with the entire group. In contrast, FE spreads information across the entire group and forces agents to cluster, which ruins the lucky agents' (whose priors are close to the truth) chance of finding the truth. Thus, extensive communication about evidence in FE-FC leads to its general loss against .65NE-FC. This loss highlights that extensively accessing diversified information is not necessarily beneficial.

Region II in Figure 2.7 can be explained in a similar fashion to Region II in Figure 2.4. FE gains a substantial advantage when the need to smooth ambiguity in information is primary: moderately large scopes (relative to the corresponding standard deviation) ensure a truth-skewed evidential distribution, but that distribution still has considerable variance. The light color in Region III is also consistent with the previous analysis: as agents' scopes become significantly large relative to the corresponding standard deviation, there is minimal difference between the average quality of evidence and conclusions. However, the upper portion of Region III warrants attention. When the truth's standard deviation is not extremely low (roughly above 10), continuously increasing agents' scopes (moving horizontally from Region II to Region III in Figure 2.7) causes the cells to shift from blue to gray rather than remaining blue. In Figure 2.4,

inquiry (5000037301 non-misleading samples out of 10^{10}).

[29] This occurs because the first asserted conclusion has no other competing ones. For further details, see Appendix B. In the literature, this phenomenon is referred to as *anchoring* (Hartmann & Rad, 2019; Braun, Rad, & Roy, 2024).

Region III reflects FE's superiority due to the benefits of accumulating large amounts of evidence. By contrast, the faint gray in the upper part of Region III in Figure 2.7 suggests that this advantage becomes less important than the ability of .65NE to preserve diversified subgroups.

> **Claim 2.4** Communicating diversified evidence prompts a compromise among the entire group of agents. The compromise impairs the performance of agents whose priors are close to the truth and may, consequently, undermine the group's overall performance.[30]

Moreover, groups that practice 0.65NE-PC or 1.5NE-PC enjoy a performance advantage over the practice of not communicating any information (PE-PC). However, depending on the extent parameter, which characterizes the neighborhood distance, the effect may be less significant than that seen in RE groups (i.e., XRE-PC compared to PE-PC). Any performance advantage from NE also decays as the maximum allowed steps increase.

2.4.4 Restricting Communication about Conclusions

There remains a question: if agents communicate evidence in a certain manner, why do they treat conclusions differently? This question is particularly salient for NE. Namely, if agents refuse to learn evidence sufficiently far from their positions, they will deny learning conclusions far from themselves as well. When agents are communicating conclusions, Claim 2.4 shows that restricted communication about evidence could lead to better results than FE. Besides the conceptual question, more important is the normative question: how or when should agents learn some but not all of the available conclusions? To answer this question, in my simulation, I examine communicative practices 1/4RE-1/4RC, 1/2RE-1/2RC, .65NE-.65NC, and 1.5NE-1.5NC.

[30] This claim appears to contradict the well-known assertion that groups of diverse problem solvers can outperform groups of high-ability problem solvers (Hong & Page, 2004). However, Grim et al. (2019) criticize that the notion of ability used by Hong and Page, arguing that it is not a robust representation of ability when diversity is said to trump so-called ability. Although Hong and Page's model is different from mine, under the situations where the notion of ability is solid, Grim et al. derive a conclusion similar to the one in my paper with a firm notion of ability.

I begin with varying extents on RC. As suggested by Claim 2.2, the extra restricted communication about conclusions still makes 1/4RE-1/4RC (or 1/2RE-1/2RC) outperform 1/4RE-PC (or 1/2RE-PC) when agents do not have access to high-quality evidence (roughly Regions I and IV in Figure 2.4 and Figure 2.7). Comparing 1/4RE-1/4RC to 1/4RE-FC shows that 1/4RE-1/4RC lags in performance in the scenarios where communication about conclusions has a positive impact (again, roughly Regions I and IV in the same two figures). This is because 1/4RE-1/4RC restricts the positive impacts of communication about conclusions on a group's performance. By contrast, the difference in performance between 1/2RE-1/2RC and 1/2RE-FC is negligible across the entire parameter space. This shows that 1/2RE-1/2RC is sufficient to achieve the performance benefits of FC, while a lower value for the extent parameter, such as 1/4, restricts the potential impact of conclusions.

The effects of restricting communication about conclusions by epistemic distance deserve more attention. Figure 2.8 compares .65NE-.65NC and .65NE-FC. The restriction on communication about conclusions in .65NE-.65NC significantly undermines the potential benefits of learning all asserted conclusions.

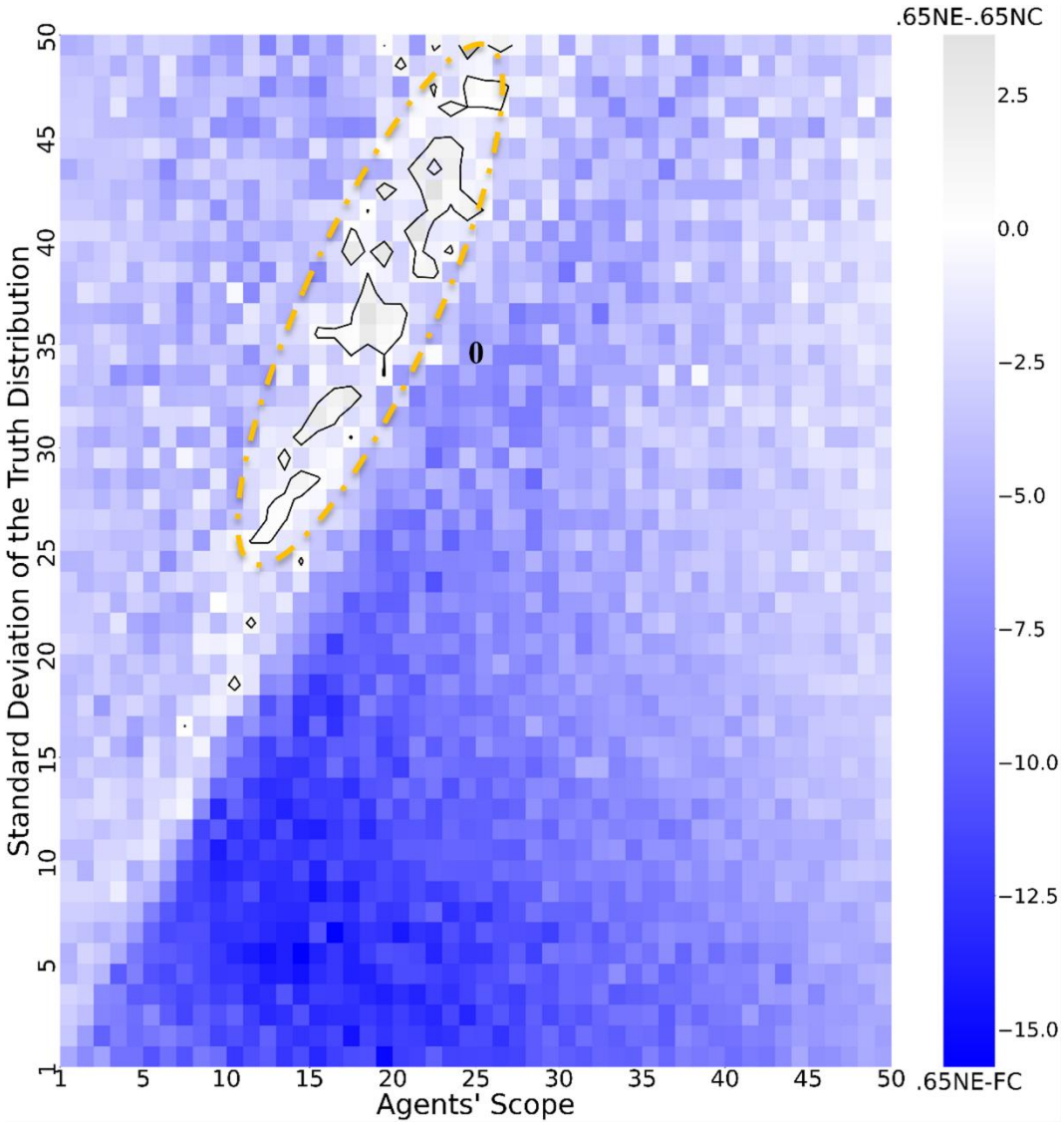

Figure 2.8 Number of groups in which the majority finds the truth: .65NE-.65NC compared to .65NE-FC at 1,500 steps. Gray indicates that .65NE-.65NC outperforms .65NE-FC, and blue indicates that .65NE-FC outperforms .65NE-.65NC. Lightness indicates the magnitude of performance difference. Abbreviations: FC, full communication of conclusions; NC, neighborhood communication of conclusions; NE, neighborhood communication of evidence. The black curves show iso-value contours of the performance difference at 0.

Unlike RC or FC, NC prevents agents from accessing conclusions that are much different from their positions. Conclusions shared through FC can help agents overcome local epistemic limitations when they lack access to high-quality evidence, whereas conclusions shared through .65NE-.65NC do not differ substantially from the evidence that agents already learn.

Although .65NE-.65NC generally performs worse than .65NE-FC, there is a small region in which .65NE-.65NC has a slight advantage (the circled region in Figure 2.8), with a margin of less than 3%. The small region[31] reflects a delicate balance between the qualities of the information exchanged through the two modes: the evidence available under .65NE-.65NC more reliably indicates the truth than the conclusions propagated through .65NE-FC[32]

> **Claim 2.5** Compared with learning all asserted conclusions, restricting communication about conclusions (whether by frequency or epistemic distance) almost never improves a group's performance. There is an exception in practice .65NE-.65NC: in a narrow set of scenarios, groups benefit slightly through .65NC when the quality of the evidence available under .65NE exceeds the quality of conclusions that would have been shared under .65NE-FC.

2.5 Identifying Optimal Practice of Communication

Although a particular practice's performance can often be explained in terms of claims from Section 2.4, mapping the overall patterns reveals deeper insights into information use under various situations. Figure 2.9 presents the best-performing practices in each of the 2,500 cells of the parameter space.[33] To avoid complicating the plot with too many colors, I omit those practices

[31] Of the 2,500 cells, .65NE-.65NC outperforms .65NE-FC in 59 cells, with only 10 of those showing a margin greater than 2%.

[32] In this comparison, both practices involve communicating evidence within neighborhoods, leading agents to polarize into subgroups. When the quality of evidence is significantly higher than that of conclusions, agents benefit from avoiding conclusions that are distant from their positions. For fixed standard deviations, increasing agents' scopes raises the quality of both accessible evidence and accessible conclusions, though not always at the same rate. To the left of the gray area, .65NE-FC performs better because agents lack access to evidence of high quality; to the right, learning all conclusions is beneficial because FC preserves any early accurate conclusions. The scenarios in which .65NE-.65NC performs slightly better indicate that, in those cases, additional conclusions from FC contribute more noise than benefit.

[33] Beyond whether the majority is correct, other evaluators of group performance may also be of interest. A detailed investigation of such evaluators deserves a separate paper. In Appendix D, I include maps of the best-performing practice among the four considered with respect to two additional evaluators: whether the entire group converges on the truth and the number of agents who find the truth. While these measures may not be the most compelling, they are illustrative of how group performance in my model can be assessed using alternative criteria.

whose performance is known to largely match or underperform another.[34] The practices being compared are PE-FC, FE-FC, .65NE-.65NC, and .65NE-FC. Each cell is colored according to which practice performs the best, and the size of the dots indicates the difference between the best-performing practice and the average performance among the four practices.

Table 2.2 cites claims from Section 2.4 to explain each region in Figure 2.9. In challenging scenarios where agents lack access to high-quality evidence (represented by Region I), PE-FC performs the best, no matter whether the difficulties are solely due to the low accessibility of truth-conducive evidence or due to both the low accessibility and ambiguity of evidence. PE-FC promotes relatively high-quality conclusions while excluding low-quality evidence. When agents' scopes are moderately large relative to the corresponding standard deviation, and the standard deviation is not extremely low (Region III), FE-FC takes advantage of accumulating a large amount of evidence in decent quality. Region IV corresponds to the most truth-conducive situations, where agents' large scopes (relative to the corresponding standard deviations) alleviate concerns of both accessibility and ambiguity. Restricted communication about evidence (e.g., .65NE-FC) limits agents' access to diversified evidence and enhances the polarized subgroups' speed in finding the truth given limited time. Indeed, in Region IV, the performance differences between FE-FC and .65NE-FC are small.[35] Finally, .65NE-.65NC appears only in a narrow band between Regions I and III, and its advantages over other practices are minor.

[34] My results rely on a limited number of simulations. For practices that are very close to each other, randomness in simulations may lead to a mosaic-looking pattern in certain regions, which makes it hard to identify a consistent pattern. Between FE-PC and FE-FC, I only include FE-FC. According to Claim 2.2 and Figure 2.5, for the scenarios where the performances of those two practices are significantly different, PE-FC performs significantly better. Hence, it does not really matter whether FE-PC or FE-FC is included in the figure.

[35] The dots are not small because the dots measure the difference between the best-performing practice and the average. In this region, PE-FC performs poorly and brings down the average.

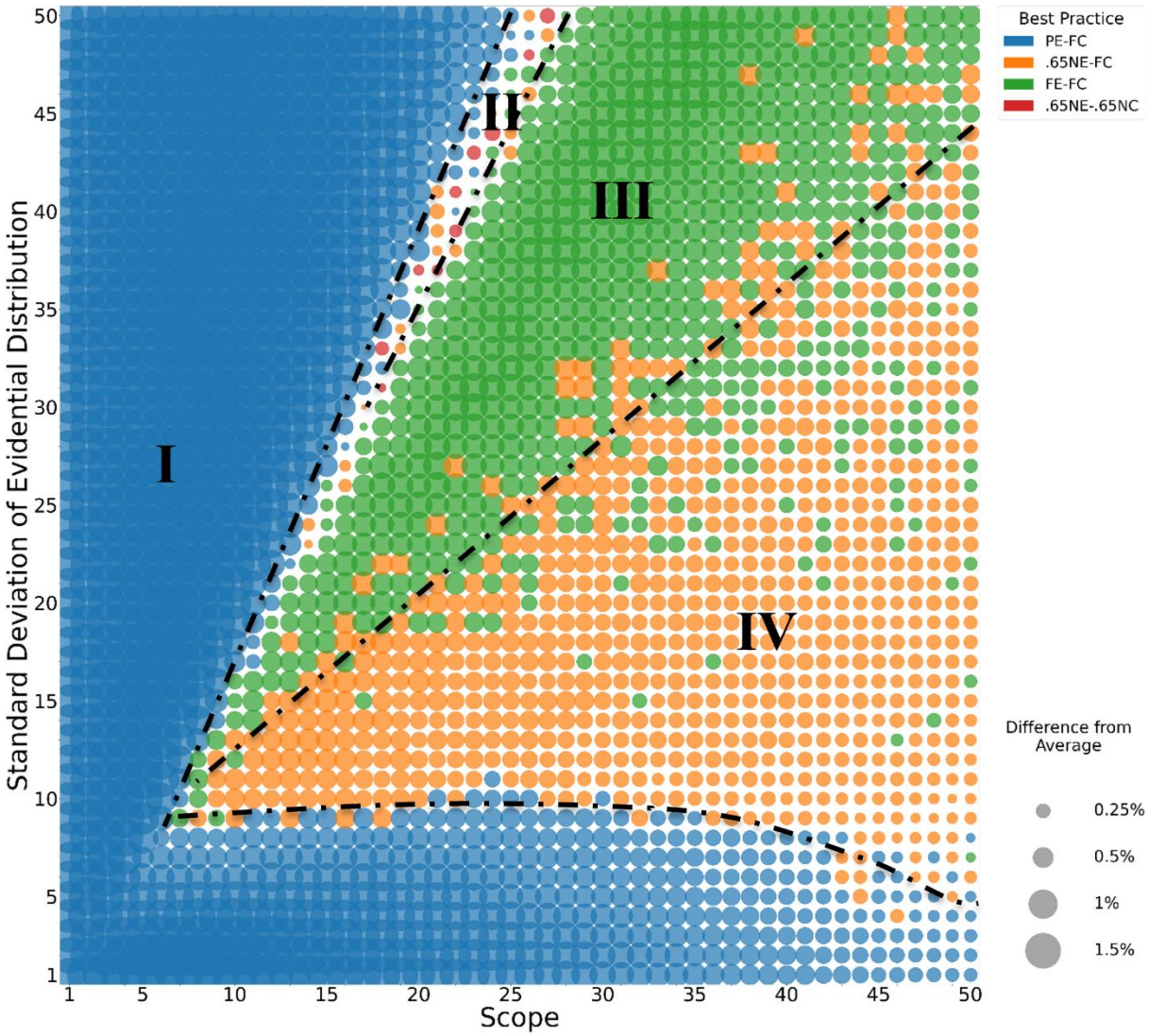

Figure 2.9 The best performance practice at 1,500 steps. Abbreviations: FC, full communication of conclusions; FE, full communication of evidence; NC, neighborhood communication of conclusions; NE, neighborhood communication of evidence; PE, private evidence. The size of each dot indicates how much the best-performing practice exceeds the average of the four practices, measured by the percentage (out of 1,200 simulations) of groups in which the majority finds the truth.

Table 2.2 Supporting explanations for regions in Figure 2.9

Region	Supporting explanation (by claims from Section 2.4)
I	2.1, 2.1′, 2.3, 2.3′
II	2.5
III	2.1, 2.1′, 2.2
IV	2.4

When there are only 750 steps in a simulation, groups practicing .65NE-FC or FE-FC do not have enough time to fully benefit from communicating high-quality evidence in a later stage (once agents have moved closer to the truth) of the simulation, resulting in PE-FC being performant in a larger area.[36] With 3,000 steps, PE-FC performs the best for a much smaller area because .65NE-FC or FE-FC have more time to benefit from communicating evidence.[37]

Figure 2.9 indicates that the Zollman effect—where less communication can lead to better outcomes—appears widely in this model. Broadly speaking, any cell where FE-FC is not optimal manifests the Zollman effect.[38] While Zollman's model focuses on how network structures restrict information flow and enable evidence accumulation, my model demonstrates that specific practices of communication can optimize the quality of information to be learned by agents. This stands in opposition to Rosenstock et al.'s (2017) view that the Zollman effect is atypical and manifests only in challenging scenarios.[39] A potential explanation for this is that my model's lack of a fixed network and its distinction between evidence and conclusions enables a wider range of strategies to be explored, providing more ways to improve the efficiency of information flow in

[36] See Figure C.1 in Appendix C.
[37] See Figure C.2 in Appendix C.
[38] Zollman's (2007) model does not distinguish between evidence and conclusions. In Zollman's model, the group stops either when everyone adopts an old method and stops discovering new evidence or when everyone is sufficiently confident about the new method. In terms of individual agents, no agent stops unless the model (the whole group) stops. So, agents who have already been sufficiently confident in the new method (the correct answer given Zollman's setup) continue to share new evidence. Because they are utilizing the new method, their evidence is likely to favor the new method. This feature makes Zollman's model comparable to practices with communication about conclusions in my model.
[39] Region IV in Figure 2.9 represents the least challenging scenarios. My results suggest that the fundamental reason for the Zollman effect is the unwise use of information rather than the challenging context.

finding the truth.

2.6 Conclusion

Strategically speaking, evidence is timely and evolves continuously; while a conclusion emerges later, remains fixed, but has been carefully considered. These features make communication about these two kinds of information effective for resolving different types of challenges. Communication about evidence is effective in addressing the problem of ambiguous evidence. However, for this approach to work, agents must have access to evidence that is, on average, sufficiently truth conducive. Communication about conclusions, on the other hand, is beneficial for overcoming low accessibility to high-quality evidence.

An ideal deliberative process maximally exploits available information by promoting high-quality information while suppressing low-quality information. The optimal practice for achieving this balance depends on the situation. When evidence is expected to be of high quality, communicating evidence is generally beneficial. Conversely, learning conclusions becomes crucial when agents lack access to high-quality evidence. In challenging scenarios where high-quality evidence is unavailable, the best strategy is to rely solely on others' conclusions without communicating evidence. Even when agents have access to high-quality evidence, extensively communicating evidence may still be suboptimal (as shown in the comparison of .65NE-FC and FE-FC in Figure 2.7), as communicating diversified evidence clusters agents into a single group and slows the overall pace towards the truth. Extensively communicating evidence reduces the chance that the first conclusion asserted is correct and increases the length of time required for the first agent to find the truth. Under time constraints, it is important to have someone find the truth quickly, allowing others to learn from them. In contrast, exhaustively learning asserted conclusions,

despite its potential for small drawbacks in rare scenarios, is almost always beneficial and is consistently better than learning only some of the conclusions.

Since my model is built on minimal assumptions beyond the relationship between agents' epistemic positions and their discovery of evidence, it enjoys broad applicability. In essence, my model provides a flexible framework for investigating scenarios involving different practices of exchanging information of varying quality. Finally, I wish to suggest a game-theoretic extension of my model. Presumably, strategically learning others' evidence could improve the truth-tracking performance of either an individual or a group. For example, an agent may choose to learn from people who have (or have not) significantly changed their positions in recent rounds. Allowing more complicated strategies in learning is likely to change the results stated in this paper. Of course, such extensions also prompt questions about different optimal learning strategies for agents and groups.

Chapter 3: A Bayesian Reflection on the Judy Benjamin Problem

3.1 Introduction

According to Bayesian orthodoxy, one should update one's beliefs by conditioning on newly acquired information. This in turn requires a well-defined probability space in which the learned information can be represented as an event. However, when the information takes the form of a conditional probability, the context may underdetermine the probability space in which conditioning is to occur. This allows for multiple plausible extended spaces to be introduced, and different extended spaces may lead to different posteriors. The Judy Benjamin (van Fraassen, 1981) problem is a case of this kind.

> **The Judy Benjamin Problem** In a war game, Judy and her team get lost in the battleground. The area is divided into four regions: Blue First camp, Blue Second camp, Red First camp, and Red Second camp. Judy's team belongs to the blue army. Assuming Judy knows nothing else about the battlefield, intuitively, her (prior) credence of being in each region should be $\frac{1}{4}$. Judy then receives information from her commander: if they are in enemy territory (i.e., the red area), the probability of being in the Red First camp is $\frac{3}{4}$.

The question is how Judy should update her belief about being in the red region given this new information. From a Bayesian perspective, Judy cannot conditionalize on the learned information within the naïve probability space (the space comprising the four regions) because the learned conditional probability does not correspond to an event in that space (Grove & Halpern, 1997; Grünwald & Halpern, 2003). Once this is recognized, two further questions arise: whether the case determines a uniquely correct posterior at all, and whether the competing answers differ in any qualitative way. Some work in the literature suggests that the Judy Benjamin problem does

not by itself determine a uniquely correct answer (Grove & Halpern, 1997; Eva, Hartmann, & Rad, 2020), because different enlarged Bayesian spaces can support different posteriors. Vasudevan (2020) proposes a qualitative contrast between the mainstream answers in terms of conservativity and epistemic charity. In this paper, I examine two influential Bayesian models, Grove and Halpern's and Vasudevan's, and defend the under-specification thesis while arguing that Vasudevan's qualitative distinction is not well supported. I then introduce a third Bayesian model as a further illustration of the under-specification thesis. The point of the third model is not to identify the uniquely correct posterior, but to show that even closely related Bayesian extensions can encode different assumptions about the structure of the situation being represented.

Grove and Halpern (1997) propose an extended space whose elements are probability distributions on the four regions. Conditioning within this space answers that Judy's credence in Red territory remains at $\frac{1}{2}$. This aligns with the intuition that the new information does not directly affect the overall probability of being in the red or blue areas (van Fraassen, 1981; Douven & Romeijn, 2011). Furthermore, this *intuitive answer* is considered conservative for incorporating the new information while making minimal adjustments to Judy's overall credence. By contrast, Vasudevan (2020) proposes a different extended space consisting of sequences of random variables that represent the four regions. Conditioning in this space yields a lower posterior in Red territory, converging to 0.468, the Minimum Relative Entropy (MRE) answer.[40] Grove and Halpern aim to show that the intuitive answer can be vindicated within a Bayesian framework, while Vasudevan shows how a different Bayesian construction recovers the MRE result. At the

[40] The concept of "relative entropy minimization" goes by various names in the literature. Relative entropy is a distance measure between two probabilistic distributions. It is also known as Kullback–Leibler divergence or cross-entropy. Regarding the Judy Benjamin problem, Relative entropy minimization is equivalent to entropy maximization if one's prior is a uniform distribution (Skyrms, 1985). In this paper, I adopt the term "relative entropy (minimization)," even when discussing literature in which other terminology is used.

same time, both models support the broader thought that the Judy Benjamin story does not itself determine a uniquely correct answer.

The remaining puzzle is that an intuitively conservative answer diverges from an answer supported by a well-established conservative statistical method. Vasudevan attempts to resolve this puzzle by appealing to the assumption built into his extended space. On Vasudevan's view, the intuitive answer reflects a conservative epistemic attitude, whereas the MRE answer reflects epistemic charity toward Judy's informant. I argue that this contrast is not well-supported. Grove and Halpern's model relies on a structurally comparable assumption in fixing its extended space. The narrative of the Judy Benjamin problem does not privilege one of these assumptions over the other. Indeed, both models are epistemically conservative for adopting the Bayesian framework, but the conservativity is relative to their chosen extended space. Thus, there is no solid basis for distinguishing the two mainstream answers in qualitative terms, such as conservativity and epistemic charity.

Given that conservativity is understood in Bayesian terms, a further question arises: what is at stake in choosing one extended Bayesian space rather than another? A choice of extended space is not merely a formal device. It reflects a way of representing the structure of the situation that gives rise to Judy's uncertainty. Grove and Halpern's model represents the relevant uncertainty as uncertainty over possible probability distributions on the four regions. Vasudevan's model represents it as uncertainty over possible sequences of region-outcomes. The third model introduced below represents it as having a two-stage structure: first concerning the Red/Blue division and then concerning the division between camps within each color. These are different structural representations of the situation, but the original Judy Benjamin story fails to determine which one is correct.

What would determine a uniquely correct choice is not merely the coherence of Judy's representation, but whether it matches the actual structure of the situation. If Judy had further information about the actual structure of the situation, e.g., information about the process by which the battlefield configuration was generated, then such information could favor one extended space over the others. In that case, the appropriate extended space would be the one whose elements and probability measure best match the relevant structure of the situation. Nevertheless, accuracy considerations cannot select a unique answer unless the story is supplemented with further structural facts.

Section 2 sets out the two Bayesian models and explains how each yields its answer to the Judy Benjamin problem. Section 3 argues that nothing in the original narrative provides a basis for qualitatively distinguishing between the two extended spaces. Together, Sections 2 and 3 support the conclusion that the Judy Benjamin problem does not contain enough information to determine a uniquely correct Bayesian posterior. Section 4 then introduces a third, two-stage model to further illustrate how different choices of extended space encode different structural assumptions. The conclusion then draws a broader moral: an accuracy-based selection among these models would require further information about the actual structure of the situation, information not supplied by the original Judy Benjamin story.

3.2 The Bayesian Approach to the Judy Benjamin Problem

In this section, I first explain why Bayes' rule cannot be straightforwardly applied to the Judy Benjamin problem. Second, I present Grove and Halpern's and Vasudevan's respective solutions to this difficulty and show how each yields a different posterior for Judy. My concern here is not whether Bayes' rule is the correct updating rule in cases like Judy's, nor whether the resulting models accurately describe Judy's or her commander's actual psychology. Rather, I focus

on the abstract problem of constructing a probability space in which Judy's learned information can be incorporated by conditioning.

Before receiving the message from her commander, Judy's prior credences are generally taken to be uniform over the four regions. Let

$$\mathcal{X} = \{R_1, R_2, B_1, B_2\},$$

where R_1 is Red First camp, R_2 is Red Second camp, B_1 is Blue First camp, and B_2 is Blue Second camp. Judy's prior satisfies

$$P_J^0(R_1) = P_J^0(R_2) = P_J^0(B_1) = P_J^0(B_2) = \frac{1}{4}.$$

Let $R = \{R_1, R_2\}$ and $B = \{B_1, B_2\}$. It follows that

$$P_J^0(R) = P_J^0(B) = \frac{1}{2}.$$

Judy then receives the message that, conditional on being in Red territory, the probability of being in Red First camp is $\frac{3}{4}$. The learned information is represented as

$$m\colon Pr(R_1|R) = \frac{3}{4}.$$

The question is how her posterior credence in R should change in light of this information.

The intuitive answer is that Judy should retain her prior credence that she is in Red territory, so that $P_J^1(R) = P_J^0(R) = \frac{1}{2}$. As Douven and Romeijn (Douven & Romeijn, 2011) argue, the conditional probability $Pr(R_1|R) = \frac{3}{4}$ should not directly affect Judy's unconditional credence that she is in Red rather than Blue territory. Such an answer reflects a "conservative attitude": Judy should not adjust her probability of R unless she receives information that bears directly on that proposition. However, as van Fraassen (1981; 1986) recognizes, the Infomin Principle[41] (i.e., MRE)

[41] Although Van Fraassen adopts the term "Infomin Principle," nowadays, the principle is known as the "principle of relative

provides a different answer. MRE selects, among the posteriors compatible with the learned information, the one that minimizes relative entropy with respect to the prior. Formally, it minimizes

$$D(P_1||P_0) = \sum_{x \in X} P_1(x) \log \frac{P_1(x)}{P_0(x)}$$

where P_0 denotes one's prior, P_1 denotes the posterior, and X is the space (Cover & Thomas, 2006; Csiszár & Körner, 2011). According to MRE, Judy's posterior should be $P_J^1(R) = 0.468$. This answer, too, is often described as conservative, since MRE chooses the admissible posterior that remains closest to the prior in the sense measured by relative entropy. Thus, the Judy Benjamin problem presents an apparent conflict between two answers that can each claim some connection to conservativity.

To see why Bayesian conditioning does not immediately resolve the problem, consider an ordinary case in which conditioning applies directly. Suppose Judy learns that she is in either Red First camp or Blue First camp, that is, $F = \{R_1, B_1\}$. In the naïve sample space $\mathcal{X}$, this learned information corresponds to the event $\{R_1, B_1\} \subseteq \mathcal{X}$. Bayesian conditioning can therefore be applied directly. For example, Judy's posterior credence in R_1 is

$$P_J^0(R_1|F) = \frac{P_J^0(R_1 \cap F)}{P_J^0(R_1 \cup F)} = \frac{1}{2}.$$

Informally, conditioning restricts the sample space to the possibilities compatible with what has been learned and then renormalizes the probability measure over that restricted space. In this example, the updated space is $\{R_1, B_1\}$, and the updated measure is again uniform on that restricted space.

entropy minimization" (or "maximum entropy").

The Judy Benjamin problem is different. The learned information $Pr(R_1|R) = 3/4$ is not itself an event in the naïve four-region space $\mathcal{X}$. The elements of $\mathcal{X}$ are regions, and no subset of these four regions says that the conditional probability of given R is $\frac{3}{4}$. Thus, Bayesian conditioning cannot be straightforwardly applied. To apply conditioning, one must introduce an extended space in which the learned information can be represented as an event.

The general structure of such an extended-space Bayesian model can be represented as follows. An extended-space model is a triple

$$(\Omega, \mu, \pi).$$

Here (Ω, μ) is a probability space of possible representations of Judy's situation (I suppress the corresponding sigma-algebra $\mathcal{F}$ here); μ is Judy's probability measure over Ω; and $\pi: \Omega \to \Delta(\mathcal{X})$ assigns to each element $\omega \in \Omega$ a first-order probability distribution π_ω over the naïve four-region space $\mathcal{X}$. The elements of Ω need not themselves be probability distributions. They need only determine a probability distribution over $\mathcal{X}$. Given such a model, Judy's credence in an event $E \subseteq \mathcal{X}$ is defined as the expectation of $\pi_\omega(E)$ with respect to μ:

$$P_J(E) = \int_\Omega \pi_\omega\,(E)\, d\mu(\omega).$$

The commander's message can then be represented as an event in the extended space:

$$L = \{\omega \in \Omega: \pi_\omega(R_1 \mid R) = \frac{3}{4}\}.$$

When L has a positive measure, Judy updates by conditioning on L:

$$\mu^1(A) = \mu(A|L) = \frac{\mu(A \cap L)}{\mu(L)}$$

for $A \subseteq \Omega$. Her posterior credence in $E \subseteq \mathcal{X}$ is then

$$P_J^1(E) = \int_\Omega \pi_\omega\,(E)\, d\mu^1(\omega).$$

The scheme above appeals to a continuous case, whereas the same reasoning applies to discrete cases.

This abstract scheme captures what Grove and Halpern's and Vasudevan's models have in common. In both cases, Judy introduces an extended space, assigns a probability measure over that space, conditions on the part of the space compatible with the commander's report, and then obtains her posterior over the naïve four-region space by taking expectations. Once an extended-space model has been chosen, an epistemic conservative requirement is that the extended space includes all admissible elements that are compatible with Judy's background information. Likewise, if there is no reason to favor one admissible element over another, the models appeal to the principle of indifference (Titelbaum, 2022): if an agent has no evidence favoring one possibility over another, they distribute their credence equally across the relevant possibilities. In the models considered below, this motivates a uniform prior measure μ^0 over the extended space. After learning, μ^1 is uniform over elements $\omega \in L$.

In Grove and Halpern's model, the elements of the extended space are themselves probability distributions over the four regions. Let

$$\Omega^G = \Delta(\mathcal{X}),$$

where $\Delta(\mathcal{X})$ is the simplex of probability distributions over $\mathcal{X}$. An element $\omega \in \Omega^G$ may be written as

$$\omega = (r_1, r_2, b_1, b_2),$$

where $r_1 + r_2 + b_1 + b_2 = 1$ and each coordinate is non-negative. The coordinates r_1, r_2, b_1, b_2 represent the probabilities assigned to R_1, R_2, B_1, B_2, respectively. In this case, the associated first-order distribution π_ω is ω itself. Given this choice of representation, and due to Judy's lack of

further information, Grove and Halpern take the prior measure μ_G^0 over Ω^G to be uniform. Judy's prior credence in an event $E \subseteq \mathcal{X}$ is then

$$P_{J,G}^0(E) = \int_{\Omega^G} \pi_\omega\,(E)\; d\mu_G^0(\omega).$$

This yields the uniform prior over the four regions. The learned information determines the following subset of Ω^G:

$$L^G = \left\{\omega \in \Omega^G \middle| \left|\frac{r_1}{r_1+r_2} - \frac{3}{4}\right| < \varepsilon\right\}.$$

Strictly speaking, L^G has measure zero under μ_G^0. Thus, conditioning on L^G should be understood as a limiting conditionalization (Grove & Halpern, 1997, p. 212). Her posterior credence in an event $E \subseteq \mathcal{X}$ is then

$$P_{J,G}^1(E) = \int_{\Omega^G} \pi_\omega\,(E)\; d\mu_G^1(\omega).$$

In particular, Grove and Halpern's model yields

$$P_{J,G}^1(R) = \frac{1}{2}.$$

Thus, Grove and Halpern's extended space vindicates the intuitive answer within a Bayesian framework.

By contrast, Vasudevan extends space to one that consists of sequences of random variables taking values in the four regions.[42] For a fixed positive integer N, let

$$\Omega_N^V = \mathcal{X}^N.$$

An element $s \in \Omega_N^V$ is a sequence

$$s = (x_1, x_2, \ldots, x_N),$$

[42] My presentation only includes those details that are necessary to show how Vasudevan's model is in keeping with the Bayesian framework. I deviate from Vasudevan's terminologies but honestly reflect the core reasoning behind his model. Details that motivate Vasudevan's overall conclusion will be discussed in the next section.

where each $x_i \in \mathcal{X}$. Each sequence determines a first-order probability distribution over $\mathcal{X}$ by relative frequency. Thus, for any event $E \subseteq \mathcal{X}$, define

$$\pi_s(E) = \frac{1}{N}\sum_{i=1}^{N} I(x_i \in E),$$

where $I(X_i \in E) = 1$ if $X_i \in E$, and $I(X_i \in E) = 0$ otherwise. Thus, the elements of Vasudevan's extended space are sequences that determine probability distributions. In the prior state, the extended space contains all sequences in Ω_N^V.[43] For purposes of comparison with Grove and Halpern's model, I reconstruct Judy's prior measure $\mu_{V,N}^0$ as uniform over Ω_N^V.[44] Judy's prior credence in an event $E \subseteq \mathcal{X}$ is then

$$P_{J,V}^0(E) = \sum_{s \in \Omega_N^V} \mu_{V,N}^0(s)\pi_s(E).$$

Since the prior measure is uniform over all sequences, this again yields a uniform prior on the four regions. Notice, however, that uniformity over sequences is not the same as uniformity over the distributions (over $\mathcal{X}$) determined by those sequences. Some distributions are realized by more sequences than others, and therefore receive greater total weight under a uniform measure over sequences. This is the central structural difference between Vasudevan's extended space and Grove and Halpern's.

The learned information restricts the extended space to those sequences in which the number of R_1-entries is three times the number of R_2-entries. Let

[43] Vasudevan essentially models Judy's case as a discrete memoryless source (DMS), even though he presents the model slightly differently. A DMS is a sequence of random variables taking values in a finite set (Csiszár & Körner, 2011). It is memoryless in the sense that values of the sequence are independent and identically distributed. Vasudevan's space consists of all the possible sequences in infinity length which is essentially modeling a space of DMS. Regarding a DMS, the REM method yields results consistent with a Bayesian method (Zellner, 1988). In modeling a DMS, Shannon's entropy measures the uncertainty embedded in an information source, and MRE is used to maximize uncertainty in the posterior (Cover & Thomas, 2006; Csiszár & Körner, 2011). This provides insight into how Vasudevan's model gives an answer consistent with REM but through a Bayesian approach.

[44] Because Vasudevan discusses the minimization of relative entropy via the maximization of Shannon entropy, he does not address Judy's prior sample space to any significant degree. When addressing the posterior, Vasudevan only needs to assume that all sequences that share a θ value are equiprobable. However, maximization of Shannon's entropy is equivalent to REM if and only if the prior is a uniform distribution, i.e., $\frac{1}{4}$ on each region (Skyrms, 1985). The speculation suggested above satisfies this condition.

$$L_N^V = \left\{ s \in \Omega_N^V \middle| \sum_{i=1}^{N} I\,(x_i = R_1) = 3 \sum_{i=1}^{N} I\,(x_i = R_2) \right\}.$$

Conditioning on L_N^V yields a uniform measure $\mu_{V,N}^1$ over L_N^V, i.e.,

$$\mu_{V,N}^1\,(\mathrm{s}) = \begin{cases} \dfrac{1}{|L_N^V|}, & \text{if } s \in L_N^V, \\ 0, & \text{otherwise.} \end{cases}$$

Then, Judy's posterior credence in an event $E \subseteq \mathcal{X}$ is

$$P_{J,V}^1(E) = \sum\nolimits_{s \in L_N^V} \mu_{V,N}^1\,(s)\pi_s(E).$$

More generally, Vasudevan defines

$$\theta(s) = \frac{\sum_{i=1}^{N} I(x_i = R_1)}{\sum_{i=1}^{N} I(x_i = R_1) + \sum_{i=1}^{N} I(x_i = R_2)},$$

that is, the ratio of R_1-outcomes to red-territory outcomes in the sequence. Thus, the model assumes that any two sequences with the same θ-value are equiprobable. In the posterior case, where $\theta = 3/4$ for every sequence, this means that the probability measure is uniform over the sequences in Ω_1^V. Vasudevan (2020, pp. 1129-33) shows that, as $N \to \infty$, Judy's posterior credence in red territory converges to the MRE answer:

$$P_{J,V}^1(R) \to 0.468.$$

Thus, Vasudevan's model also makes Bayesian conditioning applicable to the Judy Benjamin problem. The learned conditional probability is represented as a constraint on an extended space, and conditioning yields the MRE result.

The central difficulty in applying Bayes' rule to the Judy Benjamin problem is that the learned information is not an event of the naïve four-region space $\mathcal{X}$. Both models instantiate the same abstract Bayesian schema. However, the two models differ in the choice of elements to construct the space: the former uses probability distributions on the four regions, while the latter

uses sequences of random variables corresponding to the four regions. Upon building their probability spaces, both models appeal to the principle of indifference in assigning a probability measure over their corresponding spaces. The subsequent application of Bayes' rule is straightforward. In this respect, Grove and Halpern's and Vasudevan's models agree on nearly every aspect of the Bayesian methodology except for the exact constitution of the extended sample space.[45]

This is the sense in which the Judy Benjamin problem is under-specified. The original story provides Judy's prior over the four regions and the conditional probability reported by the commander. It does not provide a theoretical basis for claiming whether Judy should represent the situation by possible four-region distributions, by possible sequences of region-outcomes, or by some other structure.[46] In that sense, the original story does not determine a uniquely correct Bayesian answer.

3.3 Conservativity and Charity

Both Grove and Halpern's model and Vasudevan's model make Bayesian conditioning applicable to the Judy Benjamin problem. Vasudevan proposes a qualitative distinction between the two answers, and between the models that generate them. In his view, the intuitive answer reflects a conservative epistemic attitude, whereas the MRE answer reflects epistemic charity

[45] Bayesian conditioning in different probability spaces yields different outcomes. This mirrors the implications of treating MRE as an open framework of distance minimization. MRE posits that one should minimize the distance between their prior and posterior distributions while accommodating newly learned information. The openness lies in the fact that one may adopt different measures of the distance between two probabilistic distributions. As Douven and Romeijn (2011) and Eva, Hartmann, and Rafiee Rad (Eva, Hartmann, & Rad, 2020) point out, measuring the probabilistic distance by inverse relative entropy and minimizing the distance yields the intuitive answer.

[46] As noted in footnote 45, minimizing distance using inverse relative entropy provides a systematic approach to formalizing the intuitive answer. The Judy Benjamin problem itself does not contain any reason to favor one measure over the other. Thus, as a general framework, minimizing probabilistic distance provides an equally solid framework for justifying both answers. A similar line of reasoning applies to the Bayesian framework: the Judy Benjamin problem does not offer a basis for favoring either Grove and Halpern's space or Vasudevan's space.

toward Judy's informant. On Vasudevan's interpretation, Judy treats the commander's report as containing all the information relevant to her inquiry. She therefore assumes that, once the reported value has been fixed, there is no further probabilistically relevant distinction to be drawn among the possibilities compatible with that value. In Vasudevan's sequence model, this takes the form of assigning equal weight to all sequences with the same θ-value.

I argue that this qualitative contrast is not well-supported. The formal feature that Vasudevan associates with epistemic charity is not distinctive of the sequence model. A structurally similar no-further-discrimination feature also appears in Grove and Halpern's model. Furthermore, at the level of Bayesian machinery, the two models proceed in the same general way. Once a type of extended-space element has been chosen, each model includes the admissible elements of that type, assigns weights by an indifference rationale, and updates by conditioning on the commander's information. In this sense, both models are conservative relative to the kind of space they start with. Any qualitative contrast between them must therefore be grounded in their different choices of elements for the extended space. However, conservativity alone does not decide what the right starting space should be.

Vasudevan's discussion is motivated by a genuine puzzle. MRE updating is treated as a paradigmatically conservative method (Csiszár & Körner, 2011), since it chooses the posterior distribution closest to the prior among those satisfying the new constraint. Yet in the Judy Benjamin problem, MRE yields a posterior that conflicts with the intuitively conservative answer. Some (Seidenfeld, 1986; Dias & Shimony, 1981) have questioned the application of MRE in this context, arguing that it should not be considered as a formalization of the conservative maxim. On Vasudevan's account, Judy believes that the report from her commander contains all known information relevant to her inquiry. Recall from Section 2 that θ is the relative frequency between

R_1 and R. Since the information is conveyed through the value of θ, she treats any two sequences with the same θ-value as equiprobable. Formally,

$$\text{if } \Theta(s) = \Theta(s'), \text{ then } \mu_V^1(s) = \mu_V^1(s').$$

In the actual Judy Benjamin case, this means that all sequences in L_N^V are assigned equal posterior weight. This condition reflects a no-further-discrimination attitude. Judy does not distinguish among sequences that agree on θ-value. On Vasudevan's interpretation, this reflects an epistemic charity toward Judy's informant rather than mere conservativity.

Vasudevan formulates his condition in terms of sequences, but the epistemic charity is not tied to the use of sequences. What matters is the assignment of equal weight to elements that are indistinguishable with respect to the learned information. A structurally parallel condition may be formulated for Grove and Halpern's model, although it should not be formulated as equality of point probabilities. Their extended space is continuous, so individual elements have probability zero. Let

$$\Gamma(\omega) = \pi_\omega(R_1|R) = \frac{r_1}{r_1 + r_2}.$$

Since L^G is a measure-zero constraint surface, Grove and Halpern's posterior measure should be understood by a limiting construction. Conditioning the uniform prior on the event

$$\left|\gamma - \tfrac{3}{4}\right| < \varepsilon, \text{ as } \varepsilon \to 0$$

gives a uniform conditional measure relative to the original measure on the band. This is the continuous analogue of Vasudevan's no-further-discrimination condition. Once the value of γ is fixed, the model introduces no further distinctions among compatible elements. Thus, if Vasudevan's equiprobability condition is taken to express epistemic charity, the Grove and Halpern's model also embeds the same kind of epistemic charity. For that reason, the contrast

Vasudevan draws between the intuitive answer and the MRE answer cannot be sustained merely by such an equiprobability assumption.

I now turn to conservativity. In the present context, a Bayesian model with an extended space may be said to display a conservative attitude in two respects. First, in constructing the extended space, the model includes all possibilities of the relevant kind, given the chosen elements to represent Judy's situation. Grove and Halpern include all admissible probability distributions over the four regions. Vasudevan includes all admissible sequences of region-outcomes of length N. Second, once that space has been specified, the model assigns probabilities by an indifference rationale. Grove and Halpern use a uniform prior measure over the space of distributions, while Vasudevan's model is reconstructed as using a uniform prior measure over the sequence space.

In both respects, the models refrain from building in assumptions beyond the available information. Therefore, both models are conservative relative to their chosen representational structures. Neither model is uniquely conservative. Conservativity operates after the element type being fixed; it does not determine whether the elements should be distributions, sequences, or something else. One may say that both models are conservative in this Bayesian sense. Beyond that, if one wishes to attribute some further epistemic virtue to one answer on the basis of its Bayesian form alone, parity of reasoning requires attributing that virtue to the other as well.

3.4 Demystifying the Choice of an Extended Space

There remains a question concerning the choice of elements. If the Judy Benjamin story itself cannot determine a uniquely correct Bayesian posterior, it likewise does not determine a uniquely correct extended space. Any particular choice of extended space, therefore, carries substantive assumptions. The point is not that such choices are arbitrary. Rather, a choice of extended space reflects Judy's representation of the structure of the situation. What would make

one representation uniquely appropriate is further information about the actual structure of the situation being represented. The original story does not provide such information.

This point can be seen by comparing Grove and Halpern's model with Vasudevan's. Grove and Halpern's model represents Judy's uncertainty as uncertainty over possible probability distributions on the four regions. The extended space is $\Omega^G = \Delta(\mathcal{X})$, and its prior measure treats the admissible four-region distributions uniformly. Adopting this model amounts to representing the situation as if the relevant uncertainty were uncertainty over which four-region probability distribution obtains.

By contrast, Vasudevan's model represents the situation differently. The extended space is $\Omega_N^V = \mathcal{X}^N$ whose elements are N-sequences whose entries take values in the four regions. The prior measure treats these sequences equally. Each sequence determines a distribution over $\mathcal{X}$ by relative frequency. Adopting this model amounts to representing the situation as if the relevant uncertainty were uncertainty over possible sequences of region-outcomes. Thus, Grove and Halpern's model and Vasudevan's model differ not merely in the posterior they yield, but in the kind of structure they use to represent Judy's situation.

To further illustrate the connection between an extended space and the representation of Judy's uncertainty, I introduce a third Bayesian model. The model is purely diagnostic without aiming to offer a uniquely correct representation of Judy's situation. It shows that even closely related extended spaces can encode different structural assumptions. In particular, the model represents Judy's uncertainty in two stages. First, the model considers possible divisions between red and blue territory. Second, conditional on a given red/blue division, the model refines each color into first and second camp. Formally, define the two-stage extended space as

$$\Omega^T = \{(r, \alpha, \beta) | r, \alpha, \beta \in [0, 1]\},$$

where r represents the probability of Red territory; α represents the conditional probability of Red First camp given Red territory; and β represents the conditional probability of Blue First camp given Blue territory. Thus, each $\omega \in \Omega^T$ determines a first-order distribution π_ω over $\mathcal{X}$ as follows: $\pi_\omega(R_1) = r\alpha, \pi_\omega(R_2) = r(1-\alpha), \pi_\omega(B_1) = (1-r)\beta, \pi_\omega(B_2) = (1-r)(1-\beta)$. Accordingly, let the prior measure μ_T^0 be the product-uniform measure over Ω^T. That is, r, α, and β are treated as independent and uniformly distributed on $[0,1]$. Informally, Ω^T includes all possible first-order representations of Judy's situation, and μ_T^0 represents Judy's uncertainty over those representations. This represents Judy's uncertainty in a hierarchical way: first at the level of color, and then, conditional on a given color partition, at the level of camp.

This two-stage model may initially seem artificial. Its point is to make explicit another reasonable way of representing Judy's situation. Grove and Halpern's model treats the four regions as equally basic and places a uniform measure over distributions on those four regions. This represents Judy's prior ignorance in that she does not probabilistically distinguish among possible four-region distributions. For Judy, the distinction between enemy and friendly territory is more salient than the distinction between first and second camps. Accordingly, it is natural to represent her ignorance by assigning a uniform measure to possible Red/Blue divisions. However, a uniform measure over distributions on $\mathcal{X}$ does not in general induce a uniform marginal measure over $\{R, B\}$. The two models thus embody different but reasonable ways of modeling Judy's ignorance. The difference between the two models lies not in the first-order four-region distributions they can represent, but in how those distributions are organized and weighted at the higher-order level. Figures 3.1 and 3.2 illustrate the differences between the spaces of Grove and Halpern's model and this two-stage model.[47]

[47] The diagrams should not read as a representation of the battlefield. Instead, the diagrams use colors and regions to visually

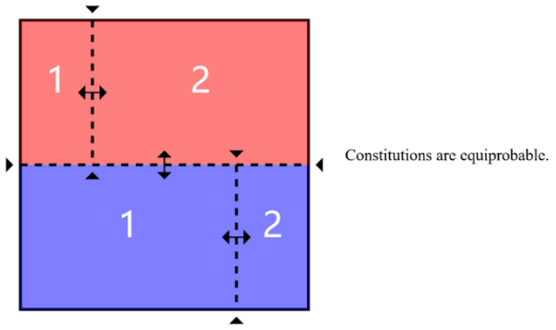

Figure 3.1. A schematic representation of Grove and Halpern's space. Each dashed line with a double arrow crossing it indicates that moving the dashed line to a different position represents a distinct partition (i.e., a different configuration of the four regions).

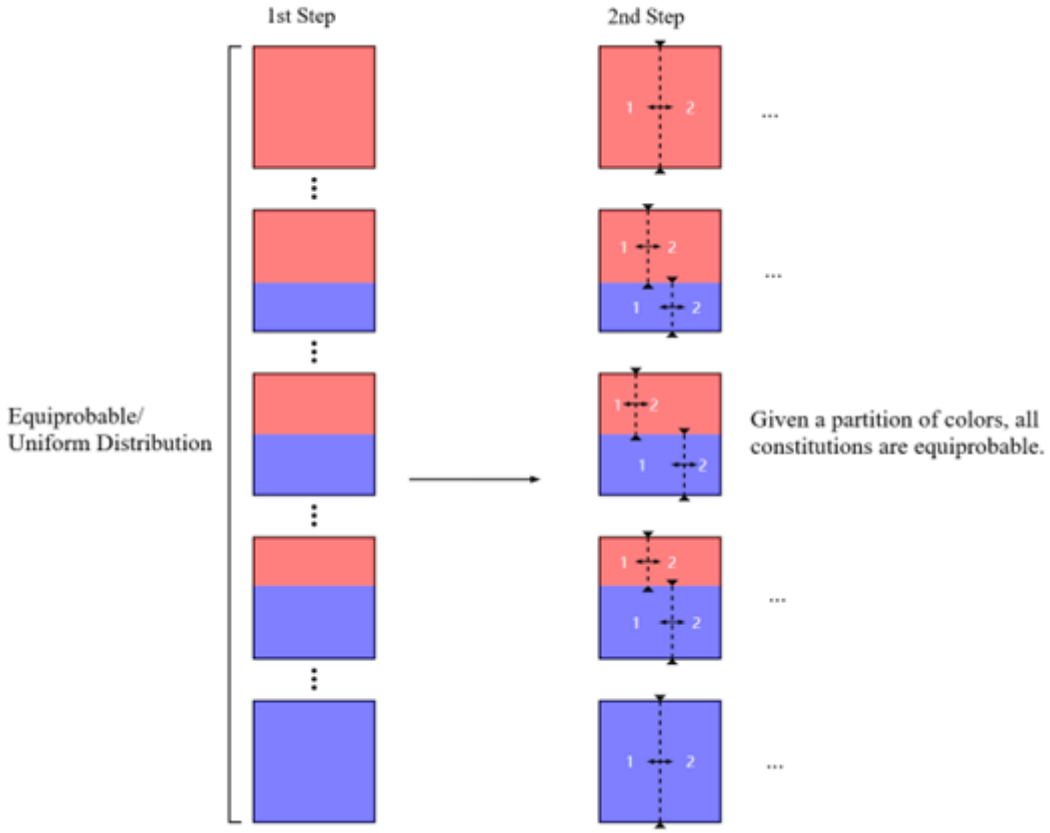

Figure 3.2. A schematic representation of the two-stage model. The left column represents all the different partitions of the two colors. In the right column, given a partition of the two colors, each dashed line with a double arrow crossing indicates that moving the dashed line to a different position represents a distinct partition (i.e., a different configuration of the four regions).

illustrate a probability distribution over the regions.

It is not hard to see that this two-stage model captures the intuitive answer to the Judy Benjamin problem. Judy's prior credence in an event $E \subseteq \mathcal{X}$ is

$$P_{J,T}^0(E) = \int_{\Omega^T} \pi_\omega\,(E) f_T^0(\omega) d\omega,$$

where f_T^0 is the density of μ_T^0. For Red territory,

$$P_{J,T}^0(R) = \int_0^1 r\, dr = \frac{1}{2}.$$

For Red First camp,

$$P_{J,T}^0(R_1) = \int_0^1 r\, dr \int_0^1 \alpha\, d\alpha = \frac{1}{4}.$$

Similarly, $P_{J,T}^0(R_2) = P_{J,T}^0(B_1) = P_{J,T}^0(B_2) = \frac{1}{4}$. Now suppose Judy learns that $Pr(R_1|R) = \frac{3}{4}$. In the two-stage model, this report fixes the within-Red parameter α at $\frac{3}{4}$. The corresponding learned event is

$$L^T = \{(r, \alpha, \beta) \in \Omega^T \,|\, \left|\alpha - \frac{3}{4}\right| < \varepsilon\}.$$

Since L^T is a lower-dimensional subset of a continuous space, conditioning on it should be understood as limiting conditionalization on increasingly narrow intervals around $\frac{3}{4}$. Under this conditional distribution, r and β remain independent and uniformly distributed on $[0, 1]$, while α is fixed at $\frac{3}{4}$. Judy's posterior credence in Red territory is therefore

$$P_{J,T}^1(R) = \int_0^1 \int_0^1 r\, d\beta\, dr = \frac{1}{2}.$$

Her posterior credence in Red First camp is

$$P_{J,T}^1(R_1) = \int_0^1 \int_0^1 \frac{3}{4} r\, d\beta\, dr = \frac{3}{8}.$$

Similarly,

$$P^1_{J,T}(R_2) = \int_0^1 \int_0^1 \frac{1}{4} r \, d\beta \, dr = \frac{1}{8}.$$

Thus, the two-stage model, like Grove and Halpern's model, vindicates the intuitive answer to Judy's original question.

This agreement on the original inquiry in the Judy Benjamin problem should not obscure the structural difference between Grove and Halpern's and the two-stage models. A further higher-order question helps illustrate the difference. In addition to asking for her posterior credence that she is in Red territory, Judy might ask how likely it is, relative to her higher-order uncertainty over possible situations, that the probability of Red territory is small.

> *Small Red* (SR): What is the probability (under Judy's higher-order uncertainty over possible situations) that the proportion of Red territory is at most some threshold p, i.e., $SR^t(p) = \mu^t(\{\omega \in \Omega | \pi_\omega(R) \leq p\})$?

This is the model-relative probability that the first-order probability of Red territory is at most p. A concrete instance could be $SR^0(\frac{1}{10})$, that is, the prior probability that Judy's chance of being in enemy territory is very low, say at most $\frac{1}{10}$. Given Judy's concern about being in enemy territory, this is at least a natural further question to ask. Under the two-stage model, since the first-stage distribution over the Red/Blue split is uniform, it follows that the answer to an SR-type question is p, i.e.,

$$SR^0_T(p) = p.$$

By contrast, Grove and Halpern's model[48] yields

[48] For further details and role of this formula, refer to Grove and Halpern's (1997) paper, pp. 211-2. In general, this result does not align with the intuition: $SR^0_T(p) = p$. For $SR^0_T\left(\frac{1}{2}\right)$, as a special case, Grove and Halpern's model gives the answer $\frac{1}{2}$. Due to the difference in presenting the problem, I am using different notations from Grove and Halpern's paper. However, the core reasoning

$$SR_G^0(p) = \mu_G^0(\{\omega : \pi_\omega(R) \leq p\}) = (3 - 2p)p^2.$$

The point is not to demonstrate that one of these answers is uniquely correct; in the present context, that would be misplaced. Rather, the contrast shows that two models can agree on Judy's posterior credence in Red territory while nevertheless representing her higher-order uncertainty differently. The choice of extended space, therefore, affects not only the answer to the original update problem but also the structure of the further questions Judy may ask within the model.

The two-stage model strengthens the under-specification thesis and further illustrates how the choice of elements represents Judy's uncertainty. Grove and Halpern's model and Vasudevan's model already show that different kinds of extended spaces can yield different posteriors. The two-stage model shows that the issue runs deeper: even models that agree on some posteriors can encode different structural representations of the situation.

What would make one of these representations uniquely appropriate is further information about the actual structure of the situation. If Judy knew, for example, that the relevant battlefield configuration was generated by first fixing the Red/Blue division and then refining each color into camps, the two-stage model would have a special claim. If she knew instead that the relevant uncertainty was generated by selecting a four-region distribution uniformly from the simplex, Grove and Halpern's model would be favored. If she knew that the situation was generated by a sequence-like process, Vasudevan's model would be favored. However, the original Judy Benjamin story supplies none of this information. Therefore, the choice of extended space is not fixed by the story itself. It reflects Judy's representation of the structure of the situation, while a uniquely correct or accuracy-favoring choice would require further facts about the situation's actual structure.

---

is the same.

3.5 Conclusion

The Judy Benjamin problem is often presented as a conflict between two answers. On the intuitive answer, Judy's credence in Red territory should remain $\frac{1}{2}$. On the MRE answer, her credence in Red territory should decrease to approximately 0.468. Both Grove and Halpern's model and Vasudevan's model make Bayesian conditioning applicable by introducing an extended space in which the learned information is an event or a constraint. Grove and Halpern's model yields the intuitive answer, while Vasudevan's model recovers, in the limit, the MRE answer. The crucial difference between the models lies not in the general Bayesian machinery but in the choice of elements used to constitute an extended space.

This point supports the under-specification thesis. The original Judy Benjamin story gives Judy a uniform prior over the four regions and conditional-probability information from her commander. Grove and Halpern's model represents Judy's situation by possible probability distributions over the four regions. Vasudevan's model represents it by possible sequences of region-outcomes. Both models adopt an indifference rationale in constructing extended spaces, and both then update by conditioning. Since the story itself does not determine which kind of representation is appropriate, it does not determine a uniquely correct Bayesian posterior.

This also undermines the proposed qualitative contrast between the two mainstream answers. Vasudevan argues that the MRE answer reflects epistemic charity toward Judy's informant, while the intuitive answer reflects a conservative attitude. However, the formal feature Vasudevan associates with epistemic charity is not unique to his sequence model. Grove and Halpern's model has a structurally analogous feature. Likewise, both models are conservative only relative to their chosen representational structures. Conservativity cannot determine whether the

elements of the extended space should be distributions, sequences, two-stage triples, or something else.

The two-stage model serves a diagnostic role. It further illustrates what is at stake in the choice of extended spaces. The two-stage model and Grove and Halpern's model both vindicate the intuitive answer to the original question, but they organize Judy's higher-order uncertainty differently. This difference becomes visible in questions such as Small Red, which asks how likely it is, relative to Judy's higher-order uncertainty, that the probability of Red territory is below a given threshold. The lesson is that an extended space does more than deliver a posterior; it represents a structure of possibilities within which further questions can be asked.

The broader moral is that a choice of extended space is a choice of representation. A particular representation reflects Judy's understanding of the structure of the situation being modeled. What would make one such representation uniquely appropriate is not merely its internal coherence, but its fit with the actual structure of the situation. If Judy had further information about how the battlefield configuration was generated, then that information could favor one extended space over the others. For example, a distributional generating structure might favor a Grove-Halpern-style space. In that case, accuracy-based comparison would be meaningful: one may ask which extended space best matches that structure and which posterior is the most accurate relative to it. However, the original Judy Benjamin story does not supply such information. Thus, accuracy considerations cannot select a unique answer in the original case. They instead show what kind of further information would be needed for a unique answer.

Chapter 4: A Calibration Theorem for Risk-Weighted Expected Utility Theory

4.1 Introduction

Expected Utility Theory (EU) faces two well-known criticisms regarding how decision-makers handle risk. The first is the Allais Paradox (1953), which shows that EU fails to account for certain seemingly reasonable preferences over lotteries. The second is Rabin's (2000) criticism, which argues that accommodating risk attitudes over small stakes commits EU to implausibly extreme attitudes towards risk at larger stakes. One important branch of alternatives to EU consists of Quiggin's (1982) Rank-Dependent Expected Utility (RDU) theory and Buchak's (2013) Risk-Weighted Expected Utility Theory (REU). Although motivated by different considerations, RDU and REU yield the same instrumental evaluations and thus recommend the same preferences.[49] Consequently, any challenge that appeals purely to preferences, such as Rabin-style arguments, must apply equally to both (Buchak, 2013, pp. 57, 70). In addition to a standard utility function, REU incorporates a risk function that transforms probabilities of events to account for a decision-maker's attitudes towards the distribution over potential outcomes.[50] A risk function assigns to each potential outcome a weight that is different from the outcome's probability. As Buchak claims, REU's ability to accommodate preferences in the Allais paradox while avoiding Rabin's criticism provides it a clear advantage over EU.[51]

Since REU's introduction, philosophers have raised two main types of objections. One line argues that REU fails to satisfy certain structural properties guaranteed by EU (Briggs, 2015;

[49] REU is formally equivalent to RDU, as REU's formula for calculating instrumental values can be rearranged into RDU's formula (Buchak, 2013, p. 57). For a detailed comparison, see Buchak (2013, pp. 56-74).
[50] As a decision theory framework, REU is open to various kinds of risk functions. Buchak adopts a power risk function $r(p) = p^a$ (where a is a constant parameter and p is the probability of an outcome) for illustration in various examples.
[51] For an example of a risk function that helps REU resolve the Allais paradox and avoid Rabin's criticism at the same time, see page 71-73 in *Risk and Rationality* (Buchak, Risk and Rationality, 2013).

Pettigrew, 2015; Joyce, 2017). Buchak (2015; 2017) responds by emphasizing that rationality need not conform to the specific standards imposed by EU theory. A second line questions whether REU's success in small-world contexts (isolated decisions) can be satisfactorily translated into grand-world contexts (extended decision-making scenarios) (Thoma & Weisberg, 2017; Thoma, 2019; Thoma & Weisberg, 2020). Buchak (2017) concedes that REU is sensitive to problem-framing but defends it as a feature of the theory rather than a flaw.[52]

This paper develops a different kind of challenge. Economists often describe such arguments as *calibration* results: they show that once a model is "calibrated" to fit seemingly reasonable behavior, it yields implausible implications in a different situation. Rabin's criticism of EU is one example. Although Rabin's calibration theorem does not apply directly to REU, I argue that the same kind of problem arises: under standard assumptions, REU cannot simultaneously accommodate certain seemingly reasonable preferences (Allais preferences) while avoiding commitments to implausibly extreme risk-averse behaviors. Because accommodating Allais preferences generates these extreme implications, REU does not genuinely resolve the Allais paradox. This undermines Buchak's claim that REU can resolve both Allais preferences and Rabin's criticism (2013, pp. 71-3). Moreover, the calibration result suggests that REU's approach of modeling attitudes towards uncertainty may be misguided.

Economists have established related calibration results for RDU/REU, but only under strong assumptions. Neilson (2001) shows that Rabin's calibration logic applies to the REU framework, but his approach relies on assuming that the agent rejects small-stakes gambles at every wealth level. Safra and Segal (2008) prove several general calibration theorems, one of which applies to REU. However, the calibration results emerge only when the agent is uncertain

[52] Thoma (2019) and Thoma & Weisberg (2020) advance the debate by proposing some new challenges in different forms. However, parts of Buchak's (2017) earlier defenses still apply here.

about their initial wealth. More precisely, the most extreme outcomes occur when initial wealth is uniformly distributed across a wide interval. In other words, their calibration theorem does not generally apply when the initial state is fixed rather than stochastic.[53] In contrast, my calibration theorem (1) does not presuppose a certain risk-averse preference across wealth levels, and (2) applies even when the agent begins with a fixed amount of wealth rather than a dispersed distribution.

My argument proceeds in two steps. The first step, though formally a special case of the second, serves to clarify the setup and illustrates how Allais preferences restrict an REU framework.

1. **Linear Utility Function.** I show that if an REU agent with a linear utility function exhibits Allais preferences, then they must undervalue small-probability gains relative to small-probability losses to an implausible degree. The proof relies on the risk function being convex on relevant probability intervals, and I argue that such convexity is well-motivated by the reasoning behind the Allais paradox.
2. **Concave Utility Function.** I then consider the more realistic case where agents have a weakly concave utility function.[54] Here, I assume Constant Relative Risk Aversion (CRRA) utility: attitudes towards uncertainty over a fixed proportion of wealth remain stable across different wealth levels. Such an assumption is both (1) reasonable in the context of the Allais paradox; and (2) consistent with standard practice in the calibration literature, where some form of constancy in the utility function is typically

[53] Safra and Segal's calibration theorems should not be undervalued. Their calibration theorems are model-independent and thus apply to a wide family of decision theories beyond RDU/REU. By contrast, my focus is on REU, with the aim of establishing a calibration theorem under more realistic assumptions.
[54] Economists widely accept diminishing marginal utility which implies a concave utility function (Reiss, 2013; Mankiw, 2016).

assumed.[55] Under CRRA, I show that no combination of a concave utility function and a risk function that accommodates Allais preferences can escape calibration.

My paper is structured as follows. Section 4.2 reviews EU, the Allais Paradox, and Rabin's theorem. Section 4.3 introduces Buchak's REU and explains how it appears to address both challenges. Section 4.4 develops the calibration result under linear utility, while Section 4.5 extends the argument to concave utility. Finally, Section 4.6 concludes the paper.

4.2 Expected Utility Theory, Rabin's Calibration, and the Allais Paradox

EU is the classical model of decision-making under uncertainty. According to EU, the only rational way to assign instrumental values to a gamble is to equate the instrumental value with the expected utility of the gamble. A gamble g is a function from a finite set of states $\{S_i\}$ to the set of outcomes and assigns an outcome x_i to each state S_i; moreover, each state occurs with probability $p(S_i)$. A utility function $u(x_i)$ assigns a utility value to each outcome.[56] The expected utility of g is:

$$\mathrm{EU}(g) = \sum_{i=1}^{n} p(S_i)\, u(x_i) = p(S_1)u(x_1) + p(S_2)u(x_2) + \cdots + p(S_n)u(x_n).$$

EU states that a rational agent should always (1) prefer gambles with higher expected utility values to those with lower values and (2) remain indifferent among gambles with equal expected utility.

Rabin's (2000) calibration theorem demonstrates that EU commits agents to unreasonable consequences over large stakes after accommodating certain risk-averse[57] preferences over small

[55] Without assuming some regularity in the rate at which the utility function flattens out, restriction on the utility function from one situation cannot be transferred effectively to another. For example, Rabin (2000) assumes agents reject certain small-stakes gambles (e.g., a 50-50 bet of losing $100 or winning $105) across a broad range of wealth. Absent such an assumption, a utility function that is concave at a single wealth level but nearly linear elsewhere could (1) reject small-stakes gambles locally while (2) escaping calibration results. My CRRA assumption plays an analogous role to Rabin's and Safra–Segal's assumptions of consistent small-stakes risk aversion. Further discussion is provided in Section 5.

[56] Here x_i represents the agent's comprehensive situation instead of merely the payoff of the gamble. This distinction matters when the utility function is non-linear.

[57] I adopt Buchak's (2013, p. 66) usage of the relevant terms. "Risk-aversion" refers to a property of an agent's behavior (or preferences) in which the agent may prefer the choice with smaller expected units of goods.

stakes. Here I appeal to some representative results from Rabin's (2000) work. Imagine a person with some mild risk-averse preferences: whenever their initial wealth is below $300,000, they always turn down a 50-50 bet that yields a $100 loss or a $110 gain. EU can accommodate this preference via a concave utility function; however, this will necessarily commit the agent to unreasonable preferences over higher stakes. For instance, with initial wealth of $290,000, a utility function that accommodates the seemingly reasonable preference (rejecting a 50-50 bet of losing $100 or winning $110) will also require them to turn down a 50-50 bet of losing $1,000 ($2,000, $4,000, or, $10,000) or winning any amounts less than $718,190 ($12,210,880, $60,528,930, or, $1.3 billion, respectively).[58] This degree of risk aversion on large stakes seems extreme.

The reasoning is straightforward: rejecting the small-stakes gamble implies that the marginal utility of an additional dollar after a gain of $110 is at most $\frac{10}{11}$ of the marginal utility of an additional dollar after the loss of $100. The risk-averse preference determines a ratio between the evaluation of a potential dollar gain and a potential dollar loss. Supposing that the same preference for risk aversion applies across a wide range of wealth levels, the reasoning extends naturally to gambles involving larger stakes.[59] Due to the diminishing marginal utility implied by the risk-averse preference, the ratio between the required winning (monetary) amount and the loss amount increases disproportionally as the stakes grow.

The Allais paradox (Allais, 1953) provides a conceptually different criticism: EU cannot simultaneously accommodate two highly intuitive lottery preferences. Consider the following

[58] For calculations, see Rabin's Appendix (2000, pp. 1289-91).
[59] For an equivalent but more intuitive explanation of the calibration results, see Hansson (1988). An apparently mild risk-averse behavior over small stakes could force the utility function to flatten extremely quickly as the agent's wealth increases. Rabin's argument appeals to the same reasoning but illustrates the problem in a more obvious way. However, Hansson's early explanation still helps understanding Rabin's argument.

lotteries as shown in Table 4.1. An agent chooses one lottery from Lottery 1 (L1) or Lottery 2 (L2) and another lottery from Lottery 3 (L3) or Lottery 4 (L4).

Table 4.1 The Allais Paradox

Ticket/ Prob.	0.01	0.1	0.89
Lottery 1	\$0	\$5 Million	\$0
Lottery 2	\$1 Million	\$1 Million	\$0
Ticket/ Prob.	0.01	0.1	0.89
Lottery 3	\$0	\$5 Million	\$1 Million
Lottery 4	\$1 Million	\$1 Million	\$1 Million

It seems intuitive to prefer L1 to L2: neither gamble is likely to yield a payoff, but L1 offers the potential to win \$5 million (higher than L2's maximum payoff). It is also intuitive to prefer L4 to L3: even though L3 could pay up to \$5 million, the small chance of getting nothing makes L3 terrible in comparison to L4 which guarantees a payoff of \$1 million.[60] By the sure-thing principle, which grounds EU,[61] if two lotteries differ only in their outcomes under some event, then the preference between them must depend entirely on those differing outcomes (Savage, 1972; Broome, 1991). Applied to the Allais paradox,[62] this yields:

$$\mathrm{EU(L1)} > \mathrm{EU(L2)}\ (<, =) \text{ if and only if } \mathrm{EU(L3)} > \mathrm{EU(L4)}\ (<, =).$$

Thus, EU forbids an agent from simultaneously holding the two seemingly reasonable preferences in the Allais paradox.[63]

[60] In the literature, many empirical studies show that people indeed have Allais preferences in various situations. To take one recent example, Blavatskyy, Ortmann, and Panchenko (2022) studied when people are more or less likely to have Allais preferences.

[61] An EU agent necessarily satisfies the sure-thing principle, since EU evaluates lotteries by summing across states linearly with respect to probabilities. Conversely, the sure-thing principle, together with other rationality axioms, yields EU as a representation theorem (Savage, 1972).

[62] In the Allais setup, Lotteries 1 and 2, and Lotteries 3 and 4, each share the same outcome under the branch with probability 0.89. The sure-thing principle implies that this common component cancels out of the comparison. Hence, the preference between L1 and L2, and between L3 and L4, must be determined entirely by the differing outcomes under the 0.01 and 0.10 branches.

[63] The expected utilities of the four lotteries are as follows: $\mathrm{EU(L1)} = 0.1u(w + 5 \times 10^6)$, $\mathrm{EU(L2)} = 0.11u(w + 1 \times 10^6)$, $\mathrm{EU(L3)} = 0.1u(w + 5 \times 10^6) + 0.89u(w + 1 \times 10^6)$, and $\mathrm{EU(L4)} = u(w + 1 \times 10^6)$, where w is the agent's initial wealth. It

Rabin's calibration and the Allais paradox reveal weaknesses of EU from different angles. Rabin's criticism argues that EU is able to accommodate small-stakes risk-averse preferences only at the cost of committing agents to implausible large-stakes risk aversion. By contrast, the Allais paradox reveals an internal deficiency of EU as it fails to accommodate two reasonable preferences simultaneously. Attempts to resolve these challenges have motivated alternatives to EU.

4.3 An Overview of Risk-Weighted Expected Utility Theory

Buchak (2013) proposes REU as a promising alternative to EU. REU introduces a risk function $r: [0,1] \rightarrow [0,1]$ which weighs each marginal improvement among payoffs. In this way, REU allows agents to weigh each improvement in outcomes differently from what its probability would suggest. The risk-weighted expected utility of a gamble, as defined by Buchak (2013, p. 53) is:

$$\mathrm{REU}(g) = u(x_1) + r\left(\sum_{i=2}^{n} p(S_i)\right)(u(x_2) - u(x_1)) + r\left(\sum_{i=3}^{n} p(S_i)\right)(u(x_3) - u(x_2)) + \cdots$$

$$+ r(p(S_n))(u(x_n) - u(x_{n-1})),$$

where: $u(x_1) \leq u(x_2) \leq \cdots \leq u(x_n)$.

An REU agent takes the worst possible outcome, x_1, as a baseline and then evaluates each successive improvement over the next-worst outcome. Each improvement is weighted by the risk function applied to the probability of achieving at least that outcome. Since nothing can be worse than x_1, the agent assigns a weight of 1 to $u(x_1)$. Then, the agent considers the improvement in utility from $u(x_1)$ to $u(x_2)$ and the probability of this improvement actualizing, $\sum_{i=2}^{n} p(S_i)$. So, the agent assigns $r(\sum_{i=2}^{n} p(S_i))$ to the improvement in utility, $u(x_2) - u(x_1)$, according to the risk

is impossible that $\mathrm{EU(L1)} > \mathrm{EU(L2)}$ and $\mathrm{EU(L4)} > \mathrm{EU(L3)}$ hold at the same time.

function. Similarly, the agent weighs the next improvement in utility, $u(x_3) - u(x_2)$, by $r(\sum_{i=3}^{n} p(S_i))$. The inclusion of a risk function is the core difference between EU and REU.[64] REU permits an agent to consider how an action's payoffs are distributed, while EU dictates that only the expected value of utility matters. As a consequence, an REU agent may prefer a gamble with lower expected utility if certain probabilistic features of the payoffs are especially appealing. As Buchak (2013, p. 66) argues, beyond expected utility, it is rationally permissible to take the probabilistic structure itself into account, and REU successfully models it.

On the surface, it appears that REU is immune to Rabin's calibration and capable of resolving the Allais paradox. With respect to Rabin, REU avoids the need for a sharply curved utility function by allowing risk aversion to be divided between the utility function and the risk function. A mildly concave (or even linear) utility function, together with an appropriate risk function, allows REU to avoid the extreme large-stakes aversion that troubles EU. With respect to the Allais paradox, REU permits agents to assign weight to a potential improvement that differs from its raw probability, thereby rejecting the sure-thing principle. In this way, an REU maximizer can prefer L4 over L3 while preferring L1 over L2.

4.4 Calibration Under a Linear Utility Function

By introducing a risk function, REU seems to resolve two significant criticisms raised against EU. In this section, I show that when the utility function is linear, accommodating Allais preferences makes REU itself subject to calibration. I proceed in three steps. First, I establish the calibration result assuming a convex risk function. Second, I show that global convexity on $[0, 1]$ is stronger than necessary; instead, various local convexity conditions suffice. Finally, I argue that

[64] In fact, EU can be thought of as a special case of REU that adopts $r(p) = p$ as its risk function.

these local convexity assumptions are philosophically well motivated by the reasoning behind Allais preferences.

The REU values of the four lotteries in the Allais paradox are:

$$\text{REU(L1)} = u(w) + r(0.1)[u(w + 5 \times 10^6) - u(w)],$$

$$\text{REU(L2)} = u(w) + r(0.11)[u(w + 1 \times 10^6) - u(w)],$$

$$\text{REU(L3)} = u(w) + r(0.99)[u(w + 1 \times 10^6) - u(w)] + r(0.1)[u(w + 5 \times 10^6) - u(w + 1 \times 10^6)],$$

$$\text{and REU(L4)} = u(w + 1 \times 10^6).$$

Given that $\text{REU(L1)} > \text{REU(L2)}$ and $\text{REU(L4)} > \text{REU(L3)}$, there are:

$$r\,(0.1)[u\,(w + 5 \times 10^6) - u(w)] > r\,(0.11)[u\,(w + 1 \times 10^6) - u(w)], \quad (4.1)$$

$$\text{and } (1 + r(0.1) - r\,(0.99))[u\,(w + 1 \times 10^6) - u(w)] > r\,(0.1)[u\,(w + 5 \times 10^6) - u(w)]. \quad (4.2)$$

Following Buchak, I assume the standard conditions for a risk-avoidant[65] agent's risk function:

$r: [0, 1] \to r[0, 1]$, $r(0) = 0$, and $r(1) = 1$ (4.c.1)

r is continuous. (4.c.2)

r is strictly increasing. (4.c.3)

r is strictly convex. (4.c.4)

The convexity in condition (4.c.4) may seem too strong.[66] I will later argue that weaker local convexity conditions are sufficient and well-motivated by the Allais reasoning.

With a linear utility function, Inequalities (4.1) and (4.2) reduce to:

$5r(0.1) > r(0.11)$, (4.c.5)

[65] By "risk-avoidant", Buchak refers to people who depreciate (assign a smaller weight to payoffs, compared to their raw probabilities) improvements in utility. Formally, risk-avoidant agents are those agents with strictly convex risk functions. Conversely, *risk-inclined* agents are those who overvalue probabilistic improvements in utility, and have strictly concave risk functions.

[66] Although a convex risk function can accommodate the Allais paradox, it is plausible that convexity restricted to certain probability intervals is sufficient.

$$4r(0.1) < 1 - r(0.99). \qquad (4.c.6)$$

Condition (4.c.6) requires a sufficiently risk-avoidant attitude in the sense that one's risk function discounts the relevant probabilistic improvement in Lottery 3 enough that the improvement cannot compensate for the small chance of the worst outcome (hence one prefers Lottery 4 over Lottery 3). By contrast, condition (4.c.5) rules out excessive risk-avoidance: one's risk function cannot discount any relevant improvement so strongly that one would instead prefer Lottery 2 to Lottery 1. Taken together, (4.c.5) and (4.c.6) therefore constrain the admissible patterns of risk-avoidance (or, the admissible risk functions) by imposing lower and upper bounds on the "weights (imposed by the risk function)" at the specific probabilities that drive Allais preferences.

Consider Gamble I (GI) (displayed in Table 4.2). Here, the agent loses \$$D$ with probability p, wins \$$H$ with probability p, and receives nothing with probability $1 - 2p$.

Table 4.2 Gamble I (GI)

State/ Prob.	Prob. p	Prob. p	Prob. $1 - 2p$
Bet I	-\$ D	\$ H	0

Suppose an REU agent is indifferent between taking or rejecting the gamble. Indifference means that accepting the gamble leaves the overall REU value unchanged:

$$\mathrm{REU}(\mathrm{GI}) = \mathrm{REU}(w),$$

where w is the agent's initial wealth. Under this indifference condition, define a function

$$h_u(p, d, w) := H',$$

where H' is the required gain to offset a loss of size $D' = d$ at the wealth level of w.[67] Then $h_u(p, d, w)$ satisfies:[68]

$$u(w + h_u(p, d, w)) - u(w) = \frac{(1-r(1-p))}{r(p)}(u(w) - u(w - d)). \quad (4.3)$$

Define the utility tradeoff ratio:

$$c_u(p, d, w) := \frac{u(w+h_u(p,d,w))-u(w)}{u(w)-u(w-d)} = \frac{1-r(1-p)}{r(p)}, \quad (4.4)$$

which represents a tradeoff ratio between the potential utility gain and the potential utility loss to sustain indifference. Define the monetary tradeoff ratio:

$$m_u(p, d, w) := \frac{h_u(p,d,w)}{d},$$

which represents a tradeoff ratio between the monetary gain and the monetary loss to sustain indifference. The term $c_u(p, d, w)$ isolates and quantifies, given a specific utility function, the contribution of the risk function to the monetary tradeoff ratio $m_u(p, d, w)$. In the special case of a linear utility function, the utility tradeoff ratio equals the monetary tradeoff ratio.[69]

For simplicity, suppose $p = 0.01$. By condition (4.c.6) [risk-aversion] and (4.c.4) [convexity],

$$c_{linear}(0.01, d, w) = \frac{h_{linear}(0.01,d,w)}{d} = \frac{1-r(0.99)}{r(0.01)} > \frac{4r(0.1)}{r(0.01)} > 40.^{70} \quad (4.5)$$

---

[67] The required gain depends on both the risk function r and the utility function u. I subscript by u and leave r implicit. In Section 4.5, r is taken to satisfy the Allais constraints, in particular, at the margin used to derive the lower bounds.

[68] Equation (4.3) follows directly from the indifference condition. Specifically,
$\text{REU}(\text{GI}) = u(w - d) + r(1 - p)\big(u(w) - u(w - d)\big) + r(p)(u\big(w + h(w, d)\big) - u(w))$ and $\text{REU}(w) = u(w)$.
Setting $\text{REU}(\text{GI}) = \text{REU}(w)$ yields Equation (4.3).

[69] Because a utility function is unique up to an affine transformation, when the utility function is linear, assuming $u(x)$=x does not reduce generality. Accordingly, $c_{linear}(p)$ equals $m_{linear}(p, d, w)$, i.e., $c_{linear}(p) = m_{linear}(p, d, w) = \frac{h_{linear}(p,d,w)}{d}$. By contrast, when the utility function is non-linear, the monetary tradeoff ratio and the utility tradeoff ratio need not coincide.

[70] By strict convexity and condition (4.c.1), for any $0 < y < x$ and $y = \lambda x$ where $0 < \lambda < 1$, there is

$$r(y) = r(\lambda x + (1 - \lambda) \cdot 0) < \lambda r(x) + (1 - \lambda)r(0) = \lambda r(x)$$

which implies $r(y) < \lambda r(x)$. Dividing both sides by y gives $\frac{r(y)}{y} < \frac{r(x)}{x}$. Substituting $y = 0.01$ and $x = 0.1$ yields $\frac{r(0.1)}{r(0.01)} > 10$.

When (4.c.6), i.e., $\mathrm{REU(L4)} > \mathrm{REU(L3)}$, holds just at the margin, $c_{linear}(0.01, d, w)$ reaches its smallest value.

Thus, to offset a \$1 loss, the winning payoff needs to be at least \$40; correspondingly, for a loss of \$50,000, the required gain is at least \$2 million. Suppose the agent's initial wealth is \$5 million. Such an agent would not accept a gamble with a probability of 0.01 of losing 1% of their wealth against an equal probability of increasing their wealth by less than 40%. Since the ratio "40" is derived under a linear utility function, the number 40 is independent of any specific level of wealth or the size of the loss. Nevertheless, this degree of risk aversion seems implausible.

The key step in inequality (4.5) is that convexity ensures $\frac{r(0.1)}{r(0.01)} > 10$. Admittedly, assuming global convexity of the risk function on $[0, 1]$ is unnecessarily strong. In fact, convexity only on the interval $[0.01, 0.1]$ suffices. Philosophically, such convexity on the range of small probability, e.g., on the interval of $[0, 0.1]$,[71] aligns with the reasoning behind the Allais paradox. The motivation for preferring L4 over L3 in the Allais paradox is that the relatively small chance (0.10) of winning \$5 million cannot compensate for the tiny (0.01) chance of receiving nothing. Put differently, the improvement from \$1 million to \$5 million is undervalued relative to avoiding the small chance of zero payoff. This undervaluation is what (a risk function's) convexity on small probabilities captures.

It is worth stressing that this point is philosophical rather than mathematical. Mathematically, condition (4.c.6) alone is sufficient for preferring L4 over L3, and it does not entail convexity on $[0, 0.1]$. For example, a risk function that is linear or even slightly concave on $[0, 0.1]$ but sharply convex near 1 could still satisfy (4.c.6). Such a risk function would treat low

[71] Convexity on $[0.01, 0.1]$ is mathematically sufficient. Hereafter, I adopt convexity on $[0, 0.1]$, since it seems reasonable to suppose that one's risk attitude does not flip precisely on $[0, 0.01]$.

probabilities in a risk-neutral or risk-inclined (overvalues improvement) manner, while showing strong risk-avoidance (undervalues improvement) at high probabilities. In the Allais context, this would mean overvaluing (being optimistic about) the 0.10 chance of \$5million while undervaluing (being pessimistic about) the 0.99 chance of receiving at least \$1 million conditional on the worst outcome of nothing. This pattern does not match standard explanations of Allais preferences and lacks philosophical plausibility. By contrast, convexity on $[0, 0.1]$ accords with the intuitive reasoning that drives those preferences and, thus, provides a well-motivated basis for the calibration result.

Moreover, if the risk function is also weakly convex (linear or convex) on $[0.99, 1]$, the inequality $c_{linear}(p, d, w) > 40$ will hold for $p < 0.01$.[72] Allais preferences indicate strong sensitivity to the 0.01 probability of the unfavorable outcome (the \$0 payoff) in L3, which makes that lottery appear especially unattractive. If an agent treats events with probability 0.99 (at least \$1 million in L3) as effectively certain, they should practically ignore 0.01 chance of nothing and cannot have Allais preferences. In other words, they should not overvalue highly likely improvements, which implies that the risk function should not be concave on $[0.99, 1]$. Weak convexity at high probabilities is therefore philosophically well motivated, even if not mathematically required.

In sum, when the utility function is linear, REU is subject to calibration under highly plausible conditions: namely, convexity of the risk function at small probabilities. If one also accepts weak convexity at high probabilities, the scope of calibration broadens even further.

[72] Instead of a formal proof, consider the following intuitive explanation. The inequality $c_{linear}(p, d, w) > 40$ holds, when $p = 0.01$. First, suppose $p < 0.01$. As p decreases from 0.01, convexity ensures that $r(p)$ decreases relatively fast. Meanwhile, if $r(p)$ is at least weakly concave, then $1 - r(1 - p)$ does not increase too slowly. Consequently, the ratio $c_{linear}(p, d, w) = \frac{1-r(1-p)}{r(p)}$ must exceed 40. Therefore, when $p < 0.01$, there must be $c_{linear}(p, d, w) > 40$ as well.

Linear Utility Calibration Theorem (LUC) Under conditions (4.c.1) through (4.c.3) with a risk function convex on $[0, 0.1]$ and weakly convex on [0.99, 1], accommodating Allais preferences (conditions (4.c.5) and (4.c.6)) commits REU with a linear utility function to extreme preferences. In Gamble I, the REU agent will not risk losing a given amount of money for an equal probability (less than 0.01) of gaining up to 40 times that loss.

4.5 The General Calibration Theorem

The results in Section 4.4 rely on the assumption of a linear utility function. Economists generally regard a concave utility function as more realistic (reflecting diminishing marginal utility). Furthermore, adopting a concave utility function appears to offer REU a potential escape from calibration. Allais preferences reflect a certain degree of risk aversion, and Section 4.4 shows that if this risk aversion is accommodated solely by the risk function, REU is subject to calibration. Intuitively, then, one might expect that splitting the burden between a concave utility function and a risk function could relieve REU from calibration.

This section shows that the above strategy does not succeed in full. I demonstrate that, within REU, once Allais preferences are imposed, any combination of a concave utility function in the constant-relative-risk-aversion family and a risk function remains subject to calibration. The linear-utility case (weakly concave) from Section 4.4 is one endpoint of the admissible range of concavity considered here. Section 4.5 extends Section 4.4 by ranging over this wider class of utility functions, but for any fixed utility function, the derivation follows the same pattern: Allais preferences yield a lower bound of the utility tradeoff ratio in Gamble I, which in turn yields a lower bound of the monetary tradeoff ratio.

To extend the calibration results beyond the linear-utility case, I adopt the Constant Relative Risk Aversion (CRRA) utility assumption: the utility function treats gains and losses that are a fixed proportion of wealth in a way that is stable across wealth levels. A key implication of CRRA is a kind of scale invariance in behaviors: under both EU and REU with CRRA utility, scaling up (or down) the agent's initial wealth and every payoff in any two lotteries by the same factor (so that all stakes remain the same proportion of wealth) does not change the preference ordering.[73] Thus, adopting CRRA ensures that an REU agent's preferences respond in a stable (scale-invariant) way to uncertainty over payoffs that are fixed proportions of wealth.

Indeed, assuming some constancy in the utility function (i.e., cross-wealth stability in preferences) is standard in establishing a calibration theorem. Both Rabin's and Safra-Segal's theorems assume that agents reject certain small-stakes gambles (e.g., a 50-50 bet with losing $100 or winning $105) across a sufficiently wide range of wealth. Without this kind of cross-wealth stability, risk aversion at a single wealth level cannot be extended to a global restriction on the utility function. Their assumption is CARA-like: under Constant Absolute Risk Aversion (CARA), absolute risk aversion is constant, so attitudes towards fixed dollar amounts of money do not vary with wealth. By contrast, the Allais paradox is typically framed as a high-stakes choice problem, and a recent empirical study suggests that incentive size can affect the strength of Allais-type preferences (Gneezy, Halevy, Hall, Offerman, & van de Ven, 2024). Given the large stakes, it is

[73] Within EU, CRRA means that the Arrow–Pratt coefficient of relative risk aversion is constant, which (in EU) provides the standard measure of "relative" risk aversion (see footnote 74). An REU agent's utility function may still be CRRA, but their overall risk attitude cannot be captured by Arrow-Pratt coefficient alone, since REU preferences also depend on the risk function. Nevertheless, within REU, CRRA utility still implies a CRRA-like scale invariance: if wealth and all payoffs are scaled by the same factor, the preference ordering will not be changed. Intuitively, CRRA makes utility "scale" with wealth. Since the risk function in REU transforms only the relevant probabilities and does not depend on payoff size, rescaling leaves the REU preference ordering unchanged. Conceptually, the utility function and the risk function capture different aspects of uncertainty: utility reflects how marginal valuations of wealth change with wealth, whereas the risk function represents attitudes towards the distribution of outcomes.

natural to adopt CRRA utility because it models risk attitudes in terms of proportional changes in wealth rather than fixed-dollar changes (Eeckhoudt, Gollier, & Schlesinger, 2005).

Accepting CRRA utility is equivalent (up to positive affine transformation) to adopting the standard CRRA utility function:

$$u_\rho(x) = \begin{cases} \frac{x^{1-\rho}-1}{1-\rho}, \rho \neq 1 \; and \; \rho \geq 0 \\ \ln(x), \rho = 1 \end{cases},^{74}$$

where ρ is the coefficient of relative risk aversion (Pratt, 1964). When $\rho = 0$, u_0 is linear; when $\rho > 0$, u_ρ is concave, and larger values of ρ correspond to greater concavity, i.e., greater Arrow-Pratt measure ($R_\rho(x) = -x\frac{{u_\rho}''(x)}{{u_\rho}'(x)}$) of relative risk aversion. CRRA thus provides a natural and general framework: it avoids the unrealistic implications of CARA by stabilizing concavity in a proportionally meaningful way.

The CRRA assumption translates the accommodation of Allais preferences to explicit constraints on the utility function. I continue with $p = 0.01$ in Gamble I. Combining Inequalities (4.1) and (4.2) yields a constraint on the utility function:

$$\frac{r(0.11)}{r(0.1)} < \frac{u_\rho(w+5\times10^6)-u_\rho(w)}{u_\rho(w+10^6)-u_\rho(w)} < 1 + \frac{1-r(0.99)}{r(0.1)}. \quad (4.6)$$

For ease of reference, let

$$t_w(\rho) := \frac{u_\rho(w+5\times10^6)-u_\rho(w)}{u_\rho(w+10^6)-u_\rho(w)}.$$

[74] The Constant Relative Risk Aversion (CRRA) utility function is the unique family of utility functions (up to affine transformation) that yields a constant Arrow–Pratt (Pratt, 1964; Arrow, 1965) measure of relative risk aversion. Absolute risk aversion is defined as $A(x) = -\frac{u''(x)}{u'(x)}$, while relative risk aversion (Arrow-Pratt) is defined as $R(x) = xA(x)$. For CRRA utility, $R(x)$ is a constant and equal to ρ. Thus, adopting a CRRA utility function entails no loss of generality once the assumption of constant relative risk aversion is in place. Technically, one could attempt to model constant relative risk aversion without specifying the CRRA functional form, i.e., by iterating a constant relative risk attitude along the utility function in the way Rabin iterates constant absolute risk aversion. However, because each iteration would have to be recalibrated to the agent's current wealth level, the procedure quickly becomes intractable. Rabin also illustrates his calibration using a specific form of CARA utility function. More generally, adopting a particular functional form that enforces some constant risk attitude does not qualitatively alter any conclusion.

Intuitively, $t_w(\rho)$ quantifies the effective curvature of u_ρ over the Allais-relevant payoff range at wealth level w. Under CRRA and by the convexity of r, accommodating Allais preferences restricts $t_w(\rho)$ to an admissible range:

$$1.1 < t_w(\rho) \leq 5,^{75}$$

with $t_w(0) = 5$ in the linear case.

With CRRA in place, each admissible value of $t_w(\rho)$ corresponds to a unique value of ρ, and thus (up to affine transformation) to a unique utility function in the CRRA family. Given such a utility function, accommodating Allais preferences imposes a lower bound on the utility tradeoff ratio in Gamble I:

$$c_\rho(0.01, d, w) = \frac{u_\rho(w + h_\rho(0.01,d,w)) - u_\rho(w)}{u_\rho(w) - u_\rho(w-d)} > 10 \cdot (t_w(\rho) - 1),^{76} \quad (4.7)$$

where $h_\rho(0.01, d, w)$ is the gain that sustains the indifference $\mathrm{REU}(\mathrm{GI}) = \mathrm{REU}(w)$. Inequality (4.7) yields a lower bound of the required gain, denoted as $\underline{h}_\rho(0.01, d, w)$. The lower bound $\underline{h}_\rho(0.01, d, w)$ determines a lower bound of the monetary tradeoff ratio:

$$\underline{m}_\rho(0.01, d, w) := \frac{\underline{h}_\rho(0.01,d,w)}{d},^{77}$$

whose magnitude represents how extreme the risk-averse behavior is to which REU is committed. This is where CRRA becomes essential: t_w constrains utility curvature over the Allais-relevant payoff range, but without a stability restriction, that constraint need not translate into any

[75] By convexity of r, on the relevant small-probability interval, $\frac{r(0.11)}{r(0.1)} > 1.1$, so Inequality (4.6) yields a lower bound of t_w. Under CRRA, t_w monotonically increases as ρ decreases (i.e., as u_ρ becomes less concave). In particular, under the linear case $\rho = 0$, t_w reaches its maximum 5.

[76] See Appendix E for the proof. For fixed w, each admissible value $t = t_w(\rho)$ corresponds to a unique ρ; thus indexing the CRRA family by ρ is equivalent to indexing it by t.

[77] Since u_ρ is strictly increasing, the solution from setting Inequality (4.7) at equality ensures $h_\rho(0.01, d, w) > \underline{h}_\rho(0.01, d, w)$, correspondingly, $m_\rho(0.01, d, w) > \underline{m}_\rho(0.01, d, w)$.

restriction on the utility tradeoff ratio that is relevant to Gamble I.[78] Intuitively, CRRA rules out such "erratic" curvature by requiring the curvature to level off at some stable rate.

Given an admissible $t_w(\rho) \in (1.1, 5]$, a potential loss d, and initial wealth w, I compute the lower bounds of the required gain $\underline{h}_\rho(0.01, d, w)$ and the monetary tradeoff ratio $\underline{m}_\rho(0.01, d, w)$ as follows:

1. For any value of $t_w(\rho) \in (1.1, 5]$, identify the corresponding value of ρ in the CRRA utility function;
2. The lower bound $\underline{c}_\rho(0.01, d, w)$ is $10 \cdot (t_w(\rho) - 1)$;
3. For each value of $t_w(\rho)$, with its corresponding ρ and the lower bound of $c_\rho(0.01, d, w)$, the lower bound of the monetary tradeoff ratio $\underline{m}_\rho(0.01, d, w)$ is determined by solving for $\underline{h}_\rho(0.01, d, w)$ in the equation:

$$\frac{u_\rho(w+\underline{h}_\rho(0.01,d,w))-u_\rho(w)}{u_\rho(w)-u_\rho(w-d)} = 10 \cdot (t_w(\rho) - 1).$$

Assuming an initial wealth of \$5 million, Figure 4.1 plots how the lower bound of the monetary tradeoff ratio $\underline{m}_\rho(0.01, d, w)$ varies with different utility functions in face of relatively small (compared to the \$5 million baseline) losses.

[78] In particular, a utility function could be nearly linear around w (so that small losses d have little marginal effect) while bending sharply near $w + 5 \times 10^6$, thereby generating the same t without imposing a substantive constraint on the utility tradeoff ratio.

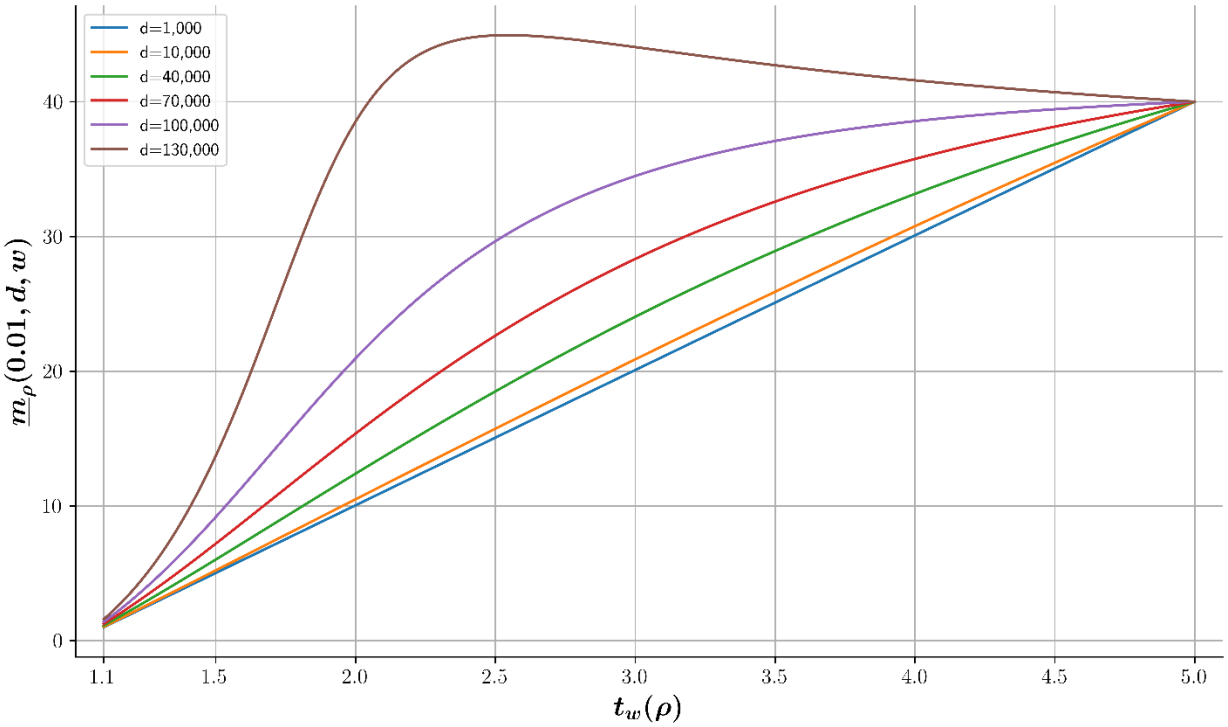

Figure 4.1. The lower bound of monetary tradeoff ratios $\underline{\boldsymbol{m}}_{\boldsymbol{\rho}}(\mathbf{0.01}, \boldsymbol{d}, \boldsymbol{w}) = \underline{\boldsymbol{h}}_{\boldsymbol{\rho}}(\mathbf{0.01}, \boldsymbol{d}, \boldsymbol{w})/\boldsymbol{d}$, where $\boldsymbol{w} = \mathbf{5} \times \mathbf{10^6}$ for $\boldsymbol{t_w}(\boldsymbol{\rho}) \in (\mathbf{1.1}, \mathbf{5}]$ for relatively small $\boldsymbol{d}$. The lower bound of $\boldsymbol{c_\rho}(\mathbf{0.01}, \boldsymbol{d}, \boldsymbol{w})$, determined by $\mathbf{REU(L4)} > \mathbf{REU(L3)}$ at the margin, is used to determine $\underline{\boldsymbol{h}}_{\boldsymbol{\rho}}(\mathbf{0.01}, \boldsymbol{d}, \boldsymbol{w})$ and $\underline{\boldsymbol{m}}_{\boldsymbol{\rho}}(\mathbf{0.01}, \boldsymbol{d}, \boldsymbol{w})$.

As suggested by the result from Section 4.4, the ratios converge to 40 as the utility function approaches linearity. When the utility function is maximally concave (correspondingly, t_w approaches 1.1) and the risk function approaches linearity, the ratios in Figure 4.1 are much smaller. For instance, when the loss d is \$130,000, the ratio remains as low as 1.595. This looks promising as a way for REU to avoid calibration. However, \$130,000 is only 2.6% of the \$5 million initial wealth, and Rabin's calibration suggests that a concave utility function tends to generate extreme behavior once the potential loss is large. So far, it is still unknown how large d must be before implausibly high monetary tradeoff ratios emerge. REU could escape calibration if extremely high

ratios appear only for extremely large losses, because it is more reasonable to demand a high tradeoff ratio when d constitutes a substantial share of wealth.

In fact, however, the potential loss d does not have to be too large to derive extremely high tradeoff ratios. Figure 4.2 plots the monetary tradeoff ratios for moderately large losses of \$300,000, \$400,000, and \$500,000.[79]

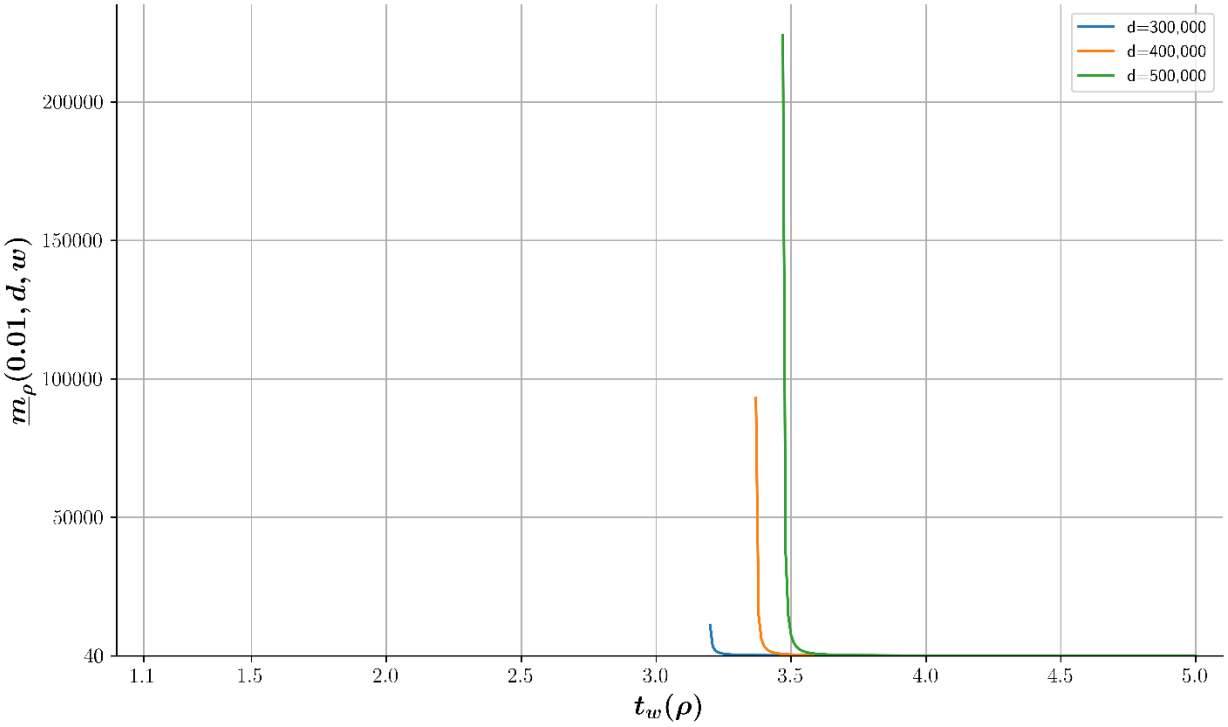

Figure 4.2. The lower bound of monetary tradeoff ratios $\underline{\boldsymbol{m}}_{\boldsymbol{\rho}}(\mathbf{0.01}, \boldsymbol{d}, \boldsymbol{w}) = \underline{\boldsymbol{h}}_{\boldsymbol{\rho}}(\mathbf{0.01}, \boldsymbol{d}, \boldsymbol{w})/\boldsymbol{d}$, where $\boldsymbol{w} = \mathbf{5} \times \mathbf{10^6}$ for $\boldsymbol{t_w}(\boldsymbol{\rho}) \in (\mathbf{1.1}, \mathbf{5}]$ with moderately large $\boldsymbol{d}$. The lower bound of $\boldsymbol{c_\rho}(\mathbf{0.01}, \boldsymbol{d}, \boldsymbol{w})$, determined by $\mathbf{REU(L4)} > \mathbf{REU(L3)}$ at the margin, is used to determine $\underline{\boldsymbol{h}}_{\boldsymbol{\rho}}(\mathbf{0.01}, \boldsymbol{d}, \boldsymbol{w})$ and $\underline{\boldsymbol{m}}_{\boldsymbol{\rho}}(\mathbf{0.01}, \boldsymbol{d}, \boldsymbol{w})$. Blank segments of curves indicate that for those $\boldsymbol{t_w}(\boldsymbol{\rho})$ values, no finite $\boldsymbol{h_\rho}(\mathbf{0.01}, \boldsymbol{d}, \boldsymbol{w})$ can offset the corresponding loss $\boldsymbol{d}$.

[79] See Appendix F for results with losses in the range of roughly \$150,000 to \$210,000. Monetary tradeoff ratios led by losses of this size are not sufficient to calibrate REU, but the results are nevertheless instructive.

Figure 4.2 illustrates that extreme results emerge under any combination of risk and utility functions.[80] For losses as small as 6% ($300,000) of initial wealth, agents with sufficiently concave utility ($t_w(\rho) < 3.2$) will reject Gamble I regardless of how large the potential gain is. As their utility function approaches linearity, the tradeoff ratios monotonically decrease and converge towards 40, which is essentially the calibration result from Section 4.4. More importantly, the combinations of utility and risk functions that yield the least problematic behavior for small stakes are the ones that yield the most extreme preferences for larger stakes, and vice versa.[81] This structural tension makes the calibration especially robust.

Furthermore, a loss of $256,805.19 (roughly 5.14% of the initial wealth) is enough to produce extreme results under any combination of risk and utility functions (i.e., a curve shaped similar to those in Figure 4.2).[82] Beyond this threshold, in the least extreme case, an REU agent, regardless of their risk function, would reject Gamble I unless the winning payoff exceeds 40 times the loss. This reproduces the problem Rabin highlights: under a concave utility function, the interaction of curvature and moderately large stakes inevitably generates implausibly risk-averse behavior. The upshot is that even a potential loss of about 5% of one's wealth does not plausibly justify the extreme tradeoff ratios that REU predicts in Gamble I.

Therefore, once Allais preferences are accommodated, no combination of risk and utility functions can shield REU from calibration. REU must either exhibit extreme risk-aversion (via the risk-avoidant risk function) or fall prey to Rabin's criticism of concave utility functions. When the

[80] All results so far concern conservative lower bounds of h given d, since they are derived by combining several inequalities (see Inequality (4.7)). Flattening each inequality introduces slack, so the actual tradeoff ratios are likely to be substantially higher than the bounds reported in Figures 41 and 4.2. In Appendix G, I use a specific risk function to illustrate more realistic ratios.

[81] Proponents of REU might argue that, for stakes equal to 6% of initial wealth, demanding a payoff 40 times the potential loss does not look especially extreme. However, as shown in Section 4, with a linear or nearly linear utility function combined with a risk-avoidant risk function, an REU agent would demand roughly 40 times the loss even for small or tiny stakes—a requirement that does seem implausibly extreme.

[82] See Appendix H for an analytic solution. The exact loss required to calibrate REU depends on initial wealth. For example, when initial wealth is $1 million, $3 million, and $7 million, the critical loss is $183,032.23 (18.33% of wealth), $239,656.92 (7.99%), and $265,114.41 (3.79%), respectively.

risk function shoulders most of the burden of accommodating the risk aversion in Allais preferences, the monetary tradeoff ratio remains bounded at 40, no matter how large the potential loss. By contrast, when the risk function takes on less of this burden, the concavity of the utility function must do more of the work. In that case, concavity itself drives extreme behavior, and REU becomes calibratable once the stakes reach only a moderately large size.

> **The General Calibration Theorem (GC)** Under conditions (4.c.1) through (4.c.3) with a risk function convex on $[0, 0.1]$ and weakly convex on [0.99, 1], accommodating Allais preferences (conditions (4.c.5) and (4.c.6)) commits REU with a CRRA utility function to extreme preferences. For an agent with an initial wealth of \$5 million, in Gamble I (with equal winning and losing probability $p \leq 0.01$),[83] the REU agent will not accept any gamble that risks a loss greater than \$256,805.19 for an equal probability of gaining up to 40 times that loss.

Informally, given a continuous and strictly increasing risk function that is convex on relatively small and sufficiently large probabilities, accommodating Allais preferences commits REU to implausible risk-averse behavior in gambles involving moderately large stakes at small probabilities (as exemplified by Gamble I).

4.6 Conclusion

This paper presents a calibration theorem for REU. The result suggests that, contrary to Buchak's advertisement, REU does not genuinely solve either the Allais paradox or Rabin's criticism. Although Rabin's calibration theorem against EU does not apply to REU, once Allais preferences are accommodated, REU itself becomes calibratable, which is precisely the type of

[83] See footnote 72.

problem Rabin highlights for EU. Unlike existing calibration theorems against RDU/REU (Neilson, 2001; Safra & Segal, 2008), my calibration does not assume that agents maintain a certain risk-averse preference at any level of wealth, nor does it require uncertainty innate to one's initial state. The assumptions underpinning my theorem are therefore weaker and, I argue, more realistic.

Proponents of REU may question whether a factor of 40 is strong enough to establish calibration, and I agree that this judgment ultimately depends on empirical intuitions. Still, modest adjustments to the Allais paradox can increase the tradeoff ratio while preserving the intuitive appeal of the preferences. For instance, if the highest payoff in L3 were \$10 million instead of \$5 million, an agent could still reasonably prefer L4 over L3. Yet, the ratio would rise from 40 to 90. Similarly, if the probability of the highest payoff \$5 million goes up from 0.1 to 0.15, the number would rise from 40 to 60. These adjustments show that Allais-like preferences, which remain intuitively rational, can yield far higher ratios. Moreover, it is worth recalling that figures like 40, 60, and 90 are conservative lower bounds derived by combining multiple inequalities; the actual ratios are likely to be considerably higher. I make no claims about what a reasonable tradeoff ratio could or ought to be. Rather, I wish to simply point out the fact that there are many possible ways to construct an intuitively acceptable set of preferences that increase the tradeoff ratio and, therefore, strengthen my approach of calibrating REU.

Despite the calibration result, I do not mean to deny the insights REU offers. Although both (1) diminishing marginal utility and (2) REU's depreciating uncertainty over marginal improvement in payoffs contribute to an agent's risk-averse behaviors, these are nevertheless two distinct concepts that capture very different types of concerns related to utility. REU highlights a

dimension that EU ignores, though whether it is rational to incorporate that dimension is a separate normative question.

Even if it is rational to care about a gamble's probabilistic structure, the calibration result raises doubts about whether REU captures people's attitudes in a reasonable way. Since the function $c_u(p, d, w)$ is strictly decreasing on $[0, 1]$,[84] an REU agent requires a higher tradeoff ratio when the winning and losing probabilities in Gamble I are small. As Inequalities (4.6) and (4.7) show, Allais preferences effectively calibrate REU because they impose a lower bound of $c_u(p, d, w)$ for small probabilities (e.g., $p = 0.01$). I argue that this consequence (that smaller probabilities require/induce higher tradeoff ratios) seems questionable. In Gamble I, the probability of winning or losing is equal, so the agent is always facing an even-chance tradeoff. It is difficult to see why one would require a much higher tradeoff ratio facing small probabilities than in a 50-50 bet. Intuitively, one might think the opposite: when both probabilities are small, the gamble feels less serious (since most likely nothing will happen), and thus the agent might demand a lower, not higher, ratio. In this sense, the mathematical structure that enables REU to accommodate Allais preferences is also what renders it calibratable. Even if REU captures something important that EU neglects, REU fails to capture that thing in a reasonable way.

Both EU and REU therefore face deep challenges, but people still seek a decision framework they can trust. REU resolves the Allais paradox only at the cost of being vulnerable to calibration, which is essentially the same kind of criticism raised by Rabin. Thus, it is difficult to claim that REU genuinely escapes either the Allais paradox or Rabin's criticism; in the end, it

[84] Suppose r is strictly increasing and strictly convex. Correspondingly, $r(1-p)$ is convex, and $1-r(1-p)$ is concave. Because concavity $1-r(1-p)$ is concave and $1-r(1-0)=0$, $\frac{1-r(1-p)}{p}$ is strictly decreasing. Since $r(p)$ is convex, $\frac{r(p)}{p}$ is strictly increasing. It follows that
$c_u(p, d, w) = \frac{(1-r(1-p))/p}{r(p)/p}$
is strictly decreasing, since its numerator is decreasing and its denominator is increasing.

stands on par with EU. In my view, EU remains preferable, because it respects the probabilistic structure of outcomes. That said, EU still requires responses to both challenges. With respect to the Allais paradox, note that the two choices are driven by different reasoning: between L1 and L2, agents care primarily about expected payoff; between L3 and L4, they care about avoiding the unlikely but disastrous outcome. Because the reasoning shifts, the fact that EU cannot accommodate both simultaneously should not necessarily count as a decisive failure of a single, uniform theory. This may simply mean that EU needs to employ different utility functions across different contexts.[85] Admittedly, any context-dependent strategy faces difficulties in preserving its normative force, such as determining which contextual shifts warrant a change in utility function.

Rabin's calibration poses the more serious difficulty. In my view, the calibration's force lies in showing that people are generally not good at evaluating small stakes in relation to their overall wealth. Consider Rabin's assumption: an agent with $290,000 turns down a 50-50 bet of losing $100 or winning $105. Here, the potential loss of $100 is only 0.0345% of their wealth. It is difficult to see why such a tiny fraction should significantly curve their utility function. Everyday cases make the point clearer. Suppose the same person can return a defective item for a $100 refund, but doing so requires an hour-long trip, disputes with staff, or tedious paperwork. In this case, it is perfectly reasonable for the person to forgo the refund, and most people would not call this irrational. For someone with substantial wealth, $100 often makes little practical difference. The refusal of a small-stakes gamble seems better explained by people's aversion to gambling or

[85] A proponent of REU could, in principle, allow either the utility function or the risk function to vary across contexts. Whether the risk function should be treated as context-dependent is a further question that I do not settle here. Plausibly, treating desires, as represented by utility, as context-sensitive raises fewer worries than treating attitudes towards risk as context-sensitive. The latter, as captured by the risk function, concerns how agents structure the means to their ends. In any case, permitting context-dependent utility functions or risk function substantially increases REU's flexibility in accommodating preferences. In that sense, REU will remain largely on a par with similarly context-sensitive versions of EU. Once such context dependence is allowed, the comparison between EU and REU may turn on other theoretical virtues. For example, EU's commitment to evaluating gambles by probabilities, rather than by context-specific probability weights. I thank an anonymous reviewer for asking whether REU can take a similar context-dependent approach, and in particular whether the risk function can vary across contexts.

their negative attitudes towards risking a loss, rather than by the precise role that $100 plays in their utility function. For these reasons, I remain sympathetic to EU. It respects the probabilistic structure of outcomes, and I have outlined several potential responses it might offer to the difficulties. After all, REU cannot resolve these difficulties any better.

Finally, recent philosophical attention has focused on alternative approaches to modeling the risk attitudes that EU ignores (Stefánsson & Bradley, 2019; Bottomley & Williamson, 2024). Future research could explore whether these theories are vulnerable to calibration.

Appendices

A. Different Extents of RE

Regarding RE, after a reasonably long time, as agents accumulate evidence and discount older information, RE prompts agents to cluster around a compromise position. This is because as agents are close to each other, they are exposed to roughly the same newly discovered evidence. This ensures a convergence that reduces the impact of extent in each round. Data from my simulation aligns with the analysis and supports the following claim regarding various extents of RE.

> **Claim A.1** When agents are communicating evidence with random others, any reasonable extent (not extremely small) of communication does not have any significant impact on whether the majority of a group finds the truth within the maximum number of steps.

Furthermore, some additional use of others' information provides an advantage over groups that refrain from communicating any evidence or conclusions (PE-PC). However, when agents have a large scope, this effect diminishes since their wider scope entails that they gain access to truth-conducive evidence on their own without relying on communication. Accordingly, learning from others becomes less important. Moreover, since agents are expected to find the truth in the long run, given a longer time, the marginal benefit of communicating evidence over not communicating it diminishes.

> **Claim A.2** Compared to the practice of not communicating any information (PE-PC), communicating evidence with random others (practices FE-PC, 1/4RE-PC, and 1/2RE-PC) always boosts a group's performance. The improvement is less significant when agents have a relatively large scope, and the relative advantage generally decreases as the number of maximum allowed steps increases.

B. Correlation Between the First Asserted Conclusion and Groups' Performance

For .65NE-FC and FE-FC, I count, in each practice, the number of groups whose first asserted conclusion (if any) is correct. In Figure B.1, I compare the percentage (out of 1,200 simulations) of groups whose first asserted conclusion is correct. The difference in percentages between the two practices determines the color lightness in each cell. A comparison between Figure B.1 and Figure 2.7 reveals a significant overlap in their overall patterns, indicating a positive correlation between the accuracy of the first asserted conclusion and the group's performance.

In addition, I count the average steps the two practices need for the first correct conclusion to be asserted. In case no agent in a group finds the truth, the duration, 1,500, is recorded. In Figure B.2, I compare the percentage of steps (out of 1,500) at which the first correct asserted conclusion emerges. The difference in these percentages between the two practices determines the color lightness in each cell. The .65NE-FC groups consistently find the truth earlier than the FE-FC groups. The areas where .65NE-FC finds the truth significantly earlier overlap with areas where .65NE-FC outperforms FE-FC, indicating a positive correlation between the *speed* of the first asserted conclusion and the group's overall performance.

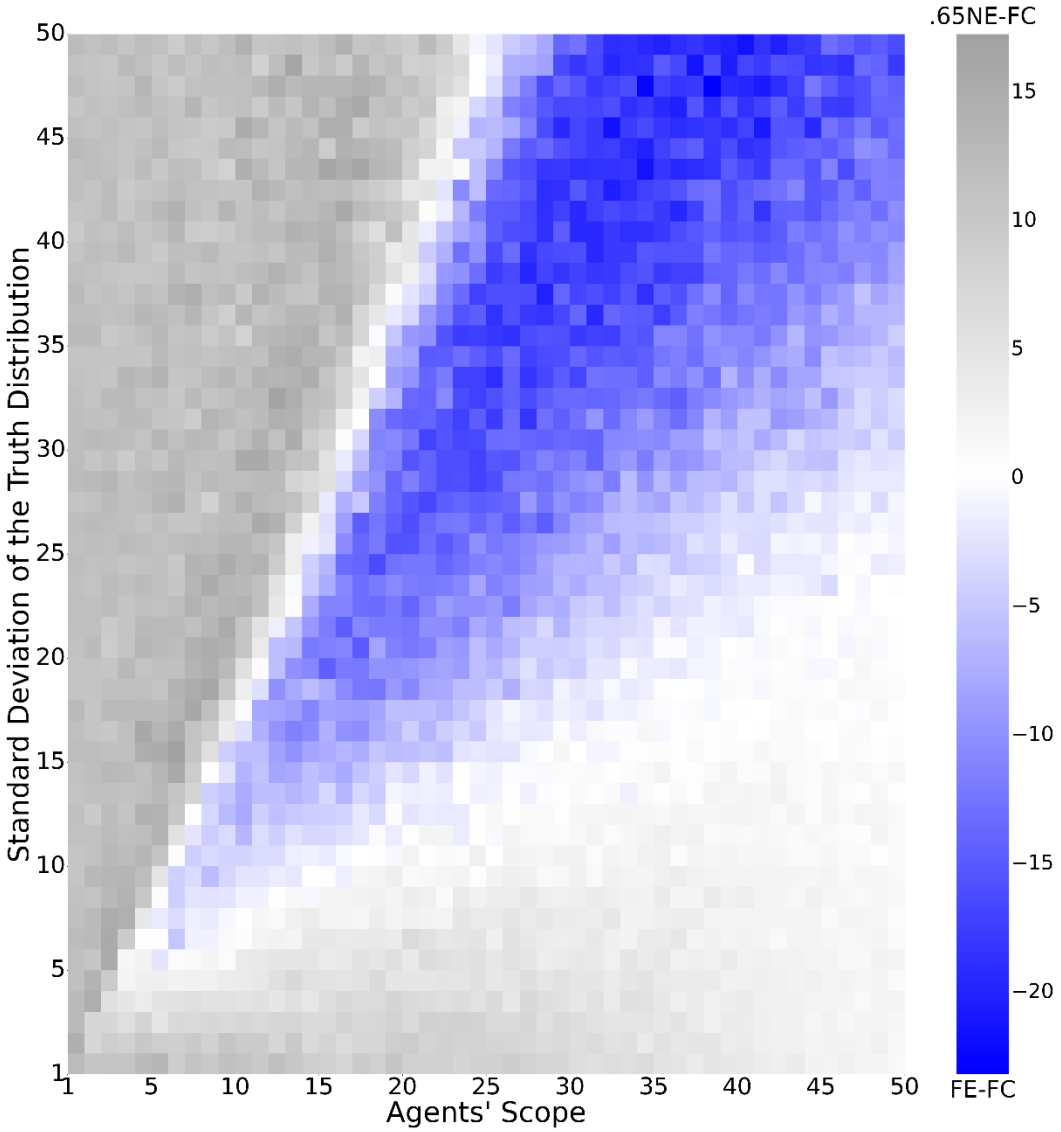

Figure B.1 Number of first asserted conclusions being correct, .65NE-FC compared to FE-FC at 1,500 steps. Gray indicates that .65NE-FC has more correct first asserted conclusions than FE-FC, and blue indicates the opposite. Lightness indicates the magnitude of the difference in the number of correct conclusions. Abbreviations: FC, full communication of conclusions; FE, full communication of evidence; NE, neighborhood communication of evidence.

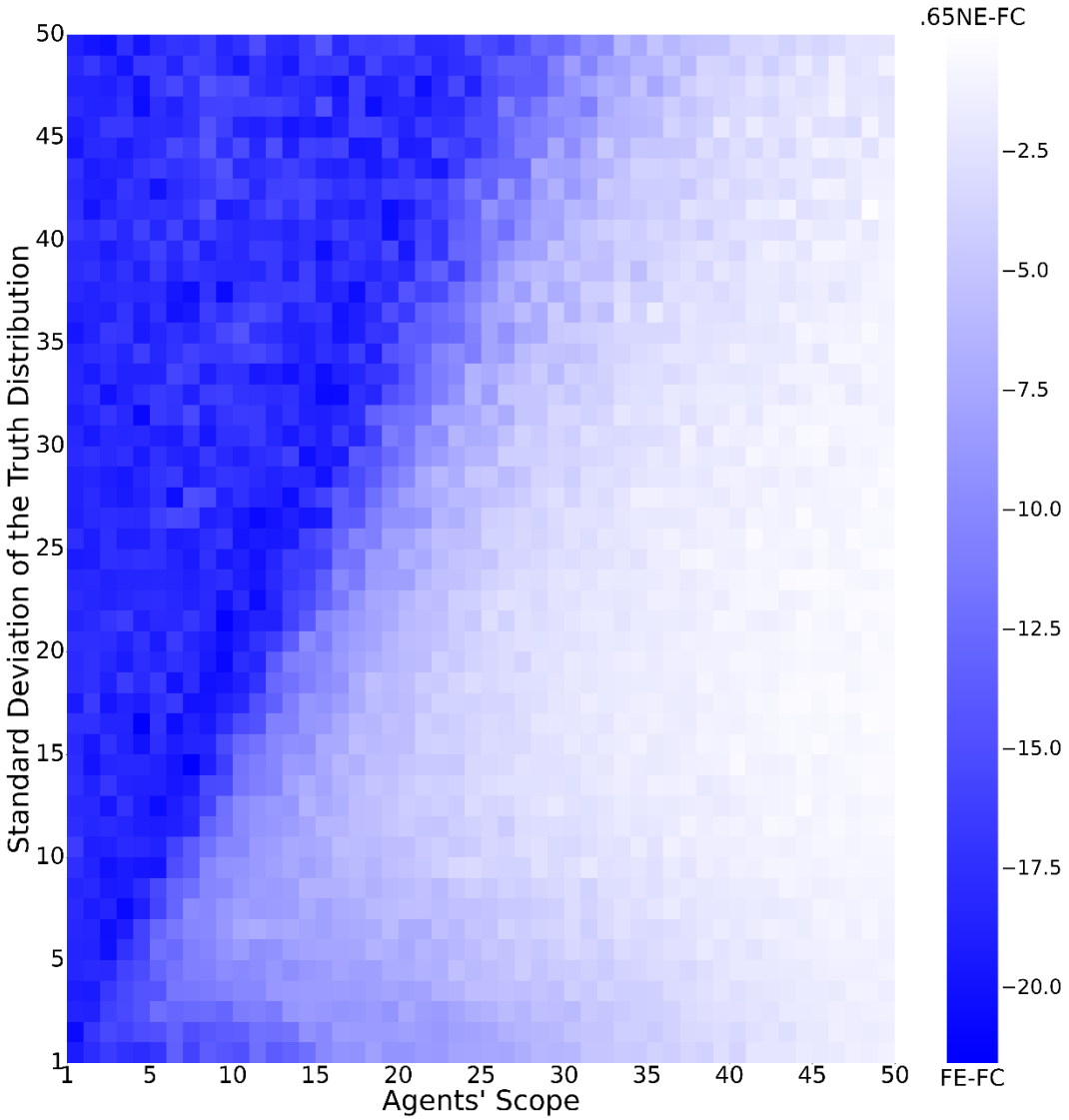

Figure B.2 Timing of the first correct conclusion's emergence, .65NE-FC compared to FE-FC at 1,500 steps. Blue indicates that the first correct conclusion appears earlier in .65NE-FC than in FE-FC, while gray indicates the opposite. Lightness indicates the magnitude of difference in the timing of the first correct conclusion. Abbreviations: FC, full communication of conclusions; FE, full communication of evidence; NE, neighborhood communication of evidence.

C. *Identifying Optimal Practices at Different Steps*

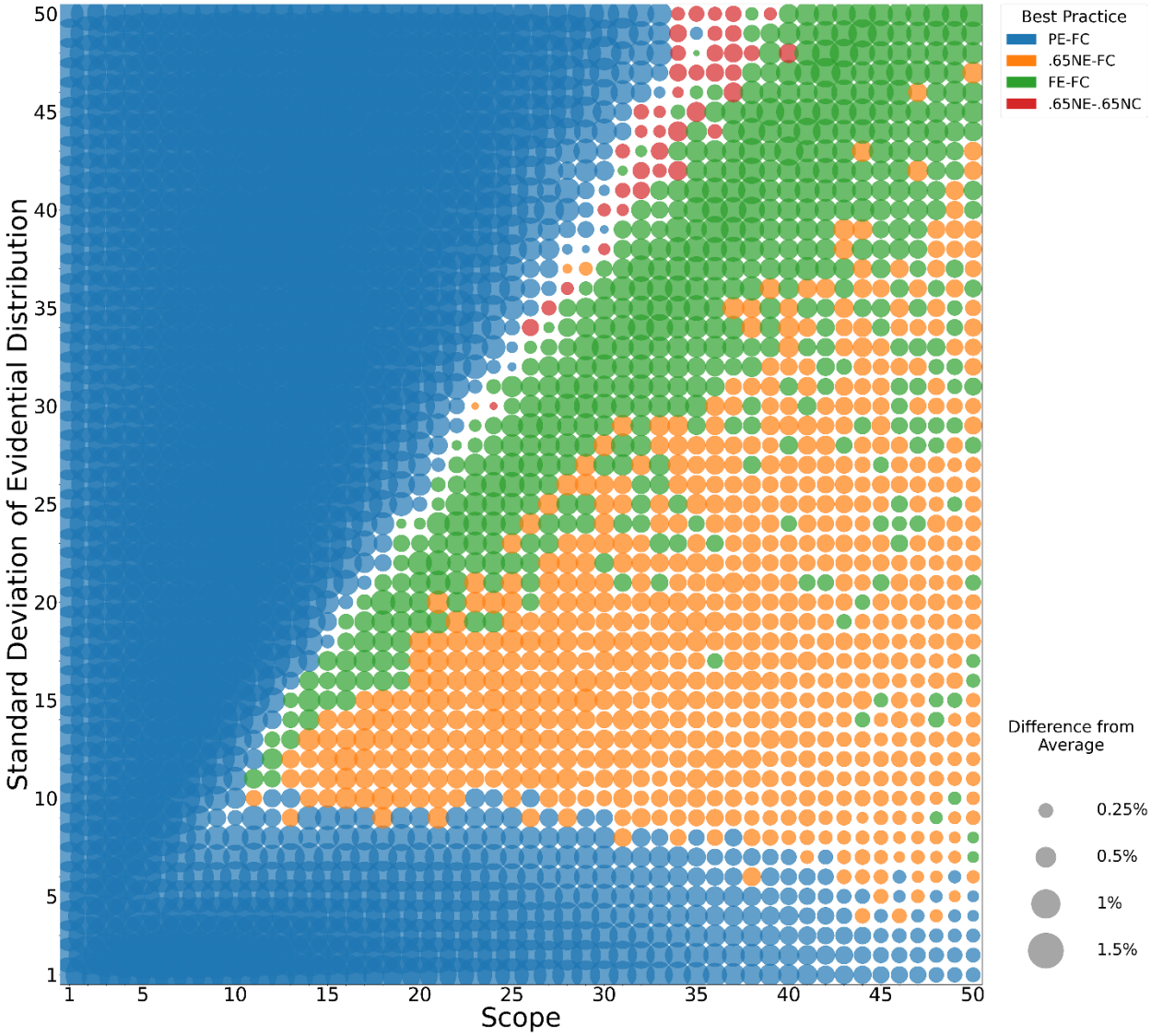

Figure C.1 The best performance practice across the parameter space at 750 steps. Abbreviations: FC, full communication of conclusions; FE, full communication of evidence; NC, neighborhood communication of conclusions; NE, neighborhood communication of evidence; PE, private evidence. The size of each dot indicates how much the best-performing practice exceeds the average of the four practices, measured by the percentage (out of 1,200 simulations) of groups in which the majority finds the truth.

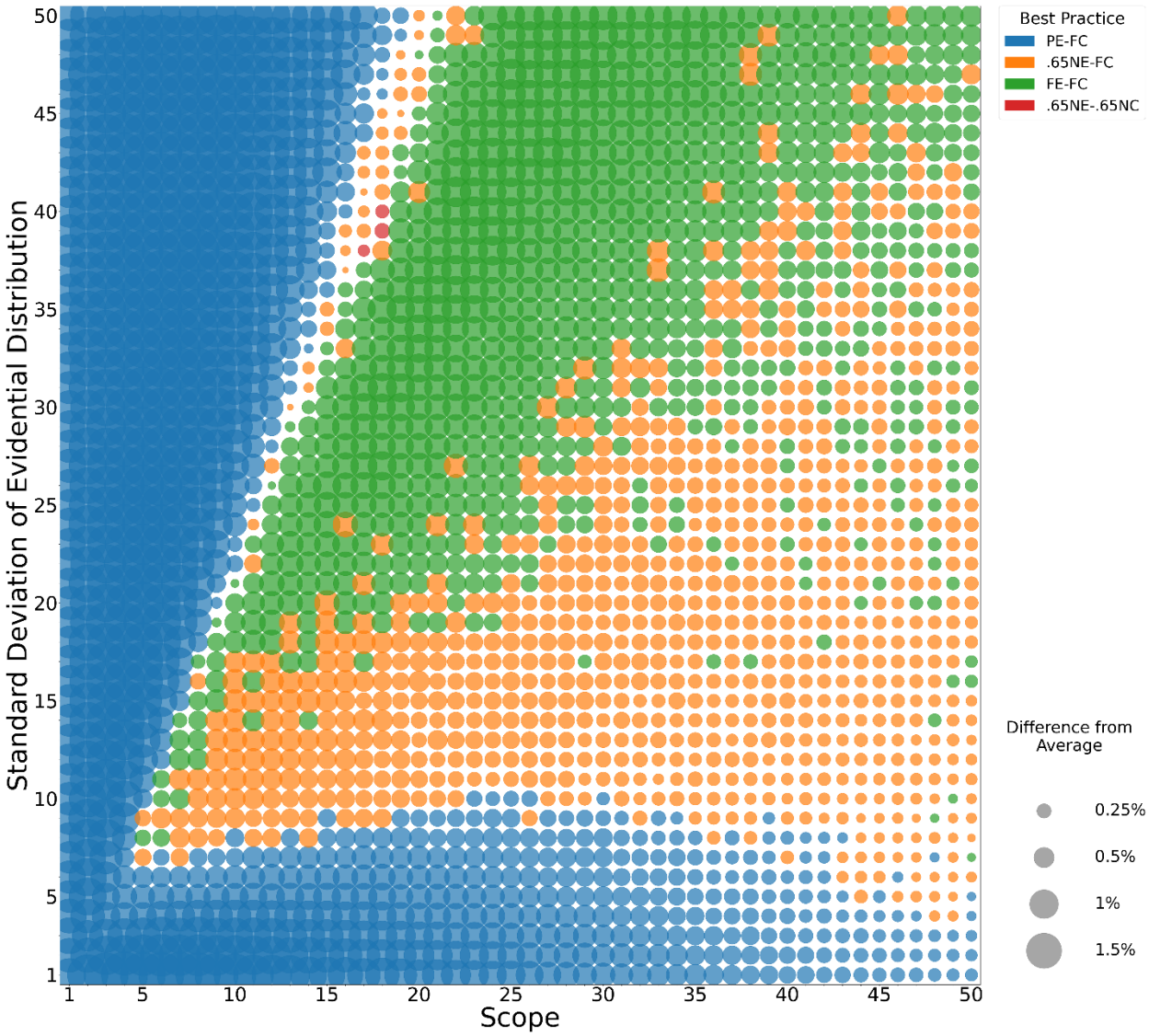

Figure C.2 The best performance practice across the parameter space at 3,000 steps. Abbreviations: FC, full communication of conclusions; FE, full communication of evidence; NC, neighborhood communication of conclusions; NE, neighborhood communication of evidence; PE, private evidence. The size of each dot indicates how much the best-performing practice exceeds the average of the four practices, measured by the percentage (out of 1,200 simulations) of groups in which the majority finds the truth.

D. *Evaluators Other than Majority Correctness*

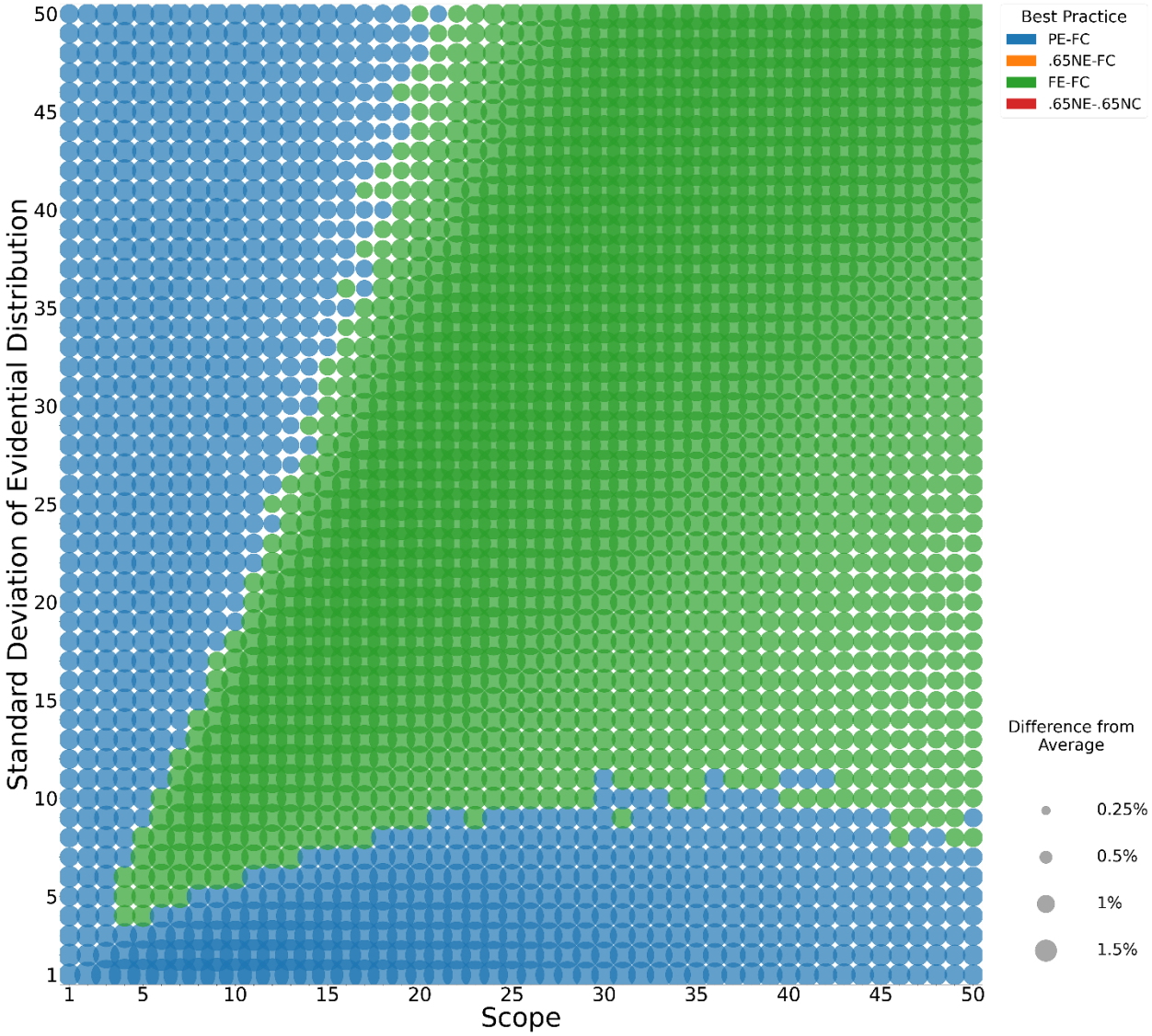

Figure D.1 The best-performing practice, evaluated by whether the entire group converges on the truth, across the parameter space at 1,500 steps. Abbreviations: FC, full communication of conclusions; FE, full communication of evidence; NC, neighborhood communication of conclusions; NE, neighborhood communication of evidence; PE, private evidence. The size of each dot indicates how much the best-performing practice exceeds the average of the four practices, measured by the percentage (out of 1,200 simulations) of groups in which the majority finds the truth.

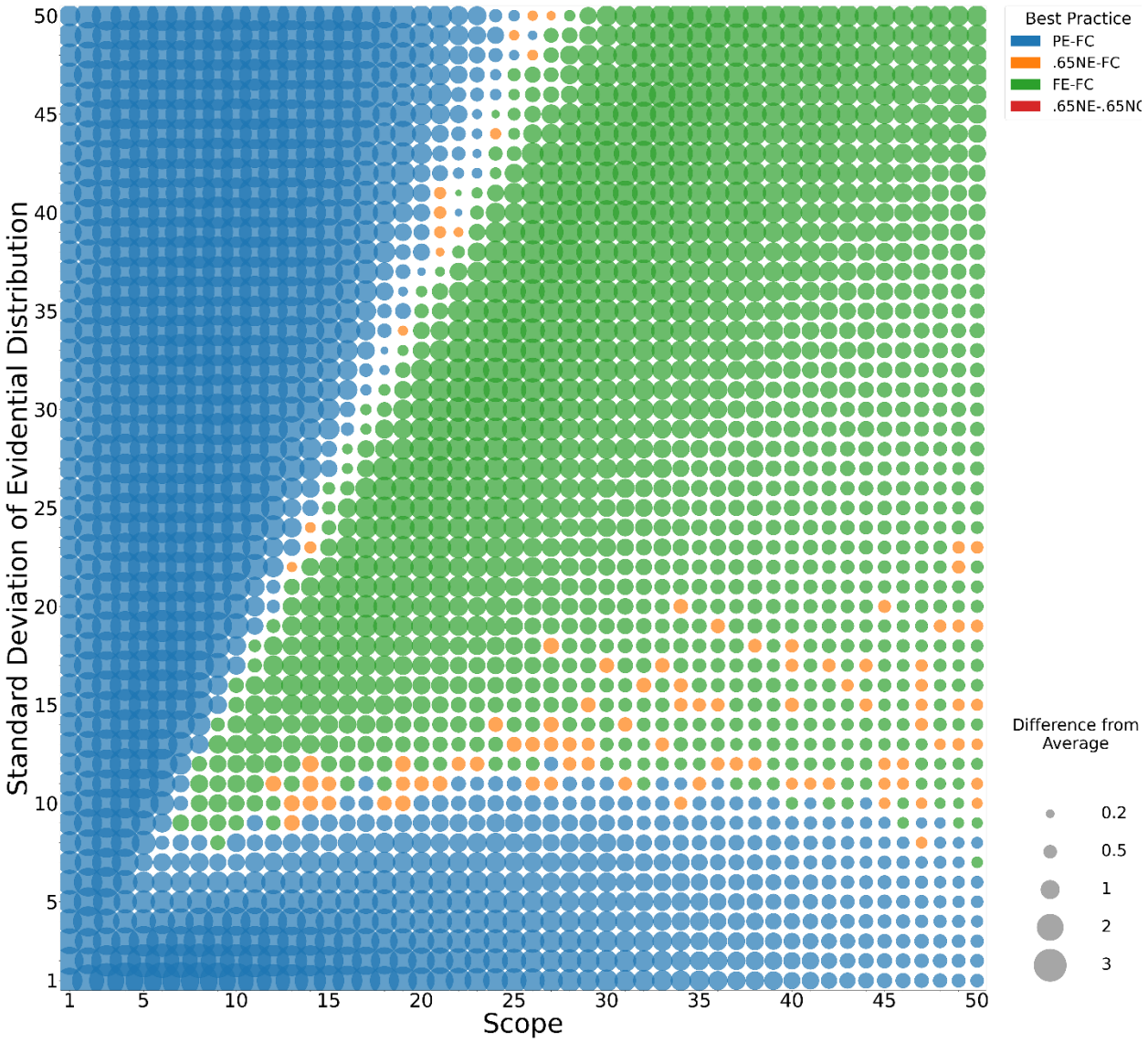

Figure D.2 The best-performing practice, evaluated by the average number of agents who converge on the truth, across the parameter space at 1,500 steps. Group size is an odd number randomly drawn from 7 to 21. Abbreviations: FC, full communication of conclusions; FE, full communication of evidence; NC, neighborhood communication of conclusions; NE, neighborhood communication of evidence; PE, private evidence. The size of each dot indicates how much the best-performing practice exceeds the average of the four practices, measured by the number of agents.

E. A Proof for Section 4.5 (Inequality 4.6 and the Algorithm)

By Equation (4.4), the balance condition $\mathrm{REU}(w + \mathrm{GI}) = \mathrm{REU}(w)$ implies the utility tradeoff ratio in Gamble I must be equal to the corresponding risk-weight ratio, such that

$$c_\rho(0.01, d, w) = \frac{u_\rho\big(w+h_\rho(0.01,d,w)\big)-u_\rho(w)}{u_\rho(w)-u_\rho(w-d)} = \frac{1-r(0.99)}{r(0.01)}.$$

By the convexity of r on the relevant intervals and Inequality (5.1), there is a lower bound:

$$c_\rho(0.01, d, w) = \frac{1-r(0.99)}{r(0.01)} > 10 \cdot \frac{1-r(0.99)}{r(0.1)} > 10 \cdot (t_w(\rho) - 1). \quad \text{(E.1)}$$

Therefore, for any CRRA utility function compatible with Allais preferences, Inequality (E.1) yields a lower bound of the utility tradeoff ratio in Gamble I such that

$$\underline{c}_\rho(0.01, d, w) = 10 \cdot (t_w(\rho) - 1).$$

Accordingly, a lower bound of the required gain, $\underline{h}_\rho(0.01, d, w)$, is obtained by setting Inequality (E.1) at equality (equivalently, by considering the boundary case where $\mathrm{REU(L4)} > \mathrm{REU(L3)}$ holds at the margin) and solving for $\underline{h}_\rho(0.01, d, w)$ in the following equation:[86]

$$\frac{u_\rho\big(w+\underline{h}_\rho(0.01,d,w)\big)-u_\rho(w)}{u_\rho(w)-u_\rho(w-d)} = 10 \cdot (t_w(\rho) - 1).$$

[86] Intuitively, under a specific utility function, the preference $\mathrm{REU(L4)} > \mathrm{REU(L3)}$ holding at the margin corresponds to a modest risk function that contribute the least to extreme risk-averse behavior. Thus, the situation defines a lower bound of h_ρ in Gamble I.

F. *Monetary Tradeoff Ratios for Mid-range Losses*

Figure F.1. shows the monetary tradeoff ratio for mid-range losses, such as \$150,000, \$180,000, and \$210,000. In the middle of the curves, blank segments appear and indicate the absence of a finite solution for $h_\rho(0.01, d, w)$. However, losses of this magnitude are insufficient to calibrate REU, since a combination of a highly concave utility function and a slightly convex risk function does not yet force an implausibly high tradeoff ratio.

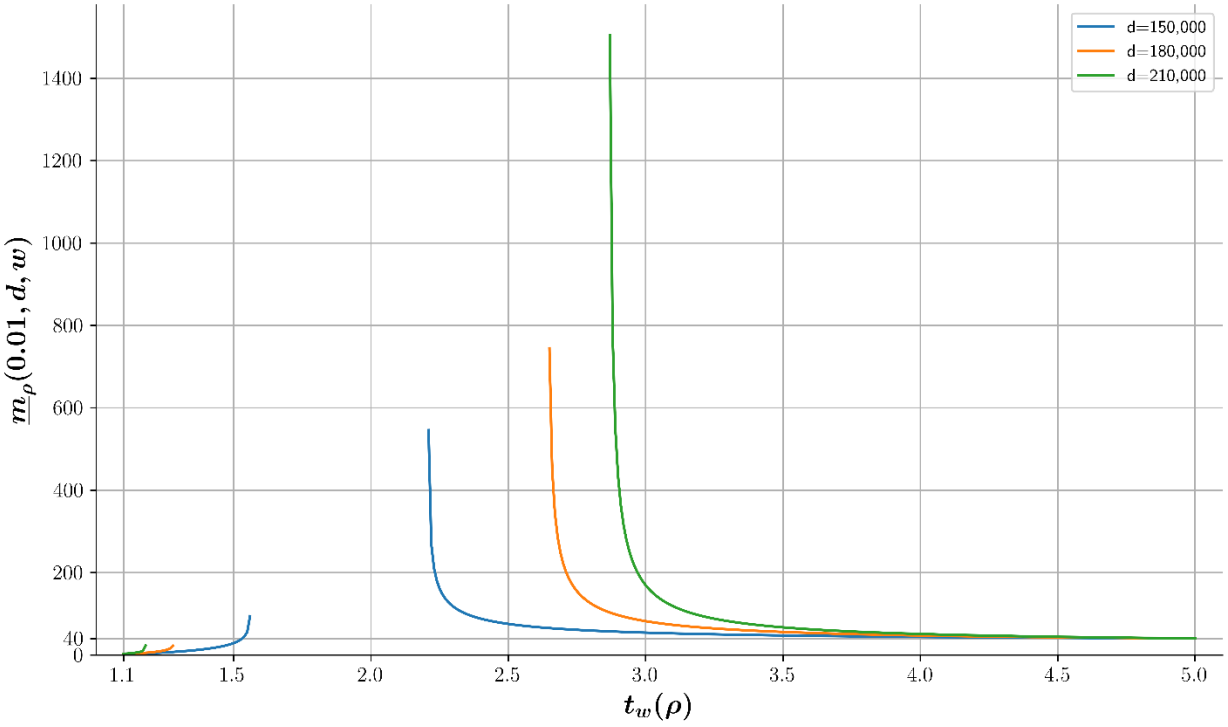

Figure F.1. The lower bound of monetary tradeoff ratios $\boldsymbol{\underline{m}_\rho(0.01, d, w) = \underline{h}_\rho(0.01, d, w)/d}$, where $\boldsymbol{w = 5 \times 10^6}$ for $\boldsymbol{t_w(\rho) \in (1.1, 5]}$ with mid-range $\boldsymbol{d}$. The lower bound of $\boldsymbol{c_\rho(0.01, d, w)}$, determined by $\boldsymbol{\mathrm{REU(L4)} > \mathrm{REU(L3)}}$ at the margin, is used to determine $\boldsymbol{\underline{h}_\rho(0.01, d, w)}$ and $\boldsymbol{\underline{m}_\rho(0.01, d, w)}$. Blank segments of curves indicate that for those $\boldsymbol{t_w}$ values, no finite $\boldsymbol{h_\rho(0.01, d, w)}$ can offset the corresponding loss $\boldsymbol{d}$.

G. *Tradeoff Ratios Under an Exponential Risk Function*

To illustrate more realistic tradeoff ratios without flattening a chain of inequalities, I adopt an exponential risk function:

$$r_e(p) = \frac{1}{\gamma-1}(\gamma^p - 1) \text{ and } \gamma > 1,$$[87]

which is globally convex. The larger the value of t, the more convex (risk-avoidant) the risk function is. A useful feature of this function is that, for small p, the utility tradeoff ratio in Gamble I closely aligns with the parameter γ:

$$c_u(p) \approx \gamma \text{ for small } p.$$

The parameter γ a flexible parameter for representing different degrees of convexity. Under a linear utility function, when $\text{REU(L4)} > \text{REU(L3)}$ holds at the margin, there is $\gamma = 49.91$. This implies that, for small probabilities, the minimal tradeoff ratio REU can sustain when facing moderately large stakes is approximately 50.

[87] Buchak adopts a power risk function $r(p) = p^a$ across her book. However, by L'Hôpital's rule, there is $\max_p c_u(p) = \lim_{p\to 0} \frac{r\prime(1-p)}{r\prime(p)} = +\infty$. Regardless of accommodating Allais preferences, a power risk function necessarily leads to extreme risk attitudes facing small probabilities. Thus, a power risk function is not a reasonable choice.

H. Thresholds for Calibration

For any CRRA utility function,

$$\lim_{x\to\infty} u_\rho(x) = \frac{1}{\rho-1}.$$

To guarantee that any combination of a utility function and a risk function demands a tradeoff ratio of at least 40, I focus on the boundary case where t_w approaches its minimum value 1.1. In this case, the lower bound of the utility tradeoff ratio approaches 1. A finite solution for $h_\rho(0.01, d, w)$ exists if and only if

$$u_\rho(w) - u_\rho(w-d) \le u_\rho(\infty) - u_\rho(w).$$

Note that

$$u_\rho(\infty) - u_\rho(w) = \frac{w^{1-\rho}}{\rho-1} \text{ and } u_\rho(w) - u_\rho(w-d) = \frac{(w-d)^{1-\rho} - w^{1-\rho}}{\rho-1}.$$

The threshold value of d is therefore determined by

$$(1 - \frac{d_*}{w})^{1-\rho_*} = 2,$$

so that $d_* = w(1 - 2^{1/(1-\rho_*)})$, where ρ_* is the solution of the parameter ρ in the equation above for the CRRA utility function satisfying

$$\frac{u_\rho(w + 5\times10^6) - u_\rho(w)}{u_\rho(w + 10^6) - u_\rho(w)} = 1.1.$$

Bibliography

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine. *Econometrica. 21 (4)*, pp. 503-546. doi:10.2307/1907921

Arrow, K. J. (1965). *Aspects of the Theory of Risk-Bearing.* Yrjö Jahnssonin Säätiö.

Baccini, E., Christoff, Z., Hartmann, S., & Verbrugge, R. (2023). The Wisdom of the Small Crowd: Myside Bias and Group Discussion. *Journal of Artificial Societies and Social Simulation, 26*(4). doi:10.18564/jasss.5184

Bala, V., & Goyal, S. (1998). Learning from Neighbours. *The Review of Economic Studies, 65*(3), 595-621. Retrieved from https://www.jstor.org/stable/2566940

Betz, G. (2013). *Debate Dynamics: How Controversy Improves Our Beliefs.* Springer Netherlands.

Blavatskyy, P., Ortmann, A., & Panchenko, V. (2022). On the Experimental Robustness of the Allais Paradox. *American Economic Journal: Microeconomics, 14 (1)*, pp. 143-163. doi:10.1257/mic.20190153

Bottomley, C., & Williamson, T. L. (2024). Rational risk-aversion: Good things come to those who weight. *Philosophy and Phenomenological Reserach, 108*(3), 697-725. doi:10.1111/phpr.13006

Braun, S. T., Rad, S. R., & Roy, O. (2024). Anchoring as a Structural Bias of Deliberation. *Erkenntnis*. doi:https://doi.org/10.1007/s10670-024-00814-7

Briggs, R. (2015). Costs of abandoning the Sure-Thing Principle. *Canadian Journal of Philosophy, 45*(5/6), 827-840. Retrieved from https://www.jstor.org/stable/26444409

Broome, J. (1991). *Weighing Goods.* Blackwell Publishers. doi:10.1002/9781119451266

Buchak, L. (2013). *Risk and Rationality.* Oxford University Press. doi:10.1093/acprof:oso/9780199672165.001.0001

Buchak, L. (2015). Revisiting Risk and Rationality: a reply to Pettigrew and Briggs. *Canadian Journal of Philosophy, 45*(5/6), 841-862. doi:10.1080/00455091.2015.1125235

Buchak, L. (2017). Replies to Commentators. *Philosophical Studies, 174*, 2397–2414. doi:10.1007/s11098-017-0907-4

Cover, T. M., & Thomas, J. A. (2006). *Elements of Information Theory* (Second ed.). Wiley. doi:10.1002/047174882X

Csiszár, I., & Körner, J. (2011). *Information Theory: Coding Theorems for Discrete Memoryless*

System (2nd ed.). Cambridge: Cambridge University Press. doi:https://doi.org/10.1017/CBO9780511921889

Deffuant, G., Neau, D., Amblard, F., & Weisbuch, G. (2000). Mixing Beliefs Among Interacting Agents. *Advances in Complex Systems, 03*(01n04), pp. 87-98.

Dias, P. M., & Shimony, A. (1981). A critique of Jaynes' maximum entropy principle. *Advances in Applied Mathematics, 2*(2), pp. 172-211. doi:10.1016/0196-8858(81)90003-8

Dietrich, F., & Spiekermann, K. (2025). Deliberation and the Wisdom of Crowds. *Economic Theory, 79*, 603-655. doi:http://dx.doi.org/10.2139/ssrn.4145408

Ding, H., & Pivato, M. (2019). Deliberation and Epistemic Democracy. *Journal of Economic Behavior and Organization, 185*, pp. 138-167.

Douven, I., & Hegselmann, R. (2021). Mis-and disinformation in a bounded confidence model. *Artificial Intelligence, 291*. Retrieved from https://doi.org/10.1016/j.artint.2020.10341

Douven, I., & Riegler, A. (2010). Extending the Hegselmann–Krause Model I. *Logic Journal of the IGPL, 18*(2), 323-335. doi:10.1093/jigpal/jzp059

Douven, I., & Romeijn, J.-W. (2011). A New Resolution of the Judy Benjamin Problem. *Mind, 120*(479), pp. 637-670. doi:https://doi.org/10.1093/mind/fzr051

Dryzek, S. J., & List, C. (2003). Social Choice Theory and Deliberative Democracy: A Reconciliation. *British Journal of Political Science, 33*(1), 1-28.

Dupuis de Tarlé, L., Michelini, M., Šešelja, D., & Straßer, C. (2025). Argumentative Agent-Based Models. *Journal of Applied Logics - IfCoLog Journal of Logics and their Applications, 12*(3), 489-547. Retrieved from https://philarchive.org/archive/DUPAAM-3

Eeckhoudt, L., Gollier, C., & Schlesinger, H. (2005). *Economic and Financial Decisions under Risk.* Princeton University Press. doi:10.2307/j.ctvcm4j15

Eva, B., Hartmann, S., & Rad, S. (2020). Learning from Conditionals. *Mind, 129*(514), pp. 461-508. doi:https://doi.org/10.1093/mind/fzz025

Fryer, R. G., Harms, P., & Jackson, M. O. (2019). Updating Beliefs when Evidence is Open to Interpretation: Implications for Bias and Polarization. *Journal of the European Economic Association, 17*(5), 1470-1501. doi:https://doi.org/10.1093/jeea/jvy025

Gneezy, U., Halevy, Y., Hall, B., Offerman, T., & van de Ven, J. (2024). *How Real is Hypothetical? A High-Stakes Test of the Allais Paradox.* Retrieved from https://EconPapers.repec.org/RePEc:tor:tecipa:tecipa-783

Grim, P., Singer, D. J., Bramson, A., Holman, B., McGeehan, S., & Berger, W. J. (2019). Diversity, Ability, and Expertise in Epistemic Communities. *Philosophy of Science,*

86(1), 98-123. doi:doi:10.1086/701070

Grove, A. J., & Halpern, J. Y. (1997). Probability Update: Conditioning vs. Cross-Entropy. *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*, (pp. 208-214). doi:https://doi.org/10.48550/arXiv.1302.1543

Grünwald, P. D., & Halpern, J. Y. (2003). Updating Probabilities. *Journal of Artificial Intelligence Research, 19*, 243-278. doi:https://doi.org/10.1613/jair.1164

Hahn, U. (2022). Collectives and Epistemic Rationality. *Topics in Cognitive Science, 14*, 602-620. doi:10.1111/tops.12610

Hahn, U., von Sydow, M., & Merdes, C. (2019). How Communication Can Make Voters Choose Less Well. *Topics in Cognitive Science, 11*, 194-206. doi:10.1111/tops.12401

Hanson, N. R. (1979). *Patterns of Discovery: An Inquiry Into the Conceptual Foundations of Science.* United Kingdom: Cambridge University Press.

Hansson, B. (1988). Risk aversion as a problem of conjoint measurement. In P. Gärdenfors, & N.-E. Sahlin, *Decision, Probability and Utility Selected Readings* (pp. 136-158). Cambridge: Cambridge University Press.

Hartmann, S., & Rad, S. R. (2018). Voting, deliberation and truth. *Synthese, 195(3)*, pp. 1273-1293. Retrieved from https://doi.org/10.1007/s11229-016-1268-9

Hartmann, S., & Rad, S. R. (2019, Feb.). Anchoring in Deliberations. *Erkenntnis, 85*(5), 1041-1069. doi:https://doi.org/10.1007/s10670-018-0064-y

Hegselmann, R., & Krause, U. (2002). Opinion dynamics and bounded confidence: models, analysis and simulation. *Journal of Artificial Societies and Social Simulation, 5(3)*.

Hegselmann, R., & Krause, U. (2006). Truth and Cognitive Division of Labour: First Steps towards a Computer Aided Social Epistemology. *Journal of Artificial Societies and Social Simulation, 9*(3). Retrieved from https://www.jasss.org/9/3/10.html

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Economic Sciences, 101*(46), 16385-16389. doi:https://doi.org/10.1073/pnas.0403723101

Joyce, J. M. (2017). Commentary on Lara Buchak's risk and rationality. *Philosophical Studies, 174*, 2385–2396. doi:10.1007/s11098-017-0905-6

Kummerfeld, E., & Zollman, K. J. (2016, Dec.). Conservatism and the Scientific State of Nature. *The British Journal for the Philosophy of Science, 67(4)*, pp. 1057-1076. doi:https://doi.org/10.1093/bjps/axv013

Mankiw, N. (2016). *Principles of Microeconomics* (8th ed.). CENGAGE Learning Custom

Publishing.

Mäs, M., & Flache, A. (2013). Differentiation without Distancing. Explaining Bi-Polarization of Opinions without Negative Influence. *PLoS ONE, 8*(11). Retrieved from https://doi.org/10.1371/journal.pone.0074516

Neilson, W. (2001). Calibration results for rank−dependent expected utility. *Economics Bulletin, 4*(10), 1-5. Retrieved from https://ideas.repec.org/a/ebl/ecbull/eb-01d80002.html

O'Connor, C., & Weatherall, J. O. (2018). Scientific polarization. *European Journal for Philosophy of Science, 8*, 855-875. doi:https://doi.org/10.1007/s13194-018-0213-9

Pettigrew, R. (2015). Risk, rationality and expected utility theory. *Canadian Journal of Philosophy, 45*(5/6), 798-826. doi:10.1080/00455091.2015.1119610

Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica, 32*(1-2), 122-136. doi:10.2307/1913738

Putnam, H. (1981). *Reason, Truth and History.* Cambridge: Cambridge University Press.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization, 3*(4), 323-343. Retrieved from 10.1016/0167-2681(82)90008-7

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica 68 (5)*, 1281-1292. Retrieved from http://www.jstor.org/stable/2999450

Reiss, J. (2013). *Philosophy of Economics: A Contemporary Introduction* (1st ed.). Routledge. doi:https://doi.org/10.4324/9780203559062

Rosenstock, S., Bruner, J., & O'Connor, C. (2017). In Epistemic Networks, Is Less Really More? *Philosophy of Science, 84*(2), 234-252. doi:https://doi.org/10.1086/690717

Safra, Z., & Segal, U. (2008). Calibration results for non-expected utility theories. *Econometrica, 76*(5), 1143-1166. Retrieved from https://www.jstor.org/stable/40056496

Savage, L. J. (1972). *The foundations of statistics.* New York: Dover Publications.

Seidenfeld, T. (1986). Entropy and Uncertainty. *Philosophy of Science, 53*(4), pp. 467-491. doi:https://doi.org/10.1086/289336

Skyrms, B. (1985). Maximum Entropy Inference as a Special Case of Conditionalization. *Synthese, 63*(1), pp. 55-74. doi:https://doi.org/10.1007/BF00485955

Stefánsson, H., & Bradley, R. (2019). What Is Risk Aversion? *The British Journal for the Philosophy of Science, 70*(1), 77-102. doi:10.1093/bjps/axx035

Thoma, J. (2019). Risk Aversion and the Long Run. *Ethics, 129*(2), 230-253. doi:10.1086/699256

Thoma, J., & Weisberg, J. (2017). Risk writ large. *Philosophical Studies, 174*, 2369–2384. doi:10.1007/s11098-017-0916-3

Thoma, J., & Weisberg, J. (2020). No escape from Allais: reply to Buchak. *Philosophical Studies, 177*, 2493–2500. doi:10.1007/s11098-019-01322-z

Titelbaum, M. (2022). *Fundamentals of Bayesian Epistemology 1: Introducing Credences.* Oxford University Press. doi:https://doi.org/10.1093/oso/9780198707608.001.0001

van Fraassen, B. (1981). A Problem for Relative Information Minimizers in Probability Kinematics. *The British Journal for the Philosophy of Science, 32*(4), 375-379. doi:https://doi.org/10.1093/bjps/32.4.375

van Fraassen, B. (1986). A Problem for Relative Information Minimizers, Continued. *British Journal for the Philosophy of Science, 37(4)*, pp. 453-463. Retrieved from http://www.jstor.org/stable/687370

Vasudevan, A. (2020). Entropy and Insufficient Reason: A Note on the Judy Benjamin Problem. *The British Journal for the Philosophy of Science, 71*(3), 1113-1141. doi:https://doi.org/10.1093/bjps/axy013

Zellner, A. (1988). Optimal Information Processing and Bayes's Theorem. *The American Statistician, 42*(4), pp. 278-280. doi:https://doi.org/10.2307/2685143

Zollman, K. J. (2007). The Communication Structure of Epistemic Communities. *Philosophy of Science, 74(5)*, pp. 574-587. doi:https://doi.org/10.1086/525605

Zollman, K. J. (2010). The Epistemic Benefit of Transient Diversity. *Erkenntnis, 72*, 17-35. doi:https://doi.org/10.1007/s10670-009-9194-6