Informativeness Orders over Ambiguous Experiments*

Zichang Wang†
Xiamen University

First version: January 21, 2022
This version: August, 2024

Abstract

We generalize Blackwell's informativeness order to ambiguous experiments. A decision maker might view a statistical experiment as ambiguous if she faces uncertainty about the data generating process for its signals. Formally, an ambiguous experiment is modeled as a mapping from an auxiliary state space to the set of unambiguous experiments. Each auxiliary state corresponds to a possible data generating process. We show that one ambiguous experiment is preferred to another by every decision maker in every decision problem if and only if they are related by a condition called *prior-by-prior dominance*, which states that for any first-order belief the decision maker entertains on the auxiliary state space, the expected experiment resulting from this belief for the first experiment is Blackwell more informative than that of the second. This equivalence is robust for any class of monotone ambiguity preferences that nests expected utility. We obtain another informativeness order when we restrict attention to decision makers who apply the maxmin criterion to evaluate ambiguous experiments and connect this informativeness order to comparisons of sets of Blackwell experiments.

Keywords: Statistical experiments, ambiguity, informativeness order
JEL classification: D81, D83

*This paper is adapted from the second chapter of my dissertation at Duke University. I am deeply indebted to my advisor, Todd Sarver, for his continued guidance and detailed comments throughout the development of this paper. I thank the editor, Faruk Gul, and two anonymous referees for comments and suggestions that lead to a significant improvement of the paper. I also thank Arjada Bardhi, Yonggyun Kim, Xiaosheng Mu, and Philipp Sadowski for many helpful discussions. All errors are my own.

†Xiamen University, Wang Yanan Institute for Studies in Economics (WISE) and the School of Economics, 422 Siming South Road, Xiamen, Fujian, 361005, China. Email: zichang.wang@xmu.edu.cn

1 Introduction

Blackwell (1951, 1953) models information as statistical experiments and establishes an elegant order for their informativeness. This order allows comparisons of information independent of decision problems or preference parameters. However, information is often ambiguous, since the exact data generating process for signals is rarely known. For example, a medical test for a certain disease can be thought of as a statistical experiment, and such tests have the possibility of both false positives and false negatives. Seldom do patients know the exact probabilities of false positives and false negatives, and they may therefore view the distribution of test results (and the test itself) as ambiguous. In this sense, Blackwell's formulation of information as unambiguous experiments is idealized.

One can make intuitive comparisons of ambiguous information. For example, patients may always prefer a medical test that has a smaller *maximum* probability of false positives or false negatives. Such intuitive comparisons cannot be formalized in Blackwell's framework due to the lack of a structure for informational ambiguity.

In this paper, we propose a model for informational ambiguity and generalize Blackwell's theorem to comparisons of ambiguous experiments. More specifically, we characterize a generalization of Blackwell's *garbling* condition that is equivalent to every decision maker preferring one ambiguous experiment to another in every decision problem. This generalization establishes an informativeness order over ambiguous experiments.

Let Ω be a finite set of payoff-relevant states. A Blackwell experiment is a mapping $p : \Omega \to \Delta(S)$ where S is a finite set of signals, and $\Delta(S)$ is the set of probability distributions over S. A Blackwell experiment p is unambiguous since the probability that signal s is generated condition on state ω is unambiguously specified for all s and ω.

We introduce an auxiliary state space Θ to capture informational ambiguity. Specifically, we model an ambiguous experiment as a mapping $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$. That is, each ambiguous experiment is identified as an indexed set of Blackwell experiments. To ease notation, we write $\mathbf{p}_\theta$ to denote the Blackwell experiment associated with state θ. That is, given ambiguous experiment $\mathbf{p}$, $\mathbf{p}_\theta$ is a mapping from Ω to $\Delta(S)$ for each θ.

Our interpretation for the auxiliary state space Θ is similar to the interpretation for the "model space" from the literature on model uncertainty (e.g., Hansen and Sargent 2001, 2008 and Marinacci 2015). A decision maker (henceforth DM) understands that signals are generated according to some data generating process (i.e., some Blackwell experiment), which is determined by some "*generative mechanism representing some natural* *or social phenomenon*" (Marinacci, 2015). DMs may not know this mechanism or have enough information to form a belief about possible mechanisms. As a result, informational ambiguity arises. Each state θ represents a mechanism and each $\mathbf{p}_\theta$ is the data generating process (Blackwell experiment) when the mechanism is θ. Even though a DM fully understands the data generating process corresponding to each generative mechanism θ, she faces ambiguity on Θ. Lack of understanding of the distribution of θ translates to lack of understanding of the probabilistic content of the experiment, making this a convenient (and in some sense canonical) way to model informational ambiguity.

Ambiguous experiments are compared according to their ex-ante value in a decision problem, which is obtained by optimizing over all action plans (which action to take after each signal). We assume that DMs adopt a two-stage criterion when evaluating action plans for an ambiguous experiment.[1] In our model, action plans are first evaluated with respect to each possible data generating process $\mathbf{p}_\theta$ through expected utility, and then these evaluations are combined together through some aggregator over the auxiliary state space Θ. These aggregators capture the DM's attitude toward informational ambiguity. Importantly, conditional on each auxiliary state θ, the DM treats the uncertainty described by $\mathbf{p}_\theta$ as objective/physical and behaves as if she is an expected utility maximizer. We believe this is a natural first assumption to make within a two-stage evaluation framework. Interested readers are referred to Marinacci (2015) for a thorough discussion on this specific type of two-stage decision criterion.

The space Θ is "auxiliary" since it does not affect the DM's ex-post payoff. When testing for a certain disease, the patient's ex-post payoff depends on the patient's true status, but not the precision of the test. The auxiliary state space influences the DM's decision and ex-ante payoff only through influencing the information content of the experiment.

Since Θ is payoff irrelevant, Blackwell dominance over $\Omega \times \Theta$ is unnecessarily strong. Specifically, we obtain informativeness orders strictly weaker than Blackwell's order. The difference is, instead of interpreting each ambiguous experiment as a single Blackwell experiment on $\Omega \times \Theta$ and compare them accordingly, we view each ambiguous experiment as a set of Blackwell experiments indexed by the elements of Θ and compare these indexed sets to assess informativeness.

Two nice features stem from using the structure of an auxiliary state space to model

[1]This criterion is similar to a common two-stage decision criterion in the model uncertainty literature. By this criterion, uncertainty over payoff-relevant states is addressed in the first stage and model uncertainty is addressed in the second stage.

informational ambiguity. First, it can accommodate very general preferences toward informational ambiguity. Second, it allows explicit correlation between ambiguous experiments. These features facilitate subtler comparisons of ambiguous experiments than simply requiring Blackwell domination over $\Omega \times \Theta$.

An alternative approach to model informational ambiguity is to consider unindexed sets of Blackwell experiments (evaluated by the maxmin criterion). We analyze this approach in Section 4 and show that comparisons of unindexed sets of Blackwell experiments evaluated by the maxmin criterion can be interpreted as comparisons of ambiguous experiments evaluated by the maxmin criterion.

1.1 Preview of Main Results

Let $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ be an ambiguous experiment. For a probability measure μ over the auxiliary state space Θ, the *expected experiment* $\mathbf{p}_\mu : \Omega \to \Delta(S)$ is defined by taking the expectation of $\mathbf{p}$ with respect to μ, that is, $\mathbf{p}_\mu := \int_\Theta \mathbf{p}_\theta \, d\mu(\theta)$. Expected experiments are unambiguous by construction. They play an important role in our results because of the two-stage decision criterion we have assumed.

We say that $\mathbf{p}$ *prior-by-prior* dominates $\mathbf{q}$ if the expected experiment $\mathbf{p}_\mu$ is Blackwell more informative than $\mathbf{q}_\mu$ for every probability measure μ over Θ. Our first result (Theorem 1 in Section 3) is that $\mathbf{p}$ is preferred to $\mathbf{q}$ (i.e., guarantees a higher ex-ante value) for every DM in every decision problem *if and only if* $\mathbf{p}$ prior-by-prior dominates $\mathbf{q}$. Prior-by-prior dominance is a direct generalization of Blackwell's informativeness notion: When Θ is a singleton, ambiguous experiments reduce to Blackwell experiments and prior-by-prior dominance reduces to being Blackwell more informative.

Our second result (Theorem 2 in Section 4) is another informativeness order over ambiguous experiments. This order is obtained by restricting attention to maxmin DMs (DMs who use Wald's maxmin criterion to evaluate ambiguous experiments). We say an ambiguous experiment $\mathbf{p}$ is *Wald more informative* than $\mathbf{q}$ if for any probability measure μ over Θ, there exists some probability measure ν over Θ such that $\mathbf{p}_\mu$ is Blackwell more informative than $\mathbf{q}_\nu$. We show that $\mathbf{p}$ is Wald more informative than $\mathbf{q}$ if and only if every maxmin DM prefers $\mathbf{p}$ to $\mathbf{q}$ in every decision problem. Since $\mathbf{p}_\mu$ and $\mathbf{q}_\nu$ can result from different μ and ν, being Wald more informative is a weaker condition than prior-by-prior dominance. That is, if $\mathbf{p}$ prior-by-prior dominates $\mathbf{q}$, then $\mathbf{p}$ is Wald more informative than $\mathbf{q}$, but the converse is not true.

There is a natural connection between ambiguous experiments and unindexed sets of Blackwell experiments when we focus on maxmin DMs. If we let $P := \{\mathbf{p}_\theta \mid \theta \in \Theta\}$ and $Q := \{\mathbf{q}_\theta \mid \theta \in \Theta\}$, then $\mathbf{p}$ is Wald more informative than $\mathbf{q}$ if and only if "for any $p \in \text{conv}(P)$, there exists $q \in \text{conv}(Q)$ such that p is more informative than q". Due to this connection, we say that an unindexed set of experiments P_1 is Wald more informative than P_2 if they can be related by the previous condition in quotation. We show (in Theorem 5) that P_1 is Wald more informative than P_2 if and only if P_1 is preferred to P_2 in every decision problem by every DM who uses the maxmin criterion.

1.2 Related Literature

We contribute to the literature on the Blackwell-style comparison of information under ambiguity. Previous studies mainly focus on comparisons of unambiguous experiments under ambiguous priors. The main finding is that the Blackwell order is still valid. That is, $\mathbf{q}$ is a garbling of $\mathbf{p}$ (viewing $\Omega \times \Theta$ as the payoff-relevant state space) if and only if $\mathbf{p}$ guarantees a higher ex-ante payoff than $\mathbf{q}$ in every decision problem (where the DM's ex-post payoff depends on $\Omega \times \Theta$) for every DM (including the ones whose preferences involve an ambiguous prior over $\Omega \times \Theta$). Çelen (2012) shows that this is true for DMs with multiple-prior preferences. Li and Zhou (2016) show that this is true for DMs with uncertainty-averse preferences (as in Cerreia-Vioglio et al. 2011). In contrast to these papers, we study comparisons of ambiguous experiments. We use a product state space structure where Θ captures informational ambiguity and is not payoff relevant. Our informativeness orders are weaker than Blackwell's order over $\Omega \times \Theta$.

The study of ambiguous information has gained increased attention in both the applied literature (e.g., Epstein and Schneider 2007, 2008, Bose and Renou 2014, Beauchêne et al. 2019, and Chen 2022) and the experimental literature (e.g., Liang 2021, Kellner et al. 2022, Epstein and Halevy 2023, Kops and Pasichnichenko 2023, and Shishkin and Ortoleva 2023). Our contribution to this rapidly growing literature is a theoretical one. We provide criteria that can robustly compare information sources that are potentially ambiguous. In particular, we impose little restriction on preferences or decision problems, in contrast to the aforementioned theoretical studies on ambiguous information.

The closest work to ours is Gensbittel et al. (2015, henceforth GRT). GRT model ambiguous information as a collection of *joint* distributions over payoff-relevant states and signals. One set $\mathcal{P} \subset \Delta(\Omega \times S)$ is said to be c-more informative than another $\mathcal{Q}$ if for any $\tilde{p} \in \mathcal{P}$, there exists $\tilde{q} \in \mathcal{Q}$ that is less informative. They show that $\mathcal{P}$ is preferred to $\mathcal{Q}$ by every DM using the maxmin criterion if and only if $\mathcal{P}$ is c-more informative than $\mathcal{Q}$. This order bears resemblance to the order induced by our Wald-more-informative condition. Despite the superficial similarity, the two orders are over different domains and not directly comparable. The key difference is that in GRT, the DM's prior over Ω is part of the information instead of a preference parameter. Due to this modeling approach, GRT's criterion does not always fully distinguish the informativeness of a set of experiments from the prior ambiguity over Ω.[2] In contrast, our order is prior-free, allowing us to isolate the informativeness of a set of experiments.

The rest of the paper is organized as follows. We present our model in Section 2. General comparisons of ambiguous experiments are studied in Section 3. Comparisons of ambiguous experiments by maxmin DMs are studied in Section 4. Section 5 concludes.

2 The Model

Let Ω be a finite set of payoff-relevant states. Let Θ be a set of auxiliary states. We call Θ the auxiliary state space and use it to capture informational ambiguity. For ease of exposition, we focus on the case where Θ is a finite set in the main text. Our results remain valid for any nonempty Θ. Results for general Θ are stated in Appendix A.

Definition 1. An **ambiguous experiment** is a mapping $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ where S is a finite set of signals.[3]

A Blackwell experiment corresponds to the special case where $\Theta = \{\theta\}$, so there is no uncertainty about the data generating process for $\mathbf{p}$. In this case, we suppress the singleton auxiliary state space. That is, a **Blackwell experiment** is a mapping $p : \Omega \to \Delta(S)$.

Fix an ambiguous experiment $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$, we write $\mathbf{p}_\theta$ to denote the Blackwell experiment associated with θ. That is, $\mathbf{p}_\theta$ is a mapping from Ω to $\Delta(S)$ for every $\theta \in \Theta$. We can treat $\mathbf{p}$ as a mapping from the auxiliary state space Θ to the set of Blackwell experiments. Each θ corresponds to a possible data generating process behind $\mathbf{p}$.

[2]For example, we can fix a Blackwell experiment $p : \Omega \to \Delta(S)$ and pair it with two (compact and convex) sets of priors, $\Pi_1 \subsetneq \Pi_2$. Then the set of joint distributions induced from p and Π_1 is strictly "c-more informative" than that induced from p and Π_2. However, that the DM obtains a lower maxmin payoff under Π_2 is not because of receiving less information, but because of larger prior ambiguity.

[3]For any finite set X, let $\Delta(X)$ denote the set of all probability measures over X.

A DM utilizes an ambiguous experiment $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ to decide on her action plans. Formally, a decision problem is a finite set of actions A.[4] An **action plan** is a mapping $\sigma : S \to \Delta(A)$. We write $\sigma(a \mid s)$ to denote the DM's (mixed) strategy after observing signal s. We use A_S to denote the collection of all action plans once the set of actions A and the set of signals S are fixed, that is, $A_S := \{\sigma \mid \sigma : S \to \Delta(A)\}$.

A DM is described by a triplet of preference parameters (u, π, V) where $u : \Omega \times A \to \mathbb{R}$ is her state-dependent utility function, π is her prior belief over Ω, and $V : \mathbb{R}^\Theta \to \mathbb{R}$ is an aggregator for vectors over Θ. The roles of u and π are similar to that in the standard expected utility framework. The aggregator V captures the DM's ambiguity attitude.

Definition 2. We say $V : \mathbb{R}^\Theta \to \mathbb{R}$ is a **monotone aggregator** if V is continuous and $V(f) \geq V(g)$ whenever two functions $f, g : \Theta \to \mathbb{R}$ satisfy $f(\theta) \geq g(\theta)$ for all $\theta \in \Theta$.

All ambiguity preferences studied in the literature can be identified with some monotone aggregator (e.g., the multiple-prior preference from Gilboa and Schmeidler 1989, and the smooth preference from Klibanoff et al. 2005). Since we only require V to be monotone, flexible ambiguity attitudes (ambiguity averse, ambiguity loving, mixed attitude) are also allowed. Let $\mathcal{V}_{Mono}$ denote the class of monotone aggregators, that is,

$$\mathcal{V}_{Mono} := \{V : \mathbb{R}^\Theta \to \mathbb{R} \mid V \text{ is monotone}\}. \tag{1}$$

$\mathcal{V}_{Mono}$ will be the largest class of aggregators we study. An important class of monotone aggregators for our subsequent analysis is the expected utility aggregators.

An **expected utility aggregator** is denoted by V_μ, where $\mu \in \Delta(\Theta)$ represents the DM's subjective belief over Θ. These aggregators correspond to DMs who are probabilistically sophisticated over Θ. For a fixed μ, V_μ is defined by $V_\mu(f) := \int_\Theta f(\theta)\, d\mu(\theta)$.

Let $\mathcal{V}_{EU}$ denote the class of expected utility aggregators, that is,

$$\mathcal{V}_{EU} := \{V_\mu \mid \mu \in \Delta(\Theta)\}. \tag{2}$$

Suppose a DM with preference parameters (u, π, V) faces a decision problem A and an ambiguous experiment $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$. The timing of events is as follows: First, the

[4]For completeness, the reader can also think about a consequence function $c : \Omega \times A \to X$ as part of a decision problem, where X is a set of possible outcomes and $c(\omega, a)$ is the outcome of action a when the payoff-relevant state is ω. For ease of exposition, we do not explicitly model the outcome space and the consequence function. They will be absorbed by the DM's state-dependent utility function.

DM commits to an action plan $\sigma \in A_S$.[5] Second, an auxiliary state θ is drawn from Θ but *not* observed by the DM and signals will be generated according to the Blackwell experiment $\mathbf{p}_\theta$. Third, a payoff relevant state ω is drawn from Ω according to π and a signal s is drawn from S according to the distribution $\mathbf{p}_\theta(\cdot \mid \omega)$. Last, the DM observes the signal realization s and acts according to her action plan $\sigma(\cdot \mid s)$.

Definition 3. For a DM with preference parameters (u, π, V) facing a decision problem A, the **ex-ante value** of an ambiguous experiment $\mathbf{p}$ is given by

$$W(\mathbf{p}) := \max_{\sigma \in A_S} V\left(\left\{U\big(\sigma, \mathbf{p}_\theta\big)\right\}_{\theta \in \Theta}\right) \tag{3}$$

where $U\big(\sigma, \mathbf{p}_\theta\big)$ is the **expected utility conditional on state** θ when the action plan is σ. Formally, for any $\theta \in \Theta$,

$$U\big(\sigma, \mathbf{p}_\theta\big) := \sum_{s \in S} \sum_{\omega \in \Omega} \sum_{a \in A} \pi(\omega) \mathbf{p}_\theta(s \mid \omega) \sigma(a \mid s) u(\omega, a). \tag{4}$$

As we have mentioned in Section 1, action plans are first evaluated through standard expected utility condition on each possible data generating process $\mathbf{p}_\theta$. These conditional expected utilities are then aggregated through V into the ex-ante payoff.

3 Comparing Ambiguous Experiments

An important notion for comparisons of ambiguous experiments is expected experiment.

Definition 4. Let $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ be an ambiguous experiment and μ be a probability measure over Θ. The **expected experiment** with respect to μ, denoted by $\mathbf{p}_\mu$, is a Blackwell experiment $\mathbf{p}_\mu : \Omega \to \Delta(S)$ defined by $\mathbf{p}_\mu := \sum_{\theta \in \Theta} \mu(\theta) \mathbf{p}_\theta$. Explicitly,

$$\mathbf{p}_\mu(s \mid \omega) := \sum_{\theta \in \Theta} \mathbf{p}_\theta(s \mid \omega) \mu(\theta), \ \forall (s, \omega) \in S \times \Omega. \tag{5}$$

For ease of exposition, we define the composition of two stochastic operators. Suppose X, Y, Z are finite sets, and $\alpha : X \to \Delta(Y)$ and $\beta : Y \to \Delta(Z)$ are two stochastic operators. Their **composition** $\beta \circ \alpha : X \to \Delta(Z)$ is defined by

$$(\beta \circ \alpha)(z \mid x) = \sum_{y \in Y} \alpha(y \mid x) \beta(z \mid y), \ \forall (x, z) \in X \times Z.$$

[5] As is well known, dynamic inconsistency may arise with belief updating when ambiguity is present. See Example 2 of Asano and Kojima (2019) for a more detailed treatment. Alternatively, we can assume that the DM is dynamically consistent.

That is, the composition of β and α gives a probability of z given x when the stochastic operator α is applied followed by β.

Let $p : \Omega \to \Delta(S)$ and $q : \Omega \to \Delta(T)$ be two Blackwell experiments. We say q is a **garbling** of p if there exists some $\gamma : S \to \Delta(T)$ such that $q = \gamma \circ p$.

Intuitively, q is a garbling of p if one can replicate q (in terms of the conditional probability of t given ω) by adding "noise" (applying a stochastic operator) to p. Our use of stochastic operators and their composition to describe garblings is based on de Oliveira (2018), which provides a simple and elegant proof for Blackwell's theorem.

We say p is **(Blackwell) more informative** than q if q is a garbling of p.

Definition 5. Let $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ and $\mathbf{q} : \Omega \times \Theta \to \Delta(T)$ be two ambiguous experiments. We say **p prior-by-prior dominates q** if $\mathbf{q}_\mu$ is a garbling of $\mathbf{p}_\mu$ for all $\mu \in \Delta(\Theta)$, and write $\mathbf{p} \trianglerighteq_{PBP} \mathbf{q}$.

In other words, $\mathbf{p} \trianglerighteq_{PBP} \mathbf{q}$ if there exists a family of garblings $\{\gamma_\mu\}$ such that

$$\mathbf{q}_\mu = \gamma_\mu \circ \mathbf{p}_\mu, \ \forall \mu \in \Delta(\Theta). \tag{6}$$

The "prior-by-prior" quantifier in the name of the condition does not refer to anything related to preference parameters. Each μ is just a probability measure over Θ and is not meant to be interpreted as anything more. Specifically, it is not a preference parameter.

Theorem 1. *Let $\boldsymbol{p} : \Omega \times \Theta \to \Delta(S)$ and $\boldsymbol{q} : \Omega \times \Theta \to \Delta(T)$ be two ambiguous experiments, and let $\mathcal{V}$ be a class of aggregators such that $\mathcal{V}_{EU} \subset \mathcal{V} \subset \mathcal{V}_{Mono}$.*

The following conditions are equivalent:

1. ***p** prior-by-prior dominates **q**.*

2. *For any A, u, π, and any action plan $\eta \in A_T$, there exists $\sigma \in A_S$ such that*

$$U\big(\sigma, \boldsymbol{p}_\theta\big) \geq U\big(\eta, \boldsymbol{q}_\theta\big), \ \forall \theta \in \Theta.$$

3. ***p** is preferred to **q** in every decision problem by every decision maker whose ambiguity preference can be represented by some $V \in \mathcal{V}$. That is, **p** has weakly higher ex-ante value than **q** for every A, u, π, and every $V \in \mathcal{V}$.*

Theorem 1 is a direct generalization of Blackwell's theorem. When Θ is a singleton, **p** and **q** reduce to Blackwell experiments, and the prior-by-prior dominance condition reduces to the garbling condition. The proof of Theorem 1 is omitted since it is a corollary of a more general theorem (Theorem 3), formally stated in Appendix A and proved in Appendix B, that covers the case in which the auxiliary state space Θ can be any non-empty set.

Condition 1 is a statistical condition that is independent of any decision problem or preference parameters. Condition 3 is an economical condition about the instrumental value of ambiguous experiments. Establishing the equivalence of conditions 1 and 3 helps us to separate comparisons of information structures from specific preference parameters or decision problems, just like Blackwell's theorem.

The necessity of prior-by-prior dominance is intuitive: Conditional on each belief over Θ, the DM is an expected utility maximizer. That is, the DM faces no informational ambiguity when she can form a belief over possible data generating processes (this includes the case where she is certain about exactly one data generating process). For such DMs, comparison of ambiguous experiments reduces to comparison of expected experiments.

The intuition for the sufficiency of the prior-by-prior dominance condition can be clarified through condition 2. Condition 2 states that for any action plan η available for $\mathbf{q}$, there exists an action plan σ for $\mathbf{p}$ that guarantees a *weakly higher* expected utility in every auxiliary state. It is straightforward to see that condition 2 implies condition 3. Since DMs are expected utility maximizers condition on each μ, prior-by-prior dominance (condition 1) guarantees the existence of a *family* of action plans $\{\sigma_\mu \mid \mu \in \Delta(\Theta)\}$ such that $U(\sigma_\mu, \mathbf{p}_\mu) = U(\eta, \mathbf{q}_\mu)$ for every μ. To close the gap between this condition and condition 2, we need to invoke a general version of the minimax theorem to prove that the existence of such an indexed family of action plans is sufficient for the existence of one single action plan σ such that $U(\sigma, \mathbf{p}_\mu) \geq U(\eta, \mathbf{q}_\mu)$ for every μ.

Intuitively, if $\mathbf{p} \trianglerighteq_{PBP} \mathbf{q}$, then a probabilistically sophisticated DM can rank them according to Blackwell's order. A DM generally lacks understanding of the distribution of θ due to the informational ambiguity. Our result states that $\mathbf{p}$ and $\mathbf{q}$ can still be ranked, independent of any decision problem or preference parameters, when more general ambiguity preferences are considered.

On the one hand, the prior-by-prior dominance condition is powerful since it is sufficient for guaranteeing higher ex-ante value for all monotone ambiguity preferences. On the other hand, it is not overly restrictive, in the sense that it is necessary for guaranteeing higher ex-ante value within the small class of expected utility preferences. Therefore, prior-by-prior dominance is a robust equivalence condition for guaranteeing higher ex-ante value for any class of monotone ambiguity preferences that nests expected utility (e.g., the multiple-prior preferences and the smooth ambiguity preferences).

4 Restricting to Wald's Maxmin Criterion

In this section, we restrict our attention to a specific class of decision makers, the maxmin DMs. These are decision makers who apply Wald's maxmin criterion when evaluating ambiguous experiments. We obtain a different informativeness order after this restriction.

The aggregator corresponding to Wald's maxmin criterion is an aggregator $V_W : \mathbb{R}^\Theta \to \mathbb{R}$ defined by $V_W(f) := \min_{\theta\in\Theta} f(\theta)$. We will refer to V_W as the Wald aggregator.

The ex-ante value of an ambiguous experiment under the Wald aggregator is

$$W(\mathbf{p}) = \max_{\sigma\in A_S} \min_{\theta\in\Theta} U(\sigma, \mathbf{p}_\theta). \tag{7}$$

where $U(\sigma, \mathbf{p}_\theta)$ is defined in the same way as in equation (4). Note that the Wald aggregator does not nest any expected utility aggregators. That is, $\mathcal{V}_{EU}$ has an empty intersection with $\{V_W\}$. Therefore, Theorem 1 does not apply directly to $\{V_W\}$.

In particular, we can construct a pair of ambiguous experiments $\mathbf{p}$ and $\mathbf{q}$ such that: (i) $\mathbf{p}$ does *not* prior-by-prior dominate $\mathbf{q}$, but (ii) $\mathbf{p}$ has higher ex-ante value than $\mathbf{q}$ for every maxmin DM in every decision problem. Therefore, for every maxmin DM to choose $\mathbf{p}$ over $\mathbf{q}$, prior-by-prior dominance is sufficient but not necessary.

Definition 6. Let $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ and $\mathbf{q} : \Omega \times \Theta \to \Delta(T)$ be two ambiguous experiments. We say $\mathbf{p}$ is **Wald more informative** than $\mathbf{q}$ if for any $\mu \in \Delta(\Theta)$, there exists $\nu \in \Delta(\Theta)$ such that $\mathbf{q}_\nu$ is a garbling of $\mathbf{p}_\mu$.

The Wald-more-informative condition induces another informativeness order.

Theorem 2. *Let $\boldsymbol{p} : \Omega\times\Theta \to \Delta(S)$ and $\boldsymbol{q} : \Omega\times\Theta \to \Delta(T)$ be two ambiguous experiments. The following conditions are equivalent:*

1. *$\boldsymbol{p}$ is Wald more informative than $\boldsymbol{q}$.*

2. *$\boldsymbol{p}$ is preferred to $\boldsymbol{q}$ in every decision problem by every decision maker who uses the maxmin criterion. That is, $\boldsymbol{p}$ has weakly higher ex-ante value than $\boldsymbol{q}$ for every A, u, π when the aggregator is V_W.*

The proof of Theorem 2 is omitted since it is a corollary of a more general theorem (Theorem 4), formally stated in Appendix A and proved in Appendix B, that covers the case in which the auxiliary state space Θ can be any non-empty set.

This new informativeness order is closely related to that induced by the prior-by-prior dominance relation. Both are obtained by comparing expected experiments associated with $\mathbf{p}$ and $\mathbf{q}$. However, they are quite different. Prior-by-prior dominance requires $\mathbf{q}_\mu$ to be a garbling of $\mathbf{p}_\mu$ for every first order belief μ over Θ. But being Wald-more-informative only requires the existence of some ν such that the resulting expected experiment $\mathbf{q}_\nu$ is a garbling of $\mathbf{p}_\mu$, and ν can be different from μ. Thus, being Wald-more informative is weaker than prior-by-prior dominance. That is, if $\mathbf{p}$ prior-by-prior dominates $\mathbf{q}$, then $\mathbf{p}$ is Wald more informative than $\mathbf{q}$, but the converse is not true.

When a DM uses the maxmin criterion to evaluate ambiguous experiments, she only cares about the worst possible data generating process for each ambiguous experiment. This makes the correlation of $\mathbf{p}$ and $\mathbf{q}$ through Θ irrelevant to her. The prior-by-prior dominance condition, which builds on the correlation between ambiguous experiments through Θ, thus becomes too strong and not necessary.

Each ambiguous experiment $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ naturally corresponds to a set of Blackwell experiments $\mathbf{p}(\Theta)$, where $\mathbf{p}(\Theta) := \{\mathbf{p}_\theta : \Omega \to \Delta(S) \mid \theta \in \Theta\}$ is the collection of all possible data generating processes for $\mathbf{p}$. This set of Blackwell experiments is not a sufficient statistic for its corresponding ambiguous experiment because we lose the information on how the set is indexed. Under general monotone ambiguity preferences, merely observing the unindexed sets of Blackwell experiemnts is not enough to compare two ambiguous experiments because there is no assumption of symmetry over the auxiliary states. Such a symmetry assumption is automatically satisfied for maxmin DMs and therefore, comparing sets of Blackwell experiments becomes sufficient. In other words, we can state the Wald-more-informative condition using two unindexed sets. Let $P = \mathbf{p}(\Theta)$ and $Q = \mathbf{q}(\Theta)$, then $\{\mathbf{p}_\mu \mid \mu \in \Delta(\Theta)\} = \mathrm{conv}(P)$ and $\{\mathbf{q}_\mu \mid \mu \in \Delta(\Theta)\} = \mathrm{conv}(Q)$. Thus, $\mathbf{p}$ is Wald more informative than $\mathbf{q}$ if and only if the following condition holds:

$$\text{For any } p \in \mathrm{conv}(P), \text{ there exists } q \in \mathrm{conv}(Q) \text{ such that } q \text{ is a garbling of } p. \qquad (\star)$$

This condition specifies an order over sets of experiments even without direct reference to any auxiliary state space Θ. We thus find it natural to say "a set of experiments P is Wald more informative than Q" when condition $(\star)$ holds. We show that P is Wald more informative than Q if and only if P guarantees a higher ex-ante value than Q for every maxmin DM in every decision problem. We refer interested readers to Theorem 5 in Appendix B.2 for a formal treatment. Orders over unindexed sets of experiments might be of interest beyond the purpose of this paper. For example, see Definition 4 of de Oliveira et al. (2017) and Definition 12 of Cheng and Börgers (2023).

5 Conclusion

In this paper, we propose a model for informational ambiguity and use it to establish informativeness orders over ambiguous experiments. These orders generalize Blackwell's garbling condition. Ambiguous experiments are modeled as mappings from an auxiliary state space to the set of Blackwell experiments. One informativeness order is induced by the prior-by-prior dominance condition. An ambiguous experiment $\mathbf{p}$ prior-by-prior dominates another $\mathbf{q}$ if the expected experiment $\mathbf{p}_\mu$ is more informative than $\mathbf{q}_\mu$ for every belief μ over the auxiliary state space. This order is robust among any monotone ambiguity preference that nest expected utility. When restricting attention to maxmin DMs, the informativeness order is characterized by the Wald-more-informative condition. An ambiguous experiment $\mathbf{p}$ is Wald more informative than $\mathbf{q}$ if for any belief μ, there exists a belief ν such that $\mathbf{p}_\mu$ is more informative than $\mathbf{q}_\nu$. An adapted version of the Wald-more-informative condition characterizes an informativeness order over sets of Blackwell experiments evaluated by the maxmin criterion.

Appendices

A General Auxiliary State Spaces

In this section, we consider auxiliary state spaces beyond finite sets and we will show that our results in the main text remain valid in this more general environment.

Let the auxiliary state space Θ be an arbitrary nonempty set.

Ambiguous experiments are defined in the same way as in Definition 1, that is, an **ambiguous experiment** is a mapping $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$, where the payoff-relevant space Ω and the set of signals S are still assumed to be finite.

Consider a DM facing a decision problem A with a state-dependent utility function

$u : \Omega \times A \to \mathbb{R}$ and a prior belief $\pi \in \Delta(\Omega)$. The **expected utility conditional on** auxiliary state θ, $U(\sigma, \mathbf{p}_\theta)$, is defined the same way as in equation (4). Note that U is bounded as a function of θ for any fixed tuple $(\mathbf{p}, A, u, \pi, \sigma)$.

We say $V : \mathbb{R}^\Theta \to \mathbb{R}$ is a **monotone aggregator** if $V(f) \geq V(g)$ whenever two functions $f, g : \Theta \to \mathbb{R}$ satisfy $f(\theta) \geq g(\theta)$ for all $\theta \in \Theta$. Comparing to Definition 2, we drop the continuity requirement. Let $\mathcal{V}_{Mono}$ denote the set of all monotone aggregators.

For a DM with preference parameters (u, π, V) in a decision problem A, the **ex-ante value** of an ambiguous experiment $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ is

$$W(\mathbf{p}) = \sup_{\sigma \in A_S} V\left(\left\{U(\sigma, \mathbf{p}_\theta)\right\}_{\theta \in \Theta}\right).$$

Comparing to its counterpart in the main text (Definition 3), we have the supremum operator instead of the maximum operator. This change has no impact on the interpretation of our comparison result: If $\mathbf{p}$ has higher ex-ante value than $\mathbf{q}$ according to the more general definition, then for any action plan η that can be made when the DM faces $\mathbf{q}$, she can find another action plan σ that guarantees

$$V\left(\left\{U(\sigma, \mathbf{p}_\theta)\right\}_{\theta \in \Theta}\right) \geq V\left(\left\{U(\eta, \mathbf{q}_\theta)\right\}_{\theta \in \Theta}\right).$$

Moreover, without the continuity requirement, we can include more aggregators that corresponds to more ambiguity preferences.

Let 2^Θ denote the set of all subsets of Θ. Let Δ_0 denote the set of all simple probability measures (i.e., probability measures with a finite support) over 2^Θ. For any $\mu \in \Delta_0$, let $\text{supp}(\mu)$ denote the support of μ. The expected experiment $\mathbf{p}_\mu$ is given by

$$\mathbf{p}_\mu(s \mid \omega) = \sum_{\theta \in \text{supp}(\mu)} \mathbf{p}_\theta(s \mid \omega)\mu(\theta), \ \forall (s, \omega) \in S \times \Omega.$$

We say that an ambiguous experiment $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ **prior-by-prior dominates** another $\mathbf{q} : \Omega \times \Theta \to \Delta(T)$ if $\mathbf{q}_\mu$ is a garbling of $\mathbf{p}_\mu$ for any $\mu \in \Delta_0$. Note that when Θ is finite, this coincides with our definition of prior-by-prior dominance in the main text (Definition 5). Finally, let $\mathcal{V}_0 := \{V_\mu \mid \mu \in \Delta_0\}$ where $V_\mu : \mathbb{R}^\Theta \to \mathbb{R}$ is the expected utility aggregator corresponding to μ. That is, $V_\mu(f) = \sum_{\theta \in \text{supp}(\mu)} \mu(\theta) f(\theta)$.

Theorem 3. *Let $\boldsymbol{p} : \Omega \times \Theta \to \Delta(S)$ and $\boldsymbol{q} : \Omega \times \Theta \to \Delta(T)$ be two ambiguous experiments and let $\mathcal{V}$ be a class of aggregators such that $\mathcal{V}_0 \subset \mathcal{V} \subset \mathcal{V}_{Mono}$. The following conditions are equivalent:*

1. $\boldsymbol{p}$ *prior-by-prior dominates* $\boldsymbol{q}$.

2. *For any* A, u, π *and any* $\eta \in A_T$, *there exists* $\sigma \in A_S$ *such that*

$$U\big(\sigma, \boldsymbol{p}_\theta\big) \geq U\big(\eta, \boldsymbol{q}_\theta\big), \ \forall \theta \in \Theta.$$

3. $\boldsymbol{p}$ *is preferred to* $\boldsymbol{q}$ *in every decision problem by every decision maker whose ambiguity preferences can be represented by some* $V \in \mathcal{V}$. *That is,* $\boldsymbol{p}$ *has weakly higher ex-ante value than* $\boldsymbol{q}$ *for every* A, u, π *and every* $V \in \mathcal{V}$.

The proof of Theorem 3 is in Appendix B.1.

When Θ is finite, $\mathcal{V}_0$ reduces to $\mathcal{V}_{EU}$ as defined in equation (2) in the main text and Theorem 3 coincides with Theorem 1.

A.1 Restricting to Wald's Maxmin Criterion

This section corresponds to Section 4 in the main text. We restrict attention to DMs who apply Wald's maxmin criterion when aggregating payoffs over the auxiliary states.

Let $V_W : \mathbb{R}^\Theta \to \mathbb{R}$ denote the Wald aggregator defined by $V_W(f) = \inf_{\theta \in \Theta} f(\theta)$.

For a maxmin DM with preference parameters (u, π) in a decision problem A, the ex-ante value of an ambiguous experiment $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ is

$$W(\mathbf{p}) = \sup_{\sigma \in A_S} \inf_{\theta \in \Theta} U(\sigma, \mathbf{p}_\theta).$$

The set of Blackwell experiments associated with $\mathbf{p}$ plays an important role in our analysis. Let $\mathbf{p}(\Theta) := \{\mathbf{p}_\theta \mid \theta \in \Theta\}$. That is, $\mathbf{p}(\Theta)$ is the set of possible data generating processes behind $\mathbf{p}$. This set $\mathbf{p}(\Theta)$ can be identified as a bounded subset of $\mathbb{R}^{|\Omega| \times |S|}$.

Since Θ is an arbitrary set and an ambiguous experiment $\mathbf{p}$ needs not be continuous in Θ, this set $\mathbf{p}(\Theta)$ could be open. To overcome this issue, we consider the closure of this set. Formally, let $\overline{\mathbf{p}(\Theta)}$ denote the closure of $\mathbf{p}(\Theta)$. This closure is taken with respect to the standard Euclidean topology over $\mathbb{R}^{|\Omega| \times |S|}$. Since $\overline{\mathbf{p}(\Theta)}$ is bounded, it must be compact.

Definition 7. Let $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ and $\mathbf{q} : \Omega \times \Theta \to \Delta(T)$ be two ambiguous experiments. We say $\mathbf{p}$ is **Wald more informative** than $\mathbf{q}$ if for every Blackwell experiment $p \in \text{conv}\left(\overline{\mathbf{p}(\Theta)}\right)$, there exists $q \in \text{conv}\left(\overline{\mathbf{q}(\Theta)}\right)$ such that q is a garbling of p.

Theorem 4. *Let* $\boldsymbol{p} : \Omega \times \Theta \to \Delta(S)$ *and* $\boldsymbol{q} : \Omega \times \Theta \to \Delta(T)$ *be two ambiguous experiments. The following conditions are equivalent:*

1. *$\boldsymbol{p}$ is Wald more informative than $\boldsymbol{q}$.*

2. *$\boldsymbol{p}$ is preferred to $\boldsymbol{q}$ in every decision problem by every decision maker who uses the maxmin criterion. That is, $\boldsymbol{p}$ has weakly higher ex-ante value than $\boldsymbol{q}$ for every A, u, π when the aggregator is V_W.*

The proof of Theorem 4 is in Appendix B.2.

When Θ is finite, both $\mathbf{p}(\Theta)$ and $\mathbf{q}(\Theta)$ are finite and equal to their respective closure. Thus, $\text{conv}\left(\overline{\mathbf{p}(\Theta)}\right) = \{\mathbf{p}_\mu \mid \mu \in \Delta(\Theta)\}$, and $\text{conv}\left(\overline{\mathbf{q}(\Theta)}\right) = \{\mathbf{q}_\nu \mid \nu \in \Delta(\Theta)\}$. Definition 7 coincides with Definition 6 and Theorem 4 coincides with Theorem 2.

B Omitted Proofs

B.1 Proof of Theorem 3

We prove Theorem 3 by showing that $1 \implies 2 \implies 3 \implies 1$.

Recall that Δ_0 is the set of all probability measures over Θ with finite support and $\mathbf{p}$ prior-by-prior dominates $\mathbf{q}$ if $\mathbf{q}_\mu$ is a garbling of $\mathbf{p}_\mu$ for all $\mu \in \Delta_0$. Part of our proof for "$1 \implies 2$" is similar in spirit for the proof of Lemma 2 in Li and Zhou (2020) and the proof of Theorem 1 in Çelen (2012).

Proof that $1 \implies 2$. Let Γ denote the set of all garblings, that is, $\Gamma := \{\gamma \mid \gamma : S \to \Delta(T)\}$. Γ is compact and convex since both S and T are finite. Fix any A, u, π, and fix any action plan $\eta \in A_T$, we define an auxiliary function $H : \Gamma \times \Delta_0 \to \mathbb{R}$ by

$$H(\gamma, \mu) := U(\eta \circ \gamma, \mathbf{p}_\mu) - U(\eta, \mathbf{q}_\mu).$$

Note that γ is a mapping from S to $\Delta(T)$ and η is a mapping from T to $\Delta(A)$. Therefore, their composition $\eta \circ \gamma$ is a mapping from S to $\Delta(A)$. In other words, $\eta \circ \gamma$ is an element in A_S. Thus, $H(\gamma, \mu)$ captures the difference between the expected payoff of action plan $\eta \circ \gamma$ under $\mathbf{p}_\mu$ and the expected payoff of action plan η under $\mathbf{q}_\mu$.

To prove 1 implies 2, it suffices to show that

$$\max_{\gamma \in \Gamma} \inf_{\mu \in \Delta_0} H(\gamma, \mu) \geq 0.$$

To show that the left hand side in the equation above is well defined and non-negative, we invoke the Kneser-Fan minimax theorem for concave-convex functions (Terkelsen, 1972).

Γ is a compact and convex subset of $\mathbb{R}^{|S|\times|T|}$ under the Euclidean topology. Δ_0 is a convex subset of the vector space $\mathbb{R}^{\Theta}$. For each $\gamma \in \Gamma$, the function $\mu \mapsto -H(\gamma,\mu)$ is linear (hence concave) on Δ_0. For each $\mu \in \Delta_0$, the function $\gamma \mapsto -H(\gamma,\mu)$ is linear (hence convex and continuous since Γ is finite dimensional) on Γ. Therefore, by the Kneser-Fan minimax theorem for concave-convex functions (Terkelsen, 1972, page 411, Corollary 2),

$$\min_{\gamma\in\Gamma}\sup_{\mu\in\Delta_0} -H(\gamma,\mu) = \sup_{\mu\in\Delta_0}\min_{\gamma\in\Gamma} -H(\gamma,\mu),$$

which further indicates that

$$\max_{\gamma\in\Gamma}\inf_{\mu\in\Delta_0} H(\gamma,\mu) = \inf_{\mu\in\Delta_0}\max_{\gamma\in\Gamma} H(\gamma,\mu).$$

The right hand side of the equation above is non-negative, since for any $\mu \in \Delta_0$, the prior-by-prior dominance condition guarantees that there exists some $\gamma_\mu \in \Gamma$ such that $\mathbf{q}_\mu = \gamma_\mu \circ \mathbf{p}_\mu$ and this garbling γ_μ guarantees that $H(\gamma_\mu,\mu) = 0$. $\square$

Proof that 2 $\Longrightarrow$ 3. Fix any (A,u,π,V) with $V \in \mathcal{V}_{Mono}$ and let $W(\mathbf{q})$ denote the ex-ante value of $\mathbf{q}$, that is,

$$W(\mathbf{q}) = \sup_{\eta\in A_T} V\left(\left\{U\big(\eta,\mathbf{q}_\theta\big)\right\}_{\theta\in\Theta}\right).$$

Then for any $\varepsilon > 0$, there exists $\eta_\varepsilon \in A_T$ such that

$$V\left(\left\{U\big(\eta_\varepsilon,\mathbf{q}_\theta\big)\right\}_{\theta\in\Theta}\right) > W(\mathbf{q}) - \varepsilon.$$

By condition 2 in Theorem 4, there exists an action plan $\sigma_\varepsilon \in A_S$ such that

$$U\big(\sigma_\varepsilon,\mathbf{p}_\theta\big) \geq U\big(\eta_\varepsilon,\mathbf{q}_\theta\big),\ \forall\theta\in\Theta.$$

By the monotonicity of V, this further indicates that

$$V\left(\left\{U\big(\sigma_\varepsilon,\mathbf{p}_\theta\big)\right\}_{\theta\in\Theta}\right) \geq V\left(\left\{U\big(\eta_\varepsilon,\mathbf{q}_\theta\big)\right\}_{\theta\in\Theta}\right) > W(\mathbf{q}) - \varepsilon.$$

Since this holds for any $\varepsilon > 0$, we have $W(\mathbf{p}) = \sup_{\sigma\in A_S} V\left(\left\{U\big(\sigma,\mathbf{p}_\theta\big)\right\}_{\theta\in\Theta}\right) \geq W(\mathbf{q})$. $\square$

Proof that 3 $\Longrightarrow$ 1. It suffices to prove the necessity of prior-by-prior dominance when $\mathcal{V} = \mathcal{V}_0$. That is, if $\mathbf{p}$ is preferred to $\mathbf{q}$ by every DM with $V \in \mathcal{V}_0$, then $\mathbf{p} \trianglerighteq_{PBP} \mathbf{q}$.

Suppose by contradiction that it is not the case $\mathbf{p} \trianglerighteq_{PBP} \mathbf{q}$, then there must exist some $\mu \in \Delta_0$ such that $\mathbf{q}_\mu$ is not a garbling of $\mathbf{p}_\mu$. Fixing this belief μ and its corresponding aggregator $V_\mu \in \mathcal{V}_0$, that is, the DM believes that μ is the correct distribution over Θ and use aggregator V_μ to evaluate action plans. Thus,

$$\begin{aligned}
W(\mathbf{p}) &= \sup_{\sigma \in A_S} \; V_\mu\left(\left\{U\left(\sigma, \mathbf{p}_\theta\right)\right\}_{\theta\in\Theta}\right) \\
&= \sup_{\sigma \in A_S} \; \sum_{\theta \in \mathrm{supp}(\mu)} U\left(\sigma, \mathbf{p}_\theta\right)\mu(\theta) \\
&= \sup_{\sigma \in A_S} \; U\left(\sigma, \sum_{\theta \in \mathrm{supp}(\mu)} \mu(\theta)\mathbf{p}_\theta\right) \\
&= \sup_{\sigma \in A_S} \; U(\sigma, \mathbf{p}_\mu) = \max_{\sigma \in A_S} \; U(\sigma, \mathbf{p}_\mu)
\end{aligned}$$

where the last equality follows from U is linear in its first argument and A_S is compact.

Since $\mathbf{q}_\mu$ is not a garbling of $\mathbf{p}_\mu$, then by Blackwell's theorem (see, for example, Theorem 1 of de Oliveira 2018), there must exist a triplet (A, u, π) such that

$$\max_{\sigma \in A_S} \; U(\sigma, \mathbf{p}_\mu) < \max_{\eta \in A_T} \; U(\eta, \mathbf{q}_\mu).$$

Combine this triplet (A, u, π) with aggregator V_μ,

$$\begin{aligned}
W(\mathbf{p}) &= \sup_{\sigma \in A_S} \; V_\mu\left(\left\{U\left(\sigma, \mathbf{p}_\theta\right)\right\}_{\theta\in\Theta}\right) \\
&= \max_{\sigma \in A_S} \; U(\sigma, \mathbf{p}_\mu) \\
&< \max_{\eta \in A_T} \; U(\eta, \mathbf{q}_\mu) \\
&= \sup_{\eta \in A_T} \; V_\mu\left(\left\{U\left(\eta, \mathbf{q}_\theta\right)\right\}_{\theta\in\Theta}\right) = W(\mathbf{q})
\end{aligned}$$

which is a direct contradiction to our assumption that $\mathbf{p}$ has weakly higher ex-ante value than $\mathbf{q}$ for any (A, u, π) and $V \in \mathcal{V}_0$. This completes the proof that prior-by-prior dominance is necessary for any class of aggregators $\mathcal{V} \supset \mathcal{V}_0$. □

B.2 Proof of Theorem 4

We start our proof by formally presenting a result about the Wald-more-informative condition over sets of Blackwell experiments.

Let (P, S) and (Q, T) be two sets of Blackwell experiments, where each $p \in P$ is a mapping $p : \Omega \to \Delta(S)$ and each $q \in Q$ is a mapping $q : \Omega \to \Delta(T)$. Suppose P is a compact subset of $\mathbb{R}^{|\Omega|\times|S|}$ and Q is a compact subset of $\mathbb{R}^{|\Omega|\times|T|}$. DMs evaluate each set of experiments by the maxmin criterion. That is, for a DM with preference parameters u and π facing decision problem A, the **ex-ante value** of P is defined by

$$W(P) := \max_{\sigma \in A_S} \min_{p \in P} U(\sigma, p).$$

Theorem 5. *Let (P, S) and (Q, T) be two compact sets of Blackwell experiments. The following are equivalent:*

1. *For every $p \in \operatorname{conv}(P)$, there exists $q \in \operatorname{conv}(Q)$ that is a garbling of p.*
2. *P is preferred to Q in every decision problem by every decision maker using maxmin criterion. That is, P has weakly higher ex-ante value than Q for every A, u, π.*

Proof of Theorem 5. To prove that $1 \implies 2$, fix any (A, u, π),

$$\begin{aligned}
W(P) &= \max_{\sigma \in A_S} \min_{p \in P} U(\sigma, p) \\
&= \max_{\sigma \in A_S} \min_{p \in \operatorname{conv}(P)} U(\sigma, p) \\
&= \min_{p \in \operatorname{conv}(P)} \max_{\sigma \in A_S} U(\sigma, p) && \text{(minimax theorem)} \\
&= \max_{\sigma \in A_S} U(\sigma, p_*) && \text{(for some } p_* \in \operatorname{conv}(P)\text{)} \\
&\geq \max_{\eta \in A_T} U(\eta, q_*) && \text{(for some } q_* \in \operatorname{conv}(Q) \text{ by condition 1)} \\
&\geq \min_{q \in \operatorname{conv}(Q)} \max_{\eta \in A_T} U(\eta, q) \\
&= \max_{\eta \in A_T} \min_{q \in \operatorname{conv}(Q)} U(\eta, q) \\
&= \max_{\eta \in A_T} \min_{q \in Q} U(\eta, q) = W(Q)
\end{aligned}$$

where the second equality follows from the fact that U is linear in its second argument, the third equality follows from von Neumann's minimax theorem since both $\operatorname{conv}(P)$ and A_S are compact and convex, the first inequality follows from condition 1 and Blackwell's theorem (see, for example, Theorem 1 of de Oliveira 2018).

Then we prove $2 \implies 1$ by proving its contrapositive.

Suppose there exists $p_0 \in \operatorname{conv}(P)$ such that no $q \in \operatorname{conv}(Q)$ is a garbling of p, we want to construct a triplet (A, u, π) in which

$$W(P) = \max_{\sigma \in A_S} \min_{p \in P} U(\sigma, p) < \max_{\eta \in A_T} \min_{q \in Q} U(\eta, q) = W(Q).$$

Fix $A = T$, that is, the action space is the set of signal realizations for Q. Then the sets of action plans are $A_S = \{\sigma \mid \sigma : S \to \Delta(T)\}$ and $A_T = \{\eta \mid \eta : T \to \Delta(T)\}$. Consider an action plan $r \in A_T$ defined by $r(t_i \mid t_j) = \mathbf{1}[i = j]$, that is, r is the action plan that reports the realized signal. Then for any $q \in Q$, $r \circ q = q$ since

$$(r \circ q)(t_i \mid \omega) = \sum_{t \in T} r(t_i \mid t) q(t \mid \omega) = q(t_i \mid \omega), \ \forall (t_i, \omega) \in A \times \Omega.$$

Let $\Lambda_{Q,r} := \{r \circ q \mid q \in \text{conv}(Q)\}$. That is, $\Lambda_{Q,r}$ is the set of all probability measures over A conditional on Ω that can be induced by action plan r together with some Blackwell experiment in $\text{conv}(Q)$. Then $\Lambda_{Q,r} = \text{conv}(Q)$. Similarly, let $\Lambda_{p_0} := \{\sigma \circ p_0 \mid \sigma \in A_S\}$.

For any $u : \Omega \times T \to \mathbb{R}$ and $\pi \in \Delta(\Omega)$,

$$\begin{aligned}
\max_{\sigma \in A_S} U(\sigma, p_0) &= \max_{\sigma \in A_S} \sum_{s \in S} \sum_{\omega \in \Omega} \pi(\omega) p_0(s \mid \omega) \sum_{a \in T} \sigma(a \mid s) u(\omega, a) \\
&= \max_{\sigma \in A_S} \sum_{\omega \in \Omega} \left(\sum_{a \in T} u(\omega, a) \underbrace{\sum_{s \in S} \sigma(a \mid s) p_0(s \mid \omega)}_{(\sigma \circ p_0)(a \mid \omega)} \right) \pi(\omega) \\
&= \max_{\lambda \in \Lambda_{p_0}} \sum_{\omega \in \Omega} \left(\sum_{a \in T} u(\omega, a) \lambda(a \mid \omega) \right) \pi(\omega)
\end{aligned}$$

where the last equality follows from the one-to-one correspondence of A_S and Λ_{p_0}.

Since there does not exist any $q \in \text{conv}(Q)$ such that q is a garbling of p_0, $\Lambda_{p_0} \cap \Lambda_{Q,r} = \emptyset$. Otherwise, there must exist λ in their intersection $\Lambda_{p_0} \cap \Lambda_{Q,r}$, which further indicates the existence of an action plan $\sigma : S \to \Delta(T)$ and an experiment $q \in \text{conv}(Q)$ such that $\sigma \circ p_0 = \lambda = q$, that is, some $q \in \text{conv}(Q)$ is a garbling of p_0. Contradiction.

But both Λ_{p_0} and $\Lambda_{Q,r}$ are compact and convex subsets of $\mathbb{R}^{|\Omega| \times |T|}$, hence we can apply the separating hyperplane theorem and conclude that there exists a nonzero vector $v \in \mathbb{R}^{|\Omega| \times |T|}$ and real numbers $c_1 < c_2$ such that

$$\sum_{\omega \in \Omega} \sum_{a \in T} v(\omega, a) \lambda(a \mid \omega) < c_1, \ \sum_{\omega \in \Omega} \sum_{a \in T} v(\omega, a) \lambda'(a \mid \omega) > c_2, \ \forall \lambda \in \Lambda_{p_0}, \ \forall \lambda' \in \Lambda_{Q,r}.$$

Consider $A = T$, $u = v$ as given above, and $\pi = \text{uniform}(\Omega)$.

$$\begin{aligned}
\max_{\sigma \in A_S} U(\sigma, p_0) &= \max_{\lambda \in \Lambda_{p_0}} \sum_{\omega \in \Omega} \left(\sum_{a \in T} v(\omega, a) \lambda(a \mid \omega) \right) \pi(\omega) \\
&= \frac{1}{|\Omega|} \cdot \max_{\lambda \in \Lambda_{p_0}} \sum_{\omega \in \Omega} \sum_{a \in T} v(\omega, a) \lambda(a \mid \omega) \\
&< \frac{1}{|\Omega|} \cdot \min_{\lambda' \in \Lambda_{Q,r}} \sum_{\omega \in \Omega} \sum_{a \in T} v(\omega, a) \lambda'(a \mid \omega) = \min_{q \in \text{conv}(Q)} U(r, q)
\end{aligned}$$

where the inequality follows from the separation result above and the last equality follows from the one-to-one correspondence of $\Lambda_{Q,r}$ and $\text{conv}(Q)$.

Therefore, with $(A, u, \pi) = (T, v, \text{uniform}(\Omega))$,

$$\begin{aligned}
W(P) &= \max_{\sigma \in A_S} \min_{p \in P} U(\sigma, p) \\
&= \max_{\sigma \in A_S} \min_{p \in \text{conv}(P)} U(\sigma, p) \\
&= \min_{p \in \text{conv}(P)} \max_{\sigma \in A_S} U(\sigma, p) \\
&\leq \max_{\sigma \in A_S} U(\sigma, p_0) \\
&< \min_{q \in \text{conv}(Q)} U(r, q) \\
&= \min_{q \in Q} U(r, q) \\
&\leq \max_{\eta \in A_T} \min_{q \in Q} U(\eta, q) = W(Q)
\end{aligned}$$

This completes the proof for Theorem 5. □

Now we can build our proof for Theorem 4 on Theorem 5. The key observation, as discussed in Section 4, is that comparisons of ambiguous experiments for maxmin DMs can be interpreted as comparisons of sets of experiments under the maxmin criterion.

For completeness, we establish the following lemma.

Lemma 1. *Let $\boldsymbol{p} : \Omega \times \Theta \to \Delta(S)$ and $\boldsymbol{q} : \Omega \times \Theta \to \Delta(T)$ be two ambiguous experiments. Let $\boldsymbol{p}(\Theta) = \{\boldsymbol{p}_\theta \mid \theta \in \Theta\}$, $\boldsymbol{q}(\Theta) = \{\boldsymbol{q}_\theta \mid \theta \in \Theta\}$. The following conditions are equivalent:*

1. *$\boldsymbol{p}$ is preferred to $\boldsymbol{q}$ in every decision problem by every decision maker who uses the maxmin criterion. That is, $\boldsymbol{p}$ has weakly higher ex-ante value than $\boldsymbol{q}$ for every A, u, π when the aggregator is V_W.*

2. *As sets of Blackwell experiments, $\overline{\boldsymbol{p}(\Theta)}$ is preferred to $\overline{\boldsymbol{q}(\Theta)}$ in every decision problem by every decision maker who uses the maxmin criterion.*

Proof of Lemma 1. Fix any A, u, π, combined with the Wald aggregator V_W,

$$
\begin{aligned}
W(\mathbf{p}) \geq W(\mathbf{q}) &\iff \sup_{\sigma \in A_S} \inf_{\theta \in \Theta} U(\sigma, \mathbf{p}_\theta) \geq \sup_{\eta \in A_T} \inf_{\theta \in \Theta} U(\eta, \mathbf{q}_\theta) \\
&\iff \sup_{\sigma \in A_S} \inf_{p \in \mathbf{p}(\Theta)} U(\sigma, p) \geq \sup_{\eta \in A_T} \inf_{q \in \mathbf{q}(\Theta)} U(\eta, q) \\
&\iff \sup_{\sigma \in A_S} \min_{p \in \overline{\mathbf{p}(\Theta)}} U(\sigma, p) \geq \sup_{\eta \in A_T} \min_{q \in \overline{\mathbf{q}(\Theta)}} U(\eta, q) \\
&\iff \max_{\sigma \in A_S} \min_{p \in \overline{\mathbf{p}(\Theta)}} U(\sigma, p) \geq \max_{\eta \in A_T} \min_{q \in \overline{\mathbf{q}(\Theta)}} U(\eta, q) \\
&\iff W\left(\overline{\mathbf{p}(\Theta)}\right) \geq W\left(\overline{\mathbf{q}(\Theta)}\right)
\end{aligned}
$$

where the third equivalence follows from facts that $\mathbf{p}(\Theta)$ and $\mathbf{q}(\Theta)$ are bounded subsets of finite-dimensional Euclidean spaces and that U is continuous in its second argument. $\square$

Proof of Theorem 4.

$\mathbf{p}$ is Wald more informative than $\mathbf{q}$

$\iff$ For any $p \in \operatorname{conv}\left(\overline{\mathbf{p}(\Theta)}\right)$, there exists $q \in \operatorname{conv}\left(\overline{\mathbf{q}(\Theta)}\right)$ that is a garbling of p

$\iff$ $\overline{\mathbf{p}(\Theta)}$ is preferred to $\overline{\mathbf{q}(\Theta)}$ in every decision problem by every decision maker who uses the maxmin criterion

$\iff$ $\mathbf{p}$ is preferred to $\mathbf{q}$ in every decision problem by every decision maker who uses the maxmin criterion

where the first equivalence follows from Definition 7, the second equivalence follows from Theorem 5, and the third equivalence follows from Lemma 1. $\square$

References

Asano, Takao, and Hiroyuki Kojima. 2019. "Consequentialism and dynamic consistency in updating ambiguous beliefs." *Economic Theory* 68 (1): 223–250.

Beauchêne, Dorian, Jian Li, and Ming Li. 2019. "Ambiguous persuasion." *Journal of Economic Theory* 179 312–365.

Blackwell, David. 1951. "Comparison of experiments." *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* 93–102.

Blackwell, David. 1953. "Equivalent comparisons of experiments." *The Annals of Mathematical Statistics* 265–272.

Bose, Subir, and Ludovic Renou. 2014. "Mechanism design with ambiguous communication devices." *Econometrica* 82 (5): 1853–1872.

Çelen, Boğaçhan. 2012. "Informativeness of experiments for MEU." *Journal of Mathematical Economics* 48 (6): 404–406.

Cerreia-Vioglio, Simone, Fabio Maccheroni, Massimo Marinacci, and Luigi Montrucchio. 2011. "Uncertainty averse preferences." *Journal of Economic Theory* 146 (4): 1275–1330.

Chen, Jaden Yang. 2022. "Biased learning under ambiguous information." *Journal of Economic Theory* 203 105492.

Cheng, Xienan, and Tilman Börgers. 2023. "Dominance and optimality." *Available at SSRN 4422054*.

Epstein, Larry G, and Yoram Halevy. 2023. "Hard-to-interpret signals." *Journal of the European Economic Association* jvad062.

Epstein, Larry G, and Martin Schneider. 2007. "Learning under ambiguity." *The Review of Economic Studies* 74 (4): 1275–1303.

Epstein, Larry G, and Martin Schneider. 2008. "Ambiguity, information quality, and asset pricing." *The Journal of Finance* 63 (1): 197–228.

Gensbittel, Fabien, Ludovic Renou, and Tristan Tomala. 2015. "Comparisons of ambiguous experiments." *HEC Paris Research Paper No. ECO/SCD-2015-1074*.

Gilboa, Itzhak, and David Schmeidler. 1989. "Maxmin expected utility with non-unique prior." *Journal of Mathematical Economics* 18 (2): 141–153.

Hansen, Lars Peter, and Thomas J Sargent. 2001. "Robust control and model uncertainty." *American Economic Review* 91 (2): 60–66.

Hansen, Lars Peter, and Thomas J Sargent. 2008. *Robustness.* Princeton university press.

Kellner, Christian, Mark T Le Quement, and Gerhard Riener. 2022. "Reacting to ambiguous messages: An experimental analysis." *Games and Economic Behavior* 136 360–378.

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji. 2005. "A smooth model of decision making under ambiguity." *Econometrica* 73 (6): 1849–1892.

Kops, Christopher, and Illia Pasichnichenko. 2023. "Testing negative value of information and ambiguity aversion." *Journal of Economic Theory* 213 105730.

Li, Jian, and Junjie Zhou. 2016. "Blackwell's informativeness ranking with uncertainty-averse preferences." *Games and Economic Behavior* 96 18–29.

Li, Jian, and Junjie Zhou. 2020. "Information order in monotone decision problems under uncertainty." *Journal of Economic Theory* 187 105012.

Liang, Yucheng. 2021. "Learning from unknown information sources." *Available at SSRN 3314789.*

Marinacci, Massimo. 2015. "Model uncertainty." *Journal of the European Economic Association* 13 (6): 1022–1100.

de Oliveira, Henrique. 2018. "Blackwell's informativeness theorem using diagrams." *Games and Economic Behavior* 109 126–131.

de Oliveira, Henrique, Tommaso Denti, Maximilian Mihm, and Kemal Ozbek. 2017. "Rationally inattentive preferences and hidden information costs." *Theoretical Economics* 12 (2): 621–654.

Shishkin, Denis, and Pietro Ortoleva. 2023. "Ambiguous information and dilation: An experiment." *Journal of Economic Theory* 105610.

Terkelsen, Frode. 1972. "Some minimax theorems." *Mathematica Scandinavica* 31 (2): 405–413.