Informativeness orders over ambiguous experiments*

Zichang Wang†

Current Draft: January 21, 2022

Abstract

We generalize Blackwell (1951)'s informativeness order to ambiguous experiments. The ambiguity in experiments is rooted in a lack of understanding about their probabilistic content. Formally, an ambiguous experiment is modeled as a mapping from an auxiliary state space to the set of unambiguous experiments. We show that one ambiguous experiment is preferred to another by every decision maker for 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 across a wide range of ambiguity preferences. Comparisons of sets of experiments evaluated using the maxmin criterion are studied as a special case and are shown to result in a weaker informativeness order called *Wald-more-informative*, which states that for any Blackwell experiment in the convex hull of the first set of experiments, there exists another in the convex hull of the second set that is Blackwell less informative.

Keywords: Statistical experiments, ambiguity, informativeness order

JEL Classification: D81, D83

*I am deeply indebted to my advisor, Todd Sarver, for his continued guidance and detailed comments throughout the development of this paper. I wish to thank Arjada Bardhi, Yonggyun Kim, Philipp Sadowski and Xiaosheng Mu for helpful discussions. I also wish to thank Tracy Lewis, Curtis Taylor and participants at the Duke Economic Theory Lunch seminar for their feedback. All errors are my own.

†Duke University, Department of Economics, 213 Social Sciences/Box 90097, Durham, NC 27708. Email: zichang.wang@duke.edu.

1 Introduction

Blackwell (1951, 1953) provides an intuitive way of modeling information by considering information as statistical experiments and establishes an elegant equivalence result on comparisons of such statistical experiments. However, information in reality is pervasively ambiguous, as we rarely know the exact probabilistic content of a statistical experiment when we receive information. 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 (the patient does not have the disease but the result is positive) and false negatives (the patient has the disease but the result is negative). Seldom do patients know the exact probabilities of false positive and false negative results, and they may therefore view the distribution of test results as ambiguous.

In this sense, Blackwell's formulation of information as unambiguous statistical experiments seems to be too ideal, and no results were provided on comparing the informativeness of experiments when ambiguity is present, even though some intuitive comparisons can be made in reality (e.g., patients may still intuitively prefer a medical test that has a smaller *maximum* probability of false positive or negative results). To accommodate such intuitive comparisons, can we generalize Blackwell's theorem on the comparison of (unambiguous) experiments to compare the "informativeness" of ambiguous experiments? In other words, can we find a meaningful generalization of Blackwell's *garbling* condition that is equivalent to every decision maker preferring one ambiguous experiment to another in every decision problem?

This paper provides a positive answer to the question above. Indeed, we can generalize the informativeness notion even when there is ambiguity in the information structures. Moreover, we can show that this new informativeness notion is robust across a wide range of ambiguity preferences.

Let Ω be a finite set of states that are directly relevant to the payoff of the decision maker (henceforth DM). An unambiguous experiment is a mapping $p : \Omega \to \Delta(S)$ where S is a finite set of signal realizations and $\Delta(S)$ is the set of all probability measures over S. The experiment p is unambiguous since the probability that signal s is observed in state ω is unambiguously specified for all s in S and ω in Ω. More than one approach is possible to model ambiguity in experiments. A natural first thought is to consider sets of unambiguous experiments (evaluated using the maxmin criterion), which we analyze in detail in Section 4. However, such a modeling approach significantly restricts the richness of ambiguity preferences that can be considered. To allow for more general ambiguity preferences (e.g., the smooth ambiguity preferences), we introduce an auxiliary state space Θ. An ambiguous experiment is modeled as a mapping $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$. For each auxiliary state $\theta \in \Theta$, $\mathbf{p}(\cdot, \theta) : \Omega \to \Delta(S)$ is the corresponding unambiguous experiment. The DM 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 ambiguity about experiments. In particular, as we will show in Section 4, this modeling approach nests the comparison of sets of experiments evaluated using the maxmin criterion as a special case. The following two examples further illustrate our modeling approach with an auxiliary state space Θ.

Example 1. Consider COVID-19 tests at different hospitals supplied by a single pharmaceutical company. The auxiliary state space Θ could be the set of all possible levels of test precision (in terms of likelihoods of false positives and false negatives). Patients may face ambiguity on Θ since there might be limited data about the precision for a relatively new test (or when applying an existing test to a new variant of the virus). The accuracy of any test clearly depends on Θ, but it can still vary across different hospitals due to different implementations (e.g., some hospitals may have more experienced testing crews than others).

Example 2. Consider financial analysis about a firm's financial status conducted by different financial service companies. The auxiliary state space Θ could be a set capturing the aspects that are common to all analysis but open to interpretation, such as the quality of public data or the representativeness (or predictive power) of past observations for future performance. Outside investors may face ambiguity on Θ due to their lack of relevant knowledge. The informational quality of any analysis clearly depends on Θ, but it can still vary across different companies.

We view the auxiliary state space as a modeling device that represents the common source of ambiguity for the ambiguous experiments to be compared. We treat this space and the dependence of the experiments on it as commonly understood by both the modeler and the DM. It is "auxiliary" since it is not directly payoff-relevant and influences the DM's decision and payoff only through influencing the information content of the experiment. Patients' payoffs only directly depend on whether or not they are infected, but not directly on the precision of the test. Investors' profits only directly depend on the financial status of the company, but not directly on the quality of the financial analysis.

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, $\mathbf{p}_\mu := \int_\Theta \mathbf{p}(\cdot, \theta)\ d\mu(\theta)$. Expected experiments are unambiguous by construction. Two ambiguous experiments $\mathbf{p}$ and $\mathbf{p}'$ are related by the *prior-by-prior dominance* condition if the expected experiment $\mathbf{p}_\mu$ is Blackwell more informative than $\mathbf{p}'_\mu$ for every μ. We show in Section 3 that $\mathbf{p}$ is preferred to $\mathbf{p}'$ (i.e., guarantees a higher ex-ante utility) by every decision maker in every decision problem *if and only if* $\mathbf{p}$ prior-by-prior dominates $\mathbf{p}'$. Prior-by-prior dominance is a direct generalization of Blackwell's informativeness notion: When Θ is a singleton, ambiguous experiments reduce to unambiguous experiments and prior-by-prior dominance reduces to being Blackwell more informative.

If a DM's ambiguity preference could be summarized by a single subjective belief μ over Θ, then the problem of comparing ambiguous experiments $\mathbf{p}$ and $\mathbf{p}'$ can be simplified as comparing their corresponding expected experiments $\mathbf{p}_\mu$ and $\mathbf{p}'_\mu$. This simplification is not possible in general since a DM facing ambiguity may have more general ambiguity preferences (e.g., the multiple prior preferences or the smooth preferences). Nonetheless, our main result implies that if $\mathbf{p}_\mu$ is Blackwell more informative than $\mathbf{p}'_\mu$ for *every* possible probability measure μ over Θ, then the DM will prefer $\mathbf{p}$ to $\mathbf{p}'$ as long as his ambiguity preference can be represented by some monotone aggregator of the auxiliary states, even if this aggregator does not correspond to a single subjective belief. This class of monotone aggregators is extremely general and includes essentially all ambiguity preference models used in the literature.

On the one hand, the prior-by-prior dominance condition is powerful since it is sufficient for guaranteeing higher ex-ante utility across a wide range of ambiguity preferences. On the other hand, it is not overly restrictive, in the sense that it is necessary for guaranteeing higher ex-ante utility within the small class of subjective expected utilities. Therefore, prior-by-prior dominance is a robust equivalence condition for guaranteeing higher ex-ante utility for any class of monotone ambiguity preferences that nests subjective expected utility.

We now give a numerical example to illustrate our comparison results:

Example 3. Consider two hospitals offering COVID-19 tests. $\mathbf{p}_1$ summarizes the test offered at Hospital 1 while $\mathbf{p}_2$ summarizes that at Hospital 2. Individuals are either infected (I) or not infected (N), and this infection status is the payoff relevant state. The outcome of a test is either positive (+) or negative (−). Probabilities of false positives and false negatives are ambiguous, modeled through an auxiliary state space $\Theta = [0, 0.02] \times [0, 0.02]$. A typical element $\theta = (\theta_+, \theta_-)$ denotes that the probability of a false positive is θ_+ and the probability of a false negative is θ_-.

In this example, $\Omega = \{I, N\}$, $S_1 = S_2 = \{+, -\}$, and $\Theta = [0, 0.02] \times [0, 0.02]$. Suppose $\mathbf{p}_1$ and $\mathbf{p}_2$ are given by

$$\mathbf{p}_1(\cdot, \theta_+, \theta_-) = \begin{array}{|c|c|c|} \hline & + & - \\ \hline I & 1-\theta_- & \theta_- \\ \hline N & \theta_+ & 1-\theta_+ \\ \hline \end{array} \quad \mathbf{p}_2(\cdot, \theta_+, \theta_-) = \begin{array}{|c|c|c|} \hline & + & - \\ \hline I & 1-1.01\theta_- & 1.01\theta_- \\ \hline N & 1.01\theta_+ & 1-1.01\theta_+ \\ \hline \end{array}$$

where $(\theta_+, \theta_-) \in \Theta$. That is, the test offered at Hospital 2 is more likely to generate false positives and false negatives in every possible state.[1] It can be checked that $\mathbf{p}_1$ prior-by-prior dominates $\mathbf{p}_2$, coinciding with our intuition that $\mathbf{p}_1$ should be regarded as "more informative." To further illustrate, consider a third test $\mathbf{p}_3$ defined by

$$\mathbf{p}_3(\cdot, \theta_+, \theta_-) = \begin{array}{|c|c|c|} \hline & + & - \\ \hline I & 1-\alpha & \alpha \\ \hline N & \alpha & 1-\alpha \\ \hline \end{array}$$

That is, $\mathbf{p}_3$ is an unambiguous test with a known false positive/negative rate of α. When $\alpha = 0$, $\mathbf{p}_3$ prior-by-prior dominates both $\mathbf{p}_1$ and $\mathbf{p}_2$, and will be considered superior by every decision maker. When $\alpha \geq 0.02$, $\mathbf{p}_1$ prior-by-prior dominates $\mathbf{p}_3$ and is thus preferred. When $\alpha \in (0, 0.02)$, $\mathbf{p}_3$ is in general not comparable with $\mathbf{p}_1$ or $\mathbf{p}_2$, as the decision maker's subjective belief over Θ becomes relevant. For example, if a decision maker believes that both θ_+ and θ_- do not exceed α with probability 1 in a certain decision scenario, then he would prefer $\mathbf{p}_1$ to $\mathbf{p}_3$ in that scenario.

Another possible approach to model the ambiguity in experiments is to consider sets of unambiguous experiments. We show in Section 4 that we can treat comparisons of sets of unambiguous experiments as a special case of comparisons of ambiguous experiments by focusing on a specific class of DMs: The DMs who apply Wald's maximin criterion when aggregating payoffs across θ. In this special case, one set of unambiguous experiments P

[1] One possible justification for the perfect correlation of the tests on Θ is as follows: Both hospitals utilize testing kits from the same pharmaceutical company but Hospital 2 has a less experienced testing crew who increases the probabilities of false positive/negative due to human error.

is preferred to another P' by every decision maker who applies the maximin criterion if and only if a *Wald-more informative* condition (W-more informative, in short) holds: We say P is W-more informative than P' if for any unambiguous experiment p in the convex hull of P, there exists p' in the convex hull of P' such that p is Blackwell more informative than p'. We show that the W-more informative condition induces an informativeness order that is weaker than that induced by prior-by-prior dominance, that is, there exists pairs of ambiguous experiments that are not comparable according to the prior-by-prior dominance condition although their corresponding sets of experiments are comparable according to the W-more informative condition. The intuition about why we get a weaker notion of informativeness is as follows: A DM who applies the maximin criterion reduces to an expected utility maximizers only when the set of experiments is a singleton. Thus, this specific class of DMs does not nest the class of DMs with subjective expected utility, which causes the prior-by-prior dominance condition to be not necessary.

For a more concrete illustration, consider Example 3 again and define

$$P_1 := \{\mathbf{p}_1(\cdot, \theta) \mid \theta \in \Theta\}, \quad P_3 := \{\mathbf{p}_3(\cdot, \theta) \mid \theta \in \Theta\}.$$

Then in cases where $\alpha = 0$ or $\alpha \geq 0.02$, W-informativeness and prior-by-prior dominance agree with each other: P_3 is W-more informative than P_1 when $\alpha = 0$ and the rank is simultaneously reversed when $\alpha \geq 0.02$. However, when $\alpha \in (0, 0.02)$, P_3 is W-more informative than P_1 even though their corresponding ambiguous experiments are not comparable according to the prior-by-prior dominance condition.

Our results hinge on several assumptions. First, we assume that the DM either has commitment power or is dynamically consistent. This assumption is needed as dynamic inconsistency may arise for some belief updating protocols when ambiguity is present. Second, we assume that the ambiguity is only on the auxiliary state space Θ while the DM has an unambiguous prior over the payoff-relevant state space Ω.[2] Third, we assume that the auxiliary states only have an indirect impact on the DM's payoff. A detailed discussion of the effects of relaxing the second and the third assumptions is in Section 5.

1.2 Related Literature

We contribute to the literature on the interaction of Blackwell's informativeness order and ambiguity preferences. Models based on the combination of ambiguous priors and

[2]Like in Example 3, the ambiguity is about the precision of the test, but not about the prior probability that a test taker is infected.

unambiguous experiments have been previously considered in the literature. Çelen (2012) shows that the Blackwell informativeness order remains valid among the class of maxmin expected utility preferences, under the assumption that the DM has commitment power. Further expanding the class of ambiguity preferences, Li and Zhou (2016) keep the commitment assumption and validate that the Blackwell informativeness order remains valid when the DM possess the very general uncertainty averse preferences as in Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011). In contrast to these studies, we consider a different problem and our study leads to a very different result: In our model, the ambiguity is in the information structures instead of in the prior beliefs over the payoff relevant state space, and our prior-by-prior dominance condition is a non-trivial generalization of the Blackwell informativeness order. Although different from existing papers, the study of ambiguous information structures could be important since information structures are indeed pervasively ambiguous in reality, and ambiguous information has gained increased attention in both the applied literature[3] and the experimental literature[4] as its importance has become recognized. We also discuss in Section 5 the combination of ambiguous priors and ambiguous experiments as an extension to our model. Chen (2020) shares our approach of modeling ambiguous experiments as a mapping from an auxiliary state space[5] to the space of Blackwell experiments, but his focus is on the learning behavior under ambiguous experiments, while our focus is on comparing their informativeness.

The closest work to ours is Gensbittel, Renou, and Tomala (2015) (henceforth GRT). In their paper, ambiguous experiments are modeled as convex and compact sets of *joint distributions* over payoff-relevant states and signal realizations. They consider DMs who apply Wald's maximin criterion and study models with and without the commitment assumption. Their case of commitment has a large overlap with the special case of our model discussed in Section 4. One major methodological difference sets us apart, both in terms of our results, and in terms of their interpretations. We model unambiguous experiments as collections of conditional distributions over signal realizations given the state, while they model them as joint distributions over states and signal realizations. Although our results are superficially similar, they are not directly comparable unless extra assumptions regarding the payoff-relevant prior are imposed. By their modeling approach,

[3]For example, Beauchêne, Li, and Li (2019), Bose and Renou (2014), Chen (2020), and Epstein and Schneider (2007, 2008).

[4]For example, Epstein and Halevy (2019), Liang (2019) and Shishkin and Ortoleva (2019).

[5]An auxiliary state space is referred to as "a model space" in his paper.

GRT treat the DM's prior belief over Ω as part of the information structure instead of part of the DM's preference parameters. Therefore, they are interpreting the theory as if the modeler can observe the DM's payoff-relevant priors. In contrast, our modeling approach treats the DM's prior belief over Ω as a preference parameter. Therefore, our comparison is prior-free in the sense that it can be applied when the modeler cannot observe DM's payoff-relevant priors.[6] In addition, GRT study a model where the commitment assumption is relaxed and the DMs still apply the maximin criterion, while we impose the commitment assumption but consider a much more general class of ambiguity preferences.

The rest of the paper is organized as follows. We review Blackwell's theorem in Section 2. General comparisons of ambiguous experiments are studied in Section 3. The special case of comparisons of sets of experiments is studied in Section 4. Variations of the key assumptions and their impact on our results are considered in Section 5. Proofs omitted in the main text are relegated to the Appendix.

2 Blackwell's Theorem

In this section, we describe some primitives for comparisons of experiments and review Blackwell's theorem. Let Ω be a finite set of states of the world. For any finite set X, we use $\Delta(X)$ to denote the set of all probability measures over X.

Definition 1. A **Blackwell experiment** (or interchageably, an **experiment**, or an **unambiguous experiment**) is a mapping $p : \Omega \to \Delta(S)$ where S is a finite set of signal realizations and p maps each state $\omega \in \Omega$ to a probability measure over S.

Slightly abusing notation, we will write $p(s \mid \omega)$ to denote the probability of observing s when the state is ω. An experiment can thus be viewed as a $|\Omega| \times |S|$ real matrix with each row representing a probability distribution over S. To define the notion of garbling, we first 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

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

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. Now we can define the notion of garbling.

[6] See Section 4.3 for additional discussion and examples.

Definition 2. Given two Blackwell experiments $p : \Omega \to \Delta(S)$ and $p' : \Omega \to \Delta(S')$, p' is a **garbling** of p if there exists some $\gamma : S \to \Delta(S')$ such that $p' = \gamma \circ p$.

Intuitively, p' is a garbling of p if one can replicate p' (in terms of the conditional probability of s' given ω) by adding "noise" (applying a stochastic operator) to p.

Consider an individual with a finite set of actions A facing a set S of signal realizations (we assume S to be finite throughout the paper). An **action plan** is a mapping $\sigma : S \to \Delta(A)$. We write $\sigma(\cdot \mid s)$ to denote the individual's (mixed) strategy after observing signal realization s. We use A_S to denote the collection of all action plans once the set of actions A and the set of signal realizations S are fixed, that is, $A_S := \{\sigma \mid \sigma : S \to \Delta(A)\}$.

Consider a Bayesian expected-utility maximizer with a finite set of actions A, a state-dependent utility function $u : \Omega \times A \to \mathbb{R}$, and a prior $\pi \in \Delta(\Omega)$. For this individual, the expected utility from action plan σ for experiment p is

$$U(\sigma, p) := \sum_{s \in S} \sum_{\omega \in \Omega} \sum_{a \in A} \pi(\omega) p(s \mid \omega) \sigma(a \mid s) u(\omega, a) \tag{1}$$

Definition 3. Facing A, u and π, the individual's **ex-ante expected utility** from the experiment p is $\max_{\sigma \in A_S} U(\sigma, p)$.

We now review Blackwell's theorem.

Theorem 1 (Blackwell (1951, 1953))**.** *Given two Blackwell experiments $p : \Omega \to \Delta(S)$ and $p' : \Omega \to \Delta(S')$, the following are equivalent:*

1. *p' is a garbling of p.*
2. *For any A, u, π and any action plan $\sigma' \in A_{S'}$, there exists an action plan $\sigma \in A_S$ such that $U\big(\sigma, p\big) = U\big(\sigma', p'\big)$.*
3. *Every Bayesian expected utility maximizer prefers p to p' for any possible decision problem. That is, p gives weakly higher ex-ante expected utility than p' for every A, u, and π.*

For a simple and elegant proof of Blackwell's theorem, see de Oliveira (2018). If any of the conditions in Theorem 1 holds, we say p is **(Blackwell) more informative** than p', and write $p \trianglerighteq p'$. Blackwell's theorem establishes the equivalence of a statistical condition (garbling) and an economical condition (higher ex-ante expected utility for any decision problem). Our goal is to study ambiguous experiments and characterize the equivalent condition (generalization of "garbling") to all decision makers with more general ambiguity preferences having higher ex-ante utility.

3 Comparing Ambiguous Experiments

As before, let Ω be a finite set of states that are directly relevant to the DM's payoff. Let Θ be a set of auxiliary states that govern the realization of Blackwell experiments. As illustrated in the introduction, Θ represents the source of ambiguity for the experiment. For ease of exposition, we focus on the case where Θ is a finite set in the main text. Our characterization result remains valid for any nonempty Θ. The more general result is stated in Appendix A and proved in Appendix B.

Definition 4. An **ambiguous experiment** is a mapping $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ where S is a finite set of signal realizations.

For each auxiliary state $\theta \in \Theta$, $\mathbf{p}(\cdot, \theta)$ can be thought of as a mapping from Ω to $\Delta(S)$, that is, $\mathbf{p}(\cdot, \theta)$ is the Blackwell experiment associated with state θ, and a natural interpretation for $\mathbf{p}$ is a mapping from the auxiliary state space to the set of Blackwell experiments. This structure helps to capture the DM's lack of understanding of the probabilistic content of the experiment.

Consider an individual with a finite set of actions A, a state-dependent utility function $u : \Omega \times A \to \mathbb{R}$ and a prior belief $\pi \in \Delta(\Omega)$. For ambiguous experiment $\mathbf{p}$ and action plan σ, the **expected utility conditional on** state θ is

$$U\big(\sigma, \mathbf{p}(\cdot, \theta)\big) = \sum_{s \in S} \sum_{\omega \in \Omega} \sum_{a \in A} \pi(\omega) \mathbf{p}(s \mid \omega, \theta) \sigma(a \mid s) u(\omega, a). \tag{2}$$

To clarify, we use $\mathbf{p}(s \mid \omega, \theta)$ to denote the probability of observing s when the pair of states is (ω, θ). This is another slight abuse of notation as we also use $\mathbf{p}(\cdot, \theta)$ to denote the Blackwell experiment associated with auxiliary state θ. These conditional expected utilities will be aggregated through an aggregator over Θ.

Definition 5. We say $V : \mathbb{R}^\Theta \to \mathbb{R}$ is a *monotone aggregator* if V is continuous[7] and $V(f) \geq V(g)$ whenever two functions $f, g : \Theta \to \mathbb{R}$ satisfy $f(\theta) \geq g(\theta)$ for all $\theta \in \Theta$.

Essentially all ambiguity preferences used in the literature correspond to some monotone aggregator, and special cases of V will be discussed in the next section to illustrate its generality. Since monotonicity is the only restriction we have on V, flexible ambiguity attitudes (ambiguity averse, ambiguity loving, mixed attitude) are allowed for essentially all classes of ambiguity preferences it nests.

[7] Since Θ is finite, we can endow $\mathbb{R}^\Theta$ with the Euclidean topology and the continuity is with respect to this standard topology.

Definition 6. Let $V : \mathbb{R}^\Theta \to \mathbb{R}$ be a monotone aggregator that captures the individuals' ambiguity preferences. Given A, u, π, and V, the individual's **ex-ante utility** from the ambiguous experiment $\mathbf{p}$ is

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

where $U\big(\sigma, \mathbf{p}(\cdot, \theta)\big)$ is the conditional expected utility defined in equation (2).

Two implicit assumptions are behind the definition above. First, we assume that the DM's has an unambiguous prior belief π over the payoff-relevant space Ω. Moreover, the payoffs $u(\omega, a)$ are first aggregated into the expected utility conditional on each auxiliary state θ, and the DM's ex-ante utility is the aggregation of these conditional expected utilities by V. Whether or not the DM has a subjective belief regarding the auxiliary state space Θ is generally irrelevant given the utility specified in equation (3).[8] Second, we assume that the auxiliary states do not directly affect the payoff of the DM. This reflects our interpretation that the auxiliary states only directly affect the information content of an experiment, but do not affect outcomes of the DM's actions.[9]

The timing of the events is as follows: First, the DM makes an action plan $\sigma \in A_S$ and we assume the DM is dynamically consistent.[10] Second, an auxiliary state θ is drawn from Θ but *not* observed by the DM and signal realizations will be generated according to the Blackwell experiment $\mathbf{p}(\cdot, \theta)$. Third, a payoff relevant state ω is drawn from Ω and a signal realization s is drawn from S according to the distribution $\mathbf{p}(\cdot \mid \omega, \theta)$. Last, the DM observes the signal realization s and acts according to his action plan $\sigma(\cdot \mid s)$.

3.1 Classes of Ambiguity Preferences

Any decision maker will be identified with his ambiguity preference, which is fully captured by its corresponding monotone aggregator V. Let $\mathcal{V}_{Mono}$ denote the class of monotone

[8] We believe this is a reasonable first step to formalize the evaluation of ambiguous information. We will discuss the impact of allowing prior ambiguity on Ω and more general procedures to aggregate $u(\omega, a)$ into an ex-ante utility function in detail in Section 5.

[9] See Section 1 for more examples. The alternative modeling approach is to consider a mapping $u : \Omega \times \Theta \times A \to \mathbb{R}$ in which the auxiliary states have a direct impact on the DM's payoff for each action. This approach will be discussed in detail in Section 5.

[10] As is well known, dynamic inconsistency may arise with belief updating when ambiguity is present. For an example that illustrates the prevalence of violation of dynamic consistency in general dynamic ambiguity preference models, see Example 2 of Asano and Kojima (2019). Alternatively, we can assume the DM can commit to any action plan he makes.

aggregators, that is, $\mathcal{V}_{Mono} := \{V : \mathbb{R}^\Theta \to \mathbb{R} \mid V \text{ is monotone}\}$. $\mathcal{V}_{Mono}$ will be the largest (in terms of set inclusion) class of aggregators we study.

Some special cases of V are listed below to illustrate the generality of our approach to represent ambiguity preferences using monotone aggregators.

Subjective expected utility. V is specified by a single prior $\mu \in \Delta(\Theta)$, which represents the DM's subjective belief over Θ. Fixing μ, $V_\mu : \mathbb{R}^\Theta \to \mathbb{R}$ is defined by

$$V_\mu(f) := \int_\Theta f(\theta)\ d\mu(\theta) = \sum_{\theta\in\Theta} f(\theta)\mu(\theta), \tag{4}$$

and as in equation (3), we apply this aggregator to $f(\theta) = U\big(\sigma, \mathbf{p}(\cdot,\theta)\big)$.[11] Let $\mathcal{V}_{EU}$ denote the class of subjective expected utility aggregators, that is, $\mathcal{V}_{EU} := \{V_\mu \mid \mu \in \Delta(\Theta)\}$.

The multiple prior preferences, introduced in Gilboa and Schmeidler (1989). V is specified by a closed set of priors $M \subset \Delta(\Theta)$, which represents the set of priors the DM is willing to entertain. Fixing M, V_M is defined by

$$V_M(f) := \min_{\mu\in M} \int_\Theta f(\theta)\ d\mu(\theta) = \min_{\mu\in M} \sum_{\theta\in\Theta} f(\theta)\mu(\theta). \tag{5}$$

In the representation of Gilboa and Schmeidler (1989), M is required to be closed and convex. We only require the closedness of M to guarantee that the minimum is well-defined. Let $\mathcal{V}_{MP}$ denote the class of aggregators corresponding to the multiple prior preferences, that is, $\mathcal{V}_{MP} := \{V_M \mid M \subset \Delta(\Theta),\ M \text{ closed}\}$.

Wald's maximin criterion. A subclass of aggregators within $\mathcal{V}_{MP}$ can be described by the maximin criterion introduced in Wald (1950), where the DM is willing to entertain all degenerate beliefs δ_θ for each $\theta \in \Theta$. The aggregator V_W is defined by

$$V_W(f) := \min_{\mu\in\{\delta_\theta|\theta\in\Theta\}} \sum_{\theta\in\Theta} f(\theta)\mu(\theta) = \min_{\theta\in\Theta}\ f(\theta). \tag{6}$$

We will sometimes refer to V_W as the Wald aggregator, and unlike the other classes of aggregators considered in this section, $\{V_W\}$ is not a superset of $\mathcal{V}_{EU}$. We will analyze this subclass in detail in Section 4.

The smooth ambiguity preferences, introduced in Klibanoff, Marinacci, and Mukerji (2005). V is specified by a function $\phi : \mathbb{R} \to \mathbb{R}$ and a second-order belief

[11] By applying V_μ to equation (3), the DM behaves as if he has a belief over $\Omega \times \Theta$ that is independent across Ω and Θ. This sense of "independence" does not carry over to the other listed ambiguity preferences, since they cannot be summarized by a single belief over Θ.

$\nu \in \Delta(\Delta(\Theta))$.[12] Fixing ϕ and ν, $V_{\phi,\nu}$ is defined by

$$V_{\phi,\nu}(f) := \int_{\Delta(\Theta)} \phi\left(\int_{\Theta} f(\theta)\, d\mu(\theta)\right) d\nu(\mu) \tag{7}$$

In the representation of Klibanoff, Marinacci, and Mukerji (2005), ϕ is required to be strictly increasing and weakly concave, where the concavity captures the DM's aversion to ambiguity. We only require ϕ to be strictly increasing to guarantee that $V_{\phi,\nu}$ is monotone and mixed ambiguity attitudes could be allowed. Let $\mathcal{V}_S$ denote the class of smooth ambiguity preference aggregators, that is, $\mathcal{V}_S := \{V_{\phi,\nu} \mid \phi \text{ is strictly increasing}, \nu \in \Delta(\Delta(\Theta))\}$.

All classes of ambiguity preferences listed above are special cases of the uncertainty averse preferences introduced in Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011), which also corresponds to a monotone aggregator over the conditional expected utilities.[13]

3.2 Prior-by-Prior Dominance

The last piece of the puzzle for establishing the informativeness order over ambiguous experiment is a statistical condition that generalize the garbling condition in Blackwell's theorem. To get to that condition, it is useful to first define expected experiments.

Definition 7. Fixing an ambiguous experiment $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ and a probability measure $\mu \in \Delta(\Theta)$, the **expected experiment** with respect to μ , denoted by $\mathbf{p}_\mu$, is a Blackwell experiment defined by $\mathbf{p}_\mu := \sum_{\theta\in\Theta} \mathbf{p}(\cdot,\theta)\mu(\theta)$, or more explicitly

$$\mathbf{p}_\mu(s \mid \omega) := \sum_{\theta\in\Theta} \mathbf{p}(s \mid \omega,\theta)\mu(\theta),\ \forall(s,\omega) \in S \times \Omega. \tag{8}$$

Now we can formally state our statistical condition.

Definition 8. Let $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ and $\mathbf{p}' : \Omega \times \Theta \to \Delta(S')$ be two ambiguous experiments. We say **$\mathbf{p}$ prior-by-prior dominates $\mathbf{p}'$** if $\mathbf{p}_\mu$ is Blackwell more informative than $\mathbf{p}'_\mu$ for all $\mu \in \Delta(\Theta)$, and write $\mathbf{p} \trianglerighteq_{PBP} \mathbf{p}'$.

[12] $\Delta(\Delta(\Theta))$ is the set of all probability measures on the Borel σ-algebra of $\Delta(\Theta)$ under the Euclidean topology.

[13] The representation is in the form of $\min_{\mu\in\Delta(\Theta)} G(U(\mu),\mu)$ where $U(\mu)$ is an expected utility index for belief μ. Thus, their aggregator G takes two arguments as inputs, the conditional expected utility and the belief itself, while our aggregator V only takes in one. Therefore, strictly speaking, the uncertainty averse preference is not a special case of our set of monotone aggregators. However, it is indeed strictly increasing in its first argument, the expected utility conditional on μ. So our analysis and main result will go through for uncertainty averse preferences as well.

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

$$\mathbf{p}'_\mu = \gamma_\mu \circ \mathbf{p}_\mu \text{ for all } \mu \in \Delta(\Theta). \tag{9}$$

The "prior-by-prior" quantifier in the name of the condition does not refer to anything related to a decision problem or preference parameters. Each prior μ is just a probability measure over Θ and is not meant to be interpreted as anything more. In this sense, this prior $\mu \in \Delta(\Theta)$ is quite different from a payoff-relevant prior belief $\pi \in \Delta(\Omega)$. The latter is a preference parameter while the former is not.

The prior-by-prior dominance condition is our desired generalization of Blackwell's garbling condition since it is the *equivalent* condition for guaranteeing a weakly higher ex-ante utility in every decision problem for every DM. Intuitively, if $\mathbf{p} \trianglerighteq_{PBP} \mathbf{p}'$, then a DM can rank them according to the Blackwell order if his preference can be represented by an aggregator that corresponds to a single subjective belief over Θ. Having such a representation is not possible in general since there is ambiguity on Θ. Our result states that $\mathbf{p}$ and $\mathbf{p}'$ can still be ranked, independent of any decision problem, when more general ambiguity preferences are considered.

Theorem 2. *Let $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ and $\mathbf{p}' : \Omega \times \Theta \to \Delta(S')$ 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. $\mathbf{p}$ *prior-by-prior dominates* $\mathbf{p}'$.
2. *For any A, u, π and any action plan $\sigma' \in A_{S'}$, there exists $\sigma \in A_S$ such that*
$$U\big(\sigma, \mathbf{p}(\cdot,\theta)\big) \geq U\big(\sigma', \mathbf{p}'(\cdot,\theta)\big) \text{ for all } \theta \in \Theta.$$
3. $\mathbf{p}$ *is preferred to* $\mathbf{p}'$ *in every decision problem by every decision maker whose ambiguity preference can be represented by some $V \in \mathcal{V}$. That is,* $\mathbf{p}$ *gives weakly higher ex-ante utility than* $\mathbf{p}'$ *for every A, u, π, and every $V \in \mathcal{V}$.*

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.

Condition 1 is a statistical condition about the ambiguous experiments and independent of any decision problem or preference parameters, and condition 3 is an economical condition about the instrumental values of ambiguous experiments. Establishing the equivalence of conditions 1 and 3 helps us achieve the separation of the preferences and information structures like Blackwell's theorem.

Condition 2 states that for any action plan σ' made for $\mathbf{p}'$, there exists an action plan σ for $\mathbf{p}$ that guarantees a *weakly higher* expected utility in every auxiliary state. It is an intermediate condition for the clarification of the difference of Theorem 2 and Blackwell's theorem. It also highlights the importance of our use of monotone aggregators to represent ambiguity preferences since we only have weak inequality for the conditional expected utilities, unlike the equality we had in its counterpart in Blackwell's theorem (condition 2 of Theorem 1).

Remark. Prior-by-prior dominance is not implied by "state-by-state dominance" (which simply means that $\mathbf{p}(\cdot,\theta)$ is Blackwell more informative than $\mathbf{p}'(\cdot,\theta)$ for all $\theta \in \Theta$). To see this, consider the following example.

Example 4. Let $\Omega = \{\omega_1, \omega_2\}$, $S = S' = \{s_1, s_2\}$, and $\Theta = \{\theta_1, \theta_2\}$, and

$\mathbf{p}(\cdot,\theta_1) =$

	s_1	s_2
ω_1	1	0
ω_2	0	1

$\mathbf{p}(\cdot,\theta_2) =$

	s_1	s_2
ω_1	0	1
ω_2	1	0

$\mathbf{p}'(\cdot,\theta_1) =$

	s_1	s_2
ω_1	0.9	0.1
ω_2	0.1	0.9

$\mathbf{p}'(\cdot,\theta_2) =$

	s_1	s_2
ω_1	0.9	0.1
ω_2	0.1	0.9

Then $\mathbf{p}(\cdot,\theta_i)$ is strictly more informative than $\mathbf{p}'(\cdot,\theta_i)$ for $i \in \{1,2\}$, but for any belief with $\mu(\theta_1) \in (0.1, 0.9)$, $\mathbf{p}'_\mu$ is strictly more informative than $\mathbf{p}_\mu$.

Another condition that is closely related to prior-by-prior dominance is *global Blackwell dominance.*

Definition 9. Let $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ and $\mathbf{p}' : \Omega \times \Theta \to \Delta(S')$ be two ambiguous experiments. We say **$\mathbf{p}$ globally Blackwell dominates $\mathbf{p}'$** if there exists a garbling $\gamma : S \to \Delta(S')$ such that $\mathbf{p}'(\cdot,\theta) = \gamma \circ \mathbf{p}(\cdot,\theta)$ for all $\theta \in \Theta$, and write $\mathbf{p} \trianglerighteq_{GB} \mathbf{p}'$.

Given the linearity of the expectation operator, $\mathbf{p} \trianglerighteq_{GB} \mathbf{p}'$ if and only there exists a single garbling $\gamma : S \to \Delta(S')$ such that

$$\mathbf{p}'_\mu = \gamma \circ \mathbf{p}_\mu \text{ for all } \mu \in \Delta(\Theta). \tag{10}$$

Comparing equations (9) and (10), we can see intuitively that prior-by-prior dominance is a weaker condition than global Blackwell dominance: To satisfy prior-by-prior dominance, different μ and μ' in $\Delta(\Theta)$ can correspond to different garblings γ_μ and $\gamma_{\mu'}$, but to satisfy global Blackwell dominance, one garbling γ has to work uniformly across all μ. The following proposition formalizes this intuition.

Proposition 1. *Let* $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ *and* $\mathbf{p}' : \Omega \times \Theta \to \Delta(S')$ *be two ambiguous experiments. If* $\mathbf{p}$ *globally Blackwell dominates* $\mathbf{p}'$*, then* $\mathbf{p}$ *prior-by-prior dominates* $\mathbf{p}'$*. The converse is not true.*

Within our current framework, global Blackwell dominance is *not* our desired generalization of Blackwell's garbling condition since it is too strong and not necessary for always guaranteeing a weakly higher ex-ante utility. It will become more relevant when we discuss the case where u can directly depend on Θ in Section 5.

3.3 Connection and Differences with Blackwell's Theorem

Theorem 2 is a direct generalization of Blackwell's theorem. When Θ is a singleton, $\mathbf{p}$ and $\mathbf{p}'$ reduce to Blackwell experiments, and the prior-by-prior dominance condition reduces to the garbling condition. Prior-by-prior dominance is necessary within the small class of $\mathcal{V}_{EU}$ and it is sufficient within the large class of ambiguity preferences with $\mathcal{V}_{Mono}$. Therefore, prior-by-prior dominance is the equivalent condition corresponding to higher ex-ante utilities within the many classes of ambiguity preferences nested between them (e.g., the class of multiple prior preferences $\mathcal{V}_{MP}$ and the class of smooth preferences $\mathcal{V}_S$).

We take Blackwell's theorem as a starting point and build our result on it, but we do not replicate its proof and our result is not its trivial implication. Taking Blackwell's theorem as given, the necessity of our prior-by-prior dominance condition is relatively easy to prove: If $\mathbf{p}_\mu$ is not Blackwell more informative than $\mathbf{p}'_\mu$ for some μ, one can fix its corresponding expected utility aggregator V_μ and Blackwell's theorem guarantees the existence of some triplet (A, u, π) in which $\mathbf{p}'_\mu$ outperforms $\mathbf{p}_\mu$.

However, proving the sufficiency of our prior-by-prior dominance condition is more subtle and requires more work.[14] The main obstacle lies in the difference of condition 2 in

[14]This highlights another difference of our result and Blackwell's theorem in terms of their proof strategy. For Blackwell's theorem, it is easier to prove the sufficiency of the garbling condition than its necessity (for example, see the proofs of Crémer (1982) and de Oliveira (2018)). But for our result, it is the sufficiency of the prior-by-prior dominance condition that is relatively hard to prove.

Theorem 2 and its counterpart in Blackwell's theorem (condition 2 in Theorem 1). One quick way to prove the sufficiency of the garbling condition in Blackwell's theorem is to realize that if $p' = \gamma \circ p$, that is, if an unambiguous experiment p' is obtained by garbling p with γ, then for any action plan σ' made for p', one can obtain exactly the same expected utility under the experiment p by following the action plan $\sigma := \sigma' \circ \gamma$, since the property of compositions of stochastic operators guarantees that

$$U(\sigma, p) = U(\sigma', p'), \text{ for all } A, u \text{ and } \pi. \tag{11}$$

This proof method could work if we were trying to show that global Blackwell dominance is sufficient for always guaranteeing a weakly higher ex-ante utility. That is, if we assume there exists a garbling γ satisfying $\mathbf{p}_\mu = \gamma \circ \mathbf{p}'_\mu$ for all $\mu \in \Delta(\Theta)$, then for any A, u, π and action plan σ' for $\mathbf{p}'$, the action plan $\sigma := \sigma' \circ \gamma$ satisfies

$$U\big(\sigma, \mathbf{p}(\cdot, \theta)\big) = U\big(\sigma', \mathbf{p}'(\cdot, \theta)\big), \text{ for all } \theta \in \Theta. \tag{12}$$

This proof method does not work when we are trying to show that prior-by-prior dominance is sufficient for guaranteeing a weakly higher ex-ante utility, since the prior-by-prior dominance condition allows the garbling to depend on μ, and as argued before, it is much weaker than requiring $\mathbf{p}$ to globally Blackwell dominate $\mathbf{p}'$. Under this weaker requirement, it is generally impossible to simultaneously obtain the same conditional expected utility under $\mathbf{p}'$ and $\mathbf{p}$ in every auxiliary state θ like in equation (12), since there is a set of garblings and it is unclear which one should be applied to an action plan for $\mathbf{p}'$ to obtain a suitable action plan for $\mathbf{p}$.

With our assumption that the aggregator is monotone, proving the sufficiency of the prior-by-prior dominance condition does not require a condition as strong as equation (12). We can replace the equality in equation (12) with a weak inequality. This is the idea behind condition 2 in Theorem 2. Two key steps in proving condition 1 implies condition 2 are: (i) a careful construction of an auxiliary function capturing the differences in conditional expected utilities under $\mathbf{p}$ and $\mathbf{p}'$ and (ii) the use of a general minimax theorem to characterize the property of said auxiliary function.

3.4 Experiments with Independent Sources of Ambiguity

Theorem 2 can be applied to comparing ambiguous experiments with independent sources of ambiguity if we assume more structure to the auxiliary state space Θ. Formally, consider an auxiliary state space Θ that can be written as $\Theta_1 \times \Theta_2$ where Θ_1 and Θ_2 can be interpreted as two different aspects of the source of ambiguity. We say two ambiguous experiments $\mathbf{p} : \Omega \times \Theta_1 \times \Theta_2 \to \Delta(S)$ and $\mathbf{q} : \Omega \times \Theta_1 \times \Theta_2 \to \Delta(T)$ have independent sources of ambiguity if $\mathbf{p}$ does not depend on Θ_2 and $\mathbf{q}$ does not depend on Θ_1. More precisely, $\mathbf{p}$ and $\mathbf{q}$ have independent sources of ambiguity if there exist ambiguous experiments $\hat{\mathbf{p}} : \Omega \times \Theta_1 \to \Delta(S)$ and $\hat{\mathbf{q}} : \Omega \times \Theta_2 \to \Delta(T)$ satisfying

$$\mathbf{p}(\cdot, \theta_1, \theta_2) \equiv \hat{\mathbf{p}}(\cdot, \theta_1) \text{ for all } \theta_2 \in \Theta_2, \text{ and } \mathbf{q}(\cdot, \theta_1, \theta_2) \equiv \hat{\mathbf{q}}(\cdot, \theta_2) \text{ for all } \theta_1 \in \Theta_1.$$

We sometimes refer to $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$ as the reduced experiment of $\mathbf{p}$ and $\mathbf{q}$, respectively. With this additional structure, the prior-by-prior dominance condition relating $\mathbf{p}$ and $\mathbf{q}$ can be simplified as follows.

Lemma 1. *Given the above construction, $\mathbf{p} \trianglerighteq_{PBP} \mathbf{q}$ if and only if $\hat{\mathbf{p}}_\mu \trianglerighteq \hat{\mathbf{q}}_\nu$ for every pair of marginal distributions $(\mu, \nu) \in \Delta(\Theta_1) \times \Delta(\Theta_2)$.*

That is, $\mathbf{p}$ prior-by-prior dominates $\mathbf{q}$ if and only if their reduced experiments are related by the following condition: the expected experiment $\hat{\mathbf{p}}_\mu$ is Blackwell more informative than $\hat{\mathbf{q}}_\nu$ for any $\mu \in \Delta(\Theta_1)$ and $\nu \in \Delta(\Theta_2)$. When Θ_1 and Θ_2 are identical copies of the same space, this condition becomes another variation of the prior-by-prior dominance condition: To compare the informativeness of $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$, the DM's lack of understanding of the correlation of $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$ requires him to compare all possible pairs of expected experiments $(\hat{\mathbf{p}}_\mu, \hat{\mathbf{q}}_\nu)$ corresponding to potentially different beliefs μ and ν, in contrast to the prior-by-prior dominance condition where the comparison is made between expected experiments corresponding to the same belief.

Suppose $\Theta = \Theta_1 \times \Theta_2$. Let $\mathcal{V}_{EU}$ and $\mathcal{V}_{Mono}$ denote the class of expected utility aggregators and the class of monotone aggregators over Θ, respectively. We can combine Theorem 2 and Lemma 1 to have the following result regarding the comparison of experiments with independent sources of ambiguity.

Corollary 1. *Let $\mathbf{p}$, $\mathbf{q}$, $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$ be constructed as above. Suppose $\mathcal{V}$ is a class of aggregators such that $\mathcal{V}_{EU} \subset \mathcal{V} \subset \mathcal{V}_{Mono}$, then the following conditions are equivalent:*

1. *$\hat{\mathbf{p}}_\mu$ is Blackwell more informative than $\hat{\mathbf{q}}_\nu$ for all $(\mu, \nu) \in \Delta(\Theta_1) \times \Delta(\Theta_2)$.*

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

In particular, the class of aggregators stated in condition 2 involves those aggregators that only aggregate over the relevant aspect of the auxiliary state space for each experiment. For example, for any subjective expected utility aggregator $V_\eta \in \mathcal{V}_{EU}$, V_η aggegrating over $\Theta_1 \times \Theta_2$ for $\mathbf{p}$ is equivalent to an expected utility aggegator over Θ_1 for $\hat{\mathbf{p}}$. Similarly, V_η aggegrating over $\Theta_1 \times \Theta_2$ for $\mathbf{q}$ is equivalent to an expected utility aggegator over Θ_2 for $\hat{\mathbf{q}}$. In this sense, condition 2 also describes the comparison of ex-ante utilities for the reduced experiments $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$.

3.5 Value of Ambiguous Information

Theorem 2 can be also be used to study the value of ambiguous information.

Let $p_U : \Omega \to \Delta(S)$ be specified by $p_U(\omega) \equiv \eta$ for all $\omega \in \Omega$ for some $\eta \in \Delta(S)$, that is, p_U is an uninformative Blackwell experiment where each ω is mapped to the same distribution η over S. Blackwell's theorem indicates that any Blackwell experiment p is more informative than p_U. Therefore, the ambiguous experiment $\mathbf{p}_U$ defined by $\mathbf{p}_U(\cdot, \theta) \equiv p_U$ for all θ will satisfy $\mathbf{p} \trianglerighteq_{PBP} \mathbf{p}_U$ for any ambiguous experiment $\mathbf{p}$. Theorem 2 then implies that any ambiguous experiment $\mathbf{p}$ can guarantee a weakly higher ex-ante utility than $\mathbf{p}_U$ in every decision problem for every decision maker. In other words, if we define the value of ambiguous information $\mathbf{p}$ as the difference in ex-ante utilities with and without it, then this value is always non-negative.

This result is analogous to the fact that unambiguous information always carries non-negative instrumental value for Bayesian expected utility maximizers. In this sense, if one perceives Blackwell's theorem as an intuitive and important decision-theoretic feature within the class of expected utility maximizers, one should also recognize our prior-by-prior dominance condition as a plausible feature for decision-theoretic frameworks studying ambiguous information.

4 Comparing Sets of Experiments

In this section, we study a special case where the comparison is between sets of Blackwell experiments and a DM's ex-ante utility is computed according to the maximin criterion. It is special in two ways. First, the auxiliary state space Θ is not explicitly needed at the outset for the comparison. Second, we focus our attention on the class of ambiguity preferences corresponding to the maximin criterion. Theorem 2 does not apply in this special case since this class of ambiguity preferences does not nest expected utility, and we will have an informativeness order that is weaker than prior-by-prior dominance. The result is stated in Section 4.1, and we clarify the connection of the special case with the general comparison of ambiguous experiments in Section 4.2. The differences between the special case we considered and the model studied in Gensbittel, Renou, and Tomala (2015) are discussed in Section 4.3.

4.1 Result

A set of Blackwell experiments is summarized as a pair (S, P) where S is a finite set of signal realizations and P is a set of Blackwell experiments where each $p \in P$ is a mapping from Ω to $\Delta(S)$. Note that we no longer have the auxiliary state space Θ as a primitive. P should be interpreted as all unambiguous experiments that are deemed possible by the DM. We identify P as a subset of $\mathbb{R}^{|\Omega|\times|S|}$ and assume it to be closed under the Euclidean topology.[15] P can be uncountable.

Consider an individual with a finite set of actions A, a state-dependent utility function $u : \Omega \times A \to \mathbb{R}$ and a prior $\pi \in \Delta(\Omega)$. We define the individual's **ex-ante maximin utility** from a set of Blackwell experiments P to be

$$\max_{\sigma \in A_S} \min_{p \in P} U(\sigma, p), \tag{13}$$

where U is defined as in equation (1). That is, when evaluating an action plan σ for P, the DM uses the Blackwell experiment in P that gives the lowest expected utility.

Let $\text{conv}(P)$ denote the convex hull of P.[16]

Theorem 3. *Fix two sets of experiments P and P'. The following are equivalent:*

1. *For any Blackwell experiment $p \in \text{conv}(P)$, there exists another Blackwell experiment $p' \in \text{conv}(P')$ such that p is Blackwell more informative than p'.*

2. *Every decision maker who applies Wald's maximin criterion prefers P to P' for any possible decision problem. That is, P gives weakly higher ex-ante maximin utility than P' for every A, u, and π.*

[15] P is a bounded set since for every element $p \in P$, every entry of p is bounded in $[0, 1]$. Hence its closedness implies its compactness.

[16] The convex combination is taken component-wise, that is, for any $p, q \in P$, $tp + (1-t)q : \Omega \to \Delta(S)$ is defined by $(tp + (1-t)q)(s \mid \omega) = tp(s \mid \omega) + (1-t)q(s \mid \omega)$.

If any of the conditions in Theorem 3 holds, we say P is W-more informative than P' and write $P \trianglerighteq_W P'$. To better compare $\trianglerighteq_W$ and $\trianglerighteq_{PBP}$, it is useful to define expected experiments for a set of Blackwell experiment. For any nonempty set Y, let $\Delta_0(Y)$ denote the set of probability measures over Y with finite support. Let (S, P) be a set of Blackwell experiments and $\nu \in \Delta_0(P)$. Then the **expected experiment** with respect to ν, denoted by $P_\nu : \Omega \to \Delta(S)$, is defined by $P_\nu := \sum_{p \in P} p\nu(p)$. With this notion, condition 1 can be rephrased as

1'. *For all $\nu \in \Delta_0(P)$, there exists some $\lambda \in \Delta_0(P')$ such that the expected experiment P_ν is Blackwell more informative than P'_λ.*

Comparing condition 1' and the prior-by-prior dominance condition, we can see intuitively that the W-more informative condition is somewhat less restrictive than the prior-by-prior dominance condition as the latter requires the Blackwell order to hold for every pair of expected experiments corresponding to the same $\mu \in \Delta(\Theta)$. We obtain this less restrictive condition mainly because we have a more restrictive class of preferences. DMs who apply the maximin criterion essentially go through all possible priors over Θ and then focus only to the worst possible ones, and thus ignoring the effect of other priors.

4.2 Connection with Ambiguous Experiments

Ambiguous experiments and sets of Blackwell experiments are different primitives. However, there is a natural correspondence between them and they are essentially equivalent when we focus our attention to DMs who use Wald's maximin criterion.

To see the correspondence, suppose we have two sets of Blackwell experiments (S, P) and (S', P'), then we can construct an auxiliary state space Θ, and two ambiguous experiments $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ and $\mathbf{p}' : \Omega \times \Theta \to \Delta(S')$ corresponding to (S, P) and (S', P'), respectively, in the following way:

$$\begin{aligned} \Theta &:= P \times P' \\ \mathbf{p}(s \mid \omega, p, p') &:= p(s \mid \omega),\ \forall (s, \omega, p, p') \in S \times \Omega \times P \times P';\ \text{and} \\ \mathbf{p}'(s' \mid \omega, p, p') &:= p'(s' \mid \omega),\ \forall (s', \omega, p, p') \in S' \times \Omega \times P \times P' \end{aligned} \tag{14}$$

That is, the situation with two sets of Blackwell experiments can be re-interpreted as following: the DM wants to compare P and P' but has no additional knowledge on how their elements are correlated. He therefore creates $\Theta = P \times P'$ as the auxiliary state space, as if he is willing to entertain each and every pair of the elements $(p, p') \in P \times P'$ to be the actual Blackwell experiments he will be facing.

Now suppose we have two ambiguous experiments $\mathbf{p}$ and $\mathbf{p}'$, then we can construct two sets of Blackwell experiments

$$P := \{\mathbf{p}(\cdot, \theta) \mid \theta \in \Theta\} \quad P' := \{\mathbf{p}'(\cdot, \theta) \mid \theta \in \Theta\} \tag{15}$$

That is, each set of Blackwell experiments is created by collecting the Blackwell experiments in each auxiliary state in its corresponding ambiguous experiment.

Given these constructions in (14) and (15), comparisons of sets of Blackwell experiments using maximin criterion are equivalent to comparisons of ambiguous experiments for DMs with the Wald aggregator V_W.[17]

Proposition 2. *If $\mathbf{p}$, $\mathbf{p}'$, P and P' satisfies one of the constructions given by (14) or (15) with P and P' being finite,[18] then for any A, u, π, and any action plans σ and σ',*

$$V_W\left(\left\{U\left(\sigma, \mathbf{p}(\cdot, \theta)\right)\right\}_{\theta \in \Theta}\right) = \min_{p \in P} U(\sigma, p), \qquad V_W\left(\left\{U\left(\sigma', \mathbf{p}'(\cdot, \theta)\right)\right\}_{\theta \in \Theta}\right) = \min_{p' \in P'} U(\sigma', p').$$

In general, comparisons of sets of Blackwell experiments is only a special case of comparisons of ambiguous experiments, and the resulting informativeness order induced by the W-more informative condition is different with that induced by the prior-by-prior dominance condition. This is because the comparisons of sets of Blackwell experiments are only possible for the specific class of decision makers who uses the maximin criterion. This class of decision makers corresponds to the Wald aggregator V_W, but $\{V_W\}$ is not a superset of $\mathcal{V}_{EU}$, that is, the class of expected utility aggregators is not a subclass of the Wald aggregator. Our argument to get the necessity of prior-by-prior dominance in the general model is through its necessity for the class of expected utility aggregators. Now that $\{V_W\}$ no longer includes $\mathcal{V}_{EU}$, the prior-by-prior dominance is no longer necessary. The following proposition states that prior-by-prior dominance is a stronger requirement than being W-more informative.

Proposition 3. *Let P and P' be two finite sets of Blackwell experiments and define $\mathbf{p}$ and $\mathbf{p}'$ as in (14), then $\mathbf{p} \trianglerighteq_{PBP} \mathbf{p}'$ implies $P \trianglerighteq_W P'$, but the converse is not true.*

[17] Recall that $V_W : \mathbb{R}^\Theta \to \mathbb{R}$ is defined by $V_W(f) = \min_{\theta \in \Theta} f(\theta)$.

[18] The finiteness assumption in Proposition 2 is only for ease of exposition, a more general proposition (Proposition 4) without this assumption is stated and proved in the Appendix.

Modeling ambiguous experiments as sets of unambiguous experiments has its advantages: The framework for Blackwell's theorem can be applied without much modification. But such advantages come with costs tightly connected with the maximin criterion: DMs' potential subjective belief over possible unambiguous experiments are completely ignored due to the somewhat extreme ambiguity attitude. Even though the result for comparisons of sets of experiments is interesting in itself, we may naturally also be interested in environments that can allow for richer ambiguity preferences.

Figure 1 summarizes the results we have so far.

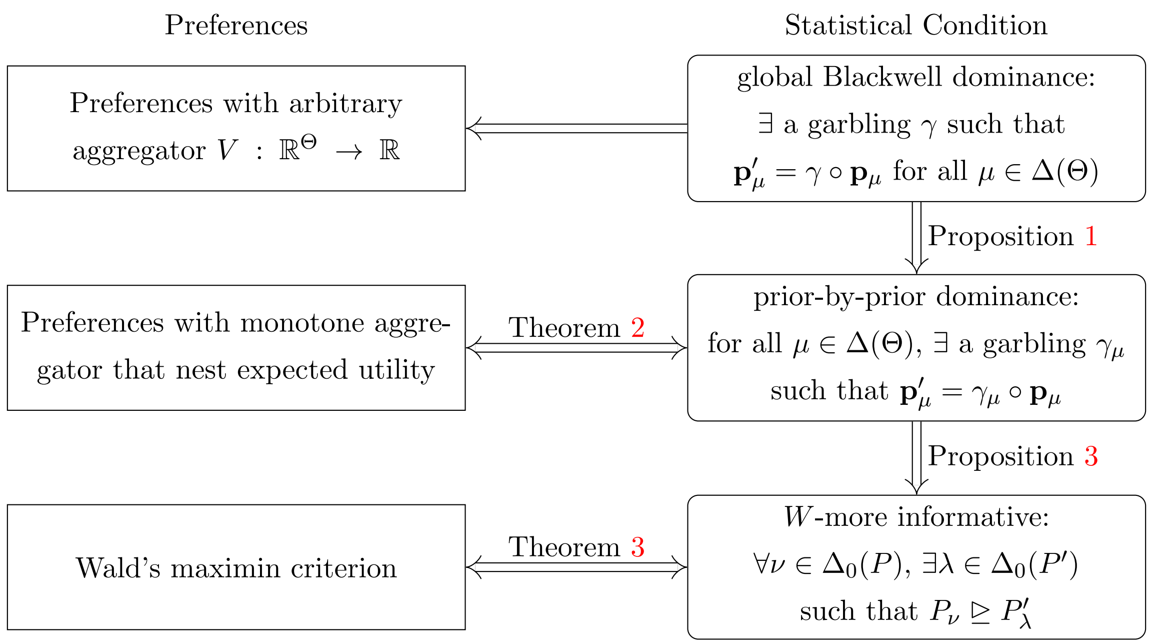

Figure 1: Summary of informativeness orders.

4.3 Connection with the Literature

Our Theorem 3 is closely related to an existing result in Gensbittel, Renou, and Tomala (2015) (henceforth GRT). Although arriving at seemingly similar results, there are subtle but important differences between our model and result for comparisons of sets of experiments (W-more informative) and the model and result presented in GRT due to a major methodological difference: We model Blackwell experiments as collections of conditional distributions $p : \Omega \to \Delta(S)$, while GRT model them as joint distributions $k \in \Delta(S \times \Omega)$ over the product space of the payoff-relevant states and the signal realizations. These two approaches are equivalent in the framework of Blackwell's theorem. However, these two approaches are no longer equivalent when ambiguous experiments are studied.

GRT model ambiguous experiments as *compact* and *convex* subsets of $\Delta(S \times \Omega)$. The benefit of this modeling approach is that one does not need to specify the source of ambiguity. That is, this method can deal with two special cases: An unambiguous prior combined with a set of Blackwell experiments (in our sense, i.e., collections of conditional distributions), or a set of priors combined with a single Blackwell experiment. To get this benefit, however, the inevitable cost is that the DM's prior over Ω is no longer a part of the DM's preference parameters, but a part of the experiment. For example, we can fix a Blackwell experiment $p : \Omega \to \Delta(S)$ and pair them with two different (compact and convex) sets of priors, $\Pi_1 \subsetneq \Pi_2$ on Ω. Then we can find examples of Π_1 and Π_2 such that the set of joint distributions induced from a Blackwell experiment p and Π_1 is deemed "more informative than" the set of joint distributions induced from p and Π_2 by GRT's criterion. For a simple example, consider a decision problem with a single action $A = \{a\}$. Then a larger set of priors necessarily result in lower payoff as the minimum is taken over a larger set. Thus, their criterion compares the informativeness and ambiguity associated with the experiment, but also the ambiguity of the specified set of priors on Ω.

Therefore, our criteria differ, and both are useful in different circumstances: Our criterion is more useful when the modeler cannot observe the decision maker's prior over Ω, while GRT's criterion is more suitable when those priors can be observed.

5 Discussion and Extensions

In this section, we consider relaxing two of the assumptions we previously made. The first is that u does not depend on Θ, that is, the auxiliary state does not directly affect the payoff of the DM. And the second is that there is no prior ambiguity, that is, the decision problems we consider all involve unambiguous prior over Ω. Table 1 is a roadmap for our exercises on relaxing these two assumptions.

	Section 3	5.1	5.2
Dependence of u on Θ	No	Yes	No
Prior ambiguity on Ω	No	Yes/No	Yes
Informativeness order	prior-by-prior dominance	global Blackwell	unclear

Table 1: Roadmap of extensions.

5.1 Direct Dependence of Payoffs on Θ

In this section, we consider individuals whose payoff may directly depend on the auxiliary state space Θ. That is, the individual's state dependent utility function is $u : \Omega \times \Theta \times A \to \mathbb{R}$. With this change in setup, an ambiguous experiment $\mathbf{p}$ is preferred to $\mathbf{p}'$ by every DM for every decision problem if and only if $\mathbf{p}$ globally Blackwell dominates $\mathbf{p}'$, which is equivalent to assuming that $\mathbf{p}$ is Blackwell more informative than $\mathbf{p}'$ viewing $\Omega \times \Theta$ as the payoff-relevant state space.

To see why this is the case, note that once the DM's payoff u depends directly on Θ, we can expand the payoff-relevant state space from Ω to $\Omega \times \Theta$. With this expansion, the ambiguity previously in the information structures is transformed into the ambiguity in the priors over Θ. That is, the decision maker faces a Blackwell experiment although there is ambiguity in his payoff-relevant prior. This problem is the center of the study in Li and Zhou (2016), where they show that Blackwell order continues to be valid when the decision makers possess uncertainty averse preferences to deal with the ambiguity in the payoff relevant space.[19] The intuition behind this result is as follows. If one experiment $\mathbf{p}$ globally Blackwell dominates another $\mathbf{p}'$, then any probability distribution over the actions *conditional* on the states ω and θ, $\lambda(a \mid \omega, \theta)$, that can be induced from some action plan for $\mathbf{p}'$ can be replicated under $\mathbf{p}$ by applying the garbling that transforms $\mathbf{p}$ to $\mathbf{p}'$, and this replication does not depend on the prior over the payoff relevant state space. Therefore, as long as the payoffs $u(\omega, \theta, a)$ are aggregated through some aggregator that respects monotonicity (e.g., the aggregator for the uncertainty averse preferences), the Blackwell informativeness order will be valid.

5.2 Prior Ambiguity

In this section, we consider decision scenarios in which prior ambiguity (i.e., ambiguity over Ω) is present. As the case where u depends on Θ and prior ambiguity over Ω are both present has already been covered in Section 5.1, we focus on the case where u does not depend directly on Θ in this section.

Formally, we consider an individual with a finite set of actions A, a state-dependent utility function $u : \Omega \times A \to \mathbb{R}$ and a monotone aggregator $\hat{V} : \mathbb{R}^{\Omega \times \Theta} \to \mathbb{R}$ that captures his ambiguity attitude. Notice that $\hat{V}$ is more general than the aggregators V we considered

[19] See Theorem 1 of Li and Zhou (2016) for more details.

in Section 3 since $\hat{V}$ is aggregating over functions in $\mathbb{R}^{\Omega\times\Theta}$ while V is only aggregating over functions in $\mathbb{R}^{\Theta}$.

For such an individual, the expected utility conditional on θ and ω for ambiguous experiment $\mathbf{p}$ and action plan σ is

$$\hat{U}(\sigma, \mathbf{p}(\cdot,\theta), \omega) := \sum_{s\in S}\sum_{a\in A} \mathbf{p}(s \mid \omega, \theta)\sigma(a \mid s)u(\omega, a) \tag{16}$$

The **ex-ante utility** for the individual is

$$\max_{\sigma\in A_S} \hat{V}\left(\left\{\hat{U}\big(\sigma, \mathbf{p}(\cdot,\theta), \omega\big)\right\}_{(\theta,\omega)\in\Theta\times\Omega}\right) \tag{17}$$

With this new formulation involving more general aggregators $\hat{V}$, the payoffs $u(\omega, a)$ can be aggregated in a more general way comparing to equation (3). For example, it can nest the case where the DM has a set of priors over Ω, first aggregating each $u(\omega, a)$ according to the multiple prior preference conditional on each θ and then aggregating the conditional utilities into the ex-ante utility.

The following proposition illustrates the results when prior ambiguity is present.

Proposition 4. *Let* $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ *and* $\mathbf{p}' : \Omega \times \Theta \to \Delta(S')$ *be two ambiguous experiments. Consider the following 4 conditions:*

1. $\mathbf{p}$ *globally Blackwell dominates* $\mathbf{p}'$*, that is,* $\mathbf{p} \trianglerighteq_{GB} \mathbf{p}'$.
2. *For any* A*,* u *and any* $\sigma' \in A_{S'}$*, there exists* $\sigma \in A_S$ *such that*
$$\hat{U}(\sigma, \mathbf{p}(\cdot,\theta), \omega) \geq \hat{U}(\sigma', \mathbf{p}'(\cdot,\theta), \omega),\ \forall(\theta, \omega) \in \Theta \times \Omega.$$
3. $\mathbf{p}$ *gives a weakly higher ex-ante utility than* $\mathbf{p}'$ *for every* A*,* u *and* $\hat{V}$.
4. $\mathbf{p}$ *prior-by-prior dominates* $\mathbf{p}'$*, that is,* $\mathbf{p} \trianglerighteq_{PBP} \mathbf{p}'$.

Then, $1 \implies 2 \iff 3 \implies 4$.

That is, in terms of guaranteeing higher ex-ante utility for every decision maker (condition 3), global Blackwell dominance is sufficient and prior-by-prior dominance is necessary. Although condition 2 is equivalent to condition 3, it is less ideal than what we desire as an equivalence condition because it is a somewhat high-order condition imposed on all possible A and u.

The following example illustrates that when prior ambiguity is present, global Blackwell dominance is not necessary for guaranteeing higher ex-ante utility. That is, condition 3 does not imply condition 1.

Example 5. Consider $\Omega = \{\omega_1, \omega_2\}$, $\Theta = \{\theta_1, \theta_2\}$, $S = S' = \{s_1, s_2\}$, with $\mathbf{p}$ and $\mathbf{p}'$ defined as follows:

$$\mathbf{p}(\cdot,\theta_1) = \begin{array}{|c|c|c|}\hline & s_1 & s_2 \\ \hline \omega_1 & 1 & 0 \\ \hline \omega_2 & 0 & 1 \\ \hline \end{array} \qquad \mathbf{p}(\cdot,\theta_2) = \begin{array}{|c|c|c|}\hline & s_1 & s_2 \\ \hline \omega_1 & 1 & 0 \\ \hline \omega_2 & 0 & 1 \\ \hline \end{array}$$

$$\mathbf{p}'(\cdot,\theta_1) = \begin{array}{|c|c|c|}\hline & s_1 & s_2 \\ \hline \omega_1 & 1 & 0 \\ \hline \omega_2 & 0 & 1 \\ \hline \end{array} \qquad \mathbf{p}'(\cdot,\theta_2) = \begin{array}{|c|c|c|}\hline & s_1 & s_2 \\ \hline \omega_1 & 0 & 1 \\ \hline \omega_2 & 1 & 0 \\ \hline \end{array}$$

That is, $\mathbf{p}$ is the unambiguous fully revealing experiment while $\mathbf{p}'$ is fully revealing in both θ_1 and θ_2 but sends opposite signals in different auxiliary states. It is clear that $\mathbf{p}$ prior-by-prior dominates $\mathbf{p}'$ but $\mathbf{p}$ does not globally Blackwell dominates $\mathbf{p}'$.

Proposition 5. *As constructed in Example 5,* $\mathbf{p}$ *gives weakly higher ex-ante utility than* $\mathbf{p}'$ *for every A, u, and $\hat{V}$.*

6 Conclusion

In this paper, we study comparisons of ambiguous experiments and establish informativeness orders over ambiguous experiments as generalizations of Blackwell's theorem. For general ambiguous experiments modeled as mappings from an auxiliary state space to the space of unambiguous experiments, the informativeness order is induced by the prior-by-prior dominance condition. One ambiguous experiment $\mathbf{p}$ prior-by-prior dominates another $\mathbf{p}'$ if their expected experiments satisfy $\mathbf{p}_\mu$ being Blackwell more informative than $\mathbf{p}'_\mu$ for every possible belief μ over the auxiliary state space. This informativeness order is robust among any monotone ambiguity preference that nests expected utility. For the special case of comparing sets of unambiguous experiments evaluated by the maxmin criterion, the informativeness order is characterized by being Wald-more informative. One set of unambiguous experiment P is Wald-more informative than another P' if for every unambiguous experiment in the convex hull of P, there exists an unambiguous experiment in the convex hull of P' that is Blackwell less informative.

An interesting question that remains open is whether or not prior-by-prior dominance is sufficient for guaranteeing higher ex-ante utility when ambiguity on Ω is present. Another potentially fruitful avenue for future research is to study the impact of relaxing the dynamic consistency assumption within the class of ambiguity preferences more general than the maxmin expected utility.

Appendices

A General Auxiliary State Space

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 the same way as in Definition 4, 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 signal realizations S are still assumed to be finite.

Consider an individual with a finite set of actions A, 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** state θ, $U\big(\sigma, \mathbf{p}(\cdot,\theta)\big)$, is defined the same way as in equation (2):

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

Note that U is bounded as a function of θ for any fixed $(\mathbf{p}, A, u, \pi, \sigma)$.

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 5, we drop the continuity requirement on the aggregator. Let $\mathcal{V}_{Mono}$ denote the set of all monotone aggregators.

Let 2^Θ denote the set of all subsets of Θ. Let Δ denote the set of all finitely additive probability measures over 2^Θ. Fixing any $\mu \in \Delta$, let V_μ denote the aggregator that corresponds to the (subjective) expected utility index, that is,

$$V_\mu\left(\big\{U\big(\sigma, \mathbf{p}(\cdot,\theta)\big)\big\}_{\theta\in\Theta}\right) := \int_\Theta U\big(\sigma, \mathbf{p}(\cdot,\theta)\big)\ d\mu(\theta). \tag{18}$$

The integral is well-defined as $U\big(\sigma, \mathbf{p}(\cdot,\theta)\big)$ is a bounded measurable function and μ is a finitely additive measure (Aliprantis and Border, 2006, Theorem 11.8). Let $\mathcal{V}_{SEU}$ denote the class of all aggregators that corresponds to some (subjective) expected utility index, that is, $\mathcal{V}_{SEU} := \{V_\mu : \mathbb{R}^\Theta \to \mathbb{R} \mid \mu \in \Delta\}$. When Θ is uncountable, $\mathcal{V}_{SEU}$ includes the set of all non-atomic (finitely additive) probability measures, and thus includes the standard subjective expected utility model à la Savage (1954).

As we no longer require any continuity of V, we slightly modify the definition of the ex-ante utility for an ambiguous experiment. Fixing an ambiguous experiment $\mathbf{p} : \Omega \times \Theta \to$

$\Delta(S)$ and (A, u, π, V), the individual's **ex-ante utility** is

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

Comparing to the definition of the ex-ante utility in the main text, we have the supremum operator instead of the maximum operator. This change has no impact on the interpretation of our comparison result: If $\mathbf{p}$ gives higher ex-ante utility than $\mathbf{p}'$, then for any action plan σ' that can be made when the DM faces $\mathbf{p}'$, he can find another action plan σ that guarantees

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

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

For any $\mu \in \Delta$, the **expected experiment** with respect to μ, $\mathbf{p}_\mu : \Omega \to \Delta(S)$, is a Blackwell experiment defined similarly as in Definition 7 by

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

Let Δ_0 denote the set of all simple probability measures (i.e., probability measures with finite supports) over 2^Θ. For any $\mu \in \Delta_0$, the expected experiment $\mathbf{p}_\mu$ is simply

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

We say that an ambiguous experiment $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ **prior-by-prior dominates** another $\mathbf{p}' : \Omega \times \Theta \to \Delta(S')$ if $\mathbf{p}_\mu$ is Blackwell more informative than $\mathbf{p}'_\mu$ 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 8). Finally, let $\mathcal{V}_0 := \{V_\mu \mid \mu \in \Delta_0\}$ where $V_\mu : \mathbb{R}^\Theta \to \mathbb{R}$ is defined as in equation (18).

Theorem 4. *Let $p : \Omega \times \Theta \to \Delta(S)$ and $p' : \Omega \times \Theta \to \Delta(S')$ be two ambiguous experiments and let $\mathcal{V}$ be a class of aggregators such that $\mathcal{V}_0 \subset \mathcal{V} \subset \mathcal{V}_{Mono}$, then the following statements are equivalent:*

1. $\mathbf{p}$ *prior-by-prior dominates* $\mathbf{p}'$.

2. *For any A, u, π and any $\sigma' \in A_{S'}$, there exists $\sigma \in A_S$ such that*

$$U\big(\sigma, \mathbf{p}(\cdot, \theta)\big) \geq U\big(\sigma', \mathbf{p}'(\cdot, \theta)\big) \text{ for all } \theta \in \Theta.$$

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

When Θ is finite, $\mathcal{V}_0$ reduces to $\mathcal{V}_{EU}$ as defined in equation (3.1) in the main text and Theorem 4 coincides with Theorem 2.

In addition to this characterization result, we also want to establish the equivalence of comparisons of ambiguous experiments and comparisons of sets of Blackwell experiments in this more general environment. Consider the Wald aggregator in this more general setting. Formally, let $V_W : \mathbb{R}^\Theta \to \mathbb{R}$ be re-defined by

$$V_W(f) := \inf_{\theta \in \Theta} f(\theta), \tag{19}$$

and we apply this aggregator to $f(\theta) = U\big(\sigma, \mathbf{p}(\cdot, \theta)\big)$. Let sets of experiments be defined in the same way as in Section 4.1, that is, a pair (S, P) where S is a finite set of signal realizations and P is a closed set (under the Euclidean topology) of Blackwell experiments, and P could be uncountable. Then we have the following proposition corresponding to Proposition 2 in the main text.

Proposition 6. *For any two sets of Blackwell experiments* (S, P) *and* (S', P')*, we can construct an auxiliary state space* Θ*, and two ambiguous experiments* $\mathbf{p} : \Omega \times \Theta \to \Delta(S)$ *and* $\mathbf{p}' : \Omega \times \Theta \to \Delta(S')$ *as in (14). Then for all* A*,* u*,* π*,* σ *and* σ'*:*

$$V_W\left(\left\{U\big(\sigma, \mathbf{p}(\cdot, \theta)\big)\right\}_{\theta \in \Theta}\right) = \min_{p \in P} \; U(\sigma, p), \;\; V_W\left(\left\{U\big(\sigma', \mathbf{p}'(\cdot, \theta)\big)\right\}_{\theta \in \Theta}\right) = \min_{p' \in P'} \; U(\sigma', p')$$

where V_W *is the Wald aggregator as defined in equation (19), and* $U\big(\sigma, \mathbf{p}(\cdot, \theta)\big)$ *is the conditional expected utility defined in equation (2).*

B Proofs

B.1 Proof of Proposition 1

Proof. Suppose $\mathbf{p}$ globally Blackwell dominates $\mathbf{p}'$, then there exists a garbling $\gamma : S \to \Delta(S')$ such that $\mathbf{p}'(\cdot, \theta) = \gamma \circ \mathbf{p}(\cdot, \theta)$ for all $\theta \in \Theta$. Then we must have $\mathbf{p}'_\mu = \gamma \circ \mathbf{p}_\mu$ for any $\mu \in \Delta(\Theta)$, since for any $(s', \omega) \in S' \times \Omega$,

$$\begin{aligned}
\mathbf{p}'_\mu(s' \mid \omega) &= \sum_{\theta \in \Theta} \mathbf{p}'(s' \mid \omega, \theta)\mu(\theta) \\
&= \sum_{\theta \in \Theta} \sum_{s \in S} \gamma(s' \mid s)\mathbf{p}(s \mid \omega, \theta)\mu(\theta) \\
&= \sum_{s \in S} \gamma(s' \mid s) \sum_{\theta \in \Theta} \mathbf{p}(s \mid \omega, \theta)\mu(\theta) = \sum_{s \in S} \gamma(s' \mid s)\mathbf{p}_\mu(s \mid \omega)
\end{aligned}$$

To prove that the converse is not true, consider the following example:

$\Omega = \{\omega_1, \omega_2\}$, $\Theta = \{\theta_1, \theta_2\}$, $S = \{s_1, s_2\}, S' = \{s'_1, s'_2\}$, and

$\mathbf{p}(\cdot, \theta_1) =$

	s_1	s_2
ω_1	1	0
ω_2	0	1

$\mathbf{p}(\cdot, \theta_2) =$

	s_1	s_2
ω_1	0.8	0.2
ω_2	0.2	0.8

$\mathbf{p}'(\cdot, \theta_1) =$

	s'_1	s'_2
ω_1	0.8	0.2
ω_2	0.2	0.8

$\mathbf{p}'(\cdot, \theta_2) =$

	s'_1	s'_2
ω_1	0.8	0.2
ω_2	0.2	0.8

Then any $\mu \in \Delta(\Theta)$ can be represented by one parameter $k \in [0, 1]$, and the expected experiments are given by $\mathbf{p}_k = k\mathbf{p}(\cdot, \theta_1) + (1 - k)\mathbf{p}(\cdot, \theta_2)$. Thus,

$\mathbf{p}_k =$

	s_1	s_2
ω_1	$0.8 + 0.2k$	$0.2 - 0.2k$
ω_2	$0.2 - 0.2k$	$0.8 + 0.2k$

and $\mathbf{p}'_k =$

	s'_1	s'_2
ω_1	0.8	0.2
ω_2	0.2	0.8

Hence $\mathbf{p}'_k = \gamma_k \circ \mathbf{p}_k$ for any $k \in [0, 1]$ and the corresponding garbling γ_k for each k is

$\gamma_k =$

	s'_1	s'_2
s_1	$\frac{3+k}{3+2k}$	$\frac{k}{3+2k}$
s_2	$\frac{k}{3+2k}$	$\frac{3+k}{3+2k}$

$\mathbf{p}$ does not globally Blackwell dominates $\mathbf{p}'$ since γ_k varies with the belief k. $\square$

B.2 Proof of Lemma 1

Proof of Lemma 1. $\mathbf{p}$ prior-by-prior dominates $\mathbf{q}$ if $\mathbf{p}_\eta$ is Blackwell more informative than $\mathbf{q}_\eta$ for every $\eta \in \Delta(\Theta_1 \times \Theta_2)$. Fix some $\eta \in \Delta(\Theta_1 \times \Theta_2)$ and let η_1 and η_2 be the marginal distributions over Θ_1 and Θ_2 induced by η, respectively.

$$\mathbf{p}_\eta = \sum_{(\theta_1,\theta_2)\in\Theta_1\times\Theta_2} \eta(\theta_1,\theta_2)\mathbf{p}(\cdot,\theta_1,\theta_2) = \sum_{\theta_1\in\Theta_1}\left(\sum_{\theta_2\in\Theta_2}\eta(\theta_1,\theta_2)\right)\hat{\mathbf{p}}(\cdot,\theta_1) = \hat{\mathbf{p}}_{\eta_1}$$
$$\mathbf{q}_\eta = \sum_{(\theta_1,\theta_2)\in\Theta_1\times\Theta_2} \eta(\theta_1,\theta_2)\mathbf{q}(\cdot,\theta_1,\theta_2) = \sum_{\theta_2\in\Theta_2}\left(\sum_{\theta_1\in\Theta_1}\eta(\theta_1,\theta_2)\right)\hat{\mathbf{q}}(\cdot,\theta_2) = \hat{\mathbf{q}}_{\eta_2}$$

For the if direction, suppose $\mathbf{p}$ does not prior-by-prior dominates $\mathbf{q}$, then there exists $\eta \in \Delta(\Theta_1 \times \Theta_2)$ such that $\mathbf{p}_\eta$ is not Blackwell more informative than $\mathbf{q}_\eta$. This gives us a pair of marginal distributions $(\eta_1, \eta_2) \in \Delta(\Theta_1) \times \Delta(\Theta_2)$ such that $\hat{\mathbf{p}}_{\eta_1}$ is not Blackwell more informative than $\hat{\mathbf{q}}_{\eta_2}$. Taking the contrapositive completes the proof.

For the only if direction, suppose there exists some pair of marginal distributions $(\mu, \nu) \in \Delta(\Theta_1) \times \Delta(\Theta_2)$ such that $\hat{\mathbf{p}}_\mu$ is not Blackwell more informative than $\hat{\mathbf{q}}_\nu$, then we can construct a joint distribution η that induces (μ, ν) such that $\mathbf{p}_\eta$ is not Blackwell more informative than $\mathbf{q}_\eta$, making it impossible for $\mathbf{p}$ to prior-by-prior dominate $\mathbf{q}$. Taking the contrapositive completes the proof. □

B.3 Proof of Theorem 3

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

$$\begin{aligned}
\max_{\sigma\in A_S}\min_{p\in P} U(\sigma,p) &= \max_{\sigma\in A_S}\min_{p\in\mathrm{conv}(P)} U(\sigma,p) \\
&= \min_{p\in\mathrm{conv}(P)}\max_{\sigma\in A_S} U(\sigma,p) && \text{(minimax theorem)} \\
&= \max_{\sigma\in A_S} U(\sigma,p_*) && \text{(for some } p_*\in\mathrm{conv}(P)) \\
&\geq \max_{\sigma'\in A_{S'}} U(\sigma',p'_*) && \text{(for some } p'_*\in\mathrm{conv}(P') \text{ by condition 2)} \\
&\geq \min_{p'\in\mathrm{conv}(P')}\max_{\sigma'\in A_{S'}} U(\sigma',p') \\
&= \max_{\sigma'\in A_{S'}}\min_{p'\in\mathrm{conv}(P')} U(\sigma',p') = \max_{\sigma'\in A_{S'}}\min_{p'\in P'} U(\sigma',p')
\end{aligned}$$

where the second equality follows from von Neumann's minimax theorem since both $\mathrm{conv}(P)$ and A_S are compact and convex. This concludes the proof for $1 \implies 2$.

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

Suppose there exists $p_0 \in \text{conv}(P)$ such that p_0 is not Blackwell more informative than any $p' \in \text{conv}(P')$, we want to construct a triplet (A, u, π) in which

$$\max_{\sigma \in A_S} \min_{p \in P} U(\sigma, P) < \max_{\sigma' \in A_{S'}} \min_{p' \in P'} U(\sigma', P').$$

Fix $A = S'$, that is, the action space is just the set of signal realizations for P'. Then the sets of action plans are $A_S = \{\sigma \mid \sigma : S \to \Delta(S')\}$, $A_{S'} = \{\sigma' \mid \sigma' : S' \to \Delta(S')\}$. Consider an action plan $r \in A_{S'}$ defined by $r(s'_i \mid s'_j) = \mathbf{1}[i = j]$, that is, r is the action plan that just reports the signal realization. Then for any $p' \in P'$, $r \circ p' = p'$ since

$$(r \circ p')(s'_i \mid \omega) = \sum_{s' \in S'} r(s'_i \mid s')p'(s' \mid \omega) = p'(s'_i \mid \omega), \ \forall s'_i \in A \text{ and } \omega \in \Omega.$$

Let $\Lambda_{P',r} := \{r \circ p' \mid p' \in \text{conv}(P')\}$. That is, $\Lambda_{P',r}$ 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. Then $\Lambda_{P',r} = \text{conv}(P')$. Similarly, let $\Lambda_{p_0} := \{\sigma \circ p_0 \mid \sigma \in A_S\}$.

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

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

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

Since p_0 is not more informative than p' for any $p' \in \text{conv}(P')$, $\Lambda_{p_0} \cap \Lambda_{P',r} = \emptyset$. If not, there must exist λ in their intersection $\Lambda_{p_0} \cap \Lambda_{P',r}$, which further indicates the existence of an action plan $\sigma : S \to \Delta(S')$ and an experiment $p' \in \text{conv}(P')$ such that $\sigma \circ p_0 = \lambda = p'$, that is, some $p' \in \text{conv}(P')$ is a garbling of p_0. Contradiction.

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

$$\sum_{\omega \in \Omega} \sum_{a \in S'} v(\omega, a) \lambda(a \mid \omega) < c_1, \ \sum_{\omega \in \Omega} \sum_{a \in S'} v(\omega, a) \lambda'(a \mid \omega) > c_2, \ \forall \lambda \in \Lambda_{p_0}, \ \forall \lambda' \in \Lambda_{P',r}.$$

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

$$\begin{aligned}
\max_{\sigma \in A_S} U(\sigma, p_0) &= \max_{\lambda \in \Lambda_{p_0}} \sum_{\omega \in \Omega} \left(\sum_{a \in S'} v(\omega, a) \lambda(a \mid \omega) \right) \pi(\omega) \\
&= \frac{1}{|\Omega|} \cdot \max_{\lambda \in \Lambda_{p_0}} \sum_{\omega \in \Omega} \sum_{a \in S'} v(\omega, a) \lambda(a \mid \omega) \\
&< \frac{1}{|\Omega|} \cdot \min_{\lambda' \in \Lambda_{P',r}} \sum_{\omega \in \Omega} \sum_{a \in S'} v(\omega, a) \lambda'(a \mid \omega) = \min_{p' \in \text{conv}(P')} U(r, p')
\end{aligned}$$

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

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

$$\begin{aligned}
\max_{\sigma \in A_S} \min_{p \in P} U(\sigma, p) &= \max_{\sigma \in A_S} \min_{p \in \text{conv}(P)} U(\sigma, p) \\
&= \min_{p \in \text{conv}(P)} \max_{\sigma \in A_S} U(\sigma, p) \\
&\leq \max_{\sigma \in A_S} U(\sigma, p_0) \\
&< \min_{p' \in \text{conv}(P')} U(r, p') = \min_{p' \in P'} U(r, p') \leq \max_{\sigma' \in A_{S'}} \min_{p' \in P'} U(\sigma'; P')
\end{aligned}$$

where the first and the last equalties follow from the facts that $U(\cdot, \cdot)$ is linear in its second argument. The second inequality follows from von Neumann's minimax theorem since both $\text{conv}(P)$ and A_S are convex and compact and $U(\cdot, \cdot)$ is linear in both arguments. This completes the proof of the contrapositive. □

B.4 Proof of Proposition 2

Proof of Proposition 2. We focus on $\mathbf{p}$ and P, the proof for $\mathbf{p}'$ and P' is the same. Fix A, u, π and σ. Under the construction in (14),

$$\begin{aligned}
V_W \left(\left\{ U\big(\sigma, \mathbf{p}(\cdot, \theta)\big) \right\}_{\theta \in \Theta} \right) &= \min_{\theta \in \Theta} U\big(\sigma, \mathbf{p}(\cdot, \theta)\big) \\
&= \min_{(p, p') \in P \times P'} U\big(\sigma, \mathbf{p}(\cdot, (p, p'))\big) = \min_{p \in P} U(\sigma, p).
\end{aligned}$$

Under the construction in (15),

$$\min_{p \in P} U(\sigma, p) = \min_{\theta \in \Theta} U\big(\sigma, \mathbf{p}(\cdot, \theta)\big) = V_W \left(\left\{ U\big(\sigma, \mathbf{p}(\cdot, \theta)\big) \right\}_{\theta \in \Theta} \right).$$

This completes the proof. □

B.5 Proof of Proposition 3

Proof. Suppose it is not the case that $P \trianglerighteq_W P'$, then there exists (A, u, π) in which P' yields strictly higher ex-ante maximin utility than P. Combine this triplet (A, u, π) with the aggregator V_W, we find a quadruple (A, u, π, V_W) in which $\mathbf{p}'$ yields strictly higher ex-ante utility than $\mathbf{p}$. Thus, by Theorem 2, $\mathbf{p}$ does not prior-by-prior dominates $\mathbf{p}'$. Taking the contrapositive completes the proof.

To see the converse is not true, consider the following example:

$\Omega = \{\omega_1, \omega_2\},\ S = S' = \{s_1, s_2\}$, and

$$P = \left\{ \begin{array}{|c|c|c|} \hline & s_1 & s_2 \\ \hline \omega_1 & 0.9 & 0.1 \\ \hline \omega_2 & 0.1 & 0.9 \\ \hline \end{array} \right\},\ P' = \left\{ \begin{array}{|c|c|c|} \hline & s_1 & s_2 \\ \hline \omega_1 & 1 & 0 \\ \hline \omega_2 & 0 & 1 \\ \hline \end{array}, \begin{array}{|c|c|c|} \hline & s_1 & s_2 \\ \hline \omega_1 & 0 & 1 \\ \hline \omega_2 & 1 & 0 \\ \hline \end{array} \right\}.$$

Then $P \trianglerighteq_W P'$, but no matter how we defined the auxiliary state space Θ, we do not get prior-by-prior dominance. $\square$

B.6 Proof or Proposition 4

Proof. To see $1 \Longrightarrow 2$: let σ' be any action plan made for $\mathbf{p}'$ and let γ be the garbling that transforms $\mathbf{p}$ to $\mathbf{p}'$. Consider $\sigma' \circ \gamma$, $\sigma' \circ \gamma$ is a feasible action plan for $\mathbf{p}$, moreover, for any $(\theta, \omega) \in \Theta \times \Omega$,

$$\begin{aligned}
\hat{U}(\sigma' \circ \gamma, \mathbf{p}(\cdot, \theta), \omega) &= \sum_{s \in S} \sum_{a \in A} \mathbf{p}(s \mid \omega, \theta)(\sigma' \circ \gamma)(a \mid s) u(\omega, a) \\
&= \sum_{s \in S} \sum_{a \in A} \mathbf{p}(s \mid \omega, \theta) \sum_{s' \in S'} \sigma'(a \mid s') \gamma(s' \mid s) u(\omega, a) \\
&= \sum_{s' \in S'} \sum_{a \in A} \left(\sum_{s \in S} \mathbf{p}(s \mid \omega, \theta) \gamma(s' \mid s) \right) \sigma'(a \mid s') u(\omega, a) \\
&= \sum_{s' \in S'} \sum_{a \in A} \mathbf{p}'(s' \mid \omega, \theta) \sigma'(a \mid s') u(\omega, a) = \hat{U}(\sigma', \mathbf{p}'(\cdot, \theta), \omega)
\end{aligned}$$

That is, by taking the action plan $\sigma' \circ \gamma$, the DM obtains the same conditional expected utilities in every pair of states (ω, θ). This together with the assumption that the aggregator $\hat{V}$ is monotone completes the proof.

The equivalence of 2 and 3 is straightforward.

To see $3 \Longrightarrow 4$: Comparing to our model in Section 3, the combination of a unique prior π over Ω and an aggregator V over $\mathbb{R}^\Theta$ is just one special case of the more general aggregator $\hat{V}$ over $\mathbb{R}^{\Omega\times\Theta}$. Therefore, the set of possible (A, u, π, V) is expanded, and prior-by-prior dominance must still be necessary for guaranteeing higher ex-ante utility. $\square$

B.7 Proof of Proposition 5

Proof. Fix any finite set of actions $A = \{a_1, a_2, \ldots, a_n\}$, and let the state-dependent utility $u : \Omega \times A \to \mathbb{R}$ be summarized by

$$u(\omega_i, a_j) = u_{ij} \text{ for } i \in \{1, 2\} \text{ and } j \in \{1, \ldots, n\}.$$

Let σ denote the DM's action plan facing $\mathbf{p}$, with $\sigma_{ij} := \sigma(a_j \mid s_i)$. That is, σ_{ij} is the probability that the DM plays a_j after observing signal s_i. Note that $\sum_j \sigma_{ij} = 1$ for $i \in \{1, 2\}$. Let σ' denote the DM's action plan facing $\mathbf{p}'$, with σ'_{ij} defined accordingly. Write $\hat{U}^{\sigma}_{\mathbf{p}}(\theta, \omega)$ to get a more compact notation for conditional expected utilities , that is,

$$\hat{U}^{\sigma}_{\mathbf{p}}(\theta, \omega) := \hat{U}\big(\sigma, \mathbf{p}(\cdot, \theta), \omega\big).$$

Then we have

$$\hat{U}^{\sigma}_{\mathbf{p}}(\theta, \omega_j) = \sum_{k=1}^{n} \sigma_{jk} u_{jk}, \ \forall \theta \in \{\theta_1, \theta_2\}, \ j \in \{1, 2\}$$

$$\hat{U}^{\sigma'}_{\mathbf{p}'}(\theta_1, \omega_j) = \sum_{k=1}^{n} \sigma'_{jk} u_{jk}, \ \forall j \in \{1, 2\}$$

$$\hat{U}^{\sigma'}_{\mathbf{p}'}(\theta_2, \omega_1) = \sum_{k=1}^{n} \sigma'_{2k} u_{1k} \quad \text{and} \quad \hat{U}^{\sigma'}_{\mathbf{p}'}(\theta_2, \omega_2) = \sum_{k=1}^{n} \sigma'_{1k} u_{2k}$$

Thus, for any $\sigma' \in A_{S'}$, we can find $\sigma \in A_S$ such that

$$\hat{U}^{\sigma}_{\mathbf{p}}(\theta, \omega) \geq \hat{U}^{\sigma'}_{\mathbf{p}'}(\theta, \omega), \text{ for all } (\theta, \omega) \in \Theta \times \Omega.$$

To achieve this, let $\bar{u}_i := \max_{j\in\{1,\ldots,n\}} u_{ij}$ and $n^*_i \in \arg\max_{j\in\{1,\ldots,n\}} u_{ij}$, and consider $\sigma^* \in A_S$ be defined by $\sigma^*_{ij} = \mathbf{1}[j = n^*_i]$. Let σ' be an arbitrary action plan in $A_{S'}$, then

$$\hat{U}^{\sigma^*}_{\mathbf{p}}(\theta_1, \omega_1) = \bar{u}_1 = \sum_{k=1}^{n} \sigma'_{1k} \bar{u}_1 \geq \sum_{k=1}^{n} \sigma'_{1k} u_{1k} = \hat{U}^{\sigma'}_{\mathbf{p}'}(\theta_1, \omega_1)$$

$$\hat{U}^{\sigma^*}_{\mathbf{p}}(\theta_1, \omega_2) = \bar{u}_2 = \sum_{k=1}^{n} \sigma'_{2k} \bar{u}_2 \geq \sum_{k=1}^{n} \sigma'_{2k} u_{2k} = \hat{U}^{\sigma'}_{\mathbf{p}'}(\theta_1, \omega_2)$$

$$\hat{U}^{\sigma^*}_{\mathbf{p}}(\theta_2, \omega_1) = \bar{u}_1 = \sum_{k=1}^{n} \sigma'_{2k} \bar{u}_1 \geq \sum_{k=1}^{n} \sigma'_{2k} u_{1k} = \hat{U}^{\sigma'}_{\mathbf{p}'}(\theta_2, \omega_1)$$

$$\hat{U}^{\sigma^*}_{\mathbf{p}}(\theta_2, \omega_2) = \bar{u}_2 = \sum_{k=1}^{n} \sigma'_{1k} \bar{u}_2 \geq \sum_{k=1}^{n} \sigma'_{1k} u_{2k} = \hat{U}^{\sigma'}_{\mathbf{p}'}(\theta_2, \omega_2)$$

Then for any monotone aggregator $\hat{V} : \mathbb{R}^{\Omega \times \Theta} \to \mathbb{R}$,

$$\sup_{\sigma \in A_S} \hat{V}\left(\hat{U}^{\sigma}_{\mathbf{p}}(\cdot,\cdot)\right) \geq \hat{V}\left(\hat{U}^{\sigma^*}_{\mathbf{p}}(\cdot,\cdot)\right) \geq \sup_{\sigma' \in A_{S'}} \hat{V}\left(\hat{U}^{\sigma'}_{\mathbf{p}'}(\cdot,\cdot)\right).$$

That is, $\mathbf{p}$ gives weakly higher ex-ante utility than $\mathbf{p}'$ for any possible $(A, u, \hat{V})$. □

B.8 Proof of Theorem 4

We prove Theorem 4 by showing that $1 \implies 2 \implies 3 \implies 1$.

Proof. Recall that Δ_0 is the set of all probability measures over Θ with finite support and $\mathbf{p} \trianglerighteq_{PBP} \mathbf{p}'$ if $\mathbf{p}_\mu$ is Blackwell more informative than $\mathbf{p}'_\mu$ for all $\mu \in \Delta_0$.

To see $1 \implies 2$:

Let Γ denote the set of all garblings, that is, $\Gamma := \{\gamma \mid \gamma : S \to \Delta(S')\}$. Γ is compact and convex since both S and S' are finite. Fix A, u, π and $\sigma' \in A_{S'}$, we define an auxiliary function $H : \Gamma \times \Delta_0 \to \mathbb{R}$ by

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

That is, $H(\gamma, \mu)$ is difference in the expected utilities of $\mathbf{p}$ and $\mathbf{p}'$ conditional on belief $\mu \in \Theta$ and garbling γ being applied to the action plan made for $\mathbf{p}'$.

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 is well defined and nonnegative, we invoke the Kneser-Fan minimax theorem for concave-convex functions (Terkelsen, 1972).

Γ is a compact and convex subset of $\mathbb{R}^{|S| \times |S'|}$ 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

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

But 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 $\mathbf{p}'_\mu = \gamma_\mu \circ \mathbf{p}_\mu$ and this garbling γ_μ guarantees that $H(\gamma_\mu, \mu) = 0$. This completes the proof that $1 \Longrightarrow 2$ since $\sigma' \circ \gamma \in A_S$ is a valid action plan for $\mathbf{p}$.

To see $2 \Longrightarrow 3$:

Fix any (A, u, π, V) with $V \in \mathcal{V}_{Mono}$ and let v^* denote the ex-ante utility for $\mathbf{p}'$, that is,

$$v^* := \sup_{\sigma' \in A_{S'}} V\left(\left\{U\big(\sigma', \mathbf{p}'(\cdot, \theta)\big)\right\}_{\theta \in \Theta}\right).$$

Then for any $\varepsilon > 0$, there exists $\sigma'_\varepsilon \in A_{S'}$ such that

$$V\left(\left\{U\big(\sigma'_\varepsilon, \mathbf{p}'(\cdot, \theta)\big)\right\}_{\theta \in \Theta}\right) > v^* - \varepsilon.$$

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

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

By the monotonicity of V, this further indicates that

$$V\left(\left\{U\big(\sigma_\varepsilon, \mathbf{p}(\cdot, \theta)\big)\right\}_{\theta \in \Theta}\right) \geq V\left(\left\{U\big(\sigma'_\varepsilon, \mathbf{p}'(\cdot, \theta)\big)\right\}_{\theta \in \Theta}\right) > v^* - \varepsilon.$$

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

To see $3 \Longrightarrow 1$:

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 $\mathbf{p}'$ by every decision maker with $V \in \mathcal{V}_0$, then $\mathbf{p} \trianglerighteq_{PBP} \mathbf{p}'$.

Suppose by contradiction that it is not the case $\mathbf{p} \trianglerighteq_{PBP} \mathbf{p}'$, then there must exist some $\mu \in \Delta_0$ such that $\mathbf{p}_\mu$ is not Blackwell more informative than $\mathbf{p}'_\mu$. 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, this DM's ex-ante utility is

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

Since $\mathbf{p}_\mu$ is not Blackwell more informative than $\mathbf{p}'_\mu$, then by Theorem 1, there must exist a triplet (A, u, π) such that

$$\max_{\sigma \in A_S} U(\sigma, \mathbf{p}_\mu) < \max_{\sigma' \in A_{S'}} U(\sigma', \mathbf{p}'_\mu),$$

which further indicates that in (A, u, π, V_μ),

$$\sup_{\sigma \in A_S} V_\mu \left(\left\{ U\big(\sigma, \mathbf{p}(\cdot, \theta)\big) \right\}_{\theta \in \Theta} \right) < \sup_{\sigma' \in A_{S'}} V_\mu \left(\left\{ U\big(\sigma', \mathbf{p}'(\cdot, \theta)\big) \right\}_{\theta \in \Theta} \right)$$

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

B.9 Proof of Proposition 6

Proof. The proof is essentially the same with that of Proposition 2. The infimum on the left hand side can be replaced by a minimum since the $\Theta = P \times P'$ is compact and our construction of $\mathbf{p}$ and $\mathbf{p}'$ guarantees the continuity of $U\big(\sigma, \mathbf{p}(\cdot, \theta)\big)$ in θ (this is because $\theta \mapsto \mathbf{p}(\cdot, \theta)$ is just $(p, p') \mapsto p$, which can be viewed as a projection map and thus continuous as we endow $P \times P'$ with the product topology). $\square$

References

Aliprantis, Charalambos D., and Kim C. Border. 2006. *Infinite dimensional analysis: A hitchhiker's guide*. Springer.

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. 2020. "Biased learning under ambiguous information." *Available at SSRN.*

Crémer, Jacques. 1982. "A simple proof of Blackwell's "comparison of experiments" theorem." *Journal of Economic Theory* 27 (2): 439–443.

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.

Epstein, Larry, and Yoram Halevy. 2019. "Hard-to-interpret signals." *Working paper.*

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.

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

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

Liang, Yucheng. 2019. "Learning from unknown information sources." *Available at SSRN 3314789.*

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

Savage, Leonard J. 1954. *The foundations of statistics.* Wiley, New York.

Shishkin, Denis, and Pietro Ortoleva. 2019. "Ambiguous information and dilation: An experiment." *Working paper.*

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

Wald, Abraham. 1950. *Statistical decision functions.*. Wiley.