Blackwell's Informativeness Ranking with Uncertainty Averse Preferences *

Jian Li † Junjie Zhou ‡

February 4, 2015

Abstract

Blackwell (1951, 1953) proposes an informativeness ranking of experiments: Experiment I is more Blackwell-informative than Experiment II *if and only if* the value of experiment I is higher than that of experiment II for all expected-utility maximizers. Under commitment and reduction, our main theorem shows that Blackwell equivalence holds for all convex and strongly monotone preferences, i.e., the uncertainty averse preferences (Cerreia-Vioglio et al. 2011b), which nest most ambiguity averse preferences commonly used in applications as special cases. Furthermore, we discuss the possibility of extending the equivalence results to the no commitment case for the maxmin expected utility and variational preferences under certain conditions.

JEL Classification Numbers: C44, D81.
KEYWORDS: Blackwell's theorem, garbling, ambiguity aversion, value of information

*We thank two anonymous referees for valuable suggestions that helped us improve the paper. We also thank Marciano Siniscalchi, Chris Shannon, Bin Miao, Maciej H.Kotowski, Haluk Ergin, John K.-H. Quah, seminar participants at SHUFE, McGill, and Concordia, and audience at FUR 2014 for helpful comments. We thank CIREQ for funding Junjie Zhou's visit to McGill in April 2014, where part of the work was conducted. Jian Li thanks McGill's economics department for support. All remaining errors are our own.

†Corresponding author. CIREQ and Department of Economics, McGill University, 855 Sherbrooke Street West, Montreal, QC, Canada, H3A 2T7. E-mail: jian.li7@mcgill.ca. Telephone: 1-514-398-4400-00830. Fax: 1-514-398-4938.

‡SIBA, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai, China 200433. E-mail: zhoujj03001@gmail.com.

1 Introduction

Consider a firm with a new product for release. There are two possible states: in one state 10% of the population like the product; in the other state 20% of the population like it. The firm manager does not know the true state, but he can always sample the population. Clearly the manager weakly prefers a larger sample to a smaller sample regardless of his utility function and his prior belief about the two states, as he can always replicate the outcome with a smaller sample by ignoring the extra samples.

A more general notion of sampling is called experiment, or information structure, which specifies the likelihoods of signal realizations conditional on each state. In his seminal papers, Blackwell (1951, 1953) defines a partial ranking of experiments, where Experiment I is more Blackwell-informative than Experiment II if the latter is a garble of the former. In other words, the less informative experiment can be considered as the more informative experiment with a noise. Blackwell's theorem establishes that the value of Experiment I is weakly higher than that of Experiment II for all expected-utility maximizers and all sets of actions *if and only if* Experiment I is more "Blackwell-informative" than Experiment II.

What if the firm manager does not have enough information to form a probabilistic belief over the two states? Recent experimental evidence suggests that when people do not know the probability of an event, they dislike betting on it.[1] The tendency, called *ambiguity aversion*, has attracted significant interest in theory and applications.[2] Thus it is of interest to study the comparison of experiment/ information acquisition for decision makers (DMs) who are ambiguity averse.

Çelen (2012) showed that Blackwell's theorem extends to maxmin expected utility (MEU) preferences (Gilboa and Schmeidler 1989).[3] In this paper, we look for broader families of ambiguity preferences whose induced value of information characterizes the Blackwell ranking. As in Çelen (2012), we consider DMs who can commit to any (ex-ante) strategy and only perceive ambiguity in the states, while treating the information structures as objectively given. We show that, with relatively mild technical assumptions, most families of ambiguity preferences commonly used in applications, such as variational preferences (Maccheroni et al. 2006a), smooth ambiguity preferences (Klibanoff et al. 2005),

[1]For a review of earlier experimental evidence, see Camerer and Weber (1992). For more recent experiments, see for instance Fox and Tversky (1995), Chow and Sarin (2001), Halevy (2007), Bossaerts et al. (2010), Abdellaoui et al. (2011).

[2]See references below for axiomatic models of ambiguity averse preferences. For the economic and financial applications of ambiguity, see Mukerji and Tallon (2004) and Epstein and Schneider (2010) and references therein.

[3]Recently, Heyen and Wiesenfarth (2014) propose a recursive calculation of the value of information; Gensbittel et al. (2015) consider ambiguous information structure and the no commitment case. Both papers focus on the MEU case.

multiplier preferences (Hansen and Sargent 2001; Strzalecki 2011), confidence preferences (Chateauneuf and Faro 2009) and second-order expected utility (Grant et al. 2009) can also induce a partial ranking of information that is equivalent to the Blackwell ranking. The largest such characterizing family we identify is the uncertainty averse preferences (Cerreia-Vioglio et al. 2011b). Our main proof suggests a link between Blackwell's equivalence and convex preferences. This also confirms the impression that the Blackwell ranking is coarse.[4]

Our paper is related to the literature on the value of information for non-EU DMs. For choice under risk, Wakker (1988), Hilton (1990), and Safra and Sulganik (1995) show how Blackwell's theorem might fail for non-EU DMs. Grant et al. (1998) study objective two-stage compound lotteries and focus on the intrinsic value of information under a fixed action.[5] For ambiguity, Siniscalchi (2011) shows how a sophisticated ambiguity averse DM might reject freely available information.[6] These earlier papers seemingly draw very different conclusions from Çelen (2012) and our paper. The key reason is that Çelen (2012) and our paper assume that a DM can commit to any signal contingent strategy and focus on the ex-ante pure decision value of information, while the earlier papers assume decisions are only made after observing the signals and consider a trade-off between the decision value and commitment value of information. Compared with the earlier findings, Celen and our paper identify a benchmark way of considering the value of information for non-EU DMs, and show that under this benchmark Blackwell's equivalence extends to all uncertainty averse preferences. Hence our results are complementary to the earlier findings.

A separate literature is motivated by the concern that the Blackwell ranking is too coarse for many applications. Some later papers study finer information rankings but impose certain structural restrictions on the Von Neumann-Morgenstern (vNM) utility indices or restrictions on the decision problems. For example, Lehmann (1988) and Persico (2000) consider utility indices satisfying the single crossing property, Athey and Levin (2001) study supermodular utility indices, and Quah and Strulovici (2009) explore interval dominance order utilities. Cabrales et al. (2013) explore non-arbitrary investment decisions and focus on ruin-averse utilities. In contrast to these papers, we do not put any restriction on the vNM utility indices, but consider the validity of the Blackwell equivalence result under non-EU preferences.

The remainder of the paper is organized as follows. We describe notation in Section 2. Section 3 introduces uncertainty averse preferences and the main assumptions. Section 4

[4]See Blackwell and Girshick (1954) and Lehmann (1988) for discussions along this line.

[5]They find that (Proposition 1, ii) intrinsic information loving implies utility function of one-stage lotteries is quasi-convex.

[6]Strzalecki (2013) and Li (2013) show if one allows for preferences for temporal resolution of uncertainties, then in some region an ambiguity aversion DM with recursive preferences might prefer late resolution of uncertainties. In this paper we assume reduction and rule out such concerns.

presents the main theorem for uncertainty averse preferences. Section 5 applies the main theorem to six well known families of ambiguity preferences. Some discussions are drawn in Section 6. The appendix includes direct proofs of smooth ambiguity preferences and second order expected utility.

2 Notation

Our notation follow that of Çelen (2012). Let $\Delta(\Omega)$ be the set of all priors on Ω. The set $int(\Delta(\Omega))$ contains priors with full support. For any matrix $\mathbf{m}_{a\times b}$ of dimension $a \times b$, m_{ij} and $\mathbf{m}'$ denote the (i,j)th entry and the transpose of $\mathbf{m}$, respectively. The inner product of two matrices of the same dimension is defined as $\langle \mathbf{m}, \mathbf{n} \rangle := \sum_i \sum_j m_{ij} n_{ij} = tr(\mathbf{m}'\mathbf{n})$. For any vector $\pi \in \mathbb{R}^n, D^\pi$ denotes the diagonal matrix such that $D^\pi_{ii} = \pi_i$. Finally $\mathbf{I}$ denotes the identity matrix.

Let $\Omega := \{\omega_1, \cdots, \omega_n\}$ be a finite set of *states*, and $A := \{a_1, \cdots, a_{|A|}\}$ be a finite set of *actions* available to a DM.[7] A DM is characterized by a *utility function* or a *vNM utility index* $u : \Omega \times A \mapsto \mathbb{R}$ and a *prior* $\pi \in \Delta(\Omega)$. We can construct a matrix $\mathbf{u}_{n\times|A|}$ with entries $u_{\omega a} = u(\omega, a)$, for all $\omega \in \Omega, a \in A$.

Experiments, or sometimes called *information structures*, are tuples $(\mathcal{S}, \mathbf{p})$ and $(\mathcal{T}, \mathbf{q})$, where $\mathcal{S} := \{s_1, \cdots, s_{|S|}\}$ and $\mathcal{T} := \{t_1, \cdots, t_{|T|}\}$ are sets of signals, and $\mathbf{p}_{n\times|S|}$ and $\mathbf{q}_{n\times|T|}$ are markov matrices.[8] In particular, $p_{\omega s} := \Pr(s|\omega)$ for $s \in \mathcal{S}$ and $q_{\omega t} := \Pr(t|\omega)$ for $t \in \mathcal{T}$.

For a DM who observes a signal s from the experiment $(\mathcal{S}, \mathbf{p})$, a strategy is a vector-valued mapping $\mathbf{f} : \mathcal{S} \mapsto \Delta(A)$. For each strategy $\mathbf{f}$ we define the matrix $\mathbf{f}_{|S|\times|A|}$, such that $(f_{j1}, \cdots, f_{j|A|}) := f(s_j)$.[9] Similarly we can define a strategy $\mathbf{g} : \mathcal{T} \mapsto \Delta(A)$. If a strategy maps every signal to the same (mixed) action $\mathbf{a}$ in $\Delta(A)$, it is identified with $\mathbf{a}$.

Blackwell (1951) defines the following ranking of two experiments.

Definition 1. *An experiment $(\mathcal{S}, \mathbf{p})$ is more Blackwell-informative than experiment $(\mathcal{T}, \mathbf{q})$ if there exists a markov matrix $\mathbf{r}$ such that $\mathbf{q} = \mathbf{pr}$.*

The matrix $\mathbf{r}$ is also called the *garbling matrix*.

We incorporate ambiguity by considering an environment in which there is ambiguity about states in Ω, while the signal-generating process, described by the likelihood matrix

[7] We assume the number of available actions is larger than the number of signals.
[8] A matrix $\mathbf{m}$ is markov if it is nonnegative and row stochastic, i.e., $m_{ij} \geq 0$ and $\sum_j m_{ij} = 1$ for all i.
[9] Strategy $\mathbf{f}$ is a markov matrix.

$\mathbf{p}$, is treated as objectively given. By focusing on unambiguous signal likelihoods, we can relate the generalized value of signals under ambiguity to a clear ranking of their informational content. Examples below illustrate situations in which this assumption is natural.

Example 1 (Partition)**.** *In many economic and financial applications, information is represented by partitions of the state space. A finer partition is more Blackwell-informative. Specifically, if $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, then the partition $\{\{\omega_1\}, \{\omega_2\}, \{\omega_3, \omega_4\}\}$ is more informative than the partition $\{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\}$. The likelihood and garbling matrices are*

$$\underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}}_{\mathbf{q}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{p}} \underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{r}}.$$

A DM *may perceive ambiguity about the states. But conditional on the true state, a partitional signal structure unambiguously describes whether it belongs to each event in the partition.*

Example 2 (Sampling)**.** *Consider the sampling example in the introduction. Clearly the larger the sample size n, the more informative the signal structure. For example, for $n = 1$ and $n = 2$, the likelihood and garbling matrices are*

$$\underbrace{\begin{bmatrix} 0.1 & 0.9 \\ 0.2 & 0.8 \end{bmatrix}}_{\mathbf{q}} = \underbrace{\begin{bmatrix} 0.01 & 0.18 & 0.81 \\ 0.04 & 0.32 & 0.64 \end{bmatrix}}_{\mathbf{p}} \underbrace{\begin{bmatrix} 1 & 0 \\ 0.5 & 0.5 \\ 0 & 1 \end{bmatrix}}_{\mathbf{r}}$$

In this case, the firm might perceive ambiguity about the preference distributions in the population. But conditional on a given proportion of the population who like the product, the sample information unambiguously follows a binomial distribution.[10]

Example 3 (Noisy communication Channel)**.** *A sender wants to transmit a piece of news to a receiver, which can be either good or bad. The news is sent via a noisy communication channel: with probability $1 - k$, the news is transmitted successfully; with probability k, the news is lost and the receiver gets an error message. A communication channel with a smaller error probability is more Blackwell-informative. For example, a message sent via email with an error probability of $1/100$ is more informative than a message sent via*

[10] For insiders of the ambiguity literature, this is reminiscent of repeated sampling from an urn with unknown compositions of black and red balls. If a state is a given composition of the urn, then it is natural to consider a DM who faces prior uncertainty about the composition of the urn, while conditional on a given composition, the likelihood of a sample history is unambiguous. See Epstein and Schneider (2007) for further discussion.

telegraph with an error probability of $1/10$. *The likelihood and garbling matrices are*

$$\underbrace{\begin{bmatrix} \frac{9}{10} & 0 & \frac{1}{10} \\ 0 & \frac{9}{10} & \frac{1}{10} \end{bmatrix}}_{\mathbf{q}} = \underbrace{\begin{bmatrix} \frac{99}{100} & 0 & \frac{1}{100} \\ 0 & \frac{99}{100} & \frac{1}{100} \end{bmatrix}}_{\mathbf{p}} \underbrace{\begin{bmatrix} \frac{10}{11} & 0 & \frac{1}{11} \\ 0 & \frac{10}{11} & \frac{1}{11} \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{r}}.$$

In this case, the content of the news might be ambiguous to the receiver, yet the error probability, which depends on the physical properties of the Internet/wire, can be viewed as objective.

3 Uncertainty Averse Preferences

To model ambiguity aversion, we take Cerreia-Vioglio et al.'s (2011b) uncertainty averse preferences, which are the most general to our knowledge and nest other ambiguity averse preferences as special cases. If a DM has uncertainty averse preferences and takes action $\mathbf{a}$, then in our notation her utility is

$$U(\mathbf{a}) = \min_{\pi \in \Delta(\Omega)} G\left(\sum_{w \in \Omega} \pi(w) \sum_{a \in A} \mathbf{a}_a u(w, a), \pi\right)$$

Here the function $G : \mathcal{X} \times \Delta(\Omega) \mapsto (-\infty, +\infty]$ is the index of uncertainty aversion, which depends on the expected utility of action $\mathbf{a}$ and prior π. The set $\mathcal{X}$ is an interval of the real line $\mathbb{R}$. Moreover, G is quasi-convex, $G(\cdot, \pi)$ is increasing for all π, and $\inf_{\pi \in \Delta(\Omega)} G(x, \pi) = x$ for all $x \in \mathcal{X}$.

We make two behavioral assumptions.

First, the DM *can commit to all signal-contingent strategies*. It is well-known that non-EU preferences can potentially be dynamically inconsistent: the ex-ante and conditional preferences might differ.[11] Under full commitment, dynamic inconsistency is not an issue: the DM will always implement her *ex-ante* optimal strategies. This allows us to focus on the pure decision value of information and to remain comparable with Blackwell (1951), which also studies the ex-ante value of information. See Section 6.3 for detailed discussion on how the analysis would change if the commitment assumption are dropped.

Second, for a given experiment $(\mathcal{S}, \mathbf{p})$, the DM faces uncertainties from two sources, Ω and $\mathcal{S}$. Since the likelihood matrix $\mathbf{p}$ is objectively given, we assume that the uncertainty

[11]Machina (1989) discusses the issue of dynamic inconsistency for non-EU DM under risk. Numerous papers discuss how an ambiguity averse DM can potentially be dynamically inconsistent. See Epstein and Schneider (2003: Section 4.1) for an example.

averse index $\tilde{G} : \mathcal{X} \times \Delta(\Omega \times \mathcal{S}) \mapsto (-\infty, +\infty]$ is related to the original index G through the following:

$$\tilde{G}(x, P) = \begin{cases} G(x, \pi), & \text{if } P = D^\pi \mathbf{p} \text{ for some } \pi; \\ +\infty, & \text{otherwise.} \end{cases}$$

This is well defined as the mapping from prior π to the joint probability $D^\pi \mathbf{p}$ is one-to-one.

Given experiment $(\mathcal{S}, \mathbf{p})$, prior π, and vNM utility index $\mathbf{u}$, the expected utility of strategy $\mathbf{f}$ is

$$\sum_{s \in \mathcal{S}} \Pr(s) \sum_{\omega \in \Omega} \Pr(\omega|s) \sum_{a \in A} f_s(a) u(\omega, a) = \sum_{\omega \in \Omega} \sum_{s \in \mathcal{S}} \sum_{a \in A} \pi_w p_{ws} f_{sa} u_{\omega a} = \langle D^\pi \mathbf{pf}, \mathbf{u} \rangle.$$

For a uncertainty averse DM with uncertainty averse index G and vNM utility index $\mathbf{u}$, her ex-ante utility from committing to strategy $\mathbf{f}$ when facing experiment $(\mathcal{S}, \mathbf{p})$ is

$$U^{UA}(\mathcal{S}, \mathbf{p}, \mathbf{f}) \quad = \quad \min_{\{D^\pi \mathbf{p} : \pi \in \Delta(\Omega)\}} \tilde{G}(\langle D^\pi \mathbf{pf}, \mathbf{u} \rangle, D^\pi \mathbf{p}) = \min_{\pi \in \Delta(\Omega)} G(\langle D^\pi \mathbf{pf}, \mathbf{u} \rangle, \pi).$$

And her value of an experiment $(\mathcal{S}, \mathbf{p})$ is $U^{UA}(\mathcal{S}, \mathbf{p}) := \max_{\mathbf{f}} U^{UA}(\mathcal{S}, \mathbf{p}, \mathbf{f})$.

Finally, we impose some technical assumptions on the uncertainty averse index G.

Assumption 1. *There exists a prior $\pi_0 \in int(\Delta(\Omega))$ and some constant $b \in int(\mathcal{X})$ such that $G(b, \pi_0) = b$. Moreover, $G(\cdot, \pi_0)$ is strictly increasing and continuous.*

The main restriction in Assumption 1 is the existence of a prior π_0 with full support. It is straightforward to check that $G(x, \pi_0) = x$ for all $x \in \mathcal{X}$ in every special case studied in Section 5. Obviously such $G(\cdot, \pi_0)$ is strictly increasing and continuous. See Section 6.2 for a discussion on necessity of the full support requirement.

To provide a better geometric intuition behind the main theorem, we look at an equivalent characterization of the uncertainty averse preferences, relying mostly on its convexity structure instead of its functional form. For each mixed strategy $\mathbf{f}$, we can construct a state-contingent mixed action $\mathbf{a}$, and a state-contingent expected utility $\mathbf{u_{pf}} \in \mathcal{X}^n$, where $\mathbf{a} = \mathbf{pf}$, and $\mathbf{u_{pf}}$ specifies $\sum_{a \in A} (\sum_{s \in S} p_{\omega s} f_{sa}) u_{\omega a}$ for each state $w \in \Omega$. An uncertainty averse DM has underlying uncertainty averse preferences over induced state-contingent actions $A^{\mathcal{S}} := \{\mathbf{a} = \mathbf{pf} : \mathbf{f} \text{ is a mixed strategy}\}$, which is represented by a utility function $U(\mathbf{pf}) = U^{UA}(\mathcal{S}, \mathbf{p}, \mathbf{f}) = \min_{\pi \in \Delta(\Omega)} G(\langle D^\pi \mathbf{pf}, \mathbf{u} \rangle, \pi)$. Let $I(\mathbf{u_{pf}}) := U(\mathbf{pf})$ then I is a function $\mathcal{X}^n \mapsto \mathbb{R}$ aggregating state-contingent utilities. By Cerreia-Vioglio et al. (2011b), if preferences admit an uncertainty averse representation, then I is (i) quasi-concave; (ii) strongly monotone: if $x_i > y_i$ for all i, $I(x) > I(y)$; (iii) normalized: for any constant $b \in \mathcal{X}$, $I(b\mathbf{1}_n) = b$; (iv) continuous.

For such an aggregator I, its Greenberg-Pierskalla superdifferential at $x \in \mathcal{X}^n$ is

$$\partial^{GP} I(x) := \{\xi : \langle \xi, y - x \rangle \leq 0 \Rightarrow I(y) \leq I(x)\}.$$

By Corollary 10.2 in Greenberg and Pierskalla (1973), Greenberg-Pierskalla superdifferential exists everywhere if I is quasi-concave, strongly monotone, and continuous. Furthermore, observe that for any $b \in \mathcal{X}$, the set $\partial^{GP} I(b\mathbf{1}_n) \cap \Delta(\Omega) = \arg\min_{\pi \in \Delta(\Omega)} G(\pi, b) = \{\pi \in \Delta(\Omega) : G(\pi, b) = b\}$.[12] So Assumption 1 is equivalent to the following assumption.

Assumption 1′. *There exists a fully supported prior π_0 and some constant $b \in int(\mathcal{X})$ such that π_0 is a Greenberg-Pierskalla superdifferential of I at $b\mathbf{1}_n$, i.e., $\pi_0 \in \partial^{GP} I(b\mathbf{1}_n)$.*

4 Main Result

Theorem 1. *Suppose that Assumption 1 holds. The following statements are equivalent:*

(i) $(\mathcal{S}, \mathbf{p})$ is more Blackwell-informative than $(\mathcal{T}, \mathbf{q})$, i.e., there exists a markov matrix $\mathbf{r}$,

$$\mathbf{q} = \mathbf{pr}.$$

(ii) $(\mathcal{S}, \mathbf{p})$ is more valuable than $(\mathcal{T}, \mathbf{q})$ for all DMs *with uncertainty averse index G, i.e.,*

$$\max_{\mathbf{f}} \min_{\pi \in \Delta(\Omega)} G(\langle D^\pi \mathbf{pf}, \mathbf{u} \rangle, \pi) \geq \max_{\mathbf{g}} \min_{\pi \in \Delta(\Omega)} G(\langle D^\pi \mathbf{qg}, \mathbf{u} \rangle, \pi), \quad \forall \mathbf{u}. \tag{1}$$

Proof. $(i) \Rightarrow (ii)$ direction. Given the set of actions A and the set of states Ω, let $A^{\mathcal{S}} = \{\mathbf{pf} | \mathbf{f} \text{ is a mixed strategy}\}$ and $A^{\mathcal{T}} = \{\mathbf{qg} | \mathbf{g} \text{ is a mixed strategy}\}$ be the set of state-contingent actions induced by experiment $(\mathcal{S}, \mathbf{p})$ and $(\mathcal{T}, \mathbf{q})$ respectively. If $\mathcal{S}$ is more Blackwell-informative than $\mathcal{T}$, Blackwell (1951: Theorem 2) shows that the set of state-contingent actions induced by $\mathcal{S}$ is larger than that by $\mathcal{T}$. Since the DM's utility of an experiment is the maximum utility of the set of state-contingent actions, a larger set is always better, and as a result the more informative experiment is always preferable. Formally,

$$\begin{aligned} \max_{\mathbf{g}} \min_{\pi \in \Delta(\Omega)} G(\langle D^\pi \mathbf{qg}, \mathbf{u} \rangle, \pi) &= \max_{\mathbf{qg} \in A^{\mathcal{T}}} U(\mathbf{qg}) \\ &\leq \max_{\mathbf{pf} \in A^{\mathcal{S}}} U(\mathbf{pf}) = \max_{\mathbf{f}} \min_{\pi \in \Delta(\Omega)} G(\langle D^\pi \mathbf{pf}, \mathbf{u} \rangle, \pi). \end{aligned}$$

[12] See Cerreia-Vioglio et al. (2011a: Section 9).

as we have $A^{\mathcal{T}} \subseteq A^{\mathcal{S}}$.[13]

$(ii) \Rightarrow (i)$ direction. We prove by contraposition. Suppose there does not exist any markov matrix $\mathbf{r}$ such that $\mathbf{q} = \mathbf{pr}$. Let $\mathcal{P} = \{D^{\pi_0}\mathbf{pr} : \mathbf{r}$ is a $\sigma \times \sigma'$ markov matrix$\}$ and $\mathcal{Q} = \{D^{\pi}\mathbf{q} : \pi \in \Delta(\Omega)\}$. Then $\mathcal{P} \cap \mathcal{Q} = \emptyset$. To see this, suppose not, then $D^{\pi_0}\mathbf{pr} = D^{\pi}\mathbf{q}$ for some markov matrix $\mathbf{r}$ and some π. Multiplying both sides by a column vector of 1's yields $\pi_0 = \pi$. Since π_0 has full support, D^{π_0} is invertible, and multiplying both sides by $(D^{\pi_0})^{-1}$ implies $\mathbf{pr} = \mathbf{q}$, contradicting the contrapositive assumption. Moreover, $\mathcal{P}$ and $\mathcal{Q}$ are nonempty, compact, and convex. By the strict separating hyperplane theorem, there exists $\mathbf{v} \neq 0$ such that $\langle \mathbf{n}, \mathbf{v} \rangle > 0 > \langle \mathbf{m}, \mathbf{v} \rangle$ for all $\mathbf{n} \in \mathcal{Q}$ and $\mathbf{m} \in \mathcal{P}$. By assumption, there is some constant $b \in int(\mathcal{X})$ such that $G(b, \pi_0) = b$. Let $A := \{a_1, \cdots, a_{|T|}\}$. Let $\mathbf{u} = \delta\mathbf{v} + b\mathbf{1}_{n\times|T|}$, where $\delta > 0$ is small enough so that $\mathbf{u} \in \mathcal{X}_{n\times|T|}$. Then

$$\langle \mathbf{n}, \mathbf{u} \rangle > b > \langle \mathbf{m}, \mathbf{u} \rangle, \qquad \forall \mathbf{n} \in \mathcal{Q} \text{ and } \mathbf{m} \in \mathcal{P}. \tag{2}$$

For the experiment $(\mathcal{S}, \mathbf{p})$ and vNM utility index $\mathbf{u}$, the LHS of inequality (1) is

$$\max_{\mathbf{f}} \min_{\pi\in\Delta(\Omega)} G(\langle D^{\pi}\mathbf{pf}, \mathbf{u} \rangle, \pi) = \max_{\mathbf{f}} I(\mathbf{u}_{\mathbf{pf}}) = I(\mathbf{u}_{\mathbf{pf}^*}).$$

The optimal strategy $\mathbf{f}^*$ exists since $\{\mathbf{u}_{\mathbf{pf}} : \mathbf{f}$ a markov matrix$\}$ is compact and I is continuous. By (2), $\langle D^{\pi_0}\mathbf{pf}^*, \mathbf{u} \rangle - b < 0$, which is equivalent to $\langle \pi_0, \mathbf{u}_{\mathbf{pf}^*} \rangle - \langle \pi_0, b\mathbf{1}_n \rangle < 0$. Since I is quasiconcave and π_0 is a Greenberg-Pierskalla superdifferential of I at $b\mathbf{1}_n$,

$$I(\mathbf{u}_{\mathbf{pf}^*}) - I(b\mathbf{1}_n) \leq 0.$$

I is normalized so $I(b\mathbf{1}_n) = b$, thus

$$\max_{\mathbf{f}} \min_{\pi\in\Delta(\Omega)} G(\langle D^{\pi}\mathbf{pf}, \mathbf{u} \rangle, \pi) = I(\mathbf{u}_{\mathbf{pf}^*}) \leq b. \tag{3}$$

For the experiment $(\mathcal{T}, \mathbf{q})$ and vNM utility index $\mathbf{u}$, the RHS of inequality (1) is

$$\max_{\mathbf{g}} \min_{\pi\in\Delta(\Omega)} G(\langle D^{\pi}\mathbf{qg}, \mathbf{u} \rangle, \pi) = \max_{\mathbf{g}} I(\mathbf{u}_{\mathbf{qg}}) \geq I(\mathbf{u}_{\mathbf{qI}})$$

where $\mathbf{I}$ is the $|T| \times |T|$ identity matrix. By (2), there exists some small enough $\epsilon > 0$ such that $b - \langle D^{\pi}\mathbf{qI}, \mathbf{u} \rangle + \epsilon < 0$. Pick any $\pi \in \partial^{GP} I(\mathbf{u}_{\mathbf{qI}} - \epsilon\mathbf{1}_n)$, then we have $\langle \pi, b\mathbf{1}_n \rangle - \langle \pi, \mathbf{u}_{\mathbf{qI}} - \epsilon\mathbf{1}_n \rangle < 0$, which implies

$$I(b\mathbf{1}_n) - I(\mathbf{u}_{\mathbf{qI}} - \epsilon\mathbf{1}_n) \leq 0.$$

[13]Note that in $(i) \Rightarrow (ii)$ we do not use quasi-concavity of $U(\cdot)$.

as I is quasiconcave. By strong monotonicity of I,

$$b = I(b\mathbf{1}_n) \leq I(\mathbf{u}_{\mathbf{qI}} - \epsilon\mathbf{1}_n) < I(\mathbf{u}_{\mathbf{qI}}). \tag{4}$$

Combining (3) and (4) yields

$$\max_{\mathbf{g}} \min_{\pi\in\Delta(\Omega)} G(\langle D^{\pi}\mathbf{qg}, \mathbf{u}\rangle, \pi) \geq I(\mathbf{u}_{\mathbf{qI}}) > b \geq \max_{\mathbf{f}} \min_{\pi\in\Delta(\Omega)} G(\langle D^{\pi}\mathbf{pf}, \mathbf{u}\rangle, \pi).$$

This is a contradiction to inequality (1). □

The intuition for the proof of $(ii) \Rightarrow (i)$ is as follows. If $(\mathcal{S}, \mathbf{p})$ is not more Blackwell-informative than $(\mathcal{T}, \mathbf{q})$, then the set $\mathcal{P}$ and the set $\mathcal{Q}$ defined above are nonempty, disjoint, convex, and compact. Thus they can be separated by a hyperplane, whose normal vector is interpreted as a utility index $\mathbf{u}$ after normalization. Figure 1 illustrates the geometric relations among state-utility vectors $\mathbf{u}_{\mathbf{qI}} - \epsilon\mathbf{1}_n$, $b\mathbf{1}_n$ and $\mathbf{u}_{\mathbf{pf}^*}$. By quasi-concavity and monotonicity of aggregator I, we can find two hyperplanes passing vectors $b\mathbf{1}_n$ and $\mathbf{u}_{\mathbf{qI}} - \epsilon\mathbf{1}_n$ and supporting their convex upper contour sets. By inequality (2), the vector $\mathbf{u}_{\mathbf{pf}^*}$ lies below the hyperplane supporting the upper contour set of vector $b\mathbf{1}_n$. Thus $I(\mathbf{u}_{\mathbf{pf}^*}) \leq I(b\mathbf{1}_n)$. Similarly, the vector $b\mathbf{1}_n$ must lie below the supporting hyperplane at vector $\mathbf{u}_{\mathbf{qI}} - \epsilon\mathbf{1}_n$ and hence $I(\mathbf{u}_{\mathbf{qI}} - \epsilon\mathbf{1}_n) \geq I(b\mathbf{1}_n)$. The value of experiment $(\mathcal{S}, \mathbf{p})$ is $I(\mathbf{u}_{\mathbf{pf}^*})$. Yet $I(\mathbf{u}_{\mathbf{qI}} - \epsilon\mathbf{1}_n)$ is strictly less than the value of experiment $(\mathcal{T}, \mathbf{q})$ as ϵ is positive and $\mathbf{I}$ is a feasible strategy under the experiment $(\mathcal{T}, \mathbf{q})$. As a result, experiment $(\mathcal{T}, \mathbf{q})$ is strictly more valuable than experiment $(\mathcal{S}, \mathbf{p})$, which contradicts inequality (1) in (ii).

5 Special Cases

Uncertainty averse preferences nest many ambiguity averse preferences as special cases. Below we give the DM's ex-ante valuation of an experiment $(\mathcal{S}, \mathbf{p})$ for six subfamilies of uncertainty averse preferences.

1. Variational preferences (Maccheroni et al. 2006a)

$$U^V(\mathcal{S}, \mathbf{p}) = \max_{\mathbf{f}} \min_{\pi\in\Delta(\Omega)} \langle D^{\pi}\mathbf{pf}, \mathbf{u}\rangle + c(\pi),$$

   where the cost function $c : \Delta(\Omega) \mapsto [0, \infty]$ is convex, lower-semi continuous, and $c^{-1}(0) \neq \emptyset$.

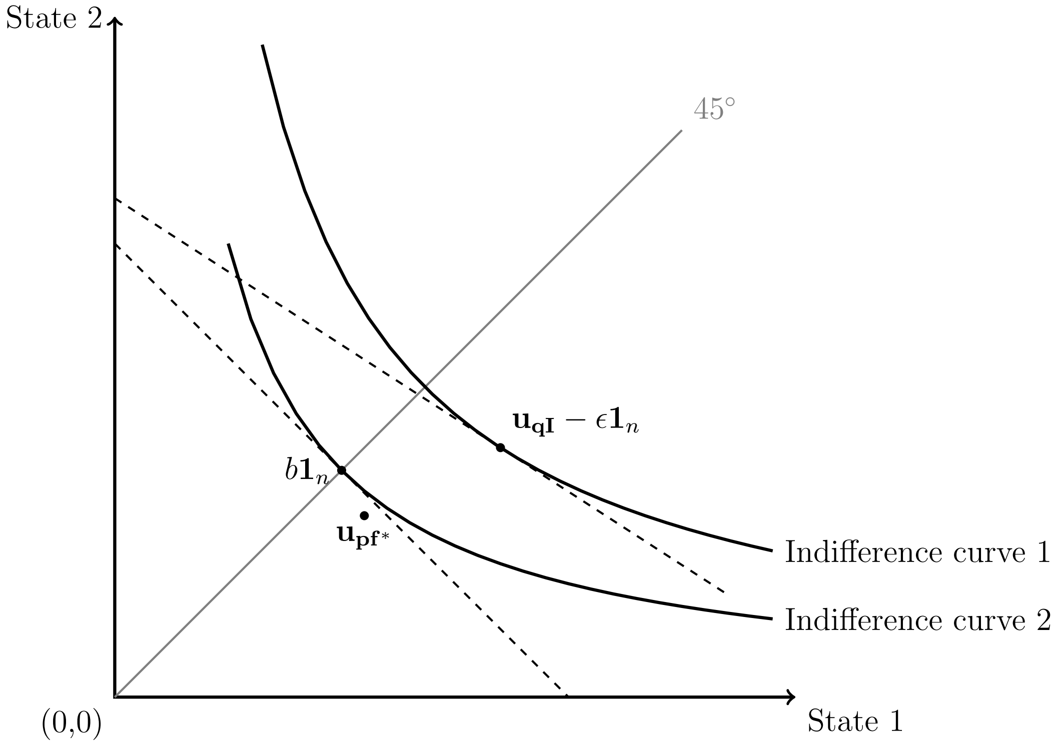

Figure 1: Supporting hyperplanes.

2. Maxmin EU (Gilboa and Schmeidler 1989)

$$U^M(\mathcal{S}, \mathbf{p}) = \max_{\mathbf{f}} \min_{\pi \in C} \langle D^\pi \mathbf{pf}, \mathbf{u} \rangle,$$

   where the prior set $C \subseteq \Delta(\Omega)$ is convex and closed.

3. Multiplier preferences (Hansen and Sargent 2001; Strzalecki 2011)

$$U^{MP}(\mathcal{S}, \mathbf{p}) = \max_{\mathbf{f}} \min_{\pi \in \Delta(\Omega)} \langle D^\pi \mathbf{pf}, \mathbf{u} \rangle + \theta R(\pi || \pi_0),$$

   where $\pi_0 \in \Delta(\Omega)$ is a reference prior, $\theta \in (0, +\infty]$ is the coefficient of ambiguity aversion, and $R(\cdot || \pi_0) : \Delta(\Omega) \mapsto [0, +\infty]$ is the relative entropy distance: $R(\pi || \pi_0) = \sum_i \pi_i \log \left(\frac{\pi_i}{\pi_{0i}} \right)$ if π is absolutely continuous with respect to π_0, and $+\infty$ otherwise.

4. Confidence preferences (Chateauneuf and Faro 2009). If $range(\mathbf{u}) = \mathbb{R}_+^{n \times k}$, then

$$U^C(\mathcal{S}, \mathbf{p}) = \max_{\mathbf{f}} \min_{\{\pi : \phi(\pi) \geq \alpha\}} \frac{1}{\phi(\pi)} \langle D^\pi \mathbf{pf}, \mathbf{u} \rangle,$$

   where the confidence level $\alpha \in (0, 1)$ and the confidence function $\phi : \Delta(\Omega) \mapsto [0, 1]$ is quasi-concave, upper semi-continuous, and $\phi(\pi) = 1$ for some $\pi \in \Delta(\Omega)$.

5. Smooth preferences (Klibanoff et al. 2005)

$$U^S(\mathcal{S},\mathbf{p}) = \max_{\mathbf{f}} \phi^{-1}\left(\int_{\Delta(\Omega)} \phi(\langle D^\pi \mathbf{pf}, \mathbf{u}\rangle) d\mu(\pi)\right),$$

where $\mu \in \Delta(\Delta(\Omega))$ is a second-order prior, and the function $\phi : \mathbb{R} \mapsto \mathbb{R}$, capturing ambiguity attitudes, is continuous, strictly increasing, and concave. Let $\Delta(\Delta(\Omega),\mu)$ denote the set of second-order priors that are absolutely continuous with respect to μ.

6. Second-order expected utility (Grant et al. 2009) [14]

$$U^{SO}(\mathcal{S},\mathbf{p}) = \max_{\mathbf{f}} \phi^{-1}\langle D^\pi \mathbf{p}, \phi(\mathbf{uf}')\rangle,$$

where $\phi : \mathbb{R} \mapsto \mathbb{R}$ is continuous, concave, and strictly increasing function. Let $\Delta(\Omega,\pi)$ be the set of priors that are absolutely continuous with respect to π.

As a direct application of Theorem 1, the following Corollary holds:

Corollary 1. *Blackwell's equivalence results hold for the following preference families: Variational preferences (VP), Maxmin EU (MEU), Multiplier preferences (MP), Confidence preferences (CP), Smooth preferences (SP), and Second-order expected utility (SOEU), with Assumption 1 taking the form specified in Table 1.*

Value of Experiment $(\mathcal{S},\mathbf{p})$	Assumption 1	Blackwell Equivalence
VP: $\max_{\mathbf{f}} \min_{\pi\in\Delta(\Omega)}\langle D^\pi \mathbf{pf}, \mathbf{u}\rangle + c(\pi)$	$\exists \pi_0 \in int(\Delta(\Omega)) \cap c^{-1}(0)$	$\surd$
MEU: $\max_{\mathbf{f}} \min_{\pi\in C}\langle D^\pi \mathbf{pf}, \mathbf{u}\rangle$	$\exists \pi_0 \in int(\Delta(\Omega)) \cap C$	$\surd$
MP: $\max_{\mathbf{f}} \min_{\pi\in\Delta(\Omega)}\langle D^\pi \mathbf{pf}, \mathbf{u}\rangle + \theta R(\pi\|\|\pi_0)$	$\pi_0 \in int(\Delta(\Omega))$	$\surd$
CP: $\max_{\mathbf{f}} \min_{\{\pi:\phi(\pi)\geq\alpha\}} \frac{1}{\phi(\pi)}\langle D^\pi \mathbf{pf}, \mathbf{u}\rangle$	$\exists \pi_0 \in int(\Delta(\Omega)) \cap \phi^{-1}(1)$	$\surd$
SP: $\max_{\mathbf{f}} \phi^{-1}\left(\int_{\Delta(\Omega)} \phi(\langle D^\pi \mathbf{pf}, \mathbf{u}\rangle) d\mu(\pi)\right)$	$\pi_0 = \int \pi d\mu(\pi) \in int(\Delta(\Omega))$	$\surd$
SOEU: $\max_{\mathbf{f}} \phi^{-1}\langle D^{\pi_0} \mathbf{p}, \phi(\mathbf{uf}')\rangle$	$\pi_0 \in int(\Delta(\Omega))$	$\surd$

Table 1: Preference families and assumptions under which Blackwell's equivalence results hold.

Proof of Corollary 1: We prove the equivalence results case by case. It suffices to verify Assumption 1 for each preferences family. Then the Blackwell's equivalence results follow from Theorem 1.

Variational preferences. Let $\mathcal{X} = \mathbb{R}$ and $G(x,\pi) = x + c(\pi)$. Then for $\pi_0 \in int(\Delta(\Omega)) \cap c^{-1}(0)$, $G(x,\pi_0) = x$ and $G(\cdot,\pi_0)$ is strictly increasing and continuous.

[14] See also Neilson (2010), Ergin and Gul (2009), Nau (2006).

Maxmin EU. Let $\mathcal{X} = \mathbb{R}$ and $G(x, \pi) = \begin{cases} x, & \text{if } \pi \in C; \\ +\infty, & \text{otherwise.} \end{cases}$ Then for $\pi_0 \in (C \cap int(\Delta(\Omega)))$, $G(x, \pi_0) = x$ and $G(\cdot, \pi_0)$ is strictly increasing and continuous.

Multiplier preferences. Let $\mathcal{X} = \mathbb{R}$ and $G(x, \pi) = x + \theta R(\pi || \pi_0)$. Since $R(\pi_0 || \pi_0) = 0$, $G(x, \pi_0) = x$ and $G(\cdot, \pi_0)$ is strictly increasing and continuous.

Confidence preferences. Let $\mathcal{X} = \mathbb{R}_+$ and $G(x, \pi) = \begin{cases} \frac{x}{\phi(\pi)}, & \text{if } \phi(\pi) \geq \alpha; \\ +\infty, & \text{otherwise.} \end{cases}$ By assumption $G(x, \pi_0) = x$, so $G(\cdot, \pi_0)$ is strictly increasing and continuous on $\mathbb{R}_+$.

Smooth preferences. Let $\mathcal{X} = \mathbb{R}$. By Cerreia-Vioglio et al. (2011b) Theorem 19, smooth ambiguity representation (ϕ, μ) is equivalent to an uncertainty averse representation with

$$G(x, \pi) = \begin{cases} x + \inf_{\nu \in \Gamma(\pi)} I_x(\nu || \mu), & \text{if } \Gamma(\pi) \neq \emptyset; \\ +\infty, & \text{if } \Gamma(\pi) = \emptyset, \end{cases}$$

where the set $\Gamma(\pi) = \{\nu \in \Delta(\Delta(\Omega), \mu) : \int \pi' d\nu(\pi') = \pi\}$. Here $I_x(\cdot || \mu) : \Delta(\Delta(\Omega)) \mapsto [0, +\infty]$ is a statistical distance function with the property that $I_x(\mu || \mu) = 0$ and $I_x(\nu || \mu) \geq 0$ for all $\nu \in \Delta(\Delta(\Omega), \mu)$.[15] Clearly $\mu \in \Gamma(\pi_0)$ and $\inf_{\nu \in \Gamma(\pi_0)} I_x(\nu || \mu) = 0$. Therefore $G(x, \pi_0) = x$ and $G(\cdot, \pi_0)$ is strictly increasing and continuous on $\mathbb{R}$.

Second-order expected utility. Let $\mathcal{X} = \mathbb{R}$. By Cerreia-Vioglio et al. (2011b) Theorem 24, the second order expected utility representation (ϕ, π_0) is equivalent to an uncertainty averse representation with

$$G(x, \pi) = \begin{cases} x + I_x(\pi || \pi_0), & \text{if } \pi \in \Delta(\Omega, \pi_0); \\ +\infty, & \text{otherwise.} \end{cases}$$

Again $I_x(\cdot || \pi_0) : \Delta(\Omega) \mapsto [0, +\infty]$ is a statistical function such that $I_x(\pi_0 || \pi_0) = 0$.[16] Thus $G(x, \pi_0) = x$ and $G(\cdot, \pi_0)$ is strictly increasing and continuous on $\mathbb{R}$. $\square$

[15]The statistical distance function is

$$I_x(\nu || \mu) \quad = \quad \phi^{-1} \left(\inf_{k \geq 0} \left[kx - \int \phi^*(k \frac{d\nu}{d\mu}) d\mu \right] \right) - x,$$

where $\phi^*(z) = \inf_{k \in \mathbb{R}} (kz - \phi(k))$ is the concave conjugate function of ϕ.

[16]Similarly, the statistical distance function is

$$I_x(\pi' || \pi) \quad = \quad \phi^{-1} \left(\inf_{k \geq 0} \left[kx - \int \phi^*(k \frac{d\pi'}{d\pi}) d\pi \right] \right) - x,$$

where $\phi^*(z) = \inf_{k \in \mathbb{R}} (kz - \phi(k))$ is the concave conjugate function of ϕ.

6 Discussion

6.1 Non-convex Preferences

Our main theorem says that Blackwell's equivalence theorem holds for all convex preferences. Below we give an example when this equivalence fails with non-convex preferences.

Consider a DM who perceives extreme ambiguity about ω and views every prior in $\Delta(\Omega)$ as possible. She is also optimistic and evaluates a strategy by the best case scenario.[17] So her ex-ante evaluation of an experiment $(\mathcal{S}, \mathbf{p})$ is described by the following maxmax EU:

$$U^{MM}(\mathcal{S}, \mathbf{p}) := \max_{\mathbf{f}} \max_{\pi \in \Delta(\Omega)} \langle D^{\pi} \mathbf{p} \mathbf{f}, \mathbf{u} \rangle = \max_{\omega, a} u(\omega, a).$$

Then for any $\mathbf{u}$, the values of two arbitrary experiments $(\mathcal{S}, \mathbf{p})$ and $(\mathcal{T}, \mathbf{q})$ are equal, i.e.,

$$U^{MM}(\mathcal{S}, \mathbf{p}) = U^{MM}(\mathcal{T}, \mathbf{q}) = \max_{\omega, a} u(\omega, a).$$

So Blackwell's equivalence theorem fails.

6.2 Full Support Assumption

Assumption 1 requires a fully supported π_0 such that $G(b, \pi_0) = b$ for some b. Although only used in proving the $(ii) \Rightarrow (i)$ direction of Theorem 1, it is not redundant. To see this, consider an MEU DM with a singleton prior set $C = \{\hat{\pi}\}$, where $\hat{\pi} = (1, \cdots, 0)$. The full support assumption is violated. For any experiment $(\mathcal{S}, \mathbf{p})$ and any utility index $\mathbf{u}$, we have

$$U^{M}(\mathcal{S}, \mathbf{p}) = \max_{\mathbf{f}} \min_{\pi \in C} \langle D^{\pi} \mathbf{p} \mathbf{f}, \mathbf{u} \rangle = \max_{\mathbf{f}} \langle D^{\hat{\pi}} \mathbf{p} \mathbf{f}, \mathbf{u} \rangle = \max_{a} u(\omega_1, a).$$

Thus *any* two experiments $(\mathcal{S}, \mathbf{p})$ and $(\mathcal{T}, \mathbf{q})$ have the same value, i.e., $U^{M}(\mathcal{S}, \mathbf{p}) = U^{M}(\mathcal{T}, \mathbf{q})$ for all $\mathbf{u}$. But we cannot say that $(\mathcal{S}, \mathbf{p})$ is more informative than $(\mathcal{T}, \mathbf{q})$, as they are arbitrarily chosen.

[17]Note that the full support assumption still holds as any interior prior is possible.

6.3 No-Commitment Case

Until now we have assumed that the DM has full commitment to all signal-contingent strategies. Without commitment, our results do not apply in general. For instance, Wakker (1988), Hilton (1990), and Safra and Sulganik (1995) show that Blackwell's theorem might fail for non-EU DMs, if choices are only made after the signal realizes. Siniscalchi (2011: Section 4.4.2) illustrates how a sophisticated MEU DM might reject freely available information if there is no commitment. The reason is that ambiguity sensitive preferences might not be dynamically consistent. When a DM is dynamically inconsistent and anticipates potential preference changes, she might reject free information in order to commit to her ex-ante optimal action. In this case, the pure decision value of information is traded off against the value of commitment. We will explore such a trade-off for general ambiguity preferences under the Blackwell-type probabilistic information in future work. On the other hand, there are well-known specifications of ambiguity sensitive and dynamically consistent preferences. For those preferences, our results apply even without commitment. In particular, Epstein and Schneider (2003) characterize dynamic consistency by a rectangularity condition for the MEU preferences. And Maccheroni et al. (2006b) show that dynamic consistency is equivalent to a "no-gain condition" for the variational preferences. Below we specify these conditions for our problem.

For a fixed state space Ω, an experiment $(\mathcal{S}, \mathbf{p})$ introduces a two-period dynamic problem with a product state space $\mathcal{S} \times \Omega$. By the end of period 1, information $\{\{s\} \times \Omega : s \in \mathcal{S}\}$ is revealed. A prior $\pi \in \Delta(\Omega)$ induces a joint probability $P = D^\pi \mathbf{p} \in \Delta(\mathcal{S} \times \Omega)$. Let $\mathbf{m} = \mathbf{p}'\pi \in \Delta(\mathcal{S})$ denote the marginal probability on signals with $m_s = \sum_\omega \pi_\omega p_{\omega s}$, and $\pi_{\cdot s} \in \Delta(\Omega)$ denote the Bayesian posterior conditional on signal s.

MEU Preferences. A convex and closed prior set $C \subseteq \Delta(\Omega)$ induces a set of joint probabilities $\mathcal{P} = \{D^\pi \mathbf{p} : \pi \in C\}$ and a set of marginal probabilities $\mathcal{M} = \{\mathbf{m} = \mathbf{p}'\pi \in \Delta(\mathcal{S}) : \pi \in C\}$. The set of signal-contingent posterior matrices is $C_{\cdot\mathcal{S}} = \{\Pi \in \mathbb{R}_{n\times|S|} :$ each column $\Pi_{\cdot s}$ equals $\pi^s_{\cdot s}$ for some $\pi^s \in C\}$. Thus the $\mathcal{S}$-rectangularized (Epstein and Schneider 2003) set of joint probabilities is

$$rect(\mathcal{P})_{\mathcal{S},\mathbf{p}} = \{\Pi D^{\mathbf{m}} : \mathbf{m} \in \mathcal{M}, \Pi \in C_{\cdot\mathcal{S}}\}.$$

We say the prior set C is $(\mathcal{S}, \mathbf{p})$-*rectangular* if

$$\mathcal{P} = rect(\mathcal{P})_{\mathcal{S},\mathbf{p}}.$$

If the prior set C is both $(\mathcal{S}, \mathbf{p})$-rectangular and $(\mathcal{T}, \mathbf{q})$-rectangular, our results extend to the case without commitment.

Variational Preferences. Given a convex, grounded, and lower semi-continuous cost function $\tilde{c} : \Delta(\mathcal{S} \times \Omega) \mapsto [0, +\infty]$. Let $c_s : \Delta(\Omega) \mapsto [0, +\infty]$ be the updated cost function conditional on signal $s \in \mathcal{S}$. Then $\tilde{c}$ and $\{c_s\}_{s \in S}$ satisfy the "no-gain condition" with discount factor 1 (Maccheroni et al. 2006b) if

$$\tilde{c}(D^{\pi}\mathbf{p}) = \min_{\{\tilde{\pi} \in \Delta(\Omega) : \mathbf{p}'\tilde{\pi} = \mathbf{m}\}} \tilde{c}(D^{\tilde{\pi}}\mathbf{p}) + \sum_s m_s c_s(\pi_{\cdot s}).$$

To further specify the conditional cost function c_s, one could use an updating formula shown by Li (2013)

$$c_s(\pi_{\cdot s}) = \min_{\{\tilde{\pi} \in \Delta(\Omega) : \tilde{\pi}_{\cdot s} = \pi_{\cdot s}\}} \frac{\tilde{c}(D^{\tilde{\pi}}\mathbf{p})}{\tilde{m}_s},$$

where $\tilde{\pi}_{\cdot s}$ and $\tilde{m}_s$ are the s-Bayesian posterior and the s-marginal probability of prior $\tilde{\pi}$.

Moreover, since the signal structure $(\mathcal{S}, \mathbf{p})$ is considered unambiguous, we assume there exists a cost function $c : \Delta(\Omega) \mapsto [0, +\infty]$ such that $\tilde{c}(D^{\pi}\mathbf{p}) = c(\pi)$ for all $\pi \in \Delta(\Omega)$. Then the "no-gain condition" and updating rule become

$$\begin{aligned} c(\pi) &= \min_{\{\tilde{\pi} \in \Delta(\Omega) : \mathbf{p}'\tilde{\pi} = \mathbf{m}\}} c(\tilde{\pi}) + \sum_s m_s c_s(\pi_{\cdot s}), \quad \text{and} \\ c_s(\pi_{\cdot s}) &= \min_{\{\tilde{\pi} \in \Delta(\Omega) : \tilde{\pi}_{\cdot s} = \pi_{\cdot s}\}} \frac{c(\tilde{\pi})}{\tilde{m}_s}. \end{aligned}$$

If the cost function c satisfies the "no-gain condition" for experiments $(\mathcal{S}, \mathbf{p})$ and $(\mathcal{T}, \mathbf{q})$, our results extend to the case without commitment.

Appendix

In Section 5, we give proofs for the smooth preferences and second-order expected utility cases via their corresponding G function representations. They rely heavily on Theorems 19 and 24 in Cerreia-Vioglio et al. (2011b), which are not obvious. Hence we provide direct proofs in the appendix, which may be of independent interest.

A Smooth preferences

Proof. We want to show that $\mathbf{q} = \mathbf{pr}$ for some markov matrix $\mathbf{r}$ if and only if

$$\max_{\mathbf{f}} \phi^{-1}\left(\int_{\Delta(\Omega)} \phi(\langle D^{\pi}\mathbf{pf}, \mathbf{u}\rangle) d\mu(\pi)\right) \geq \max_{\mathbf{g}} \phi^{-1}\left(\int_{\Delta(\Omega)} \phi(\langle D^{\pi}\mathbf{qg}, \mathbf{u}\rangle) d\mu(\pi)\right), \quad \forall \mathbf{u}. \tag{5}$$

"Only if" part: For any μ, let $\mathbf{g}^*$ be the strategy that maximizes the RHS of inequality 5. Let $\mathbf{f}^* = \mathbf{r}\mathbf{g}^*$. $\mathbf{f}^*$ is a strategy for experiment $(\mathcal{S}, \pi)$ because $\mathbf{r}$ is markov. Thus

$$\begin{aligned}
\max_{\mathbf{g}} \phi^{-1}\left(\int_{\Delta(\Omega)} \phi(\langle D^\pi \mathbf{q}\mathbf{g}, \mathbf{u}\rangle) d\mu(\pi)\right) &= \phi^{-1}\left(\int_{\Delta(\Omega)} \phi(\langle D^\pi \mathbf{q}\mathbf{g}^*, \mathbf{u}\rangle) d\mu(\pi)\right) \\
&= \phi^{-1}\left(\int_{\Delta(\Omega)} \phi(\langle D^\pi \mathbf{p}\mathbf{r}\mathbf{g}^*, \mathbf{u}\rangle) d\mu(\pi)\right) \\
&= \phi^{-1}\left(\int_{\Delta(\Omega)} \phi(\langle D^\pi \mathbf{p}\mathbf{f}^*, \mathbf{u}\rangle) d\mu(\pi)\right) \\
&\leq \max_{\mathbf{f}} \phi^{-1}\left(\int_{\Delta(\Omega)} \phi(\langle D^\pi \mathbf{p}\mathbf{f}, \mathbf{u}\rangle) d\mu(\pi)\right).
\end{aligned}$$

"If" part: Suppose there is no $\mathbf{r}$ such that $\mathbf{p} = \mathbf{q}\mathbf{r}$. By assumption, $\pi_0 = \int \pi d\mu(\pi)$ has full support. Define $\mathcal{P} = \{D^{\pi_0}\mathbf{p}\mathbf{r} : \text{for some markov matrix } \mathbf{r}\}$ and $\mathcal{Q} = \{D^{\hat{\pi}}\mathbf{q} : \hat{\pi} \in \Delta(\Omega)\}$. Then $\mathcal{P}$ and $\mathcal{Q}$ are nonempty, convex, compact and $\mathcal{P} \cap \mathcal{Q} = \emptyset$. By the Separating Hyperplane Theorem, there exists $\mathbf{v} \neq 0$ such that $\langle \mathbf{n}, \mathbf{v}\rangle > 0 > \langle \mathbf{m}, \mathbf{v}\rangle$ for all $\mathbf{n} \in \mathcal{Q}$ and $\mathbf{m} \in \mathcal{P}$. We can show that for this $\mathbf{v}$,

$$\max_{\mathbf{f}} \langle D^{\pi_0}\mathbf{p}\mathbf{f}, \mathbf{v}\rangle < 0 < \max_{\mathbf{g}} \langle D^\pi \mathbf{q}\mathbf{g}, \mathbf{v}\rangle, \quad \forall \pi. \tag{6}$$

Assume $\phi(0) = 0$ and $\phi'(0) = 1$. This is WLOG because $\phi(\cdot)$ is unique up to a positive affine transformation.[18] Let $M_0 := \max_{i,j} |v_{ij}|$. We first claim that there exists a positive constant M_1 such that

$$|\phi(t) - t| \leq M_1 t^2, \qquad \forall t \in [-M_0, M_0].$$

(For example, pick $M_1 = \frac{1}{2}\max_{t\in[-M_0,M_0]} |\phi''(t)|$.)

For any strategy $\mathbf{f}$ and any $\epsilon \in (0,1)$, $|\langle D^\pi \mathbf{p}\mathbf{f}, \epsilon\mathbf{v}\rangle| \leq \epsilon \max_{ij} |v_{ij}| = \epsilon M_0$, therefore

$$\begin{aligned}
&|\max_{\mathbf{f}} \int_{\Delta(\Omega)} \phi(\langle D^\pi \mathbf{p}\mathbf{f}, \epsilon\mathbf{v}\rangle) d\mu(\pi) - \max_{\mathbf{f}} \int_{\Delta(\Omega)} \langle D^\pi \mathbf{p}\mathbf{f}, \epsilon\mathbf{v}\rangle d\mu(\pi)| \qquad (7) \\
\leq\ & \max_{\mathbf{f}} \int_{\Delta(\Omega)} |\phi(\langle D^\pi \mathbf{p}\mathbf{f}, \epsilon\mathbf{v}\rangle) - \langle D^\pi \mathbf{p}\mathbf{f}, \epsilon\mathbf{v}\rangle| d\mu(\pi) = \max_{\mathbf{f}} \int_{\Delta(\Omega)} M_1(\epsilon M_0)^2 d\mu(\pi) = M_1(\epsilon M_0)^2.
\end{aligned}$$

Similarly,

$$|\max_{\mathbf{g}} \int_{\Delta(\Omega)} \phi(\langle D^\pi \mathbf{q}\mathbf{g}, \epsilon\mathbf{v}\rangle) d\mu(\pi) - \max_{\mathbf{g}} \int_{\Delta(\Omega)} \langle D^\pi \mathbf{q}\mathbf{g}, \epsilon\mathbf{v}\rangle d\mu(\pi)| \leq M_1(\epsilon M_0)^2. \tag{8}$$

[18] We assume ϕ is twice continuously differentiable around 0.

Moreover,

$$\max_{\mathbf{f}} \int_{\Delta(\Omega)} \langle D^{\pi}\mathbf{pf}, \epsilon\mathbf{v}\rangle d\mu(\pi) = \max_{\mathbf{f}} \langle D^{\int_{\Delta(\Omega)} \pi d\mu(\pi)}\mathbf{pf}, \epsilon\mathbf{v}\rangle = \epsilon \max_{\mathbf{f}} \langle D^{\pi_0}\mathbf{pf}, \mathbf{v}\rangle \tag{9}$$

and similarly

$$\max_{\mathbf{g}} \int_{\Delta(\Omega)} \langle D^{\pi}\mathbf{pg}, \epsilon\mathbf{v}\rangle d\mu(\pi) = \epsilon \max_{\mathbf{g}} \langle D^{\pi_0}\mathbf{qg}, \mathbf{v}\rangle. \tag{10}$$

Define $\delta := \max_{\mathbf{g}} \langle D^{\pi_0}\mathbf{qg}, \mathbf{v}\rangle - \max_{\mathbf{f}} \langle D^{\pi_0}\mathbf{pf}, \mathbf{v}\rangle$. Clearly $\delta > 0$ by equation (6). Let

$$\bar{\epsilon} := \min\left(1, \frac{\delta}{2M_1 M_0^2}\right) > 0.$$

Pick any ϵ satisfying $0 < \epsilon < \bar{\epsilon}$. Then by the triangular inequality and equations (7)–(10), we have

$$\begin{aligned}
& \max_{\mathbf{f}} \int_{\Delta(\Omega)} \phi(\langle D^{\pi}\mathbf{pf}, \epsilon\mathbf{v}\rangle) d\mu(\pi) - \max_{\mathbf{g}} \int_{\Delta(\Omega)} \phi(\langle D^{\pi}\mathbf{qg}, \epsilon\mathbf{v}\rangle) d\mu(\pi) \qquad (11) \\
\leq \quad & \max_{\mathbf{f}} \int_{\Delta(\Omega)} \langle D^{\pi}\mathbf{pf}, \epsilon\mathbf{v}\rangle d\mu(\pi) - \max_{\mathbf{g}} \int_{\Delta(\Omega)} \langle D^{\pi}\mathbf{pg}, \epsilon\mathbf{v}\rangle d\mu(\pi) \\
& + \left| \max_{\mathbf{f}} \int_{\Delta(\Omega)} \phi(\langle D^{\pi}\mathbf{pf}, \epsilon\mathbf{v}\rangle) d\mu(\pi) - \max_{\mathbf{f}} \int_{\Delta(\Omega)} \langle D^{\pi}\mathbf{pf}, \epsilon\mathbf{v}\rangle d\mu(\pi) \right| \\
& + \left| \max_{\mathbf{g}} \int_{\Delta(\Omega)} \phi(\langle D^{\pi}\mathbf{qg}, \epsilon\mathbf{v}\rangle) d\mu(\pi) - \max_{\mathbf{g}} \int_{\Delta(\Omega)} \langle D^{\pi}\mathbf{qg}, \epsilon\mathbf{v}\rangle d\mu(\pi) \right| \\
\leq \quad & -\epsilon\delta + M_1 M_0^2 \epsilon^2 + M_1 M_0^2 \epsilon^2 = -\epsilon(\delta - 2M_1 M_0^2 \epsilon) < 0.
\end{aligned}$$

Define $A := \{a_1, \cdots, a_{|T|}\}$ and $\mathbf{u} := \epsilon\mathbf{v}$. Then

$$\max_{\mathbf{f}} \phi^{-1}\left(\int_{\Delta(\Omega)} \phi(\langle D^{\pi}\mathbf{pf}, \mathbf{u}\rangle) d\mu(\pi)\right) < \max_{\mathbf{g}} \phi^{-1}\left(\int_{\Delta(\Omega)} \phi(\langle D^{\pi}\mathbf{qg}, \mathbf{u}\rangle) d\mu(\pi)\right),$$

which contradicts equation (5). □

B Second-order expected utility

Proof. We want to show that $\mathbf{q} = \mathbf{pr}$ for some markov matrix $\mathbf{r}$ if and only if

$$\max_{\mathbf{f}} \phi^{-1} \langle D^{\pi_0}\mathbf{p}, \phi(\mathbf{uf}')\rangle \geq \max_{\mathbf{g}} \phi^{-1} \langle D^{\pi_0}\mathbf{q}, \phi(\mathbf{ug}')\rangle, \quad \forall \mathbf{u}. \tag{12}$$

“Only if” part: Suppose $\mathbf{q} = \mathbf{pr}$. For any strategy $\mathbf{g} : \mathcal{T} \mapsto \Delta(X)$, define $\mathbf{f} = \mathbf{rg}$. First, we show that

$$\phi^{-1}\langle D^{\pi_0}\mathbf{p}, \phi(\mathbf{uf}')\rangle|_{\mathbf{f}=\mathbf{rg}} \geq \phi^{-1}\langle D^{\pi_0}\mathbf{q}, \phi(\mathbf{ug}')\rangle$$

This follows from the fact that

$$\begin{aligned}
\langle D^{\pi_0}\mathbf{p}, \phi(\mathbf{uf}')\rangle|_{\mathbf{f}=\mathbf{rg}} &= \langle D^{\pi_0}\mathbf{p}, \phi(\mathbf{u}(\mathbf{rg})')\rangle = \langle D^{\pi_0}\mathbf{p}, \phi(\mathbf{ug}'\mathbf{r}')\rangle \\
&\geq \langle D^{\pi_0}\mathbf{p}, \phi(\mathbf{ug}')\mathbf{r}'\rangle = \langle D^{\pi_0}\mathbf{pr}, \phi(\mathbf{ug}')\rangle = \langle D^{\pi_0}\mathbf{q}, \phi(\mathbf{ug}')\rangle,
\end{aligned}$$

where in the inequality step we use Jensen’s inequality as ϕ is concave, and each column of $\mathbf{r}'$ is nonnegative and adds up to one (recall that $\mathbf{r}$ is row-stochastic, so its transpose is column-stochastic).

As a result,

$$\begin{aligned}
\max_{\mathbf{f}} \phi^{-1}\langle D^{\pi_0}\mathbf{p}, \phi(\mathbf{uf}')\rangle &\geq \phi^{-1}\langle D^{\pi_0}\mathbf{p}, \phi(\mathbf{uf}')\rangle|_{\mathbf{f}=\mathbf{r}\hat{\mathbf{g}}} \\
&\geq \phi^{-1}\langle D^{\pi_0}\mathbf{q}, \phi(\mathbf{u}\hat{\mathbf{g}}')\rangle = \max_{\mathbf{g}} \phi^{-1}\langle D^{\pi_0}\mathbf{q}, \phi(\mathbf{ug}')\rangle,
\end{aligned}$$

where $\hat{\mathbf{g}} = \arg\max_{\mathbf{g}} \phi^{-1}\langle D^{\pi_0}\mathbf{q}, \phi(\mathbf{ug}')\rangle$.

“If” part: This part is similar to that in the proof of the smooth preference case.

Suppose there is no $\mathbf{r}$ such that $\mathbf{p} = \mathbf{qr}$. By assumption, the reference prior π_0 has full support. Define $\mathcal{P} = \{D^{\pi_0}\mathbf{pr} : \text{for some markov } \mathbf{r}\}$ and $\mathcal{Q} = \{D^{\hat{\pi}}\mathbf{q} : \hat{\pi} \in \Delta(\Omega)\}$. These are nonempty, convex, compact and $\mathcal{P} \cap \mathcal{Q} = \emptyset$. By the Separating Hyperplane Theorem, there exists $\mathbf{v} \neq 0$ such that $\langle \mathbf{n}, \mathbf{v}\rangle > 0 > \langle \mathbf{m}, \mathbf{v}\rangle$ for all $\mathbf{n} \in \mathcal{Q}$ and $\mathbf{m} \in \mathcal{P}$. By a similar argument, we can show that for this $\mathbf{v}$,

$$\max_{\mathbf{f}}\langle D^{\pi_0}\mathbf{p}, \mathbf{vf}'\rangle < 0 < \max_{\mathbf{g}}\langle D^{\pi_0}\mathbf{q}, \mathbf{vg}'\rangle. \tag{13}$$

Assume $\phi(0) = 0$ and $\phi'(0) = 1$. This is WLOG because $\phi(\cdot)$ is unique up to a positive affine transformation. Let $M_0 := \max_{i,j}|v_{ij}|$. We first claim that there exists a positive constant M_1 such that

$$|\phi(t) - t| \leq M_1 t^2, \qquad \forall t \in [-M_0, M_0].$$

(For example, pick $M_1 = \frac{1}{2}\max_{t\in[-M_0,M_0]}|\phi''(t)|$.)

For any strategy $\mathbf{f}$ and any $\epsilon \in (0,1)$, each entry of $\epsilon\mathbf{vf}'$ is bounded by ϵM_0, and therefore

$$\max_{i,j}|\phi(\epsilon\mathbf{vf}') - \epsilon\mathbf{vf}'|_{ij} \leq M_1(\max_{i,j}|\epsilon\mathbf{vf}'|_{ij})^2 \;\; \leq \;\; M_1 M_0^2\epsilon^2. \tag{14}$$

Thus

$$|\langle D^{\pi_0}\mathbf{p}, \phi(\epsilon\mathbf{vf'})\rangle - \langle D^{\pi_0}\mathbf{p}, \epsilon\mathbf{vf'}\rangle| \leq \sum_{ij} \pi_{0i}p_{ij}|\phi(\epsilon\mathbf{vf'}) - \epsilon\mathbf{vf'}|_{ij} \leq M_1 M_0^2 \epsilon^2 \sum_{ij} \pi_{0i}p_{ij} = M_1 M_0^2 \epsilon^2. \tag{15}$$

Similarly, for any strategy $\mathbf{g}$,

$$|\langle D^{\pi_0}\mathbf{q}, \phi(\epsilon\mathbf{vg'})\rangle - \langle D^{\pi_0}\mathbf{q}, \epsilon\mathbf{vg'}\rangle| \leq M_1 M_0^2 \epsilon^2. \tag{16}$$

Define $\delta := \max_{\mathbf{g}}\langle D^{\pi_0}\mathbf{q}, \mathbf{vg'}\rangle - \max_{\mathbf{f}}\langle D^{\pi_0}\mathbf{p}, \mathbf{vf'}\rangle$. Clearly $\delta > 0$ by equation (13). Let $\bar{\epsilon} := \min\left(1, \frac{\delta}{2M_1M_0^2}\right) > 0$. Then for any ϵ satisfying $0 < \epsilon < \bar{\epsilon}$, we have

$$\begin{aligned}
& \max_{\mathbf{g}}\langle D^{\pi_0}\mathbf{q}, \phi(\epsilon\mathbf{vg'})\rangle - \max_{\mathbf{f}}\langle D^{\pi_0}\mathbf{p}, \phi(\epsilon\mathbf{vf'})\rangle \\
= & \left(\max_{\mathbf{g}}\langle D^{\pi_0}\mathbf{q}, \phi(\epsilon\mathbf{vg'})\rangle - \max_{\mathbf{g}}\langle D^{\pi_0}\mathbf{q}, \epsilon\mathbf{vg'}\rangle\right) + \left(\max_{\mathbf{f}}\langle D^{\pi_0}\mathbf{p}, \epsilon\mathbf{vf'}\rangle - \max_{\mathbf{f}}\langle D^{\pi_0}\mathbf{p}, \phi(\epsilon\mathbf{vf'})\rangle\right) \\
& + \left(\max_{\mathbf{g}}\langle D^{\pi_0}\mathbf{q}, \epsilon\mathbf{vg'}\rangle - \max_{\mathbf{f}}\langle D^{\pi_0}\mathbf{p}, \epsilon\mathbf{vf'}\rangle\right).
\end{aligned}$$

By equation (16), the first term satisfies

$$\left(\max_{\mathbf{g}}\langle D^{\pi_0}\mathbf{q}, \phi(\epsilon\mathbf{vg'})\rangle - \max_{\mathbf{g}}\langle D^{\pi_0}\mathbf{q}, \epsilon\mathbf{vg'}\rangle\right) \geq -\max_{\mathbf{g}}|\langle D^{\pi_0}\mathbf{q}, \phi(\epsilon\mathbf{vg'})\rangle - \langle D^{\pi_0}\mathbf{q}, \epsilon\mathbf{vg'}\rangle| \geq -M_1M_0^2\epsilon^2.$$

Similarly, by equation (15), the second term satisfies

$$\left(\max_{\mathbf{f}}\langle D^{\pi_0}\mathbf{p}, \epsilon\mathbf{vf'}\rangle - \max_{\mathbf{f}}\langle D^{\pi_0}\mathbf{p}, \phi(\epsilon\mathbf{vf'})\rangle\right) \geq -M_1M_0^2\epsilon^2.$$

And the third term is

$$\left(\max_{\mathbf{g}}\langle D^{\pi_0}\mathbf{q}, \epsilon\mathbf{vg'}\rangle - \max_{\mathbf{f}}\langle D^{\pi_0}\mathbf{p}, \epsilon\mathbf{vf'}\rangle\right) = \epsilon\left(\max_{\mathbf{g}}\langle D^{\pi_0}\mathbf{q}, \mathbf{vg'}\rangle - \max_{\mathbf{f}}\langle D^{\pi_0}\mathbf{p}, \mathbf{vf'}\rangle\right) = \epsilon\delta.$$

As a consequence,

$$\max_{\mathbf{g}}\langle D^{\pi_0}\mathbf{q}, \phi(\epsilon\mathbf{vg'})\rangle - \max_{\mathbf{f}}\langle D^{\pi_0}\mathbf{p}, \phi(\epsilon\mathbf{vf'})\rangle \geq -M_1M_0^2\epsilon^2 - M_1M_0^2\epsilon^2 + \epsilon\delta = \epsilon(\delta - 2M_1M_0^2\epsilon) > 0.$$

Define $A := \{a_1, \cdots, a_{|T|}\}$, $\mathbf{u} := \epsilon\mathbf{v}$, then

$$\max_{\mathbf{g}} \phi^{-1}\langle D^{\pi_0}\mathbf{q}, \phi(\mathbf{ug'})\rangle > \max_{\mathbf{f}} \phi^{-1}\langle D^{\pi_0}\mathbf{p}, \phi(\mathbf{uf'})\rangle,$$

which contradicts inequality (12). $\square$

References

Abdellaoui, M., A. Baillon, L. Placido, and P. P. Wakker (2011). The rich domain of uncertainty: Source functions and their experimental implementation. *American Economic Review 101*(2), 695–723.

Athey, S. and J. Levin (2001). The value of information in monotone decision problems. Working Paper, MIT.

Blackwell, D. (1951). The comparison of experiments. In *Proceedings, Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 93102. Univ. of California Press, Berkeley.

Blackwell, D. (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics 24*(2), 265–272.

Blackwell, D. and M. A. Girshick (1954). *Theory of games and statistical decisions*. New York: Dover Publications.

Bossaerts, P., P. Ghirardato, S. Guarnaschelli, and W. R. Zame (2010). Ambiguity in asset markets: Theory and experiment. *Review of Financial Studies 23*(4), 1325–1359.

Cabrales, A., O. Gossner, and R. Serrano (2013). Entropy and the value of information for investors. *American Economic Review 103*(1), 360–77.

Camerer, C. and M. Weber (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty 5*(4), 325–70.

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

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2011a). Complete monotone quasiconcave duality. *Mathematics of Operations Research 36*(2), 321–339.

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2011b). Uncertainty averse preferences. *Journal of Economic Theory 146*(4), 1275–1330.

Chateauneuf, A. and J. H. Faro (2009). Ambiguity through confidence functions. *Journal of Mathematical Economics 45*(9-10), 535–558.

Chow, C. C. and R. K. Sarin (2001). Comparative ignorance and the Ellsberg paradox. *Journal of Risk and Uncertainty 22*(2), 129–139.

Epstein, L. and M. Schneider (2003). Recursive multiple-priors. *Journal of Economic Theory 113*(1), 1–31.

Epstein, L. G. and M. Schneider (2007). Learning under ambiguity. *Review of Economic Studies 74*(4), 1275–1303.

Epstein, L. G. and M. Schneider (2010). Ambiguity and asset markets. *Annual Review of Financial Economics 2*(1), 315–346.

Ergin, H. and F. Gul (2009). A subjective theory of compound lotteries. *Journal of Economic Theory 144*(3), 899–929.

Fox, C. R. and A. Tversky (1995). Ambiguity aversion and comparative ignorance. *The Quarterly Journal of Economics 110*(3), 585–603.

Gensbittel, F., L. Renou, and T. Tomala (2015). Comparisons of ambiguous experiments. SSRN Working Paper.

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

Grant, S., A. Kajii, and B. Polak (1998). Intrinsic preference for information. *Journal of Economic Theory 83*(2), 233–259.

Grant, S., B. Polak, and T. Strzalecki. (2009). Second-order expected utility. Working Paper, Harvard University.

Greenberg, H. J. and W. P. Pierskalla (1973). Quasi-conjugate functions and surrogate-duality. *Cahier Centre d'Etude de Recherche Operationelle 15*, 437–448.

Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica 75*(2), 503–536.

Hansen, L. P. and T. J. Sargent (2001). Robust control and model uncertainty. *The American Economic Review , Papers and Proceedings 91*(2), 60–66.

Heyen, D. and B. R. Wiesenfarth (2014). Informativeness of experiments for MEU — a recursive definition. *Journal of Mathematical Economics*, Accepted manuscript.

Hilton, R. W. (1990). Failure of Blackwell's theorem under Machina's generalization of expected-utility analysis without the independence axiom. *Journal of Economic Behavior & Organization 13*(2), 233–244.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005). A smooth model of decision making under ambiguity. *Econometrica 73*(6), 1849–1892.

Lehmann, E. L. (1988). Comparing location experiments. *The Annals of Statistics 16*(2), 521–533.

Li, J. (2013). Preferences for information and ambiguity. Working Paper, McGill University.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006a). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica 74*(6), 1447–1498.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006b). Dynamic variational preferences. *Journal of Economic Theory 128*(1), 4–44.

Machina, M. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature 27*(4), 1622–1668.

Mukerji, S. and J.-M. Tallon (2004). An overview of economic applications of david schmeidler's models of decision making under uncertainty. In *Uncertainty in Economic Theory,* (I. Gilboa, ed.), New York: Routledge.

Nau, R. F. (2006). Uncertainty aversion with second-order utilities and probabilities. *Management Science 52*(1), 136–145.

Neilson, W. (2010). A simplified axiomatic approach to ambiguity aversion. *Journal of Risk and Uncertainty 41*(2), 113–124.

Persico, N. (2000). Information acquisition in auctions. *Econometrica 68*(1), 135–148.

Quah, J. K.-H. and B. Strulovici (2009). Comparative statics, informativeness, and the interval dominance order. *Econometrica 77*(6), 1949–1992.

Safra, Z. and E. Sulganik (1995). On the nonexistence of blackwell's theorem-type results with general preference relations. *Journal of Risk and Uncertainty 10*(3), 187–201.

Siniscalchi, M. (2011). Dynamic choice under ambiguity. *Theoretical Economics 6*(3), 379–421.

Strzalecki, T. (2011). Axiomatic foundations of multiplier preferences. *Econometrica 79*(1), 47–73.

Strzalecki, T. (2013). Temporal resolution of uncertainty and recursive models of ambiguity aversion. *Econometrica 81*(3), 1039–1074.

Wakker, P. (1988). Nonexpected utility as aversion of information. *Journal of Behavioral Decision Making 1*(3), 169–175.