Ranking of Experiments

Edi Karni[*,†]

March 20, 2021

Abstract

Departing from the reduction of compound lotteries axiom on multi-stage lotteries induced by experiments, this paper shows that Blackwell's (1953) definition of the relation "more informative" on the set of information structures is equivalent to experiments being more valuable to a class of nonexpected utility preferences. This result extends Blackwell's (1953) theorem and suggests new method of evaluating experiments.

Keyword: Blackwell's theorem; Comparison of experiments; Reduction of compound lotteries; Value of information;

JEL classification: D81, D83

1 Introduction

From a decision making point of view, experiments are valuable because they provide information that helps decision makers choose courses of actions whose payoffs are higher in the states that are more likely to obtain. Blackwell (1953) formalized this perception as a binary relation, "more informative than" on the set of experiments. According to Blackwell one experiment is more informative than another if, for every set of feasible actions, it yields a richer menu of experiment-wise expected payoffs (i.e., expected-loss vectors) each of which corresponds to an action taken contingent on the experimental observations. Blackwell characterized this relation by proving that one experiment is more informative than another if and only if the information content of the latter is obtained by garbling the information content of the former. Equivalently, an experiment is more informative if it allows choices that have higher expected utility. In this paper, I refer to the equivalence between being more informative in this sense and containing clearer information as Blackwell's theorem.[1]

[*] Johns Hopkins University, Department of Economics, e-mail: karni@jhu.edu

[†] Useful comments by Edward Schlee are gratefull acknowledged.

[1] Cremer (1982) and Leshno and Spector (1992) provide simple proofs of Blackwell's theorem.

As being better informed seems unambiguously beneficial, the equivalence between ranking experiments by their information content and their ranking by the expected utility criterion seems oddly restrictive. This equivalence is particularly disconcerting in view of experimental evidence suggesting that subjects violate the tenets of expected utility theory–the sure thing principle and the independence axiom–systematically, and the proliferation, over the last 40 years, of non-expected utility models of decision making under risk and under uncertainty.[2]

Experimentation followed by signal-contingent choice of actions may be regarded as a multi-stage compound lottery. In this paper I argue that the critical aspect of the expected utility model that is underlying Blackwell's theorem is a postulate, known as reduction of compound lotteries, that requires that the probability of the ultimate payoff is equal to the product of the probabilities on the events that lead to it. Moreover, I show that replacing the reduction of compound lotteries postulate with an alternative procedure, analogous to the certainty equivalent reduction proposed by Segal (1990), implies that experiments whose information content is superior are preferred according to a large class of non-expected utility models.

To grasp this point, consider a decision maker who faces a choice among feasible actions whose consequences depend on the realization of some underlying states. Suppose that the likelihoods of the various states to obtain is quantified by a (prior) probability distribution function. Before choosing an action, the decision maker receives a signal (i.e., an observation), produced by an experiment, that informs him about the likely realization of the states. Upon receiving such signal, the decision maker invokes Bayes' rule to update the prior state probabilities, and then proceeds to choose an action from the feasible set. This process may be thought of as two-stage compound lottery. In the first stage, the experiment produces a signal, according to some probability distribution on the set of signals, following which the decision maker chooses an action. In the second stage, a state is selected (according to the posterior distribution), and the decision maker is awarded the prize that corresponds to the image of the selected state under the chosen action.[3] The question is how do decision makers perceive this two-stage lottery. According to the reduction of compound lotteries postulate, this two-stage lottery is reduced to a single-stage lottery by attributing to each prize the probability of the signal multiplied by the posterior probabilities of the states to which the chosen course of action assigns that prize. Analysis that treat the two-stage process as equivalent to its one-stage reduction runs the risk of disregarding subtleties that beset the extensive form decision process.

I contend that because the first and second stages are separated by a decision, they merit distinct treatments. Application of the reduction of compound lotteries which seems

[2]Wakker (1988), Schlee (1990) and Safra and Sulganik (1995), demonstrated that non-expected utility theories imply that information may have negative value.

[3]If the prize itself is a lottery ticket than the procedure described above amounts to three-stage lottery in which, in the third and final stage, the lottery corresponding to the image of the selected state under the chosen action is played out to determine the prizes.

compelling when the transition between the stages is automatic, seems less so when the two stages are separated by an intermediate decision. In this paper I pursue this line of reasoning and propose to analyze the decision process in its extensive form. I show that, by replacing the reduction of compound lotteries in the first stage, it is possible to identify a class of non-expected utility preferences that unambiguously values one experiment over another if and only if the information content of the latter is obtained by garbling that of the former. The proposed extension rules out the possibility of negatively valued information which, I regard as a powerful argument in favor of the proposed model. The expected utility model is a special case of this class–the only model in this class that is consistent with the reduction of the compound lotteries postulate.

The surprising (difficult) aspect of Blackwell's theorem is that a more informative experiment (that is, experiment that affords better decisions by the expected utility criterion) implies sufficiency (that is, clearer signal). In this paper, informativeness corresponds to an experiment being more valuable in the sense of affording better decisions for a broader set of preferences, including expected utility preferences. Consequently, this direction of the proof relies on Blackwell' Theorem. The novelty of this paper is the observation that the full power of expected utility, in particular, the reduction of compound lotteries which is implied by the independence axiom, is not needed for Blackwell's result.

The next section includes a brief review of Blackwell's (1953) theorem. Section 3 reviews the reduction procedures. Section 4 extends Blackwell's theorem. Section 5 provides an axiomatic characterization of preference relations that underlie the extension of Blackwell's theorem. Section 6 discusses the value of information and briefly reviews the related literature.

2 The Analytical Framework and Blackwell's Theorem

2.1 The analytical framework

Let $S = \{s_1, ..., s_n\}$ be a finite set of *states* and denote by $\Delta(S)$ the simplex in $\mathbb{R}^n$. Subsets of S are *events.* Let X a set of *outcomes* and denote by $\Delta(X)$ the set of simple probability distributions on X referred to as *lotteries*.[4] Mappings on S to $\Delta(X)$ are referred to as *acts*, representing potential courses of action. The set of all acts is denoted by $\mathcal{H}$. For all $f, g \in \mathcal{H}$ and $\alpha \in [0,1]$ define $\alpha f + (1-\alpha) g \in \mathcal{H}$ by $(\alpha f + (1-\alpha) g)(s) = \alpha f(s) + (1-\alpha) g(s)$, for all $s \in S$. Thus, $\mathcal{H}$ is convex. Constant acts (i.e., acts that assign the same image to every state) are identified with elements of $\Delta(X)$, thus, $\Delta(X) \subset \mathcal{H}$. For all $p, q \in \Delta(X)$ and $\alpha \in [0,1]$ define $\alpha p + (1-\alpha) q \in \Delta(X)$ by $(\alpha p + (1-\alpha) q)(x) = \alpha p(x) + (1-\alpha) q(x)$, for all $x \in X$. Throughout I denote by δ_x the distribution function that assigns x the unit probability mass and by $\Delta(\Delta(X))$ the set of simple probability distributions with supports in $\Delta(X)$.

[4] Simple probability distributions are probability distribution functions with finite supports.

Complete and transitive binary relations $\succcurlyeq$ on $\mathcal{H}$ are referred to as *preference relations*. The *strict preference relation*, $\succ$, and the *indifference relation,* $\sim$, are the asymmetric and the symmetric parts of $\succcurlyeq$, respectively. A real-valued function U on $\mathcal{H}$ is said to *represent* the preference relation $\succcurlyeq$ if, for all $f, g \in \mathcal{H}$, $f \succcurlyeq g$ if and only if $U(f) \geq U(g)$.

Experiments are random variables, $\widetilde{y}^j$, taking values in the sets $Y^j = \{y_1^j, ..., y_{k(j)}^j\}$, $j \in \mathbb{N}$, where $\mathbb{N}$ denotes the set of natural numbers. Let $\mathcal{Y}$ denote the set of all experiments. An experiment $\widetilde{y}^j$ is identified with a joint probability distribution μ^j on $Y^j \times S$. Given $\widetilde{y}^j \in \mathcal{Y}$ and $s_i \in S$, the conditional probability on Y^j is depicted by a vector of probabilities, $\mu^j(\cdot \mid s_i) := \left(\mu^j\left(y_1^j \mid s_i\right), ..., \mu^j\left(y_{k(j)}^j \mid s_i\right)\right)$. Without loss of generality assume that for each $y_k^j \in Y^j$ there exists $s_i \in S$ such that $\mu^j\left(y_k^j \mid s_i\right) > 0$. Thus, for all $\pi \in \Delta(S)$, $\mu^j\left(y_k^j\right) := \Sigma_{s_i \in S}\mu^j\left(y_k^j \mid s_i\right)\pi(s_i) > 0$.

Applied to constant acts, expected utility theory provides necessary and sufficient conditions for the preference relations to have affine representation. Consequently, there exist a real-valued function u on X such that, for all $p \in \Delta(X)$, $U(p) = \Sigma_{x \in X} u(x) p(x)$. Denote by $\mathcal{U}$ the set of all such functions.

2.2 Blackwell's theorem

Consider an expected utility maximizing decision maker characterized by a utility function $u \in \mathcal{U}$. Given a probability distribution $\pi \in \Delta(S)$ and a non-empty set, $B \subseteq \mathcal{H}$, of feasible acts, denote by $f_B^*(\pi, u)$ the solution to the problem:

$$\max_{f \in B} \Sigma_{s \in S}\pi(s)\Sigma_{x \in X} u(x) f(s)(x). \tag{1}$$

Define $U(\pi, u, B) = \Sigma_{s \in S}\pi(s)\Sigma_{x \in X} u(x) f_B^*(\pi, u)(x)$.

Suppose that, before choosing an act from a feasible set B, the decision maker observes the outcome $y_k^j \in Y^j$ of an experiment $\widetilde{y}^j \in \mathcal{Y}$ and updates the prior distribution π according to Bayes' rule to obtain the posterior probability distribution $\pi\left(s_i \mid y_k^j\right) = \mu^j\left(y_k^j \mid s_i\right)\pi(s_i)/\mu^j\left(y_k^j\right)$, for all $s_i \in S$. The decision maker's ex-post problem is:

$$\max_{f \in B} \Sigma_{s \in S}\pi\left(s \mid y_k^j\right)\Sigma_{x \in X} u(x) f(s)(x). \tag{2}$$

Letting $U\left(\pi\left(\cdot \mid y_k^j\right), u, B\right) = \Sigma_{s \in S}\pi\left(s \mid y_k^j\right)\Sigma_{x \in X} u(x) f_B^*\left(\pi\left(\cdot \mid y_k^j\right), u\right)(x)$, the expected utility associated with the experiment $\widetilde{y}^j$ is:

$$\bar{U}\left(\widetilde{y}^j; \pi, u, B\right) := \Sigma_{k=1}^{k(j)} U\left(\pi\left(\cdot \mid y_k^j\right), u, B\right)\mu^j\left(y_k^j\right). \tag{3}$$

Definition 1: An experiment $\widetilde{y}^j$ is *more informative* than another experiment $\widetilde{y}^h$ if $\bar{U}\left(\widetilde{y}^j; \pi, u, B\right) \geq \bar{U}\left(\widetilde{y}^h; \pi, u, B\right)$, for all $(\pi, u) \in \Delta S \times \mathcal{U}$ and $B \subseteq \mathcal{H}$.

An *information structure,* is an $n\times\ k\left(j\right)$ right-stochastic matrix, $P\left(\widetilde{y}^j\right)$, whose generic element is $\mu^j\left(y_k^j \mid s_i\right)$. Let $\mathcal{M}$ be the set of $k\left(j\right)\times k\left(h\right)$ Markov matrices dubbed *garbling matrices.* The garbling matrix, M, introduces noise that blurs the information in $P\left(\widetilde{y}^j\right)$.

Definition 2: An experiment $\widetilde{y}^j$ is *sufficient* for $\widetilde{y}^h$ if the corresponding information structures satisfy $P\left(\widetilde{y}^j\right)M = P\left(\widetilde{y}^h\right)$, for some $M\in\mathcal{M}$.

With these definitions in mind Blackwell's theorem is stated as follows:

Blackwell's Theorem: *An experiment $\widetilde{y}^j$ is more informative than another experiment $\widetilde{y}^h$ if and only if $\widetilde{y}^j$ is sufficient for $\widetilde{y}^h$.*

According to Blackwell's theorem the relation "being sufficient" is equivalent to being ranked higher by all preference relations that admit expected utility representations.

3 Informative Signals and Reduction Procedures

3.1 Signals

According to Blackwell's (1953) theorem more informative experiments produce clearer signals in the following sense: When comparing two experiments every signal produced by the less informative experiment is an average of signals produce by the more informative experiment (i.e., the sufficient experiment). Formally, let $\widetilde{y},\widetilde{y}'\in\mathcal{Y}$ with supports Y and Y', respectively. Suppose that $\widetilde{y}$ is sufficient for $\widetilde{y}'$. By definition, $P\left(\widetilde{y}\right)M = P\left(\widetilde{y}'\right)$, for some $M\in\mathcal{M}$. Hence, for every $y_r'\in Y'$ and $s_i\in S$,

$$\mu\left(y_r' \mid s_i\right) = \Sigma_{j\in\{k|y_k\in Y\}}^n \mu\left(y_j \mid s_i\right)m_{jr}, \tag{4}$$

where $\left(m_{jr}\right)_{j\in\{k|y_k\in Y\}}$ is the $r-th$ column of the garbling matrix M. By Bayes' rule, for every $y_r'\in Y'$ and $s_h\in S$,

$$\pi\left(s_h \mid y_r'\right) = \frac{\pi\left(s_h\right)\Sigma_{\{j|y_j\in Y\}}\mu\left(y_j \mid s_h\right)m_{jr}}{\Sigma_{i=1}^n\pi\left(s_i\right)\Sigma_{\{j|y_j\in Y\}}\mu\left(y_j \mid s_i\right)m_{jr}}. \tag{5}$$

$$= \frac{\Sigma_{\{j|y_j\in Y\}}m_{jr}\pi\left(s_h\right)\mu\left(y_j \mid s_h\right)}{\Sigma_{\{j|y_j\in Y\}}m_{jr}\Sigma_{i=1}^n\pi\left(s_i\right)\mu\left(y_j \mid s_i\right)} = \Sigma_{\{j|y_j\in Y\}}\pi\left(s_h \mid y_j\right)m_{jr},$$

where, to obtain the last equality divide the numerator and denominator by $\Sigma_{i=1}^n\pi\left(s_i\right)\mu\left(y_j \mid s_i\right)$ and use the fact that $\Sigma_{\{j|y_j\in Y\}}m_{jr} = 1$. Consequently, for each act-posterior probability pair, $\left(f,\pi\left(\cdot \mid y_r'\right)\right)\in\mathcal{H}\times\Delta\left(S\right)$, that is feasible under the less informative experiment corresponds a set $\left\{\left(f,\pi\left(\cdot \mid y_j\right)\right) \mid y_j\in Y\right\}\subset\mathcal{H}\times\Delta\left(S\right)$ of act-posterior probability pairs of the more informative experiment. In other words, from an ex-ante viewpoint, the more informative experiment offers a richer set of opportunities to match feasible acts to the perceived likelihood of the states depicted by their posterior probabilities.

3.2 Reduction procedures

The axiomatic structure underlying expected utility theory implies a property known as *reduction of compound lotteries*. Formally, reduction of compound lotteries maintains that every compound lottery, $\ell \in \Delta(\Delta(X))$ is equivalent (indifferent) to the one stage lottery $p_\ell \in \Delta(X)$, where $p_\ell(x) = \Sigma_{p \in Supp\ell}\ell(p)\, p(x)$, for all $x \in X$, where $Supp\ell$ denotes the support of ℓ.

Given $p \in \Delta(X)$ the certainty equivalent of p is an outcome $c(p) \in X$ such that $\delta_{c(p)} \sim p$. The certainty equivalent may or may not exist depending on the richness of the set of outcomes X. If X includes the certainty equivalents of all the elements of $\Delta(X)$, then implicit in expected utility theory is another reduction procedure maintaining that every compound lottery, $\ell \in \Delta(\Delta(X))$ is equivalent to $q_\ell \in \Delta(X)$, where $q_\ell(c(p)) = \Sigma_{p \in Supp\ell}\ell(p)\,\delta_{c(p)}$, for all $c(p)$ such that $p \in Supp\ell$. I refer to this form of reduction as *certainty equivalence reduction.*[5]

Every $(f, \pi) \in \mathcal{H} \times \Delta(S)$ may be regarded as a two-stage lottery in which, in the first stage, a state $s \in S$, is drawn at random according to the distribution π and, in the second stage, an outcome $x \in X$ is determined by the lottery $f(s) \in \Delta(X)$. Applying reduction of compound lotteries, (f, π) is equivalent to $\Sigma_{s \in S}\pi(s)\, f(s) \in \Delta(X)$.

Let $B \subset \mathcal{H}$ denote a set of feasible acts and consider the a decision problem depicted by a quadruplet $(B, \pi, u, \widetilde{y}) \in 2^{\mathcal{H}} \backslash \varnothing \times \Delta(S) \times \mathcal{U} \times \mathcal{Y}$. Given the parameters $u \in \mathcal{U}$ and $\pi \in \Delta(S)$, the decision problem requires a plan of choosing acts in B contingent on the realization of signals produced by $\widetilde{y}$. The signals produced by $\widetilde{y}$ are drawn at random from the set Y according to a distribution μ. This problem induces a three-stage compound lottery. In the first stage, a signal, $y \in Y$ is drawn at random according to a distribution μ on Y. Contingent on the signal, the posterior distribution $\pi(\cdot \mid y) \in \Delta(S)$ is calculated using Bayes' rule and an act, $f_B^*(\pi(\cdot \mid y), u) \in B$, is chosen. In the second stage a state, $s \in S$ is selected according to the posterior distribution $\pi(\cdot \mid y)$ and the lottery $f_B^*(\pi(\cdot \mid y), u)(s) \in \Delta(X)$ is awarded as a prize. In the third and final stage, the lottery $f_B^*(\pi(\cdot \mid y), u)(s)$ determines the final outcome $x \in X$.

If the reduction of compound lotteries is applied then the compound lottery induced by a decision problem $(B, \pi, u, \widetilde{y})$ is equivalent to the one-stage lottery

$$\Sigma_{y \in Y}\mu(y)\, \Sigma_{s \in S}\pi(s \mid y)\, f_B^*(\pi(\cdot \mid y), u)(s) \in \Delta(X).$$

If the certainty equivalence reduction is applied then the one-stage lottery induced by the same decision problem is:

$$\Sigma_{y \in Y}\mu(y)\, \delta_{c\left(\Sigma_{s \in S}\pi(s|y)\delta_{c\left(f_B^*(\pi(\cdot|y),u)(s)\right)}\right)}.$$

In expected utility theory, these two reduction procedures are equivalent.

[5]For more detailed discusssion and application see Segal (1987).

An alternative approach, which I propose and study in this paper, applies the reduction of compound lotteries to the second stage and a procedure analogous to certainty equivalence reduction to the first stage. Under this *hybrid procedure*, for each $y \in Y$, the second-stage reduction induces a one-stage lottery $\Sigma_{s \in S} \pi(s \mid y) f_B^* (\pi(\cdot \mid y), u)(s) \in \Delta(X)$. Denote the valuation of this lottery by $u(\Sigma_{s \in S} \pi(s \mid y) f_B^* (\pi(\cdot \mid y_j), u)(s))$. According to the hybrid procedure, the lottery induced by a decision problem $(B, \pi, u, \widetilde{y})$ is $\Sigma_{y \in Y} \mu(y) \delta_{u\left(\Sigma_{s \in S} \pi(s|y) f_B^*(\pi(\cdot|y),u)(s)\right)} \in \Delta(\mathbb{R})$ whose value is

$$\Sigma_{y \in Y} \mu(y) u(\Sigma_{s \in S} \pi(s \mid y) f_B^* (\pi(\cdot \mid y), u)(s)).$$[6]

The justification for applying distinct procedures to the different stages is the nature of the uncertainties involved. In the second stage, given the act and the (updated) state probabilities, the outcome is selected "algorithmically" without interference by the decision maker. By contrast, after the first stage, corresponding to each signal there is an interim stage at which, the decision maker interferes by updating the state probabilities and choosing an act. This aspect of the process suggests that decision makers may regard the first stage as qualitatively distinct from the later stages and, consequently, treat them differently. Specifically, according to the hybrid procedure, assessing the value of experiments, decision makers envision the acts that they would choose contingent on the signals, assign these acts utility values and take the mean utility values as the value of the experiment.

4 Blackwell's Theorem Extended

4.1 Utility representation on $\mathcal{H} \times \Delta(S)$.

The spaces $\mathcal{H}$ and $\Delta(S)$ are a convex subset of $\mathbb{R}^n$ and, as such, they are connected separable topological spaces. Let $\succcurlyeq$ be a preference relation (that is, complete and transitive binary relation) on $\mathcal{H} \times \Delta(S)$.

Definition 3. The preference relation $\succcurlyeq$ on $\mathcal{H} \times \Delta(S)$ is continuous if the sets $\{(f, \pi') \in \mathcal{H} \times \Delta(S) \mid (f, \pi') \succcurlyeq (\pi, g)\}$ and $\{(f, \pi') \in \mathcal{H} \times \Delta(S) \mid (\pi, g) \succcurlyeq (f, \pi')\}$ are closed in the topology of $\mathbb{R}^n$, for all $(g, \pi) \in \mathcal{H} \times \Delta(S)$.

Definition 4. A preference relation is said to satisfy the reduction of compound lotteries axiom if $(f, \pi) \sim \Sigma_{s \in S} \pi(s) f(s)$, for all $(f, \pi) \in \mathcal{H} \times \Delta(S)$.

With this in mind we have:

Proposition 1: A continuous preference relation $\succcurlyeq$ on $\mathcal{H} \times \Delta(S)$ satisfies reduction of compound lotteries if and only if there exists a continuous function v on $\Delta(X)$ such that,

[6] If the certainty equivalents exist this is equivalent to

$$\Sigma_{y \in Y} \mu(y) u(c(\Sigma_{s \in S} \pi(s \mid y) f_B^* (\pi(\cdot \mid y), u)(s))).$$

for all $(f, \pi'), (g, \pi) \in \mathcal{H} \times \Delta(S)$,

$$\left(f, \pi'\right) \succcurlyeq (g, \pi) \Leftrightarrow v\left(\Sigma_{s \in S} \pi'(s) f(s)\right) \geq v\left(\Sigma_{s \in S} \pi(s) g(s)\right). \tag{6}$$

Proof. The existence of a continuous, real-valued, function v on $\mathcal{H} \times \Delta(S)$ representing $\succcurlyeq$ is implied by a theorem of Debreu (1954). Thus, for all $(f, \pi'), (g, \pi) \in \mathcal{H} \times \Delta(S)$,

$$\left(f, \pi'\right) \succcurlyeq (g, \pi) \Leftrightarrow v\left(f, \pi'\right) \geq v(g, \pi).$$

By reduction of compound lotteries, for all $(h, \bar{\pi}) \in \mathcal{H} \times \Delta(S)$, $(h, \bar{\pi}) \sim \Sigma_{s \in S} \bar{\pi}(s) h(s)$. Hence, by transitivity, $(f, \pi') \succcurlyeq (g, \pi)$ if and only if $\Sigma_{s \in S} \pi'(s) f(s) \succcurlyeq \Sigma_{s \in S} \pi(s) f(s)$. Since $\Delta(X) \subset \mathcal{H}$, by the representation,

$$\Sigma_{s \in S} \pi'(s) f(s) \succcurlyeq \Sigma_{s \in S} \pi(s) f(s) \Leftrightarrow v\left(\Sigma_{s \in S} \pi'(s) f(s)\right) \geq v\left(\Sigma_{s \in S} \pi(s) g(s)\right).$$

Hence, $(f, \pi') \succcurlyeq (g, \pi)$ if and only if $v\left(\Sigma_{s \in S} \pi'(s) f(s)\right) \geq v\left(\Sigma_{s \in S} \pi(s) g(s)\right)$. ■

Henceforth, I denote by $\mathcal{V}$ the set of functions representing continuous preference relations on $\mathcal{H} \times \Delta(S)$ that satisfy reduction of compound lotteries.

4.2 Compound lotteries reductions and their representations

For every $(f, \pi) \in \mathcal{H} \times \Delta(S)$, define

$$f^{\pi}(x) = \Sigma_{s \in S} \pi(s) f(s)(x), \ \forall x \in X. \tag{7}$$

Similarly, given an experiment, $\widetilde{y}^j \in \mathcal{Y}$, with support Y^j, for every signal $y_k^j \in Y^j$ let

$$f^{\pi\left(\cdot | y_k^j\right)}(x) := \Sigma_{s \in S} \pi\left(s \mid y_k^j\right) f(s)(x), \ \forall x \in X. \tag{8}$$

For every $(\pi, v) \in \Delta(S) \times \mathcal{V}$ and $B \subseteq 2^{\mathcal{H}} \backslash \varnothing$, define

$$f_B^*(\pi, v) \in \arg\max_{f \in B} v\left(f^{\pi}\right). \tag{9}$$

Given $(B, \pi, v, \widetilde{y}) \in 2^{\mathcal{H}} \backslash \varnothing \times \Delta S \times \mathcal{V} \times \mathcal{Y}$, define

$$V(\widetilde{y}; \pi, v, B) = \Sigma_{y \in Y} \mu(y) v\left(f_B^*(\pi(\cdot \mid y), v)\right). \tag{10}$$

Let

$$p_{(\pi, v)}^B(\widetilde{y}) := \Sigma_{y \in Y} \mu(y) \Sigma_{s \in S} f_B^*(\pi(\cdot \mid y), v)(s). \tag{11}$$

By definition, $p_{(\pi, v)}^B(\widetilde{y}) \in \Delta(X)$ is the lottery obtained by the application of the reduction of compound lotteries to the three stages of the lottery. Expected utility representation admits reduction of compound lotteries, hence,

$$V(\widetilde{y}; \pi, v, B) = v\left(p_{(\pi, v)}^B(\widetilde{y})\right) = \Sigma_{x \in X} p_{(\pi, v)}^B(\widetilde{y})(x) u(x). \tag{12}$$

This is the reduction procedure that underlies Blackwell's theorem.

4.3 The main result

Let $\succcurlyeq$ be a continuous preference relation on $\mathcal{H} \times \Delta(S)$ that admits hybrid reduction representation. Formally, given $(\pi, v, B) \in \Delta(S) \times \mathcal{V} \times 2^{\mathcal{H}} \backslash \varnothing$, the value of the experiment $\widetilde{y} \in \mathcal{Y}$ whose support is Y is:

$$\hat{V}(\widetilde{y}; \pi, v, B) := \Sigma_{y \in Y} \mu(y) \max_{f \in B} v\left(f^{(\pi(\cdot|y))}\right), \tag{13}$$

where $f^{(\pi(\cdot|y))} = \Sigma_{s \in S} f_B^*(\pi(\cdot \mid y), v)(s) \pi(s) \in \Delta(X)$. Consequently, given $(v, \pi, B) \in \Delta(S) \times \mathcal{V} \times 2^{\mathcal{H}} \backslash \varnothing$, for all $\widetilde{y}, \widetilde{y}' \in \mathcal{Y}$,

$$\widetilde{y} \succcurlyeq \widetilde{y}' \Leftrightarrow \hat{V}(\widetilde{y}; \pi, v, B) \geq \hat{V}\left(\widetilde{y}'; \pi, v, B\right). \tag{14}$$

With this in mind I make the following definition:

Definition 5: Experiment $\widetilde{y}$ is *more valuable* than experiment $\widetilde{y}'$ if $\hat{V}(\widetilde{y}; \pi, v, B) \geq \hat{V}(\widetilde{y}'; \pi, v, B)$, for all $(v, \pi, B) \in \Delta(S) \times \mathcal{V} \times 2^{\mathcal{H}} \backslash \varnothing$.

Theorem 1: *An experiment $\widetilde{y}$ is more valuable than another experiment $\widetilde{y}'$ if and only if $\widetilde{y}$ is sufficient for $\widetilde{y}'$.*

Proof. (Sufficiency) Suppose that $\widetilde{y}$ is sufficient for $\widetilde{y}'$. Let Y and Y' denote the supports of $\widetilde{y}$ and $\widetilde{y}'$, respectively. Fix $(\pi, v, B) \in \Delta(S) \times \mathcal{V} \times 2^{\mathcal{H}} \backslash \varnothing$ then, by (5), for every $y_r' \in Y'$ and $s \in S$, $\pi(\cdot \mid y_r') = \Sigma_{\{j \mid y_j \in Y\}} \pi(\cdot \mid y_j) m_{jr}$. Hence, for every $y_r' \in Y'$ and $r \in \{k \mid y_k' \in Y'\}$,

$$\max_{f \in B} v\left(f^{\pi(\cdot|y_r')}\right) = \max_{f \in B} v\left(f^{\Sigma_{\{j|y_j \in Y\}} \pi(\cdot|y_j) m_{jr}}\right) \leq v\left(f^*\left(\pi\left(\cdot \mid y_j\right), v\right)\right), \text{ for all } y_j \in Y. \tag{15}$$

By (4), $\mu(y_r') = \Sigma_{s \in S} \mu(y_r' \mid s) \pi(s) = \Sigma_{j \in \{k|y_k \in Y\}}^n (\Sigma_{s \in S} \mu(y_j \mid s) \pi(s)) m_{jr}$. Consequently,

$$\hat{V}\left(\widetilde{y}'; \pi, v, B\right) = \Sigma_{y_r' \in Y'} \mu\left(y_r'\right) \max_{f \in B} v\left(f^{\pi(\cdot|y_r')}\right) \tag{16}$$

$$= \Sigma_{\{r|y_r' \in Y'\}} \Sigma_{\{j|y_j \in Y\}} \left(\Sigma_{s \in S} \mu\left(y_j \mid s\right) \pi(s)\right) m_{jr} \max_{f \in B} v\left(f^{\Sigma_{\{k|y_k \in Y\}} \pi(\cdot|y_k) m_{kr}}\right)$$

$$\leq \Sigma_{\{j|y_j \in Y\}} \mu(y_j) \left(\Sigma_{\{r|y_r \in Y'\}} m_{jr}\right) v\left(f^* \pi\left(\cdot \mid y_k\right), v\right)$$

$$= \Sigma_{y_j \in Y} \mu(y_j) \max_{f \in B} v\left(f^{\pi(\cdot|y_j)}\right) = \hat{V}(\widetilde{y}; \pi, v, B),$$

where the inequality is implied by (15) and the last equality follows from the fact that, for each $j \in \{k \mid y_k \in Y\}$, $\Sigma_{\{r|y_r' \in Y'\}} m_{jr} = 1$.

(Necessity) Since expected utility representations are a subset of the set of preference relations that have hybrid representations, necessity is implied by the necessity part of Blackwell's theorem. ♠

5 Characterization of the Hybrid Representation

5.1 Utility representations

The primitives of the hybrid model consist of a *choice set*, $\Delta(\Delta(X))$, whose elements are simple two-stage compound lotteries on X, and a preference relation $\succcurlyeq$ on $\Delta(\Delta(X))$. A generic element μ of $\Delta(\Delta(X))$ is a distribution function with finite support in $\Delta(X)$. A generic element p of $\Delta(X)$ is a distribution function with finite support in X. Identifying $p \in \Delta(X)$ with $\delta_p \in \Delta(\Delta(X))$ implies that $\Delta(X) \subset \Delta(\Delta(X))$. For all $\mu, \mu' \in \Delta(\Delta(X))$ and $\alpha \in [0,1]$ define $\alpha\mu + (1-\alpha)\mu' \in \Delta(\Delta(X))$ by $(\alpha\mu + (1-\alpha)\mu')(p) = \alpha\mu(p) + (1-\alpha)\mu'(p)$, for all $p \in \Delta(X)$.

The structure of the preference relation is depicted by the following axioms:

(A.1) (**Weak order**) $\succcurlyeq$ is complete and transitive binary relation on $\Delta(\Delta(X))$.

(A.2) (**Archimedean**) For all $\mu, \mu', \mu'' \in \Delta(\Delta(X))$ such that $\mu \succ \mu' \succ \mu''$ there are $\alpha, \beta \in (0,1)$ such that $\alpha\mu + (1-\alpha)\mu'' \succ \mu'$ and $\mu' \succ \beta\mu + (1-\beta)\mu''$.

(A.3) (**First-Stage Independence**) For all $\mu, \mu', \mu'' \in \Delta(\Delta(X))$ and $\alpha \in (0,1]$ $\mu \succcurlyeq \mu'$ if and only if $\alpha\mu + (1-\alpha)\mu'' \succcurlyeq \alpha\mu' + (1-\alpha)\mu''$.

By the expected utility theorem we have:

Proposition 2. *A preference relation $\succcurlyeq$ on $\Delta(\Delta(X))$ satisfies (A.1)-(A.3) if and only if there exists a real-valued function w on $\Delta(X)$ such that, for all $\mu, \mu' \in \Delta(\Delta(X))$,*

$$\mu \succcurlyeq \mu' \Leftrightarrow \Sigma_{p \in Supp\mu}\mu(p)\,w(p) \geq \Sigma_{p \in Supp\mu'}\mu'(p)\,w(p).$$

Moreover, w is unique up to positive affine transformation.

5.2 Consistency and hybrid representation

Consider next a binary relation, $\hat{\succcurlyeq}$ on $\mathcal{H} \times \Delta(S)$. By Proposition 1, $\hat{\succcurlyeq}$ is a continuous weak-order satisfying reduction of compound lotteries if and only if there is a continuous, real-valued, function on $\Delta(X)$ such that, for all $(f, \pi'), (g, \pi) \in \mathcal{H} \times \Delta(S)$,

$$\left(f, \pi'\right) \hat{\succcurlyeq} (g, \pi) \Leftrightarrow v\left(\Sigma_{s \in S}\pi'(s)\,f(s)\right) \geq v\left(\Sigma_{s \in S}\pi(s)\,g(s)\right). \tag{17}$$

The next axiom links the preference relations $\succcurlyeq$ on $\Delta(\Delta(X))$ and $\hat{\succcurlyeq}$ on $\mathcal{H} \times \Delta(S)$. It asserts that the preference between two act-probability pairs in $\mathcal{H} \times \Delta(S)$ is the same as that between their corresponding reductions to one stage lotteries in $\Delta(X)$. Formally,

(A.4) (**Consistency**) For all $(f, \pi'), (g, \pi) \in \mathcal{H} \times \Delta(S)$, $(f, \pi') \hat{\succcurlyeq} (g, \pi)$ if and only if $\delta_{f^{\pi'}} \succcurlyeq \delta_{g^{\pi}}$.

To state the hybrid representation theorem I introduce the following additional notations and definitions. Given $F \times \Pi \subseteq \mathcal{H} \times \Delta(S)$, define $T(F \times \Pi) := \{f^{\pi} \mid (f, \pi) \in F \times \Pi\} \subseteq \Delta(X)$ and denote by $\mu_T \in \Delta(\Delta(X))$ the generic compound lottery whose support is $T \subseteq \Delta(X)$. Let $\mathcal{T} := \{\, T(F \times \Pi) \mid F \times \Pi \subseteq \mathcal{H} \times \Delta(S)\}$.

Theorem 2. *The following two conditions are equivalent:*

(i) $\succcurlyeq$ *on* $\Delta(\Delta(X))$ *is an Archimedean weak-order satisfying first-stage independence;* $\hat{\succcurlyeq}$ *on* $\mathcal{H} \times \Delta(S)$ *is a continuous weak order satisfying reduction of compound lotteries; jointly* $\succcurlyeq$ *and* $\hat{\succcurlyeq}$ *satisfy consistency.*

(ii) There exist real-valued continuous function v *on* $\Delta(X)$ *such that, for all* $T, T' \in \mathcal{T}$,

$$\mu_T \succcurlyeq \mu_{T'} \Leftrightarrow \Sigma_{f^\pi \in T} \mu_T(f^\pi) v(f^\pi) \geq \Sigma_{f^\pi \in T'} \mu_{T'}(f^\pi) v(f^\pi).$$

Moreover, v *is unique up to positive affine transformations.*

Proof. To prove sufficiency, let $(f,\pi), (g,\pi) \in \mathcal{H} \times \Delta(S)$. By reduction of compound lotteries, $(f,\pi) \hat{\sim} f^\pi$ and $(g,\pi) \hat{\sim} g^\pi$. By Proposition 1, $(f,\pi) \hat{\succcurlyeq} (g,\pi)$ if and only if $v(f^\pi) \geq v(g^\pi)$. Thus, by transitivity of $\hat{\succcurlyeq}$, $f^\pi \hat{\succcurlyeq} g^\pi$ if and only if $v(f^\pi) \geq v(g^\pi)$.

By Proposition 2, $\mu_{\{f^\pi\}} \succcurlyeq \mu_{\{g^\pi\}}$ if and only if $w(f^\pi) \geq w(g^\pi)$. Consistency, (A.5), implies that $v(f^\pi) \geq v(g^\pi)$ if and only if $w(f^\pi) \geq w(g^\pi)$. Hence, by the uniqueness of w, $v = bw + a$, $b > 0$. The conclusion follows from Proposition 2.

Necessity is immediate. ♣

Consider next the application of the hybrid representation to the ranking of experiments. Quadruplets $(\widetilde{y}, \pi, v, B) \in \mathcal{Y} \times \Delta(S) \times 2^{\mathcal{H}} \backslash \varnothing$ translate into $(\mu(\widetilde{y}), \pi(\widetilde{y}), f_B^*(\pi, v; \widetilde{y})) \in \Delta(Y) \times \Delta(S)^{|Y|} \times \mathcal{H}^{|Y|}$ as follows: $\mu(\widetilde{y}) := \left(\mu(y_1), ..., \mu\left(y_{|Y|}\right)\right)$, $\pi(\widetilde{y}) = \left(\pi(\cdot \mid y_1), ..., \pi\left(\cdot \mid y_{|Y|}\right)\right)$ and $f_B^*(\pi, v; \widetilde{y}) = \left(f_B^*(\pi(\cdot \mid y_1), v), ..., f_B^*\left(\pi\left(\cdot \mid y_{|Y|}\right), v\right)\right)$. Hence, the hybrid representation of experiments is as follows:

$$(\widetilde{y}, \pi, v, B) \to \hat{V}(\widetilde{y}; \pi, v, B) = \Sigma_{j=1}^k \mu(y_j) v\left(\Sigma_{s \in S} f_B^*(\pi(\cdot \mid y_k), v)(s) \pi(s \mid y_j)\right). \quad (18)$$

This is the representation in (14) and Definition 5.

6 Concluding Remarks

6.1 The value of information

Lurking in the background of Blackwell's theorem are two tacit properties of expected utility theory – *consequentialism* and *reduction of compound lotteries*. The former maintains that, facing sequential decisions involving risky choices, decision makers are "forward looking" in the sense that, at every decision node, their preferences are unaffected by outcomes that did not materialize, or "roads not taken," along the decision making path. The latter asserts that decision makers evaluate acts solely by the ultimate probability distributions they induce on outcomes regardless of whether the outcome is drawn, in a single step, from a known distribution or is arrived at by more convoluted trajectory that includes chance and decision nodes.

Safra and Sulganik (1995) demonstrated that, maintaining consequentialism and reduction of compound lotteries, a more informative experiment à la Blackwell is ranked below a less informative one by convex non-expected utility preferences.[7] The present result "explains" this finding arguing that Blackwell's theorem implies that the only decision theory that maintains the properties of reduction of compound lotteries and consequentialism, according to which being better informed is necessarily valuable is expected utility theory.

6.2 Related literature

Segal (1990) was the first to propose a model of decision making under risk in which he replaced the reduction of compound lotteries with certainty equivalence reduction, to characterized rank-dependent utility model.[8] Seo (2009) obtains smooth ambiguity averse representations of choice under uncertainty that departs from the reduction of compound lotteries axiom. Halevy (2007) presented experimental evidence suggesting that subjects whose behavior violates reduction of compound lotteries under risk are more likely to exhibit ambiguity aversion when facing decision making under uncertainty.

The present work argues that being better informed is unambiguously better. This natural and intuitive presumption justifies the departure from the reduction of compound lotteries in sequential decision situations in which, at the interim stages, information may be exploited by choosing acts that better match the underlying data.

Following Blackwell (1953), the primitives of the hybrid model include objective probability distributions on the state space. The Bayesian tradition in the theory of decision making under uncertainty maintains that the prior state probabilities are derived from the underlying reference relations. In view of these distinct outlooks, it is worth mentioning that the representation in Proposition 1 of preference relation on $\mathcal{H} \times \Delta(S)$ can be obtained by the application of the probabilistic sophisticated choice models of Machina and Schmeidler (1992, 1995). The primitives of these models consist of a choice sets, $\mathcal{F}$ and a preference relation $\succsim$ on $\mathcal{F}$. A probability $\pi \in \Delta(S)$ and a utility function, v on the set $\{f^\pi \mid f \in \mathcal{F},\, \pi \in S\}$ are derived concepts. In particular, Machina and Schmeidler show that a preference relation $\succsim$ on $\mathcal{F}$ satisfies the axioms of the probabilistic sophisticated choice if and only if there is a utility function v on $\{f^\pi \mid f \in \mathcal{F},\, \pi \in S\}$ and a probability distribution $\pi \in \Delta(S)$ such that for all $f, g \in \mathcal{F}$, $f \succsim g$ if and only if $v(f^\pi) \geq v(g^\pi)$. This result may be regarded as the Bayesian version of Proposition 1. It implies the reduction of subjective compound lotteries corresponding $(f, \pi) \in \mathcal{F} \times \Delta(S)$. In this interpretation, the hybrid model of this paper is a model of choice of experiments under uncertainty.

[7] Wakker (1988) showed that departing from the independence axiom while maintiaining consequentiolism and reduction of compound lotteries necessarily result in situations in which the decision maker refuses free information. Schlee (1990) demonstrate that the same point in the context of the the rank-dependent utility model.

[8] Since the independence axiom of expected utility theory implies reduction of compound lotteries, departure from the latter property implies the departure from the former axiom.

References

[1] Blackwell, D. (1953) "Equivalent Comparisons of Experiments," *Annals of Mathematics and Statistics*, 74, 265-272.

[2] Cremer, J. (1982) "A Simple Proof of Blackwell's "Comparison of Experiments" Theorem." *Journal of Economic Theory*, 27, 439-443.

[3] Debreu, G. (1954) "Representation of a Preference Ordering by a Numerical Function," in: R. M. Thrall, C. H. Coombs, and R. L. Davis, eds., *Decision Processes*, New York: Wiley, 159-165.

[4] Halevy, Y. (2007) "Ellsberg Revisited: An Experimental Study," *Econometrica*, 75, 503–536.

[5] Leshno, M. and Spector, Y. (1992) "An Elementary Proof of Blackwell' Theorem." *Mathematical Social Sciences*, 25, 95-98.

[6] Machina, M. and Schmeidler D. (1992) "A More Robust Definition of Subjective Probability." *Econometrica*, 60, 745–780.

[7] Machina, M. and Schmeidler D. (1995) "Bayes without Bernoulli: Simple Conditions for Probabilistically Sophisticated Choice." *Journal of Economic Theory,* 67, 106–128.

[8] Safra, Z. and Sulganik, E. (1995) "On the Nonexistence of Blackwell's Theorem-Type Results with General Preference Relations." *Journal of Risk and Uncertainty* 10, 187-201.

[9] Schlee, E. (1990) "The Value of Information in Anticipated Utility Theory," *Journal of Risk and Uncertainty,* 3, 83-92.

[10] Segal, U. (1990) "Two Stage Lotteries without the Reduction Axiom." *Econometrica*, 58 , 349–377.

[11] Seo, K. (2009) "Ambiguity and Second-Order Beliefs," *Econometrica*, 77, 1575–1605

[12] Wakker, P. (1988) "Nonexpected Utility as Aversion to Information," *Journal of Behavioral Decision Making* 1, 169-175.