Informativeness of Experiments for $\min\max$ and $\max\min$ Decision Rules

Boğaçhan Çelen*

Columbia University

September 12, 2011

Abstract

This note shows that the celebrated Blackwell's theorem extends to $\max\min$ and $\min\max$ decision rules.

JEL Classification Numbers: C44, C65, D81.

Keywords: Blackwell, informativeness, decision rules, minmax, maxmin.

1 Introduction

Blackwell [1, 2] shows the equivalence of the following orderings: (*i*) a statistical experiment is more *valuable* than another one, if for each expected-utility-maximizing decision maker (DM), the ex ante expected utility under the former is higher that the one under the latter; and (*ii*) a statistical experiment is more *informative* than another one, if the latter can be obtained by garbling the former.

Incidentally, the connection between the two orderings is far from obvious, both from technical and conceptual view points. In fact, *valuableness* is defined with reference to DMs' preferences and subjective beliefs about the underlying states of the world; hence, is an economically relevant and intuitive concept. However, *informativeness* is a pure statistical property that is free of all economic variables. This is what makes Blackwell's theorem an important economic result because informativeness of an experiment is equivalent to its value for *any* DM, who may hold *any* prior belief.

*Columbia Business School, 3022 Broadway, 602 Uris Hall, New York, NY 10027. E-mail: bc2132@columbia.edu, url: http://celen.gsb.columbia.edu/. I would like to acknowledge helpful comments of Paolo Siconolfi, Lones Smith and Junjie Zhou. I appreciate Junjie Zhou for his comments which helped me to generalize the results. Also, the question originates from a discussion with Efe Ok, which I would like to acknowledge.

This note shows that Blackwell's theorem in fact extends to a larger set of decision rules including min max or the max min. To be precise, a DM who uses min max or max min decision rule considers a set of priors rather than a unique one. Moreover, a min max DM makes his decision under the assumption that his prior is the one that minimizes the maximum ex ante expected utility he can attain, whereas a max min DM maximizes assuming that his prior is the one that minimizes the ex ante expected utility he could obtain.

This result is of particular interest since the decision theory literature addressing uncertainty (or ambiguity) aversion (viz. Gilboa and Schmeidler [4]) demonstrates that preferences that exhibit uncertainty aversion can be represented by the minimum expected utility a DM gets. The result that max min preferences obey an important and natural property as Blackwell's theorem, provides further support to the study of uncertainty aversion, which has become an important tool in many applied fields of economics over the last two decades.

The next section introduces the notation and the key concepts. Section 3 states the main result, and the last section discusses the contribution of the result.

2 Notation and Preliminaries

For any matrix $\boldsymbol{m}_{a\times b}$ of dimension $a \times b$, m_{ij} and $\boldsymbol{m}'$ denote the (i,j)th entry and the transpose of $\boldsymbol{m}$, respectively. The inner product of two matrices of the same dimension is defined as $\langle \boldsymbol{m}, \boldsymbol{n} \rangle := \sum_j \sum_i m_{ij} n_{ij} = tr(\boldsymbol{m}'\boldsymbol{n})$. For any vector $\boldsymbol{\pi} \in \mathbb{R}^n$, $\boldsymbol{D}^{\boldsymbol{\pi}}_{n\times n}$ denotes the diagonal matrix such that $D^{\boldsymbol{\pi}}_{ii} = \pi_i$, and finally $\boldsymbol{I}$ denotes the identity matrix.

Let $\Omega := \{w_1, \cdots, w_n\}$ be a finite set of the *states*, and $X := \{a_1, \cdots, a_x\}$ be a finite set of *actions* available to a decision-maker. A decision-maker is characterized by a *utility function* $u : \Omega \times X \mapsto \mathbb{R}$ and a *prior* $\boldsymbol{\pi} \in \Delta(\Omega)$—i.e. $\pi_i = \Pr(\omega_i)$.[1] For a utility function u, we define the matrix $\boldsymbol{u}_{n\times x}$ such that $u_{ij} := u(\omega_i, a_j)$ and denote a Bayesian DM by $(\boldsymbol{\pi}, \boldsymbol{u})$.

An *experiment* is a tuple $(S, \boldsymbol{p})$, where $S := \{s_1, \cdots, s_\sigma\}$ is a set of signals, and $\boldsymbol{p}_{n\times\sigma}$ is a Markov matrix such that $p_{ij} := \Pr(s_j|\omega_i)$ for $s_j \in S$.

For a DM $(\boldsymbol{\pi}, \boldsymbol{u})$ who observes a signal s from the experiment $(S, \boldsymbol{p})$, a strategy is a vector valued map $f : S \mapsto \Delta(X)$. For each strategy f we define the matrix $\boldsymbol{f}_{\sigma\times x}$, such that $(f_{i1}, \cdots, f_{ix}) := f(s_i)$.

[1] For a finite set A, $\Delta(A)$ denotes the set of all probability distributions over A.

The value of $(S, \boldsymbol{p})$ for a given strategy $\boldsymbol{f}$ is

$$\begin{aligned} \mathcal{U}^{\boldsymbol{f}}_{(\pi,u)}(S, \boldsymbol{p}) &:= \sum_j \Pr(s_j) \sum_i \Pr(\omega_i | s_j) \sum_k f_{jk} u(\omega_i, a_k), \\ &= \sum_j \sum_i p_{ij} \pi_i \sum_k f_{jk} u_{ik}, \qquad \text{(by Bayes' rule)} \\ &= \langle \boldsymbol{D}^{\boldsymbol{\pi}} \boldsymbol{p} \boldsymbol{f}, \boldsymbol{u} \rangle. \end{aligned}$$

For $\boldsymbol{f}^* \in \arg\max_{\boldsymbol{f}} \langle \boldsymbol{D}^{\boldsymbol{\pi}} \boldsymbol{p} \boldsymbol{f}, \boldsymbol{u} \rangle$, we will write $\mathcal{U}^*_{(\boldsymbol{\pi},\boldsymbol{u})}(S, \boldsymbol{p}) := \mathcal{U}^{\boldsymbol{f}^*}_{(\boldsymbol{\pi},\boldsymbol{u})}(S, \boldsymbol{p})$.

We are ready to define the two orderings that are central in Blackwell's theorem.

DEFINITION 1. (i) $(S, \boldsymbol{p})$ *is more valuable than* $(T, \boldsymbol{q})$ *if* $\mathcal{U}^*_{(\boldsymbol{\pi},\boldsymbol{u})}(S, \boldsymbol{p}) \geq \mathcal{U}^*_{(\boldsymbol{\pi},\boldsymbol{u})}(T, \boldsymbol{q})$ *for all* DMS $(\boldsymbol{\pi}, \boldsymbol{u})$. (ii) $(S, \boldsymbol{p})$ *is more informative than* $(T, \boldsymbol{q})$ *if there exists a Markov matrix* $\boldsymbol{r}$ *such that* $\boldsymbol{q} = \boldsymbol{p}\boldsymbol{r}$.

Blackwell [1] establishes the equivalence of two orderings.

THEOREM 1 (BLACKWELL 1951). $(S, \boldsymbol{p})$ *is more informative than* $(T, \boldsymbol{q})$ *if and only if* $(S, \boldsymbol{p})$ *is more valuable than* $(T, \boldsymbol{q})$.

For a short proof of the theorem we refer the reader to Crémer [3].

3 MAIN RESULT

In order to define the value of an experiment for max min or min max decision rules, we posit that a DM is identified by a convex and compact subset A of the set of all priors $\Delta(\Omega)$, such that $int(A) \neq \varnothing$.[2] Therefore, we denote a DM by a tuple $(A, \boldsymbol{u})$.

Let us start with the min max decision rule, which asserts that a DM has the prior belief that minimizes the maximum expected payoffs he can attain. Hence, we can define the value of an experiment for a DM (A, u) as

$$\mathcal{V}^*_{(A,\boldsymbol{u})}(S, \boldsymbol{p}) := \min_{\boldsymbol{\pi} \in A} \mathcal{U}^*_{(\boldsymbol{\pi},\boldsymbol{u})}(S, \boldsymbol{p}).$$

Next, we define the ordering that ranks experiments in terms of their value for min max decision rule.

DEFINITION 2. $(S, \boldsymbol{p})$ *is more* $\mathcal{V}$*-valuable than* $(T, \boldsymbol{q})$ *if* $\mathcal{V}^*_{(A,\boldsymbol{u})}(S, \boldsymbol{p}) \geq \mathcal{V}^*_{(A,\boldsymbol{u})}(T, \boldsymbol{q})$ *for all* DMS $(A, \boldsymbol{u})$.

[2] $int(A)$ denotes the interior of the set A.

The main result, which we will state next, shows the equivalence of informativeness and valuableness of experiments for the min max decision rule.

THEOREM 2. *$(S, \boldsymbol{p})$ is more informative than $(T, \boldsymbol{q})$ if and only if $(S, \boldsymbol{p})$ is more $\mathcal{V}$-valuable than $(T, \boldsymbol{q})$.*

Proof. In order to show the "if" part, suppose that there exists a Markov matrix $\boldsymbol{r}$ such that $\boldsymbol{q} = \boldsymbol{pr}$. Note that $\mathcal{V}^*_{(A,\boldsymbol{u})}(S, \boldsymbol{p}) = \mathcal{U}_{(\hat{\boldsymbol{\pi}},\boldsymbol{u})}(S, \boldsymbol{p})$ for some $\hat{\boldsymbol{\pi}} \in A$. Then, by Theorem 1, we have

$$\mathcal{V}^*_{(A,\boldsymbol{u})}(S, \boldsymbol{p}) \geq \mathcal{U}^*_{(\hat{\boldsymbol{\pi}},\boldsymbol{u})}(T, \boldsymbol{p}) \geq \mathcal{V}^*_{(A,\boldsymbol{u})}(T, \boldsymbol{p}).$$

Now, we will show the "only if" part. Let $(S, \boldsymbol{p})$ and $(T, \boldsymbol{q})$ be two experiments where $|S| = \sigma$ and $|T| = \tau$. By contrapositive, suppose that $(S, \boldsymbol{p})$ is not more informative than $(T, \boldsymbol{q})$, i.e. there is no $\boldsymbol{r}$ such that $\boldsymbol{q} = \boldsymbol{pr}$.

Let A be a convex and compact subset of $\Delta(\Omega)$ such that $int(A) \neq \varnothing$. Let $\tilde{\boldsymbol{\pi}} \in int(A)$ and define $\mathcal{P} := \{\boldsymbol{D}^{\tilde{\boldsymbol{\pi}}}\boldsymbol{pr} : \boldsymbol{r}_{\sigma\times\tau}$ is a Markov matrix$\}$, and $\mathcal{Q} := \{\boldsymbol{D}^{\boldsymbol{\pi}}\boldsymbol{q} : \boldsymbol{\pi} \in \Delta(\Omega)\}$. Note that $\mathcal{P} \cap \mathcal{Q} = \varnothing$ because $\boldsymbol{D}^{\tilde{\boldsymbol{\pi}}}\boldsymbol{pr} \in \mathcal{P} \cap \mathcal{Q}$ contradicts that there is no $\boldsymbol{r}$ for which $\boldsymbol{q} = \boldsymbol{pr}$ since $\boldsymbol{D}^{\tilde{\boldsymbol{\pi}}}$ is not singular.

Since, $\mathcal{P}$ and $\mathcal{Q}$ are disjoint, nonempty, convex, and compact sets, by separating hyperplane theorem, there exists a $\boldsymbol{v}$ such that $\langle \boldsymbol{m}, \boldsymbol{v}\rangle < 0$ and $\langle \boldsymbol{n}, \boldsymbol{v}\rangle > 0$ for all $\boldsymbol{m} \in \mathcal{P}, \boldsymbol{n} \in \mathcal{Q}$. Let us define $X := \{a_1, \cdots a_\tau\}$ and $\boldsymbol{u} := \boldsymbol{v}$. Then

$$\mathcal{V}^*_{(A,\boldsymbol{v})}(S, \boldsymbol{p}) \leq \mathcal{U}^*_{(\tilde{\boldsymbol{\pi}},\boldsymbol{v})}(S, \boldsymbol{p}) = \langle \boldsymbol{D}^{\tilde{\boldsymbol{\pi}}}\boldsymbol{pf}^*, \boldsymbol{v}\rangle.$$

Since $\boldsymbol{f}^*$ is a Markov matrix $\boldsymbol{D}^{\tilde{\boldsymbol{\pi}}}\boldsymbol{pf}^* \in \mathcal{P}$; thus, $\mathcal{V}^*_{(A,\boldsymbol{v})}(S, \boldsymbol{p}) < 0$.

Now, suppose that $\mathcal{V}^*_{(A,\boldsymbol{v})}(T, \boldsymbol{q}) = \mathcal{U}^*_{(\hat{\boldsymbol{\pi}},\boldsymbol{v})}(T, \boldsymbol{q})$ for some $\hat{\boldsymbol{\pi}} \in A$. Then for $\boldsymbol{f} = \boldsymbol{I}$, we have

$$\mathcal{U}^*_{(\hat{\boldsymbol{\pi}},\boldsymbol{v})}(T, \boldsymbol{q}) \geq \langle \boldsymbol{D}^{\hat{\boldsymbol{\pi}}}\boldsymbol{qI}, \boldsymbol{v}\rangle > 0$$

since $\boldsymbol{D}^{\hat{\boldsymbol{\pi}}}\boldsymbol{q} \in \mathcal{Q}$. Therefore, $\mathcal{V}^*_{(A,\boldsymbol{v})}(S, \boldsymbol{p}) < 0 < \mathcal{V}^*_{(A,\boldsymbol{v})}(T, \boldsymbol{q})$.

□

In order to show the analogous result for the max min decision rule, let us write

$$\mathcal{W}^*_{(A,\boldsymbol{u})}(S, \boldsymbol{p}) := \max_{\boldsymbol{f}} \min_{\boldsymbol{\pi}\in A} \mathcal{U}^{\boldsymbol{f}}_{(\boldsymbol{\pi},\boldsymbol{u})}(S, \boldsymbol{p}),$$

and define the following ordering.

DEFINITION 3. *$(S, \boldsymbol{p})$ is more $\mathcal{W}$-valuable than $(T, \boldsymbol{q})$ if $\mathcal{W}^*_{(A,\boldsymbol{u})}(S, \boldsymbol{p}) \geq \mathcal{W}^*_{(A,\boldsymbol{u})}(T, \boldsymbol{q})$ for all* DMS *$(A, \boldsymbol{u})$.*

The next theorem extends the result for the max min decision rule.

THEOREM 3. *$(S, \boldsymbol{p})$ is more informative than $(T, \boldsymbol{q})$ if and only if $(S, \boldsymbol{p})$ is more $\mathcal{W}$-valuable than $(T, \boldsymbol{q})$.*

Proof. For a given experiment $(S, \boldsymbol{p})$, $\boldsymbol{u}$ and X, let us define a map $h : \Delta(\Omega) \times M \mapsto \mathbb{R}$, where M is the set of all $\sigma \times x$ Markov matrices. h is defined by $h(\boldsymbol{\pi}, \boldsymbol{f}) := \langle \boldsymbol{D}^{\boldsymbol{\pi}} \boldsymbol{p} \boldsymbol{f}, \boldsymbol{u} \rangle$. Observe that both $\Delta(\Omega)$ and M are convex and compact. Also, $h(\boldsymbol{\pi}, \cdot)$ and $h(\cdot, \boldsymbol{f})$ are linear for all $\boldsymbol{f} \in M$, and for all $\boldsymbol{\pi} \in \Delta(\Omega)$, respectively. Therefore, by Sion [5]'s generalized min max theorem,

$$\mathcal{V}^*_{(A,\boldsymbol{u})}(S, \boldsymbol{p}) = \min_{\boldsymbol{\pi} \in A} \max_{\boldsymbol{f}} h(\boldsymbol{\pi}, \boldsymbol{f}) = \max_{\boldsymbol{f}} \min_{\boldsymbol{\pi} \in A} h(\boldsymbol{\pi}, \boldsymbol{f}) = \mathcal{W}^*_{(A,\boldsymbol{u})}(S, \boldsymbol{p}).$$

Hence, the result follows from Theorem 2. □

4 CONCLUDING REMARKS

The economic value of information has been discussed in various contexts under a number of alternative definitions. Yet, Blackwell's approach to the question has been, perhaps, the most widely accepted and useful one. This note contributes to the discussion by showing that Blackwell's theorem extends in fact to a larger domain of decision rules. In addition, the result reinforces the value of preferences—which has become increasingly important in a variety of applications—that exhibit uncertainty aversion (max min expected utility).

REFERENCES

[1] Blackwell, D. (1951). "Comparison of Experiments." *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, pp. 93-102.

[2] Blackwell, D. (1953). "Equivalent Comparison of Experiments." *Annals of Mathematical Statistics*, 24, pp. 265-272.

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

[4] Gilboa, I. and D. Schmeidler. (1989). "Maxmin Expected Utility with a Non-unique Prior." *Journal of Mathematical Economics*, 18, pp. 141-153.

[5] Sion, M. (1959). “On General Minimax Theorems.” *Pacific Journal of Mathematics*, 8, pp. 171-176.