Beyond the Blackwell Order in Dichotomies

Andrew Kosenko*

April 23, 2024

Abstract

I establish a translation invariance property of the Blackwell order for dichotomies, show that garbling experiments reduces the norm of their difference, and show that the norm of the distance from the identity matrix may be interpreted as a measure of informativeness. The better experiment is closer to the fully revealing experiment; this measure extends the Blackwell order, is complete, and prior-independent.

Keywords: Blackwell experiments, information orders.
JEL Classification: D81, D83, C44, C65.

1 Introduction

In a bedrock contribution (Blackwell (1951, 1953)), David Blackwell established the equivalence of two notions of ordinal rankings of experiments - those of informativeness, and payoff-richness (as well as the related notion of sufficiency). Here I first ask whether the Blackwell order is preserved when both the better and the worse experiments are garbled using the same garbling, and then show that the matrix norm of the difference between a fully revealing experiment and another one is a convenient and appealing completion of the Blackwell order. An application illustrating the usefulness of this completion concludes. Throughout, I focus only on dichotomies: experiments with two states and two signal realizations.

*Assistant Professor of Economics, Department of Economics, Accounting, and Finance, School of Management, Marist College. 51 Fulton St., Poughkeepsie NY, 12601. kosenko.andrew@gmail.com.
I am grateful to Navin Kartik and Nate Neligh for discussions and suggestions, and to two very helpful anonymous referees.

Section 2 first asks: Given two Blackwell-ranked experiments, is the order preserved if signal realizations from both experiments are subjected to noise? More precisely, suppose both experiments undergo yet another stochastic transformation, say, M. If A Blackwell-dominates B, does experiment MA always dominate experiment MB? Theorem 1 answers in the affirmative, highlighting a curious translation invariance property.

Equally important is the question of completing the (notoriously partial) Blackwell order. Theorem 2 shows that *all* dichotomous experiments are ranked by taking the infinity norm of the difference between any experiment and the fully revealing experiment. The interpretation is that more informative experiments are "closer" to the fully revealing experiment (represented by the identity matrix). This measure completes Blackwell's order within this class of experiments, is complete, and prior-independent. Two counterexamples follow each of theorem 1 and theorem 2, showing that neither result can be extended beyond two states or two signal realizations.

Restricting attention to such dichotomies is common; the underlying state in interesting problems often is binary (the product, project, firm, or match, is truly good or bad) and thus the assumption of two states is common in this literature (Keppo et al. (2008), de Oliveira et al. (2021), Mu et al. (2021)). Assuming binary signal realizations (studied also in Birnbaum's (1961) "simple binary experiments," Torgersen's (1970) "double dichotomies," and Blackwell and Girshik's (1979) "binomial dichotomies") reflects the fact that much of the relevant evidence (passing or failing a test or an audit, presence or absence of a pathogen or biomarker) in these settings is also binary. In addition, many economic *decisions* are binary (convict/acquit, purchase/not, approve/disapprove, vote yes/no, tests of simple hypotheses); with binary decisions, and multiple signal realizations, many of those signals would lead to one of the two decisions,[1] effectively acting as one signal realization.

Notation

Throughout, the state space $\Omega = \{\omega_0, \omega_1\} = \{0, 1\}$ and the signal space $S = \{s_0, s_1\}$ are fixed. A *Blackwell experiment* is a 2×2 stochastic matrix $P = \{p_{ij}\}$ (i.e. $p_{ij} \geq 0$, and for each j, $p_{1j} + p_{2j} = 1$; the matrix is column-stochastic, with entries representing the probabilities of signal realizations in each state). Denote by $\mathbb{1}$ the identity matrix, interpreted as a fully revealing experiment. Experiment A *Blackwell dominates* experiment B, (written $A \succeq_B B$), if an only if an expected utility maximizing decision maker (DM) will prefer A over B, or if and only if there exists a stochastic matrix Γ (a *garbling*), with $\Gamma A = B$.

[1] An insightful anonymous referee points out that with two states, two actions, and many signal realizations, the composition of an experiment with a strategy is itself a dichotomy.

2 Translation Invariance and a Cardinal Measure of Informativeness

Given two Blackwell-ranked experiments, is the ordering preserved if signal realizations from both experiments are subjected to noise? More precisely, suppose both experiments undergo yet another garbling M. If A Blackwell-dominates B, does experiment MA Blackwell-dominate experiment MB?

The question of noise added to signal realizations is animated by the growing research program grappling with the impact of noise, errors, inattention, and other imperfections in communication and interpretation, on established results. There are at least two reasons why such a second-order garbling my occur. First, the DM may be inattentive, and not recognize some signal realizations, merge, or misinterpret them (Bloedel and Segal (2021)). Second, the DM may observe signal realizations with transmission noise (Hernandez and von Stengel (2014), Blume, Board, and Kawamura (2007)).

More formally, consider the following:

Definition 1. $A = \begin{pmatrix} a_1 & 1-a_2 \\ 1-a_1 & a_2 \end{pmatrix}$ *is a diagonally dominant experiment if* $\{a_1, a_2\} \in [\frac{1}{2}, 1]^2$.

Diagonally dominant experiments do not change the signal labels on average and focusing on diagonally dominant experiments or garblings involves no loss of generality if the object of interest is the distribution of posterior beliefs. Theorem 1 given the translation invariance result:

Theorem 1 (Translation invariance of $\succeq_B$). *Let Γ_1 be a diagonally dominant garbling matrix, and take a non-singular experiment A. Let $B = \Gamma_1 A$ (i.e. $A \succeq_B B$). For any non-singular matrix M, we have that:*

i) MA Blackwell-dominates MB, and furthermore,

ii) Since there exists Γ_1 with $\Gamma_1 A = B$, there exists a matrix Γ_2, with Γ_2 similar to Γ_1 such that $\Gamma_2 MA = MB$

In other words, the diagram in figure 1 commutes.

$$\begin{array}{ccc} A & \xrightarrow{\Gamma_1} & B \\ \downarrow{\scriptstyle M} & & \downarrow{\scriptstyle M} \\ MA & \xrightarrow{\Gamma_2} & MB \end{array}$$

Figure 1: Translation invariance of $\succeq_B$

Theorem 1 has two takeaways. One is that the Blackwell order is partially *translation invariant* - the garbling M "shifts" any experiment by an amount "proportional" to the initial distance, because the resulting matrices are still ranked. The second takeaway is that Γ_1 and Γ_2 are *similar* matrices - in other words, they represent the same linear transformation, but in different bases. Thus, the features of the linear transformation that have to do with the characteristic polynomial (which does not depend on the choice of basis), such as the determinant, trace and eigenvalues, but also the rank and the normal forms, are preserved. The fact that the linear operator mapping A into B, and the linear operator mapping MA into MB, turn out to be the *same* linear operator is thought-provoking.

For a minimal counterexample (showing that the theorem does not extend beyond two signal realizations),[2] let $|\Omega| = 2$, and $|S| = 3$ and suppose A is a fully revealing experiment, and Γ_1 with $p, q \in (\frac{1}{2}, 1]$ is given below, so that $B (= \Gamma_1 A)$ is partially revealing, and take M as below.[3] Then $MA = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, a fully uninformative experiment, and $MB = \begin{pmatrix} p & 1 \\ 1-p & 0 \end{pmatrix}$. Clearly, $A \succ_B B$, yet $MA \prec_B MB$.

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \Gamma_1 = \begin{pmatrix} p & 1-q & 0 \\ 0 & q & 0 \\ 1-p & 0 & 1 \end{pmatrix}, B = \begin{pmatrix} p & 1-q \\ 0 & q \\ 1-p & 0 \end{pmatrix}, M = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1}$$

For a 3×3 counterexample (showing that the theorem does not extend beyond two states), consider (letting $B = \Gamma_1 A$):

$$A = \begin{pmatrix} 0.9 & 0.25 & 0.15 \\ 0.05 & 0.5 & 0.15 \\ 0.05 & 0.25 & 0.7 \end{pmatrix}, \Gamma_1 = \begin{pmatrix} 0.51 & 0 & 0 \\ 0.49 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, M = \begin{pmatrix} 0.7 & 0.1 & 0.2 \\ 0.1 & 0 & 0.4 \\ 0.2 & 0.9 & 0.4 \end{pmatrix} \tag{2}$$

Here MA does not Blackwell-dominate MB (the required Γ_2 is not stochastic).

Going beyond theorem 1, and restricting attention to a particular norm - the infinity norm, denoted by $||\cdot||_\infty$ - we obtain a completion of Blackwell's order, and a cardinal informativeness result.

Theorem 2 (A Cardinal Measure of Informativeness)**.** *Let A and B be two 2×2 experiments, and suppose that A is diagonally dominant. Then $A \succeq_B B$ implies $\|\mathbb{1} - A\|_\infty \leq \|\mathbb{1} - B\|_\infty$.*

[2] A version of this counterexample was suggested by Alex Frankel.

[3] Note that for such minimal counterexamples diagonal dominance is undefined.

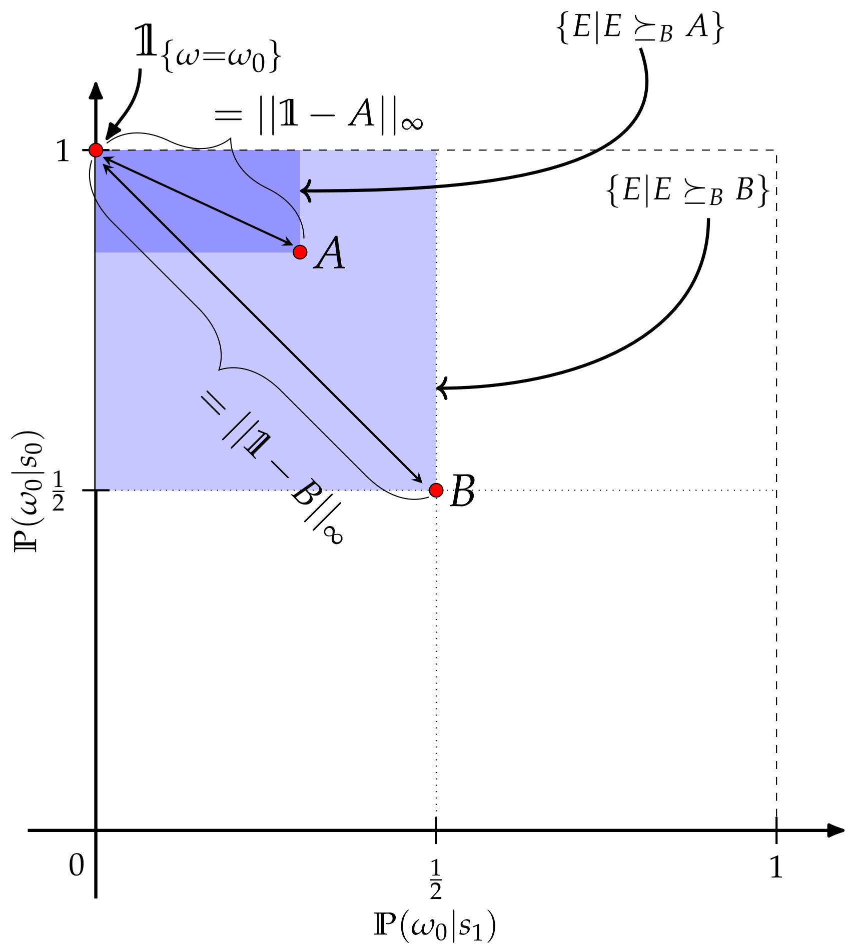

Figure 2: $A \succeq_B B \Rightarrow A \succeq_{||\cdot||_\infty} B$: Blackwell informativeness and norm differences.

The states are ω_0 and ω_1, and signal realizations are s_0 and s_1. The prior of $\omega = \omega_0$ is $\frac{1}{2}$, the true state is ω_0, and A and B are (with abuse of nomenclature) two pairs of posterior beliefs resulting from the eponymous experiments. The possible posterior beliefs after a signal realization are on the axes; in light blue is the set of experiments and posterior belief distributions that are Blackwell better than B (and a mean-preserving spread of posteriors), while in dark blue is the corresponding set for A. E is a generic experiment (and associated posterior belief distribution).

Thus, the "closer" a matrix is to full revelation, the "better" it is. The norm is a continuous function, and thus, if $A \succeq_B B$ are Blackwell ranked experiments, this completion assigns "nearby" unranked experiments values that are "close" to the values for A and B. Its interpretation also has the intuitively attractive features that relate this order to Blackwell's, and to mean preserving spreads; figure 3 illustrates.

Unfortunately, theorem 2 also does not extend beyond dichotomies. For a minimal counterexample with two states and three signal realizations, consider

$$A = \begin{pmatrix} 0.4870 & 0.5984 \\ 0.4386 & 0.2385 \\ 0.0744 & 0.1631 \end{pmatrix}, \Gamma = \begin{pmatrix} 0.2328 & 0.3042 & 0.1225 \\ 0.0644 & 0.2672 & 0.3710 \\ 0.7028 & 0.4286 & 0.5065 \end{pmatrix}, \Gamma A = B = \begin{pmatrix} 0.2559 & 0.2318 \\ 0.1761 & 0.1628 \\ 0.5680 & 0.6054 \end{pmatrix} \quad (3)$$

Using $E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ in place of the identity for the norm computations we obtain: $||E - A||_\infty - ||E - B||_\infty = 0.0268$. For an example with three states, let

$$A = \begin{pmatrix} 0.55 & 0 & 0 \\ 0.45 & 0.55 & 0.45 \\ 0 & 0.45 & 0.55 \end{pmatrix}, \Gamma = \begin{pmatrix} 0.5 & 1 & 0 \\ 0.5 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4)$$

where we let $B = \Gamma A$. In this case, $\|\mathbb{1} - A\|_\infty - \|\mathbb{1} - B\|_\infty = 0.075$.

Application

Suppose[4] a von Neumann-Morgenstern DM faces a choice between experiments $A_1 = \begin{pmatrix} 0.7 & 0.4 \\ 0.3 & 0.6 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{pmatrix}$, and a third experiment, $B = \begin{pmatrix} 0.85 & 0.49 \\ 0.15 & 0.51 \end{pmatrix}$, which is more informative than the other two in one state, and nearly uninformative in the other. A_2 Blackwell-dominates A_1, yet B is not ranked vis-à-vis either A_1 or A_2. In terms of norm distances, we have:

$$||\mathbb{1} - A_1||_\infty = 0.7 > ||\mathbb{1} - B||_\infty = 0.64 > ||\mathbb{1} - A_2||_\infty = 0.5 \quad (5)$$

How should a DM who cares about action in *both* states choose? From a (non-Blackwell-ranked) menu $\mathcal{M}_1 = \{A1, B\}$ the infinity norm difference order says that a DM should chose B (B is closer to full revelation as evidenced by a *smaller* distance to full revelation than A_1), and from a menu $\mathcal{M}_2 = \{B, A2\}$ they should choose A_2.

Notably, completing the Blackwell order using norms in dichotomies is (unlike other completions of the order) prior-independent, stated without reference to a decision problem (and thus not tied to a utility specification), simple and easy to compute, and easily interpretable in terms of mean-preserving spreads

[4]Generalizing this example is beyond the scope of this note.

Appendix: Proofs

Proof of theorem 1. $\Gamma_1 A = B$ by assumption; if a Γ_2 with the stated properties, exists, we would have $\Gamma_2 MA = MB$. But then

$$\Gamma_2 MA = MB \iff \Gamma_2 MA = M\Gamma_1 A \tag{6}$$

$$\Rightarrow \Gamma_2 M = M\Gamma_1 \tag{7}$$

$$\Rightarrow \Gamma_2 = M\Gamma_1 M^{-1} \tag{8}$$

Substituting the resulting Γ_2 verifies what was needed to show; the last equation confirms that the fact that Γ_1 and Γ_2 are similar matrices and gives an explicit formula for Γ_2. It remains to show that Γ_2 is a garbling - stochastic - matrix. Computing explicitly we obtain

$$M\Gamma_1 M^{-1} = \underbrace{\begin{pmatrix} m_1 & 1-m_2 \\ 1-m_1 & m_2 \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} \gamma_1 & 1-\gamma_2 \\ 1-\gamma_1 & \gamma_2 \end{pmatrix}}_{\Gamma_1} \underbrace{\frac{1}{|M|}\begin{pmatrix} m_2 & m_2-1 \\ m_1-1 & m_1 \end{pmatrix}}_{M^{-1}} = \tag{9}$$

$$= \begin{pmatrix} \gamma_2 + m_1 - m_2 + \gamma_1 m_2 - \gamma_2 m_1 & m_1 - \gamma_1 - m_2 + \gamma_1 m_2 - \gamma_2 m_1 + 1 \\ m_2 - m_1 - \gamma_2 - \gamma_1 m_2 + \gamma_2 m_1 + 1 & \gamma_1 - m_1 + m_2 - \gamma_1 m_2 + \gamma_2 m_1 \end{pmatrix} \tag{10}$$

and $|M| = m_1 m_2 - (1-m_2)(1-m_1) = m_1 + m_2 - 1$ and $\gamma_1, \gamma_2 \in [\frac{1}{2}, 1]$, by assumption. The columns sum to unity, and the restriction on γ_1 and γ_2 ensures that that each entry is non-negative. □

Proof of theorem 2. Let $A = \begin{pmatrix} a_1 & 1-a_2 \\ 1-a_1 & a_2 \end{pmatrix}$. Because A Blackwell-dominates B by supposition, there exists some $\Gamma = \begin{pmatrix} \gamma_1 & 1-\gamma_2 \\ 1-\gamma_1 & \gamma_2 \end{pmatrix}$ such that $B = \Gamma A$. Computing directly, $\|\mathbb{1} - A\|_\infty = 2 - a_1 - a_2$, therefore

$$\|\mathbb{1} - \Gamma A\|_\infty - \|\mathbb{1} - A\|_\infty = (2 - \gamma_1 - \gamma_2)(a_1 - a_2 - 1) \tag{11}$$

The term $2 - \gamma_1 - \gamma_2$ is always nonnegative (since Γ is column-stochastic), and the term $a_1 - a_2 - 1$ is nonnegative because of the supposition that A is diagonally dominant. Thus,

$$\|\mathbb{1} - B\| - \|\mathbb{1} - A\| = \|\mathbb{1} - \Gamma A\| - \|\mathbb{1} - A\| \geq 0 \tag{12}$$

which completes the proof.

□

References

[1] Birnbaum, A. (1961). On the Foundations of Statistical Inference: Binary Experiments. *The Annals of Mathematical Statistics*, 32(2), 414–435. http://www.jstor.org/stable/2237753

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

[3] Blackwell, D. (1953). Equivalent Comparisons of Experiments. *Ann. Math. Statist.* 24(2): 265-272.

[4] Blackwell, D. & Girshik. (1979). *Theory of Games and Statistical Decisions*. Dover Books on Mathematics.

[5] Bloedel, A. & I. Segal. (2021). Persuading a Rationally Inattentive Agent. Working paper.

[6] Blume, A., Board, O. & K. Kawamura. (2007). Noisy talk. *Theoretical Economics* 2(4), 395–440.

[7] de Oliveira, H., Y. Ishii, & X. Lin. (2021). Robust Merging of Information. arXiv:2106.00088v1, https://doi.org/10.48550/arXiv.2106.00088

[8] Hernandez, P., & B. von Stengel. (2014). Nash Codes for Noisy Channels. *Operations Research* 62:6, 1221-1235. https://doi.org/10.48550/arXiv.1202.1547

[9] Keppo, J, G. Moscarini & L. Smith. (2008). The demand for information: More heat than light. *Journal of Economic Theory* 138(1): 21-50, https://doi.org/10.1016/j.jet.2007.03.003

[10] Mu, X., Pomatto, L., Strack, P. & O. Tamuz. (2021). From Blackwell Dominance in Large Samples to Rényi Divergences and Back Again. *Econometrica*, 89: 475-506. https://doi.org/10.3982/ECTA17548

[11] Torgersen, E.N. (1970) Comparison of experiments when the parameter space is finite. Z. *Wahrscheinlichkeitstheorie und Verwandte Gebiete* 16, 219–249 . https://doi.org/10.1007/BF00534598