A proof of Blackwell's theorem

By Juan Dodyk

In these notes I prove the following version of Blackwell's theorem (keep reading for the definitions of the terms).

Theorem 1 (Blackwell). *Let Ω, S, S' be separable metrizable, $\mu_0 \in \Delta(\Omega)$, $\sigma : \Omega \to \Delta(S)$, $\sigma' : \Omega \to \Delta(S')$ signals with μ_0, $\sigma_0 = \sigma \circ \mu_0$ and $\sigma_0' = \sigma' \circ \mu_0$ regular.*[1] *Then the following are equivalent:*

(1) σ is more informative than σ' relative to μ_0,
(2) $p_\sigma = \tau \circ p_{\sigma'}$ for a mean-preserving spread τ,
(3) $\int f \, dp_\sigma \geqslant \int f \, dp_{\sigma'}$ for every $f \in C^\ell(\Delta(\Omega))$ convex,
(4) σ is more valuable than σ' relative to μ_0.

Blackwell (1951, 1953) proved a stronger version of this theorem for Ω finite. The equivalence (1) $\Leftrightarrow$ (4) was proved for general Ω by Charles Boll in the 50s, and the equivalence (2) $\Leftrightarrow$ (3) was proved independently by Pierre Cartier and Volker Strassen for Ω compact metrizable in the 60s (see Phelps, 2001, Ch. 15 for a proof based on Cartier's ideas, and Aliprantis and Border, 2006, Th. 19.40 for a proof based on Strassen's). Khan, Yu and Zhang (2020) claim that the full equivalence for general Ω is missing in the literature, and they present a proof of a version of the theorem that is stronger in some respects (part of the result doesn't require Ω to have a topology) but weaker in others (they require the signals to be continuous in some sense and absolutely continuous with respect to a given measure, and the full equivalence requires Ω compact). Their proof relies heavily on a Prokhorov theorem for random measures on Polish spaces taken from Crauel (2002), which has a very long proof, and also an approximation theorem of measures by martingales taken from Khan et al. (2008).

Taking inspiration from the simple proof of Blackwell's theorem for finite Ω, S, S' by de Oliveira (2018), I present a proof that is, I believe, significantly more self-contained than that in Khan, Yu and Zhang (2020), and simpler than Cartier's and Strassen's. The main ingredient is a result on disintegration of measures taken from Dellacherie and Meyer (1978) that I interpret as about the existence of posterior beliefs. This result is, I believe, fundamental for game theory and is not hard to prove.

The idea of the proof is as follows. First, I prove that for any signal $\sigma : \Omega \to \Delta(S)$ there is a *direct* signal $\tilde{\sigma} : \Omega \to \Delta(\Delta(\Omega))$, i.e., a signal such that its realizations are the posterior

[1] Note that every (Borel) probability measure in a Polish space is regular (Aliprantis and Border, 2006, 12.7), so if we assume Ω, S, S' Polish we can drop the assumption that $\mu_0, \sigma_0, \sigma_0'$ are regular.

beliefs they induce, such that σ and $\tilde{\sigma}$ are equally informative in the Blackwell sense, and they induce the same distribution of posteriors p_σ. For (1) $\Rightarrow$ (2), I note that σ is more informative than σ' iff $\tilde{\sigma}$ is more informative than $\tilde{\sigma}'$, where $\tilde{\sigma}$ and $\tilde{\sigma}'$ are direct. Then the fact that $\tilde{\sigma}$ is more informative than $\tilde{\sigma}'$ immediately implies that there is a mean preserving spread τ with $p_\sigma = \tau \circ p_{\sigma'}$. (2) $\Rightarrow$ (3) follows immediately by a version of Jensen inequality. (3) $\Rightarrow$ (4) follows by noting that, given a decision problem, we can write the expected payoff of an agent who observes the realization of a signal σ as the expected value of $f(\mu)$, where μ follows the distribution of posteriors induced by σ and $f(\mu)$ is the maximum expected payoff when the agent chooses her action under the posterior belief μ; f is convex, so (3) implies that σ is more valuable than σ'. Finally, (4) $\Rightarrow$ (1), the most difficult step, follows from a separation argument. If (1) doesn't hold, i.e., σ is not more informative than σ', then σ', viewed as a strategy (where the signal realizations are actions), must be outside of the set of strategies available given that the agent observes the realization of σ. Choosing the right topology, this set is closed, so the Hahn-Banach theorem applies, and we obtain a decision problem where the agent does better observing the realization of σ' than that of σ, proving that (4) is false.

Notation.—If (X, Σ_X) is a measurable space (i.e., Σ_X is a σ-algebra), $B(X)$ is the set of bounded Σ_X-measurable functions $X \to \mathbb{R}$, $\mathcal{M}(X)$ the set of finite signed measures, and $\Delta(X)$ the set of probability measures. If $x \in X$, $\delta_x \in \Delta(X)$ is given by $\delta_x(E) = \mathbb{1}(x \in E)$ for any $E \in \Sigma_X$. If Σ_X, Σ_Y are two σ-algebras, $\Sigma_X \otimes \Sigma_Y$ is the product σ-algebra, i.e., the σ-algebra generated by $\{E \times F : E \in \Sigma_X, F \in \Sigma_Y\}$. If $\mu, \nu \in \mathcal{M}(X)$, $\nu \ll \mu$ means that ν is absolutely continuous with respect to μ. If X is metrizable, I automatically endow it with the Borel σ-algebra $\mathcal{B}_X$, and I endow $\Delta(X)$ with the weak* topology (the minimal topology that makes $\mu \mapsto \int f \, d\mu$ continuous for every $f \in C_b(X)$), which makes $\Delta(X)$ metrizable if X is separable (Aliprantis and Border, 2006, 15.12); $\mu \in \Delta(X)$ is regular iff $\sup\{\mu(K) : K \subset X, K \text{ compact}\} = 1$ (see Aliprantis and Border, 2006, 12.5 and 12.6). When I say "by Riesz" I mean by the Riesz Representation Theorem for compact Hausdorff spaces (Aliprantis and Border, 2006, 14.14), and "by Hahn-Banach" means by the Strong Separating Hyperplane Theorem (Aliprantis and Border, 2006, 5.79). I will use several times the following fundamental duality result without mention. It can be proved the same way as Aliprantis and Border (2006, 5.93).

THEOREM 2. *Let X be a real vector space, and L a set of linear functions $X \to \mathbb{R}$. We endow X with the minimal topology that makes the functions in L continuous. Then X is a locally convex topological vector space, and X^*, the set of linear continuous functions $X \to \mathbb{R}$, is the vector space generated by L.*

A. *Stochastic maps*

Given two measurable spaces (X, Σ_X), (Y, Σ_Y), a *stochastic map* is a function $f : X \to \Delta(Y)$ such that, for each $g \in B(Y)$, the function $x \in X \mapsto \int g \, df(x)$, which we denote $f(g|\cdot)$, is measurable. Given $f : X \to \Delta(Y)$ and $g : Y \to \Delta(Z)$ stochastic, we define $g \circ f : X \to \Delta(Z)$ by $g \circ f(E|x) = \int g(E|\cdot)\, df(x)$. We can prove easily that it's stochastic, and $\circ$ is associative: $(h \circ g) \circ f = h \circ (g \circ f)$. Also, if $\{x\}$ is measurable for every $x \in X$, then X has an identity $\mathrm{id}_X : X \to \Delta(X)$ given by $\mathrm{id}_X(x) = \delta_x$, which is stochastic. Thus the measurable spaces with measurable singletons form a category whose morphisms are the stochastic maps. We can view measures $\mu \in \Delta(X)$ as stochastic maps $\mu : \bullet \to \Delta(X)$ from a one-point space, so $f : X \to \Delta(Y)$ stochastic induces $f \circ \mu \in \Delta(Y)$.

Given a *prior* $\mu_0 \in \Delta(\Omega)$ and a *signal* $\sigma : \Omega \to \Delta(S)$ (stochastic), we can define the ex ante measure $\sigma_0 = \sigma \circ \mu_0$. We say that the stochastic map $\beta : S \to \Delta(\Omega)$ is a *posterior map* if $\int_F \beta(E|\cdot)\, d\sigma_0 = \int_E \sigma(F|\cdot)\, d\mu_0$ for all $E \in \Sigma_\Omega$, $F \in \Sigma_S$, or equivalently $\int f\beta(g|\cdot)\, d\sigma_0 = \int \sigma(f|\cdot) g\, d\mu_0$ for all $f \in B(S)$, $g \in B(\Omega)$. I represent this with the following diagram.

$$\begin{tikzcd} \bullet \arrow[r, "\mu_0"] \arrow[dr, "\sigma_0"'] & \Omega \arrow[d, "\sigma"] \\ & S \arrow[u, dashed, bend right=60, "\beta"'] \end{tikzcd}$$

Theorem 3 (Existence of posteriors). *If Ω is separable metrizable, (S, Σ_S) is measurable, $\mu_0 \in \Delta(\Omega)$ is regular, and $\sigma : \Omega \to \Delta(S)$ is stochastic, then there is a posterior map β, and if β, β' are posterior maps then $\beta = \beta'$ σ_0-a.e.*

Proof.[2] Assume first that Ω is compact. Let $D \subset C(\Omega)$ be dense, numerable, containing $q\mathbb{1}_\Omega$ for $q \in \mathbb{Q}$, and closed under addition (it exists by Aliprantis and Border, 2006, 9.14). Every $f \in D$ induces a measure $\nu_f(F) = \int f\sigma(F|\cdot)\, d\mu_0$ for $F \in \Sigma_S$ such that $\nu_f \ll \sigma_0$, so let $\beta(f|\cdot) = \frac{d\nu_f}{d\sigma_0}$, the Radon-Nikodym derivative. We have that $\beta(\cdot|s)$ is additive and monotone ($\beta(f|s) \geqslant \beta(g|s)$ if $f \geqslant g$), and $\beta(q\mathbb{1}_\Omega|s) = q$ for all $s \in \tilde{S} \in \Sigma_S$, where $\sigma_0(\tilde{S}) = 1$, and we get $|\beta(f|\cdot) - \beta(g|\cdot)| \leqslant \|f - g\|_\infty$ in $\tilde{S}$ (we have $f \leqslant g + q\mathbb{1}_\Omega$ for any $q \in \mathbb{Q}$, $q \geqslant \|f - g\|_\infty$, so it follows by additivity and monotonicity). For each $s \in \tilde{S}$ and $f \in C(\Omega)$ we define $\beta(f|s) = \lim_n \beta(f_n|s)$ for any $f_n \to f$ with $f_n \in D$, clearly well-defined; clearly $\beta(\cdot|s)$ is linear and continuous, hence by Riesz it is a measure. For $s \notin \tilde{S}$ we set $\beta(s)$ constant. Let $\mathcal{F}$ be the set of $f \in B(\Omega)$ such that $\beta(f|\cdot)$ is measurable; it is a vector space closed by pointwise dominated limits that contains D, hence it contains $C(\Omega)$, and therefore every $f \in B(\Omega)$. We have that $\int_F \beta(f|\cdot)\, d\sigma_0 = \int f\sigma(F|\cdot)\, d\mu_0$ for every $F \in \Sigma_S$, $f \in D$, hence it's true for every $f \in B(\Omega)$, and therefore β is a posterior distribution, as desired. If β' is another posterior, for every $f \in D$ it agrees σ_0-a.e. with β, so $\beta = \beta'$ σ_0-a.e.

Now, let Ω be separable. By regularity there is $K = \bigcup_{n\in\mathbb{N}} K_n$ with K_n compact, $\mu_0(K) = 1$. We embed Ω in a compact metric space $\tilde{\Omega}$ (see the proof of Aliprantis and Border, 2006, 15.12). Let $s \in S$. We define $\tilde{\mu}_0 \in \Delta(\tilde{\Omega})$ by $\tilde{\mu}_0(E) = \mu_0(E \cap K)$ and $\tilde{\sigma} : \tilde{\Omega} \to \Delta(S)$ by $\tilde{\sigma}(x) = \sigma(x)$ if $x \in K$, and δ_s otherwise; $\tilde{\sigma}$ is stochastic since $K \in \mathcal{B}_{\tilde{\Omega}}$. We apply the result and obtain $\tilde{\beta} : S \to \Delta(\tilde{\Omega})$ measurable. Let $\tilde{S} = \{s \in S : \tilde{\beta}(K|s) = 1\}$, $\omega \in K$, and $\beta : S \to \Delta(\Omega)$ be given

[2] I adapt the proof in Dellacherie and Meyer (1978, p. 78), which establishes a very similar result.

by $\beta(E|s) = \tilde{\beta}(E \cap K|s)$ if $s \in \tilde{S}$ and $\beta(s) = \delta_\omega$ otherwise. It's easy to see that $\tilde{\sigma}_0(\tilde{S}) = 1$, and β is a posterior of σ. For uniqueness, if β' is another posterior map let $\tilde{S} = \{s \in S : \beta'(K|s) = 1\}$; we have $\sigma_0(\tilde{S}) = 1$, so define $\tilde{\beta}' : S \to \Delta(\tilde{\Omega})$ by $\tilde{\beta}'(E|s) = \beta(K \cap E|s)$ if $s \in \tilde{S}$, and $\tilde{\beta}'(s) = \delta_\omega$ otherwise. We see that $\tilde{\beta}'$ is another posterior map of $\tilde{\sigma}$, so $\tilde{\beta} = \tilde{\beta}'$ holds $\tilde{\sigma}_0$-a.e., and $\beta = \beta'$ holds σ_0-a.e. ∎

The following is immediate but worth recording.

Proposition 1. *If $\beta : S \to \Delta(\Omega)$ is a posterior for $\sigma : \Omega \to \Delta(S)$ over prior $\mu_0 \in \Delta(\Omega)$ then for every $f \in B(S \times \Omega)$ we have $\iint f(s,\omega)\, d\sigma(s|\omega)\, d\mu_0(\omega) = \iint f(s,\omega)\, d\beta(\omega|s)\, d\sigma_0(s)$.*

Proof. Let $\mathcal{F}$ be the set of functions that satisfy the conclusion. It is a vector space and it's closed under pointwise dominated limits. The fact that β is a posterior implies that $fg \in \mathcal{F}$ for every $f \in B(S), g \in B(\Omega)$. Therefore $\mathbb{1}_{E \times F} \in \mathcal{F}$ for every $E \in \Sigma_S, F \in \Sigma_\Omega$, so $\mathcal{A} = \{E \in \Sigma_S \otimes \Sigma_\Omega : \mathbb{1}_E \in \mathcal{F}\}$ is a monotone class that includes an algebra including $E \times F$ with $E \in \Sigma_S, F \in \Sigma_\Omega$, hence $\mathcal{A} = \Sigma_S \otimes \Sigma_\Omega$, and every simple function $\sum_{i=1}^n c_i \mathbb{1}_{E_i}$ with $E_i \in \Sigma_S \otimes \Sigma_\Omega$ is in $\mathcal{F}$. We are done, since $\iint f(s,\omega)\, d\sigma(s|\omega)\, d\mu_0(\omega) = \sup\{\iint g(s,\omega)\, d\sigma(s|\omega)\, d\mu_0(\omega) : 0 \leqslant g \leqslant f \text{ simple}\}$ for $f \geqslant 0$, and similarly for the other one. ∎

Definition 1 (Distribution of posteriors). Given a regular prior $\mu_0 \in \Delta(\Omega)$ and a signal $\sigma : \Omega \to \Delta(S)$, a posterior map β_σ of σ induces a measure $p_\sigma \in \Delta(\Delta(\Omega))$, the *distribution of posteriors*, given by $p_\sigma(E) = \sigma_0(\beta_\sigma^{-1}(E))$ for $E \in \mathcal{B}_{\Delta(\Omega)}$. Notice that it is independent of the choice of β_σ.

Proposition 2. *Let Ω be separable metrizable, $\mu_0 \in \Delta(\Omega)$ regular, and $\sigma : \Omega \to \Delta(S)$ a signal. Then p_σ is regular.*

Proof. Using the construction and notation of the proof of Theorem 3, we have $\tilde{p} \in \Delta(\tilde{\Omega})$ defined by $\tilde{p}(E) = \tilde{\sigma}_0(\tilde{\beta}^{-1}(E))$ for $E \in \mathcal{B}_{\Delta(\tilde{\Omega})}$ is regular, since every finite Borel measure on a compact metric space is regular (Aliprantis and Border, 2006, 12.7). Now $p(\Delta(K)) = 1$ since $\beta(S) \subset \Delta(K)$, so $p(E) = \tilde{p}(E \cap \Delta(K))$ if $E \in \mathcal{B}_{\Delta(\Omega)}$, and p is regular as well. ∎

We have $\int \mu(E)\, dp_\sigma(\mu) = \mu_0(E)$ for every $E \in \mathcal{B}_\Omega$, i.e., the mean of the posteriors is the prior, since $\int \mu(E)\, dp_\sigma(\mu) = \int \beta_\sigma(E|\cdot)\, d\sigma_0 = \int_E \sigma(\Omega|\cdot)\, d\mu_0 = \mu_0(E)$. Conversely, if $p \in \Delta(\Delta(\Omega))$ regular is such that $\int \mu\, dp(\mu) = \mu_0$, define $S = \Delta(\Omega)$, $\beta : S \to \Delta(\Omega)$ by $\beta(\mu) = \mu$, and let $\sigma : \Omega \to \Delta(S)$ be a posterior map of β with prior p (it exists by Theorem 3). We have $\beta_0 = \int \beta(\cdot|\mu)\, dp(\mu) = \int \mu\, dp(\mu) = \mu_0$, so the definition of σ implies $\int \sigma(E|\cdot)\, d\mu_0 = \int_E \beta(\Omega|\cdot)\, dp = p(E)$, and we get $\sigma_0(E) = \int \sigma(E|\cdot)\, d\mu_0 = p(E)$ and $p_\sigma(E) = \sigma_0(\beta^{-1}(E)) = \sigma_0(E) = p(E)$. In other words, any distribution of posteriors $p \in \Delta(\Delta(\Omega))$ such that $\int \mu\, dp(\mu) = \mu_0$ is in fact the distribution of posteriors of some signal. Let's record this.

Proposition 3. *Let Ω be separable metrizable, $\mu_0 \in \Delta(\Omega)$ and $p \in \Delta(\Delta(\Omega))$ regular. There is a signal $\sigma : \Omega \to \Delta(S)$ such that $p = p_\sigma$ iff $\int \mu\, dp(\mu) = \mu_0$.*

In the reasoning above, note that we can induce the posteriors' distribution p_σ using a signal $\tilde{\sigma} : \Omega \to \Delta(\Delta(\Omega))$ such that its posterior map is $\mu \mapsto \mu$, i.e., its realizations are their own posteriors. This motivates the following definition.

Definition 2 (Direct signals). A signal $\sigma : \Omega \to \Delta(\Delta(\Omega))$ is *direct* with respect to $\mu_0 \in \Delta(\Omega)$ if $\mu \mapsto \mu$ is a posterior map with prior μ_0.

We want to formalize the intuition that any signal is "informationally equivalent" (in some sense) to a direct signal.

Definition 3 (Blackwell informativeness). Given two signals $\sigma : \Omega \to \Delta(S)$ and $\sigma' : \Omega \to \Delta(S')$, we say that σ is *Blackwell-more informative* than σ' relative to μ_0, $\sigma \geqslant_B^{\mu_0} \sigma'$, if there is $\tau : S \to \Delta(S')$ stochastic such that $\int \sigma'(f|\cdot) g \, d\mu_0 = \int \tau \circ \sigma(f|\cdot) g \, d\mu_0$ holds for every $f \in B(S')$, $g \in B(\Omega)$. It's easy to see that this relation is symmetric and transitive. We say that $\sigma \sim_B^{\mu_0} \sigma'$, i.e., σ and σ' are *equally informative*, if $\sigma \geqslant_B^{\mu_0} \sigma'$ and $\sigma' \geqslant_B^{\mu_0} \sigma$.

Proposition 4. *If S' is separable metrizable then $\sigma \geqslant_B^{\mu_0} \sigma'$ iff $\sigma' = \tau \circ \sigma$ holds μ_0-a.e.*

Proof. Following the proof of Aliprantis and Border (2006, 15.12), there is a countable set $D \subset C_b(S')$ such that $\mu, \mu' \in \Delta(S')$ are equal iff $\mu(f) = \mu'(f)$ for every $f \in D$. For each $f \in D$, $\sigma \geqslant_B^{\mu_0} \sigma'$ implies that $\sigma'(f|\cdot) = \tau \circ \sigma(f|\cdot)$ holds μ_0-a.e., hence taking the intersection of those sets, the equation holds for every $f \in D$ in a μ_0-full measure set, and in that set we get that $\sigma' = \tau \circ \sigma$. ■

Proposition 5. *Let Ω, S be separable metrizable, $\mu_0 \in \Delta(\Omega)$ regular, $\sigma : \Omega \to \Delta(S)$ stochastic, and $\sigma_0 = \sigma \circ \mu_0$ regular. There is a $\tilde{\sigma} : \Omega \to \Delta(\Delta(\Omega))$ direct with $\sigma \sim_B^{\mu_0} \tilde{\sigma}$, and $p_{\tilde{\sigma}} = p_\sigma$.*

Proof. Given $\sigma : \Omega \to \Delta(S)$ we can define a stochastic map $\tilde{\beta}_\sigma : S \to \Delta(\Delta(\Omega))$ by $\tilde{\beta}_\sigma(s) = \delta_{\beta_\sigma}$ that maps each signal realization to the posterior it induces. We can define a new signal $\tilde{\sigma} : \Omega \to \Delta(\Delta(\Omega))$ by $\tilde{\sigma} = \tilde{\beta}_\sigma \circ \sigma$, so $p_\sigma = \tilde{\sigma} \circ \mu_0$. Let's verify that $\tilde{\sigma}$ is direct, i.e., the map $\beta_{\tilde{\sigma}} : \mu \mapsto \mu$ is a posterior for $\tilde{\sigma}$. To prove this, we have to verify that $\int f \beta_{\tilde{\sigma}}(g|\cdot) \, dp_\sigma = \int g \tilde{\sigma}(f|\cdot) \, d\mu_0$ for any $f \in B(\Delta(\Omega))$, $g \in B(\Omega)$. Now,

$$\begin{aligned}
\int f \beta_{\tilde{\sigma}}(g|\cdot) \, dp_\sigma &= \iint f(\mu)\mu(g) \, d\tilde{\sigma}(\mu|\cdot) \, d\mu_0 = \iiint f(\mu)\mu(g) \, d\tilde{\beta}_\sigma(\mu|s) \, d\sigma(s|\cdot) \, d\mu_0 \\
&= \int f(\beta_\sigma(s))\beta_\sigma(g|s) \, d\sigma_0(s) = \int \sigma(f \circ \beta_\sigma|\cdot) g \, d\mu_0 \\
&= \iint f(\beta_\sigma(s)) \, d\sigma(s|\cdot) g \, d\mu_0 = \iint \tilde{\beta}_\sigma(f|s) \, d\sigma(s|\cdot) g \, d\mu_0 \\
&= \int \tilde{\sigma}(f|\cdot) g \, d\mu_0,
\end{aligned}$$

as desired.

Clearly $\sigma \geqslant_B \tilde{\sigma}$. We have to prove that $\tilde{\sigma} \geqslant_B^{\mu_0} \sigma$. Let τ be the posterior of $\tilde{\beta}_\sigma$ over the prior

σ_0, which exists by Theorem 3 since σ_0 is assumed regular. See the following diagram.

$$\begin{tikzcd}
\bullet \arrow[r, "\mu_0"] \arrow[rr, bend left, "\sigma_0"] \arrow[drr, bend right, "p_\sigma"'] & \Omega \arrow[r, "\sigma"] \arrow[dr, "\tilde{\sigma}"'] & S \arrow[d, "\tilde{\beta}_\sigma"] \\
& & \Delta(\Omega) \arrow[u, dashed, bend right=60, "\tau"']
\end{tikzcd}$$

We have to verify that $\int \tau \circ \tilde{\sigma}(f|\cdot) g \, d\mu_0 = \int \sigma(f|\cdot) g \, d\mu_0$ for any $f \in B(S)$, $g \in B(\Omega)$. To start, let's write down the definition of τ. This is that $\int \hat{g}\tau(f|\cdot) \, dp_\sigma = \int f\tilde{\beta}_\sigma(\hat{g}|\cdot) \, d\sigma_0$ for any $f \in B(S)$, $\hat{g} \in B(\Delta(\Omega))$. We apply it to $\hat{g}(\mu) = \mu(g)$, which is in $B(\Delta(\Omega))$ by Aliprantis and Border (2006, 15.13). The LHS is

$$\begin{aligned}
\int \mu(g)\tau(f|\mu) \, dp_\sigma(\mu) &= \iint \mu(g)\tau(f|\mu) \, d\tilde{\sigma}(\mu|\cdot) \, d\mu_0 = \iiint \mu(g)\tau(f|\mu) \, d\tilde{\beta}_\sigma(\mu|s) \, d\sigma(s|\cdot) \, d\mu_0 \\
&= \int \beta_\sigma(g|s)\tau(f|\beta_\sigma(s)) \, d\sigma_0(s) = \int g\sigma(\tau(f|\beta_\sigma(\cdot))|\cdot) \, d\mu_0 \\
&= \int g \int \tau(f|\beta_\sigma(s)) \, d\sigma(s|\cdot) \, d\mu_0 = \int g \iint \tau(f|\cdot) \, d\tilde{\beta}_\sigma(s) \, d\sigma(s|\cdot) \, d\mu_0 \\
&= \int g\tau \circ \tilde{\sigma}(f|\cdot) \, d\mu_0
\end{aligned}$$

and the RHS is $\int f\tilde{\beta}_\sigma(\tilde{g}|\cdot) \, d\sigma_0 = \int f\beta_\sigma(g|\cdot) \, d\sigma_0 = \int \sigma(f|\cdot) g \, d\mu_0$, so we get that $\int \tau \circ \tilde{\sigma}(f|\cdot) g \, d\mu_0 = \int \sigma(f|\cdot) g \, d\mu_0$, hence $\tilde{\sigma} \geqslant_B^{\mu_0} \sigma$, as desired. ∎

Hence if $\sigma : \Omega \to \Delta(S)$ and $\sigma' : \Omega \to \Delta(S')$ are two signals with S, S' separable metrizable such that their ex ante measures σ_0 and σ_0' are regular, we have that $\sigma \geqslant_B^{\mu_0} \sigma'$ iff $\tilde{\sigma} \geqslant_B^{\mu_0} \tilde{\sigma}'$, where $\tilde{\sigma}, \tilde{\sigma}' : \Omega \to \Delta(\Delta(\Omega))$ are direct.

I will use a version of Jensen's inequality. First, I need the following.

Definition 4 ($C^\ell(\Delta(X))$). If X is metrizable, $C^\ell(\Delta(X))$ is the set of functions $f \in B(\Delta(X))$ that are lower-semicontinuous, i.e., the epigraph $\{(\mu, \lambda) \in \Delta(X) \times \mathbb{R} : f(\mu) \leqslant \lambda\}$ is closed, when $\Delta(X)$ has the minimal topology that makes $\hat{g}(\mu) = \mu(g)$ continuous for every $g \in B(X)$.

Proposition 6 (Jensen). *Let X be metrizable, $p \in \Delta(\Delta(X))$, $\bar{p} \in \Delta(X)$ such that $\bar{p}(f) = \int \mu(f) \, dp(\mu)$ for every $f \in B(X)$, and $f \in C^\ell(\Delta(X))$ convex. Then $\int f \, dp \geqslant f(\bar{p})$.*

Proof.[3] The epigraph $C = \{(\mu, \lambda) \in \Delta(X) \times \mathbb{R} : f(\mu) \leqslant \lambda\}$ is convex and closed in $\mathcal{M}(X) \times \mathbb{R}$ with the topology induced by $\{\hat{g}\}_{g \in B(X)}$. If $(\bar{p}, \int f \, dp) \notin C$, by Hahn-Banach there is $\phi : \mathcal{M}(X) \times \mathbb{R} \to \mathbb{R}$ linear continuous that separates them. Now $\phi(\mu, \lambda) = \hat{g}(\mu) + a\lambda$ for $g \in B(X)$, $a \in \mathbb{R}$. Hence there is $t \in \mathbb{R}$ such that $\hat{g}(\mu) + af(\mu) > t > \hat{g}(\bar{p}) + a \int f \, dp$ for every $\mu \in \Delta(X)$. Integrating we get $\int (\hat{g}(\mu) + af(\mu)) \, dp(\mu) = \hat{g}(\bar{p}) + a \int f \, dp \geqslant t$, absurd. Hence $\int f \, dp \geqslant f(\bar{p})$, as claimed. ∎

We are ready to prove (1) ⇒ (2) ⇒ (3) in Theorem 1. First, a definition.

[3] This is folklore, but it's easier to include a proof than to find a proof of this exact version in the literature.

Definition 5 (Mean-preserving spread). A stochastic map $\tau : \Delta(\Omega) \to \Delta(\Delta(\Omega))$ is a *mean-preserving spread* iff for every $\mu \in \Delta(\Omega)$ we have that $\int \nu \, d\tau(\nu|\mu) = \mu$, i.e., $\int \nu(f) \, d\tau(\nu|\mu) = \mu(f)$ for every $f \in B(\Delta(\Omega))$.

Proposition 7. *Let Ω, S, S' be separable metrizable, $\mu_0 \in \Delta(\Omega)$ regular, $\sigma : \Omega \to \Delta(S)$, $\sigma' : \Omega \to \Delta(S')$ signals with σ_0, σ_0' regular. If $\sigma \geqslant_B^{\mu_0} \sigma'$ then there is a mean-preserving spread $\gamma : \Delta(\Omega) \to \Delta(\Delta(\Omega))$ such that $p_\sigma = \gamma \circ p_{\sigma'}$. If the latter is the case, then $\int f \, dp_\sigma \geqslant \int f \, dp_{\sigma'}$ for every $f \in C^\ell(\Delta(\Omega))$ convex.*

Proof. By Proposition 5 we can assume that $\sigma, \sigma' : \Omega \to \Delta(\Delta(\Omega))$ are direct. We have $\sigma \geqslant_B^{\mu_0} \sigma'$, so $\sigma' = \tau \circ \sigma$ holds μ_0-a.e., where $\tau : \Delta(\Omega) \to \Delta(\Delta(\Omega))$ is stochastic. Let $\gamma : \Delta(\Omega) \to \Delta(\Delta(\Omega))$ be the posterior map of τ with prior p_σ, which exists by Proposition 2 and Theorem 3. This means that $\int f\gamma(g|\cdot) \, dp_{\sigma'} = \int \tau(f|\cdot) g \, dp_\sigma$ for every $f, g \in B(\Delta(\Omega))$. See the diagram.

$$\begin{tikzcd}
\bullet \arrow[r, "\mu_0"] \arrow[rr, bend left, "p_\sigma"] \arrow[drr, bend right, "p_{\sigma'}"'] & \Omega \arrow[r, "\sigma"] \arrow[dr, "\sigma'"'] & \Delta(\Omega) \arrow[d, "\tau"] \\
& & \Delta(\Omega) \arrow[u, dashed, bend right, "\gamma"']
\end{tikzcd}$$

We note that $p_\sigma = \gamma \circ p_{\sigma'}$ by taking $f = \mathbb{1}_{\Delta(\Omega)}$ in the equation above. Moreover, $\int \nu \, d\gamma(\nu|\mu) = \mu$ holds for $p_{\sigma'}$-almost all μ. To see this, take any $f \in B(\Delta(\Omega))$, $g \in B(\Omega)$, and note that

$$\begin{aligned}
\int f \int \nu(g) \, d\gamma(\nu|\cdot) \, dp_{\sigma'}(\mu) = \int f\gamma(\nu \mapsto \nu(g)|\cdot) \, dp_{\sigma'} &= \int \tau(f|\mu)\mu(g) \, dp_\sigma(\mu) \\
&= \int g\sigma(\tau(f|\cdot)|\cdot) \, d\mu_0,
\end{aligned}$$

and $\int f(\mu)\mu(g) \, dp_{\sigma'}(\mu) = \int \sigma'(f|\cdot) g \, d\mu_0 = \int g \int \tau(f|\mu) \, d\sigma(\mu|\cdot) \, d\mu_0 = \int g\sigma(\tau(f|\cdot)|\cdot) \, d\mu_0$, so they are equal. Noting that the definition of γ and the equation $p_\sigma = \gamma \circ p_{\sigma'}$ still hold if we change γ in a set of $p_{\sigma'}$ measure zero, we can set $\gamma(\mu) = \delta_\mu$ for μ in the set where $\int \nu \, d\gamma(\nu|\mu) = \mu$ doesn't hold, so it holds everywhere. Therefore γ is a mean-preserving spread, as desired.

Now, let $f \in C^\ell(\Delta(\Omega))$ be convex, and $p_\sigma = \gamma \circ p_{\sigma'}$ where γ is a mean-preserving spread. We have

$$\int f \, dp_\sigma = \iint f(\nu) \, d\gamma(\nu|\cdot) \, dp_{\sigma'} \geqslant \int f\left(\int \nu \, d\gamma(\nu|\cdot)\right) dp_{\sigma'} = \int f \, dp_{\sigma'}$$

by Proposition 6. ■

B. *Decision problems and Blackwell's theorem*

Let Ω be separable metrizable and $\mu_0 \in \Delta(\Omega)$. A *decision problem* is a pair (A, u) where A compact metrizable is a set of actions and $u : A \times \Omega \to \mathbb{R}$ bounded, the payoff function, is a

Carathéodory function, i.e., $u(\cdot,\omega)$ is continuous for each $\omega \in \Omega$ and $u(a,\cdot)$ is measurable for each $a \in A$. (This implies $u \in B(A \times \Omega)$ by Aliprantis and Border, 2006, 4.51.) A *strategy* is a stochastic map $\alpha : \Omega \to \Delta(A)$, and let $\mathcal{A}$ be the set of strategies. We can define the *expected payoff* of a strategy as $U(\alpha) = \iint u(a,\omega)\, d\alpha(a|\omega)\, d\mu_0(\omega)$. If an agent observes the realization of a signal $\sigma : \Omega \to \Delta(S)$, she can choose an action $\tilde{\alpha} : S \to \Delta(A)$ that induces a strategy $\alpha = \tilde{\alpha} \circ \sigma$, so let $\mathcal{A}_\sigma = \{\tilde{\alpha} \circ \sigma \mid \tilde{\alpha} : S \to \Delta(A) \text{ stochastic}\}$ be the set of strategies available given signal σ, and let $V(\sigma) = \sup\{U(\alpha) : \alpha \in \mathcal{A}_\sigma\}$ be the value of signal σ. We say that σ is *more valuable* than σ', $\sigma \geqslant_D^{\mu_0} \sigma'$, if $V(\sigma) \geqslant V(\sigma')$ for every decision problem (A,u). It's a symmetric and transitive relation.

Proposition 8. *Let Ω be separable metrizable, $\mu_0 \in \Delta(\Omega)$, and $\sigma : \Omega \to \Delta(S)$, $\sigma' : \Omega \to \Delta(S')$ signals. Then $\sigma \geqslant_B^{\mu_0} \sigma'$ implies $\sigma \geqslant_D^{\mu_0} \sigma'$.*

Proof. Assume that $\sigma \geqslant_B^{\mu_0} \sigma'$, so let $\tau : S \to \Delta(S')$ stochastic such that $\int \sigma'(f|\cdot) g\, d\mu_0 = \int \tau \circ \sigma(f|\cdot) g\, d\mu_0$ for each $f \in B(S')$, $g \in B(\Omega)$. Using the same argument as in Proposition 1 we obtain that $\iint f(s',\omega)\, d\sigma'(s'|\omega)\, d\mu_0(\omega) = \iint f(s',\omega)\, d\tau \circ \sigma(s'|\omega)\, d\mu_0(\omega)$ for every $f \in B(S' \times \Omega)$. Now let (A,u) be a decision problem. We want to show that $V(\sigma) \geqslant V(\sigma')$. It's enough to show that $\{U(\alpha) : \alpha \in \mathcal{A}_{\sigma'}\} \subset \{U(\alpha) : \alpha \in \mathcal{A}_\sigma\}$. So let $\alpha : S' \to \Delta(A)$ be stochastic, and let $\alpha' = \alpha \circ \tau$. Define $f(s',\omega) = \alpha(u(\cdot,\omega)|s')$, which is in $B(S' \times \Omega)$. Now $\int u(a,\omega)\, d\alpha \circ \sigma'(a|\omega) = \int \alpha(u(\cdot,\omega)|s')\, d\sigma'(s'|\omega) = \int f(s',\omega)\, d\sigma'(s'|\omega)$, and $\int u(a,\omega)\, d\alpha' \circ \sigma(s|\omega)\, d\mu_0(\omega) = \iint f(s',\omega)\, d\tau \circ \sigma(s'|\omega)\, d\mu_0(\omega)$. Hence $U(\alpha' \circ \sigma) = U(\alpha \circ \sigma')$, and we are done. ∎

In fact, we can prove something stronger.

Proposition 9. *Let Ω, S, S' separable metrizable, $\mu_0 \in \Delta(\Omega)$ regular, $\sigma : \Omega \to \Delta(S)$, $\sigma' : \Omega \to \Delta(S')$ signals with σ_0, σ_0' regular. Then if $\int f\, dp_\sigma \geqslant \int f\, dp_{\sigma'}$ for all $f \in C^\ell(\Delta(\Omega))$ convex then $\sigma \geqslant_D^{\mu_0} \sigma'$.*

Proof. Using Proposition 5 and the above we can assume that $\sigma, \sigma' : \Omega \to \Delta(\Delta(\Omega))$ are direct. Now using Proposition 1 we have

$$\begin{aligned} V(\sigma) &= \sup\left\{\iint u(\alpha(\mu),\omega)\, d\sigma(\mu|\omega)\, d\mu_0(\omega) \mid \alpha : \Delta(\Omega) \to \Delta(A) \text{ stochastic}\right\} \\ &= \sup\left\{\iint u(\alpha(\mu),\omega)\, d\mu\, dp_\sigma(\mu) \mid \alpha : \Delta(\Omega) \to \Delta(A) \text{ stochastic}\right\}. \end{aligned}$$

By the Measurable Maximum Theorem (Aliprantis and Border, 2006, Theorem 18.19) applied to $\tilde{u}(a,\mu) = \int u(a,\cdot)\, d\mu$, we have that there is $\alpha : \Delta(\Omega) \to \Delta(A)$ measurable such that $\alpha(\mu)$ maximizes $\int u(a,\cdot)\, d\mu$ for each μ, hence $V(\sigma) = \int f\, dp_\sigma$, where $f(\mu) = \max_{a \in \Delta(A)} \int u(a,\cdot)\, d\mu$. Clearly f is convex, bounded and measurable, and $f \in C^\ell(\Delta(\Omega))$, since the epigraph is $\bigcap_{a \in \Delta(A)} \{(\mu,\lambda) \in \Delta(\Omega) \times \mathbb{R} : \int u(a,\cdot)\, d\mu \leqslant \lambda\}$, and each of those sets is closed in the relevant topology since $u(a,\cdot) \in B(\Omega)$, so we obtain $V(\sigma) \geqslant V(\sigma')$. ∎

We are ready to prove the main theorem.

Theorem 1 (Blackwell). *Let Ω, S, S' be separable metrizable, $\mu_0 \in \Delta(\Omega)$, $\sigma : \Omega \to \Delta(S)$, $\sigma' : \Omega \to \Delta(S')$ signals with μ_0, $\sigma_0 = \sigma \circ \mu_0$ and $\sigma_0' = \sigma' \circ \mu_0$ regular.*[4] *Then the following are equivalent:*

(1) σ is more informative than σ' relative to μ_0,

(2) $p_\sigma = \tau \circ p_{\sigma'}$ for a mean-preserving spread τ,

(3) $\int f\, dp_\sigma \geqslant \int f\, dp_{\sigma'}$ for every $f \in C^\ell(\Delta(\Omega))$ convex,

(4) σ is more valuable than σ' relative to μ_0.

Proof. We proved (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) in Propositions 7 and 9, so we have to prove (4) $\Rightarrow$ (1). Suppose that (1) doesn't hold, so assume that $\sigma : \Omega \to \Delta(S)$, $\alpha : \Omega \to \Delta(A)$ are stochastic, and σ is not more informative than α relative to μ_0. This means that there is no $\tilde{\alpha} : S \to \Delta(A)$ stochastic such that $\alpha = \tilde{\alpha} \circ \sigma$ μ_0-a.e. We want to prove that (4) doesn't hold. Let's assume first that A is compact.

Let M_S be the vector space of measurable bounded functions $a : S \to \mathcal{M}(A)$ with the minimal topology that makes the linear functions $L_{f,g}(a) = \int a(f|\cdot) g\, d\sigma_0$ continuous, where $f \in C(A)$ and $g \in B(S)$. The set M_S^* of linear continuous functions $\phi : M_S \to \mathbb{R}$ is the vector space generated by $\{L_{f,g} : f \in C(A), g \in B(S)\}$. Let A_S be the subset of M_S composed by $a : S \to \Delta(A)$. I claim that it is compact. Let $I = C(A) \times B(S)$ and

$$\iota : A_S \to K = \prod_{(f,g) \in I} [-\|f\|_\infty \|g\|_\infty, \|f\|_\infty \|g\|_\infty]$$

given by $\iota(a) = (\int a(f|\cdot) g\, d\sigma_0)_{(f,g)\in I}$. Let $a_\alpha \in A_S$ be a net. By Tychonoff, K is compact, hence $\iota(a_\alpha)$ has a convergent subnet $\iota(a_{\alpha_\beta}) \to \bar{a} \in K$. We proceed as in the proof of Theorem 3 and construct $a : S \to \Delta(A)$ measurable such that $\int a(f|\cdot) g\, d\sigma_0 = \bar{a}(f, g)$ for all $f \in C(A)$, $g \in B(S)$. Let $D \subset C(A)$ be dense, numerable, containing $q\mathbb{1}_A$ for every $q \in \mathbb{Q}$, and closed under addition. Every $f \in D$ induces a function $\nu_f : C(S) \to \mathbb{R}$ given by $\nu_f(g) = \bar{a}(f, g)$; given that $\iota(a_{\alpha_\beta}) \to \bar{a}$, we have that ν_f is linear and continuous, hence it's a measure, and we see that $\nu_f \ll \sigma_0$. Hence we can define $a(f|\cdot) = \frac{d\nu_f}{d\sigma_0}$. In a σ_0-full measure set $\tilde{S}$ we have that $a(\cdot|s)$ is additive, monotone, and $a(q\mathbb{1}_A|s) = q$ for every $q \in \mathbb{Q}$, hence $|a(f|\cdot) - a(g|\cdot)| \leqslant \|f - g\|_\infty$ in $\tilde{S}$. For each $s \in \tilde{S}$ and $f \in C(A)$ we define $a(f|s) = \lim_n a(f_n|s)$ for any $f_n \to f$ with $f_n \in D$; it's well defined and $a(\cdot|s)$ is linear continuous, hence $a(\cdot|s) \in \Delta(A)$ by Riesz. We extent a to S by setting $a(\cdot|s) \in \Delta(A)$ constant (arbitrary) for $s \in \tilde{S}^c$. Let $\mathcal{F} = \{f \in B(A) : a(f|\cdot)$ is measurable$\}$; we have $D \subset \mathcal{F}$ and $\mathcal{F}$ is a vector space closed by pointwise limits, hence $C(A) \subset \mathcal{F}$ and $\mathcal{F} = B(A)$. Therefore $a \in A_S$, and $\int a(f|\cdot) g\, d\sigma_0 = \int \lim_n a(f_n|\cdot) g\, d\sigma_0 = \lim_n \int a(f_n|\cdot) g\, d\sigma_0 = \lim_n \bar{a}(f_n, g) = \bar{a}(f, g)$ since $\bar{a}(\cdot, g)$ is continuous for every $g \in B(S)$, hence $\iota(a) = \bar{a}$, and $\iota(a_{\alpha_\beta}) \to \iota(a)$, which implies $a_{\alpha_\beta} \to a$ in A_S. Hence A_S is compact, as claimed.

[4]Note that every (Borel) probability measure in a Polish space is regular (Aliprantis and Border, 2006, 12.7), so if we assume Ω, S, S' Polish we can drop the assumption that $\mu_0, \sigma_0, \sigma_0'$ are regular.

We define M_Ω and A_Ω similarly, except we use μ_0 instead of σ_0 when defining the topology. Let $\hat{\sigma} : A_S \to A_\Omega$ be given by $\hat{\sigma}(a) = a \circ \sigma$. It is continuous: if $a_\alpha \to a$ in A_S, and $f \in C(A)$, $g \in B(\Omega)$, we have

$$\begin{aligned}\int \hat{\sigma}(a_\alpha)(f|\cdot)g\, d\mu_0 = \int a_\alpha \circ \sigma(f|\cdot)g\, d\mu_0 = \int \sigma(a_\alpha(f|\cdot)|\cdot)g\, d\mu_0 \\ = \int a_\alpha(f|\cdot)\beta_\sigma(g|\cdot)\, d\sigma_0 \to \int a(f|\cdot)\beta_\sigma(g|\cdot)\, d\sigma_0 \\ = \int \hat{\sigma}(a)(f|\cdot)g\, d\mu_0,\end{aligned}$$

so $\hat{\sigma}(a_\alpha) \to \hat{\sigma}(a)$. Its image is $\hat{\sigma}(A_S) = \{a \circ \sigma : a \in A_S\}$, which is compact since A_S is compact.

Let $N = \{a \in M_\Omega : a = 0\ \mu_0\text{-a.e.}\}$, a closed vector subspace of M_Ω, and $\pi : M_\Omega \to M_\Omega/N$ the projection to the quotient space. For each $f \in M_\Omega^*$ we have $f|_N = 0$ (since that is true for every $L_{f,g}$ with $f \in C(A), g \in B(\Omega)$), hence there is $\hat{f} : M_\Omega/N \to \mathbb{R}$ linear given by $\hat{f}(\pi(a)) = f(a)$. We endow M_Ω/N with the minimal topology that makes them continuous; π is continuous, and M_Ω/N is Hausdorff. The assumption that σ is not more informative than α relative to μ_0 means that $\pi(\alpha) \notin \pi(\hat{\sigma}(A_S))$. Now both are convex and compact, hence by Hahn-Banach there is $\hat{\phi} : M_\Omega \to \mathbb{R}$ linear continuous and $t \in \mathbb{R}$ such that $\hat{\phi}(\pi(\alpha)) > t > \hat{\phi}(\pi(a \circ \sigma))$ for all $a \in A_S$. Let $\phi = \hat{\phi} \circ \pi$. It is in M_Ω^*, therefore $\phi(a) = \sum_{i=1}^n c_i \int a(f_i|\cdot)g_i\, d\mu_0$ for some $c_i \in \mathbb{R}$, $f_i \in C(A)$, $g_i \in B(\Omega)$, so $\phi(a) = \int u(a(\omega), \omega)\, d\mu_0(\omega)$ for $u(a, \omega) = \sum_{i=1}^n c_i f_i(a) g_i(\omega)$, which is bounded, continuous in a and measurable in ω. Hence (A, u) is a decision problem and $U(\alpha) > t > U(\alpha')$ for every $\alpha' \in \mathcal{A}_\sigma$, so $V(\alpha) > V(\sigma)$ and (4) doesn't hold, as desired.

Now, let's prove it for A separable. We proceed as in the proof of Aliprantis and Border (2006, 15.12). There is $\tilde{A}$ compact metrizable such that A is a topological subspace of $\tilde{A}$. By regularity of α_0 there is $K = \bigcup_{n\in\mathbb{N}} K_n$ with $K_n \subset A$ compact (therefore also compact in $\tilde{A}$) such that $\alpha_0(K) = 1$, so by modifying α in a μ_0-null set we can take $\alpha : \Omega \to \Delta(K) \subset \Delta(\tilde{A})$, since $K \in \mathcal{B}_{\tilde{A}}$. Suppose that there is $\tilde{a} : S \to \Delta(\tilde{A})$ such that $\alpha = \tilde{a} \circ \sigma$ μ_0-a.e. Again, we can modify $\tilde{a}$ in a σ_0-null set so that $\tilde{a} : S \to \Delta(K)$ and $\alpha = \tilde{a} \circ \sigma$ still holds μ_0-a.e. This is absurd since we are assuming that (1) doesn't hold. Therefore we can apply what we proved for $\tilde{A}$, and we obtain a decision problem $(\tilde{A}, u)$ such that $V(\alpha) > V(\sigma)$, hence (4) doesn't hold. ■

Comment.—If Ω is compact, the theorem still holds if we change (3) by a weaker version, namely, that $\int f\, dp_\sigma \geqslant \int f\, dp_{\sigma'}$ for every $f \in C(\Delta(\Omega))$ convex (3′). Aliprantis and Border (2006, 19.40) proves (2) ⇔ (3′), but my proof of (3) ⇒ (4) breaks down with (3′). To complete the proof it's enough to show (2) ⇒ (1), since we proved that (1) ⇒ (2) in Proposition 7 and (1) ⇔ (4) in Proposition 8 and Theorem 1. The proof of (2) ⇒ (1) is in fact very easy.

Proposition 10. *Let Ω, S, S' separable metrizable, $\mu_0 \in \Delta(\Omega)$, $\sigma : \Omega \to \Delta(S)$, $\sigma' : \Omega \to \Delta(S')$ signals with $\mu_0, \sigma_0, \sigma'_0$ regular. If there is a mean-preserving spread $\gamma : \Delta(\Omega) \to \Delta(\Delta(\Omega))$ such that $p_\sigma = \gamma \circ p_{\sigma'}$ then $\sigma \geqslant_B^{\mu_0} \sigma'$.*

Proof. We can assume that σ, σ' are direct by Proposition 5. Take the posterior map of γ with prior $p_{\sigma'}$. See the diagram:

$$\begin{tikzcd}
\bullet \arrow[r, "\mu_0"] \arrow[rr, bend left, "p_\sigma"] \arrow[drr, bend right, "p_{\sigma'}"'] & \Omega \arrow[r, "\sigma"] \arrow[dr, "\sigma'"] & \Delta(\Omega) \arrow[d, dashed, bend left=60, "\tau"] \\
 & & \Delta(\Omega) \arrow[u, "\gamma"]
\end{tikzcd}$$

Take $\beta : \mu \mapsto \mu$, the posterior of both σ and σ', $f \in B(\Delta(\Omega))$ and $g \in B(\Omega)$. By the definition of τ we have $\int \tau(f|\cdot)\beta(g|\cdot)\, dp_\sigma = \int f\gamma(\beta(g|\cdot)|\cdot)\, dp_{\sigma'}$. The LHS is $\int \sigma(\tau(f|\cdot)|\cdot)g\, d\mu_0 = \int \tau\circ\sigma(f|\cdot)g\, \mu_0$, and the RHS is $\int f\gamma(\beta(g|\cdot)|\cdot)\, dp_{\sigma'} = \int f\beta(g|\cdot)\, dp_{\sigma'} = \int \sigma'(f|\cdot)g\, d\mu_0$, since $\gamma(\beta(g|\cdot)|\mu) = \int \nu(g)\, d\gamma(\nu|\mu) = \mu(g) = \beta(g|\mu)$ by the definition of γ. Hence $\int \sigma'(f|\cdot)g\, d\mu_0 = \int \tau\circ\sigma(f|\cdot)g\, \mu_0$, i.e., $\sigma \geqslant_B^{\mu_0} \sigma'$. ∎

References

Aliprantis, Charalambos D. and Kim C. Border. 2006. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. 3rd [rev. and enl.] ed ed. Berlin ; New York: Springer.

Blackwell, David. 1951. "Comparison of Experiments." *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* 2:93–103.

Blackwell, David. 1953. "Equivalent Comparisons of Experiments." *The Annals of Mathematical Statistics* 24(2):265–272.

Crauel, H. 2002. *Random Probability Measures on Polish Spaces*. Number v. 11 *in* "Stochastics Monographs" London ; New York: Taylor & Francis.

de Oliveira, Henrique. 2018. "Blackwell's Informativeness Theorem Using Diagrams." *Games and Economic Behavior* 109:126–131.

Dellacherie, Claude and Paul-Andre Meyer. 1978. *Probabilities and Potential. A*. North-Holland.

Khan, M Ali, Haomiao Yu and Zhixiang Zhang. 2020. "A Missing Equivalence Theorem on Information Structures on a General State Space.".

Khan, M. Ali, Yeneng Sun, Rabee Tourky and Zhixiang Zhang. 2008. "Similarity of Differential Information with Subjective Prior Beliefs." *Journal of Mathematical Economics* 44(9):1024–1039.

Phelps, Robert R. 2001. *Lectures on Choquet's Theorem*. Number 1757 *in* "Lecture Notes in Mathematics" 2nd ed ed. Berlin ; New York: Springer.