Online Appendix

for the paper "Elicitation and Evaluation of Statistical Forecasts" by N.S. Lambert

A Proofs of Section 3

In the proofs that follow, $L^1(\mu)$ is the collection of μ-integrable real functions on $(\Omega, \mathcal{G})$. Unless mentioned otherwise, the topology used on $L^1(\mu)$, $\mathcal{C}(\Omega)$ and $\Delta\Omega$ (viewed as a subset of $\mathcal{C}(\Omega)$), will be that induced by the L^1 norm, denoted by $\|\cdot\|$ or sometimes $\|\cdot\|_1$ to avoid ambiguity. $L^\infty(\mu)$ is the collection of all μ-measurable, essentially bounded real functions on $(\Omega, \mathcal{G})$. The L^∞ norm is defined by $\|f\|_\infty = \sup\{m \mid \mu(f > m) = 0\}$. For members of $L^1(\mu)$ and $L^\infty(\mu)$ I make the usual μ-almost everywhere identification. Given a linear topological space $\mathcal{E}$, $\mathcal{E}^*$ represents the topological dual endowed with the operator norm. For an arbitrary function f, $\{f = \alpha\}$ denotes the set $\{x \mid f(x) = \alpha\}$, and $f_{|\mathcal{D}}$ the restriction of f to domain $\mathcal{D}$. Given a subset $\mathcal{S}$ of a linear (topological) space, $\mathcal{S}^\circ$ denotes the interior of $\mathcal{S}$, $\langle \mathcal{S} \rangle$ its linear span, and $\langle \mathcal{S} \rangle_a$ its affine span. For a scoring rule S, I use the short notation

$$\bar{S}_p(t) = \mathop{\mathsf{E}}_{\omega \sim p}[S(t, \omega)] \ .$$

$\mathcal{L}$ is the Lebesgue measure.

The proofs use the following elementary lemmas.

Lemma 4. *If Φ is a linear functional on $\mathcal{C}(\Omega)$ such that $\ker \Phi \cap \Delta\Omega \neq \emptyset$, then $\ker \Phi$ is the linear span of its intersection with $\Delta\Omega$.*

Proof. Let $f_0 \in \Delta\Omega$ with $\Phi(f_0) = 0$. Take any $f \in \ker \Phi$. As Ω is compact, $\inf f_0 > 0$ so that, if α is chosen large enough, $f + \alpha f_0 > 0$. So, defining $\beta = \int_\Omega (f + \alpha f_0) \mathrm{d}\mu$ and $f_1 = (f + \alpha f_0)/\beta$, we have that $\Phi(f_1) = 0$ and $f_1 \in \Delta\Omega$. Hence $f = \beta f_1 - \alpha f_0 \in \langle \ker \Phi \cap \Delta\Omega \rangle$. $\square$

Lemma 5. *If $h : [a, b] \mapsto \mathbb{R}_+$ is a Lebesgue measurable function with $\int h > 0$, then there exists $\epsilon > 0$ such that $\{h \geq \epsilon\}$ has strictly positive measure.*

Proof. As $\{h > 0\}$ is the limit of the monotone increasing sequence of sets $\{h \geq 1/k\}$ and $\{h > 0\}$ has strictly positive measure, the sets $\{h \geq 1/k\}$ must have strictly positive measure as k grows large enough. $\square$

Lemma 6. *If $h : [a,b] \mapsto \mathbb{R}_+$ is a Lebesgue measurable function that is strictly positive almost everywhere, and $A \subset [a,b]$ is a measurable set of strictly positive measure, then $\int_A h > 0$.*

Proof. As $A \cap \{h > 0\}$ is the limit of the monotone increasing sequence of sets $(A \cap \{h \geq 1/k\})$, for k large enough, the set $A \cap \{h \geq 1/k\}$ must have strictly positive measure, and $\int_A h \geq \lambda(A \cap \{h \geq 1/k\})/k > 0$. □

Lemma 7. *If Φ is a linear functional on $\mathcal{C}(\Omega)$ that is not continuous, then there exits a sequence $(f_n)_n$ of nonnegative functions of $\mathcal{C}(\Omega)$ converging to zero such that either ($\Phi(f_n) > 1$ for all n) or ($\Phi(f_n) < -1$ for all n).*

Proof. If, for all $\delta > 0$, there exists $\epsilon > 0$ such that for all nonnegative $f \in \mathcal{C}(\Omega)$, $\|f\| \leq \epsilon$ implies $|\Phi(f)| \leq \delta$, then Φ is continuous at 0.

Indeed, take any $\delta > 0$. The condition on Φ implies the existence of some $\epsilon > 0$, such that for all nonnegative $f \in \mathcal{C}(\Omega)$ with $\|f\| \leq \epsilon$, $|\Phi(f)| \leq \delta/2$. Take any $f \in \mathcal{C}(\Omega)$, with $\|f\| \leq \epsilon$. Then $f = f_+ - f_-$, where $f_+ = \max(f, 0)$ and $f_- = \max(-f, 0)$. Noting that $f_+, f_- \in \mathcal{C}(\Omega)$ and that $\|f_+\|, \|f_-\| \leq \|f\| \leq \epsilon$, we get that $|\Phi(f)| \leq |\Phi(f_+)| + |\Phi(f_-)| \leq \delta$. Hence f is continuous at 0, and, by linearity, continuous on $\mathcal{C}(\Omega)$.

Therefore, if Φ is not continuous, there must exist some $\delta > 0$ and a sequence $(f_n)_n$ of nonnegative elements of $\mathcal{C}(\Omega)$ converging to zero such that, for all n, $\Phi(f_n) < -\delta$ or $\Phi(f_n) > -\delta$. We get the lemma by rescaling the whole sequence and extracting a subsequence on which the same inequality holds. □

Proof of Proposition 1

Let Γ be the associated statistic function. If S is order sensitive, then, for all p, $\bar{S}_p$ is (weakly) increasing on $\{t \leq \Gamma(p)\}$ and is (weakly) decreasing on $\{t \geq \Gamma(p)\}$. Therefore if S was not order sensitive, there would exist $p_0 \in \Delta\Omega$ and $x, y \in \Theta^\circ$, with, for example, $y > x > \Gamma(p_0)$, such that $\bar{S}_{p_0}(y) > \bar{S}_{p_0}(x)$. Let $p_1 \in \{\Gamma > y\}$ and define $p_\lambda = \lambda p_1 + (1-\lambda)p_0$. By continuity, there exists λ_x such that $\Gamma(p_{\lambda_x}) = x$ and $\lambda_y > \lambda_x$ such that $\Gamma(p_{\lambda_y}) = y$. Then

$$0 \geq \bar{S}_{p_{\lambda_x}}(y) - \bar{S}_{p_{\lambda_x}}(x) = \lambda_x(\bar{S}_{p_1}(y) - \bar{S}_{p_1}(x)) + (1-\lambda_x)(\bar{S}_{p_0}(y) - \bar{S}_{p_0}(x)) \ .$$

As $\bar{S}_{p_0}(y) - \bar{S}_{p_0}(x) > 0$, we must have $\bar{S}_{p_1}(y) - \bar{S}_{p_1}(x) < 0$. Hence,

$$\begin{aligned}\bar{S}_{p_{\lambda_y}}(y) - \bar{S}_{p_{\lambda_y}}(x) &= \lambda_y(\bar{S}_{p_1}(y) - \bar{S}_{p_1}(x)) + (1-\lambda_y)(\bar{S}_{p_0}(y) - \bar{S}_{p_0}(x)) \ , \\ &< \lambda_x(\bar{S}_{p_1}(y) - \bar{S}_{p_1}(x)) + (1-\lambda_x)(\bar{S}_{p_0}(y) - \bar{S}_{p_0}(x)) \ .\end{aligned}$$

But, as S is proper, we should have $\bar{S}_{p_{\lambda_y}}(y) - \bar{S}_{p_{\lambda_y}}(x) \geq 0$. By the same logic, if S is strictly proper, it is also strictly order sensitive.

Proof of Theorem 1

Let (Θ, F) the a regular real-valued continuous statistic and Γ be the associated statistic function.

<u>**Part (1) $\Rightarrow$ (2) :**</u>

Let S be a strictly proper scoring rule. Take $p, q \in \Delta\Omega$, and $0 < \alpha < 1$. Suppose $p, q \in F(\theta)$. Then, for all $\hat{\theta} \neq \theta$,

$$\mathop{\mathsf{E}}_{\omega\sim p}[S(\hat{\theta}, \omega)] \leq \mathop{\mathsf{E}}_{\omega\sim p}[S(\theta, \omega)] \ ,$$

and

$$\mathop{\mathsf{E}}_{\omega\sim q}[S(\hat{\theta}, \omega)] \leq \mathop{\mathsf{E}}_{\omega\sim q}[S(\theta, \omega)] \ ,$$

and so, by linearity of the expectation operator,

$$\begin{aligned}\mathop{\mathsf{E}}_{\omega\sim\alpha p+(1-\alpha)q}[S(\hat{\theta}, \omega)] &= \alpha \mathop{\mathsf{E}}_{\omega\sim p}[S(\hat{\theta}, \omega)] + (1-\alpha) \mathop{\mathsf{E}}_{\omega\sim q}[S(\hat{\theta}, \omega)] \\ &\leq \alpha \mathop{\mathsf{E}}_{\omega\sim p}[S(\theta, \omega)] + (1-\alpha) \mathop{\mathsf{E}}_{\omega\sim q}[S(\theta, \omega)] \\ &= \mathop{\mathsf{E}}_{\omega\sim\alpha p+(1-\alpha)q}[S(\theta, \omega)] \ ,\end{aligned}$$

which, by strict properness, implies $\alpha p + (1-\alpha)q \in F(\theta)$. Hence the convexity of the sets $F(\theta)$.

<u>**Part (2) $\Rightarrow$ (3) :**</u>

First remark that, as Γ is continuous, the set of values taken by the statistic, Θ, is an interval of the real line. This can be seen by applying the intermediate value theorem to the continuous function $\alpha \mapsto \Gamma(\alpha p + (1-\alpha)q)$, defined on $[0, 1]$ for any $p, q \in \Delta\Omega$.

<u>Step 1.</u> Let us start by showing that if, for all θ, $\{\Gamma = \theta\}$ is convex, then it is also

the case that $\{\Gamma \geq \theta\}$, $\{\Gamma > \theta\}$, $\{\Gamma \leq \theta\}$ and $\{\Gamma < \theta\}$ are convex. I prove the first case, the other three work in a similar fashion.

Let $\theta \in \Theta^\circ$, and $p, q \in \Delta\Omega$, with $\Gamma(p) \geq \Gamma(q) \geq \theta$. Consider the function $f(\alpha) = \Gamma(\alpha p + (1-\alpha)q)$ defined on $[0,1]$. Note that f is continuous. To prove that $\{\Gamma \geq \theta\}$ is convex, it suffices to show that the image of f is the interval $[\Gamma(q), \Gamma(p)]$. We already know that $[\Gamma(q), \Gamma(p)] \subseteq f([0,1])$ by continuity of f, observing that $f(0) = \Gamma(q)$ and $f(1) = \Gamma(p)$. So let

$$
\begin{aligned}
a &= \sup\{\alpha \in [0,1] \mid f(\alpha) = \Gamma(q)\} \ , \\
b &= \inf\{\alpha \in [0,1] \mid f(\alpha) = \Gamma(p)\} \ .
\end{aligned}
$$

By continuity of f, the above two sets are closed and nonempty, so $f(a) = f(0) = \Gamma(q)$ and $f(b) = f(1) = \Gamma(p)$. Also, by convexity of the level sets of Γ, $f([0,a]) = \{\Gamma(q)\}$ and $f([b,1]) = \{\Gamma(p)\}$. Besides, if, for some $\alpha^* > a$, $f(\alpha^*) < f(0)$ then by continuity $f(\alpha) = f(0)$ for some $\alpha > \alpha^*$, violating a's definition. Similarly, there does not exist α^* with $f(\alpha^*) > f(1)$, and $f([0,1]) = [\Gamma(q), \Gamma(p)]$. So $\{\Gamma \geq \theta\}$ is convex.

<u>Step 2.</u> Let $\theta \in \Theta^\circ$. Let's start by showing the existence of a nonzero linear functional Φ on $\mathcal{C}(\Omega)$, continuous in the L^∞-norm topology, such that

$$
\begin{aligned}
\{\Gamma < \theta\} &\subset \{\Phi \leq 0\} \ , \\
\{\Gamma \geq \theta\} &\subset \{\Phi \geq 0\} \ .
\end{aligned}
$$

First note that, as Γ is continuous in the L^1-norm topology, it is also continuous in the richer, stronger L^∞-norm topology (recall that Ω was chosen to be compact). Using the L^∞-norm topology on $\mathcal{C}(\Omega)$ will be useful, because it makes the relative interior of $\Delta\Omega$ nonempty (by compactness of Ω) so as to be able to apply the Hahn-Banach separating hyperplane theorem.

Indeed, each $p \in \{\Gamma < \theta\}$ belongs to the (L^∞) relative interior of $\{\Gamma < \theta\}$, while each $p \in \{\Gamma > \theta\}$ belongs to the (L^∞) relative interior of $\{\Gamma \geq \theta\}$. By the previous step both $\{\Gamma < \theta\}$ and $\{\Gamma \geq \theta\}$ are convex, and since they are disjoint with nonempty (L^∞) relative interior, we can apply the Hahn-Banach separating hyperplane theorem and find a nonconstant affine function Φ on the affine span of $\Delta\Omega$, continuous in the L^∞-norm topology and such that $\Phi(\{\Gamma < \theta\}) \leq 0$ and $\Phi(\{\Gamma \geq \theta\}) \geq 0$.

Let us write Φ as

$$\Phi(f) = \Phi_0 + V(f - p_0)$$

for some $p_0 \in \Delta\Omega$, where $\Phi_0 = \Phi(p_0)$ and V is a linear functional on $\{f \in \mathcal{C}(\Omega) \mid \int_\Omega f \mathrm{d}\mu = 0\}$. V is continuous with respect to the L^∞ norm. We can extend Φ to a linear functional on the whole space $\mathcal{C}(\Omega)$, also continuous with respect to the L^∞ norm, by defining

$$\Phi(f) = \Phi_0 \int_\Omega f \mathrm{d}\mu + V\left(f - \left(\int_\Omega f \mathrm{d}\mu\right) p_0\right) .$$

<u>Step 3.</u> Now we show that Φ is also continuous with respect to the L^1 norm, *i.e.*, $\Phi \in (\mathcal{C}(\Omega))^*$. Suppose by contradiction it is not the case. Then, by Lemma 7, there exists a sequence $(\tilde{f}_n)_n$ of nonnegative functions of $\mathcal{C}(\Omega)$ that converge to zero, and such that, for example, $\Phi(\tilde{f}_n) < -1$ for all n (the case $\Phi(\tilde{f}_n) > 1$ is treated in a similar fashion).

Let $p_0 \in \{\Gamma > \theta\}$. Then, $\Phi(p_0) \geq 0$. If α is chosen large enough, for all n, we get

$$\Phi\left(\frac{p_0 + \alpha\tilde{f}_n}{\|p_0 + \alpha\tilde{f}_n\|}\right) = \frac{1}{\|p_0 + \alpha\tilde{f}_n\|}(\Phi(p_0) + \alpha\Phi(\tilde{f}_n)) < 0 , \tag{6}$$

noting that $\|p_0 + \alpha\tilde{f}_n\| > 0$ as p_0 is strictly positive and $\tilde{f}_n$ is nonnegative.

By continuity of Γ, there exists $\epsilon > 0$ such that, if $p \in \Delta\Omega$ with $\|p - p_0\| \leq \epsilon$, then $|\Gamma(p) - \Gamma(p_0)| \leq \frac{\Gamma(p_0)-\theta}{2}$. In particular, we get

$$\{p \in \Delta\Omega \mid \|p - p_0\| \leq \epsilon\} \subset \{\Gamma > \theta\} \subset \{\Phi \geq 0\} . \tag{7}$$

Note that $p_0 + \alpha\tilde{f}_n > 0$ and so $\frac{p_0+\alpha\tilde{f}_n}{\|p_0+\alpha\tilde{f}_n\|} \in \Delta\Omega$. Besides,

$$\begin{aligned}\left\|\frac{p_0 + \alpha\tilde{f}_n}{\|p_0 + \alpha\tilde{f}_n\|} - p_0\right\| &= \frac{1}{1 + \alpha\|\tilde{f}_n\|}\|\alpha\tilde{f}_n - \alpha\|\tilde{f}_n\|p_0\| , \\ &\leq \|\tilde{f}_n\|\alpha(1 + \|p_0\|) ,\end{aligned}$$

and so, as $\|\tilde{f}_n\| \to 0$, there exists N such that $\|\tilde{f}_N\|\alpha(1 + \|p_0\|) \leq \epsilon$. By (7), $\Phi((p_0 + \alpha\tilde{f}_n)/\|p_0 + \alpha\tilde{f}_n\|) \geq 0$, contradicting (6). Hence $\Phi \in (\mathcal{C}(\Omega))^*$.

<u>Step 4.</u> Using the same θ as in the preceding step, as $\{\Gamma < \theta\}$ and $\{\Gamma > \theta\}$ are open

sets of $\Delta\Omega$, we have that $\{\Gamma < \theta\} \subset \{\Phi < 0\}$ and $\{\Gamma > \theta\} \subset \{\Phi > 0\}$. In summary, we have shown the existence of a continuous linear functional Φ on $\mathcal{C}(\Omega)$ satisfying

$$\{\Gamma < \theta\} \subset \{\Phi < 0\} \ ,$$
$$\{\Gamma \geq \theta\} \subset \{\Phi \geq 0\} \ ,$$
$$\{\Gamma > \theta\} \subset \{\Phi > 0\} \ .$$

By a symmetric argument, there exists a continuous linear functional Ψ that satisfies

$$\{\Gamma < \theta\} \subset \{\Psi < 0\} \ ,$$
$$\{\Gamma \leq \theta\} \subset \{\Psi \leq 0\} \ ,$$
$$\{\Gamma > \theta\} \subset \{\Psi > 0\} \ .$$

We show that Φ and Ψ are positively collinear. If they are not collinear, then $\ker \Phi \cap \Delta\Omega \neq \ker \Psi \cap \Delta\Omega$ by Lemma 4. As $\ker \Phi \cap \Delta\Omega \subseteq \{\Gamma = \theta\}$ and $\ker \Psi \cap \Delta\Omega \subseteq \{\Gamma = \theta\}$, there exist $p, q \in \{\Gamma = \theta\}$ such that $\Phi(p) = 0$ with $\Psi(p) < 0$, and $\Phi(q) > 0$ with $\Psi(q) = 0$. So,

$$\Phi\left(\frac{p+q}{2}\right) > 0 \quad \text{and} \quad \Psi\left(\frac{p+q}{2}\right) < 0 \ .$$

By continuity of Φ and Ψ, there exists an open ball $\mathcal{B}$ centered on $(p+q)/2$, in the L^1 norm, such that $\Phi(\mathcal{B})$ contains only strictly positive values and $\Psi(\mathcal{B})$ contains only strictly negative values. Since $\mathcal{B} \cap \Delta\Omega \neq \emptyset$, these assertions imply that Γ is both greater than or equal to θ and less than or equal to θ on $\mathcal{B} \cap \Delta\Omega$, and so equals θ on this open set of $\Delta\Omega$. This contradicts the regularity assumption on Γ. So Ψ and Φ are collinear, and, by their sign properties above, are positively collinear, implying $\{\Gamma = \theta\} = \ker \Phi \cap \Delta\Omega$.

In conclusion, for all $\theta \in \Theta^\circ$, there exists $L_\theta \in (\mathcal{C}(\Omega))^*$ such that $\{\Gamma = \theta\} = \ker L_\theta \cap \Delta\Omega$.

Part (3) $\Rightarrow$ (1) :

Suppose that, for all $\theta \in \Theta^\circ$, there exists $\Phi_\theta \in (\mathcal{C}(\Omega))^*$ such that, for all $p \in \Delta\Omega$, $\Phi_\theta(p) = 0$ if and only if $\Gamma(p) = \theta$. Any nonzero factor of Φ_θ possesses the same property, so we can choose without loss $\|\Phi_\theta\| = 1$, and orient Φ_θ such that $\Phi_\theta(p) > 0$ for some given $p \in \{\Gamma > \theta\}$. By continuity of Γ and convexity of $\Delta\Omega$, Φ_θ has the following properties:

$$\{\Gamma < \theta\} = \{\Phi < 0\} \cap \Delta\Omega \ ,$$
$$\{\Gamma = \theta\} = \{\Phi = 0\} \cap \Delta\Omega \ ,$$
$$\{\Gamma > \theta\} = \{\Phi > 0\} \cap \Delta\Omega \ .$$

By the Banach extension theorem, Φ_θ can be extended to an element of $(L^1(\mu))^*$ with $\|\Phi_\theta\| = 1$. Applying a version of the Riesz Representation theorem (Theorem 1.11 of Megginson, 1998)), there exists a function $g_\theta \in L^\infty(\Omega)$ such that, for all $f \in L^1(\mu)$,

$$\Phi_\theta(f) = \int_\Omega f g_\theta \, \mathrm{d}\mu \ .$$

Note that Ω being compact, $L^\infty(\mu) \subset L^1(\mu)$.

<u>Step 1.</u> This steps shows that the function $\theta \mapsto g_\theta$ is uniformly continuous on every segment of Θ°, with respect to the L^1 norm.

Let us begin by showing that, for all $\theta_0 \in \Theta^\circ$, $\lim_{\theta \to \theta_0} \Phi_\theta(f) = 0$ whenever $f \in \ker \Phi_{\theta_0} \cap \Delta\Omega$. To see this, let $f \in \{\Gamma = \theta_0\}$, and, for any $\epsilon > 0$, consider the open ball $\mathcal{B}_\epsilon$ in $L^1(\mu)$ of radius ϵ that is centered on f. Note that Φ_{θ_0} takes both strictly positive and strictly negative values on $\mathcal{B}_\epsilon$, meaning that Γ takes values that are both above and below θ_0. By continuity of Γ, there exists some $\delta > 0$ such that $(\theta_0 - \delta, \theta_0 + \delta) \subset \Gamma(\mathcal{B} \cap \Delta\Omega)$. In particular, for all $\theta \in (\theta_0 - \delta, \theta_0 + \delta)$, there is $g \in \mathcal{B}_\epsilon \cap \Delta\Omega$ with $\Gamma(g) = \theta$, hence $|\Phi_\theta(f)| = |\Phi_\theta(f - g) + \Phi_\theta(g)| \leq \|\Phi_\theta\| \|f - g\| \leq \epsilon$. Therefore, we have that $\lim_{\theta \to \theta_0} \Phi_\theta(f) = 0$.

Observing that $\ker \Phi_{\theta_0 | \mathcal{C}(\Omega)} = \langle \ker \Phi_{\theta_0} \cap \Delta\Omega \rangle$ by Lemma 4, the above limit remains valid whenever $f \in \ker \Phi_{\theta_0 | \mathcal{C}(\Omega)}$. Now we can extend the limit to all members of $\mathcal{C}(\Omega)$. To do so, take some function $v \in \mathcal{C}(\Omega)$ such that $\Phi_{\theta_0}(v) = 1$. Any $f \in \mathcal{C}(\Omega)$ can be written $f = \Phi_{\theta_0}(f) v + w$ where $w \in \ker \Phi_{\theta_0 | \mathcal{C}(\Omega)}$. Then $|\Phi_\theta(f) - \Phi_{\theta_0}(f)\Phi_\theta(v)| \to 0$ as $\theta \to \theta_0$. Besides, for all θ, $\|\Phi_\theta\| = 1$, hence $|\Phi_\theta(v)| \to 1$ as $\theta \to \theta_0$ and the orientation that was decided of Φ_θ yields $\Phi_\theta(v) \to 1$. So, $\Phi_\theta(f) \to \Phi_{\theta_0}(f)$ whenever $f \in \mathcal{C}(\Omega)$.

Finally, we extend the limit on the whole space $L^1(\mu)$. Let $f \in L^1(\mu)$ and fix $\epsilon > 0$. Choose $g \in \mathcal{C}(\Omega)$ such that $\|f - g\| < \epsilon/4$. Then there exists $\delta > 0$ such that if $|\theta - \theta_0| < \delta$, $|\Phi_\theta(g) - \Phi_{\theta_0}(g)| < \epsilon/2$. Writing $\Phi_\theta(f) = \Phi_\theta(g) + \Phi_\theta(f - g)$, and similarly for $\Phi_{\theta_0}(f)$, we get $|\Phi_\theta(f) - \Phi_{\theta_0}(f)| \leq |\Phi_\theta(g) - \Phi_{\theta_0}(g)| + \|f - g\|(\|\Phi_\theta\| + \|\Phi_{\theta_0}\|) = \epsilon$. Consequently, Φ_θ converges pointwise to Φ_{θ_0} on $L^1(\mu)$.

That $\lim_{\theta\to\theta_0}\Phi_\theta(f) = \Phi_{\theta_0}(f)$ for all $f \in L^1(\mu)$ implies that the application $\theta \mapsto g_\theta$ is continuous with respect to the L^1 norm. By the Heine-Cantor theorem the application is uniformly continuous on every segment of Θ°.

<u>STEP 2.</u> Now let $[a,b] \subset \Theta^\circ$. We shall construct a bounded function $G : [a,b]\times\Omega \mapsto \mathbb{R}$ that is $(\mathcal{L}\otimes\mu)$-measurable and that offers a nearly perfect approximation of the g_θ's, in the sense that for almost every θ, $\|G(\theta,\cdot) - g_\theta\| = 0$.

We construct the function G as a limit of functions G_n. Let $d_{k,n} = a + k\frac{b-a}{2^n}$. Let τ_n be the function defined on $[a,b]$ by $\tau_n(\theta) = d_{k,n}$ when $\theta \in [d_{k,n}; d_{k+1,n})$. Define $G_n(\theta,\omega) = g_{\tau_n(\theta)}(\omega)$. Recall that $\theta \mapsto g_\theta$ is uniformly continuous on $[a,b]$ with respect to the L^1 norm. Let $\epsilon > 0$. Choose δ such that, for $\theta_1,\theta_2 \in [a,b]$,

$$|\theta_1 - \theta_2| < \delta \Rightarrow \|g_{\theta_1} - g_{\theta_2}\| < \frac{\epsilon}{2} \ .$$

Choose N such that $1/2^N < \delta$. Take any $n,m > N$.

$$\begin{aligned}|G_n(\theta,\omega) - G_m(\theta,\omega)| &= |g_{\tau_n(\theta)}(\omega) - g_{\tau_m(\theta)}(\omega)| \ ,\\ &\le |g_{\tau_n(\theta)}(\omega) - g_\theta(\omega)| + |g_{\tau_m(\theta)}(\omega) - g_\theta(\omega)| \ ,\end{aligned}$$

but, as $n,m > N$, $|\tau_n(\theta) - \theta|, |\tau_m(\theta) - \theta| \le 2^{-N} < \delta$. And so, for all $\theta \in [a,b]$,

$$\begin{aligned}\|G_n(\theta,\cdot) - G_m(\theta,\cdot)\| &\le \|g_{\tau_n(\theta)} - g_\theta\| + \|g_{\tau_m(\theta)} - g_\theta\| \ ,\\ &\le \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \ ,\end{aligned}$$

which shows that $(G_n)_n$ is a Cauchy sequence in the Banach space $L^1(\mathcal{L}\otimes\mu)$, and thus converges to some function $G \in L^1(\mathcal{L}\otimes\mu)$. Besides, as each G_n is such that $|G_n| \le 1$, we have that $|G| \le 1$.

Moreover, for almost every θ, $\|G(\theta,\cdot) - g_\theta\| = 0$. This can be seen by writing

$$\|G(\theta,\cdot) - g_\theta\| \le \|G(\theta,\cdot) - G_n(\theta,\cdot)\| + \|G_n(\theta,\cdot) - g_\theta\| \ ,$$

for all n. First we observe that

$$\|G_n(\theta,\cdot) - g_\theta\| = \|g_{\tau_n(\theta)} - g_\theta\| \underset{n\infty}{\longrightarrow} 0 \ .$$

Second, as a consequence of G_n converging to G and Fubini's theorem, the function

$\theta \mapsto \|G(\theta,\cdot) - G_n(\theta,\cdot)\|$ converges to zero in the L^1 norm. Since convergence in L^1 implies convergence almost everywhere, for almost every θ,

$$\|G(\theta,\cdot) - G_n(\theta,\cdot)\| \underset{n\infty}{\longrightarrow} 0 \ .$$

<u>STEP 3.</u> We then extend our function G to the whole interval Θ and construct a bounded function $H : \Theta \times \Omega \mapsto \mathbb{R}$ that is $(\mathcal{L} \otimes \mu)$-measurable and such that for almost every θ, $\|H(\theta,\cdot) - g_\theta\| = 0$.

Θ is an interval and so is an (at most) countable union of closed segments of the form $[a,b]$, whose interiors are pairwise disjoint. Let $\mathcal{F}$ be such a family. For any $[a,b] \in \mathcal{F}$, there exists, by the preceding step, a function $H_{[a,b]}$ integrable on $[a,b] \times \Omega$ and such that, for almost every θ, $\|H_{[a,b]}(\theta,\cdot) - g_\theta\|$. Besides, $|H_{[a,b]}| \leq 1$.

Define H as follows:

- if θ is an extremity of an interval of $\mathcal{F}$, let $H(\theta,\cdot) = 0$,
- otherwise, let $H(\theta,\omega) = H_{[a,b]}(\theta,\omega)$.

H inherits the nice properties of the $H_{[a,b]}$'s: $|H| \leq 1$, H is integrable on every segment of Θ, and, for almost every θ, $\|H(\theta,\cdot) - g_\theta\| = 0$. This last point implies that, whenever $p \in \Delta\Omega$, $\int_\Omega H(\theta,\cdot)p\mathrm{d}\mu$ equals $\Phi_\theta(p)$ for almost every θ.

<u>STEP 4.</u> At last we can construct a strictly proper scoring rule. Choose any $\theta_0 \in \Theta$ and let

$$S(\theta,\omega) = \int_{\theta_0}^{\theta} H(t,\omega)\mathrm{d}t \ .$$

As $|H|$ is bounded, for all $p \in \Delta\Omega$, a direct application of Fubini's theorem yields

$$\begin{aligned}
\underset{\omega\sim p}{\mathsf{E}}[S(\theta,\omega)] &= \int_\Omega \left(\int_{\theta_0}^{\theta} H(t,\omega)\mathrm{d}t\right) p(\omega)\mathrm{d}\mu(\omega) \ , \\
&= \int_{\theta_0}^{\theta} \left(\int_\Omega H(t,\omega)p(\omega)\mathrm{d}\mu(\omega)\right) \mathrm{d}t \ .
\end{aligned}$$

Suppose for example that $\Gamma(p) > \theta$, then

$$\begin{aligned}
\underset{\omega\sim p}{\mathsf{E}}[S(\Gamma(p),\omega)] - \underset{\omega\sim p}{\mathsf{E}}[S(\theta,\omega)] &= \int_{\theta}^{\Gamma(p)} \left(\int_\Omega H(t,\omega)p(\omega)\mathrm{d}\mu(\omega)\right) \mathrm{d}t \ , \\
&> 0
\end{aligned}$$

since, for almost all $t < \Gamma(p)$, $\int_\Omega H(t,\cdot)p\mathrm{d}\mu = \Phi_t(p) > 0$. And similarly for $t > \Gamma(p)$. Hence S is strictly proper.

Proof of Theorem 2

If part:

In the proof of Theorem 1, we constructed a function H that is $(\mathcal{L}\otimes\mu)$-measurable, satisfies $|H| \leq 1$, and such that, for almost every θ, and all $p \in \Delta\Omega$,

$$\int_\Omega H(\theta,\omega)p(\omega)\mathrm{d}\mu(\omega)$$

is strictly positive when $\Gamma(p) > \theta$, strictly negative when $\Gamma(p) < \theta$, and zero when $\Gamma(p) = \theta$. Choose $S_0 = H$. Assume that for almost every ω and all θ, scoring rule S takes the form

$$S(\theta,\omega) = \kappa(\omega) + \int_{\theta_0}^{\theta} \xi(t)S_0(t,\omega)\mathrm{d}t \ ,$$

for some $\theta_0 \in \Theta$, $\kappa : \Omega \mapsto \mathbb{R}$, and $\xi : \mathcal{I} \mapsto \mathbb{R}_+$ a Lebesgue measurable bounded function.

For all $p \in \Delta\Omega$,

$$\bar{S}_p(\theta) = \int_\Omega \kappa\mathrm{d}\mu + \int_\Omega\int_{\theta_0}^{\theta} \xi(t)S_0(t,\omega)\mathrm{d}t\mathrm{d}\mu(\omega) \ .$$

Take for example $\theta < \Gamma(p)$. By Fubini's theorem,

$$\bar{S}_p(\Gamma(p)) - \bar{S}_p(\theta) = \int_\theta^{\Gamma(p)} \xi(t)\left(\int_\Omega S_0(t,\omega)p(\omega)\mathrm{d}\mu(\omega)\right)\mathrm{d}t \ .$$

As, for almost all $t < \Gamma(p)$, $\int_\Omega S_0(t,\omega)p(\omega)\mathrm{d}\mu(\omega) > 0$, we get $\bar{S}_p(\Gamma(p)) - \bar{S}_p(\theta) \geq 0$, implying that S is proper. If, in addition, $\int_\theta^{\Gamma(p)} \xi > 0$, then by Lemma 5, there is $\epsilon > 0$ such that $A = \{\xi \geq \epsilon\}$ is of strictly positive Lebesgue measure. Hence,

$$\bar{S}_p(\Gamma(p)) - \bar{S}_p(\theta) \geq \epsilon \int_A \left(\int_\Omega S_0(t,\omega)p(\omega)\mathrm{d}\mu(\omega)\right)\mathrm{d}t$$

which is strictly positive by Lemma 6, making S strictly proper.

Only if part:

Let S be a regular scoring rule for Γ, and $\theta_0 \in \Theta$. If S is (strictly) proper, $(\theta, \omega) \mapsto S(\theta, \omega) - S(\theta_0, \omega)$ is also (strictly) proper. Thus we can assume with loss of generality that $S(\theta_0, \cdot) = 0$.

As $S(\cdot, \omega)$ is Lipschitz continuous, it is also absolutely continuous and there is a function $G : \Theta \times \Omega \mapsto \mathbb{R}$ such that, for all θ, ω,

$$S(\theta, \omega) = \int_{\theta_0}^{\theta} G(t, \omega) \mathrm{d}t \ .$$

Moreover, for all ω, $\theta \mapsto S(\theta, \omega)$ is differentiable except possibly on a measure zero set (that may depend on ω), and

$$\frac{S(\theta, \omega)}{\partial \theta} = G(\theta, \omega) \ .$$

G can be chosen such that, if $S(\cdot, \omega)$ is not differentiable at θ, $G(\theta, \omega) = 0$. Besides, G is $(\mathcal{L} \otimes \mu)$-measurable as a limit of the measurable functions

$$(\theta, \omega) \mapsto \begin{cases} n(S(\theta + 1/n, \omega) - S(\theta, \omega)) & \text{if } S(\cdot, \omega) \text{ is differentiable at } \theta \ , \\ 0 & \text{otherwise} \ . \end{cases}$$

Finally, as S is Lipschitz continuous, G is bounded.

For all $\theta \in \Theta$, define Ψ_θ, a continuous functional on $L^1(\mu)$, as

$$\Psi_\theta(f) = \int_\Omega G(\theta, \omega) f(\omega) \mathrm{d}\mu(\omega)$$

<u>Step 1.</u> I start by proving the existence of a set $\mathcal{Z}$ of Lebesgue measure zero such that, whenever $\theta \notin \mathcal{Z}$, $\bar{S}_p$ is differentiable at θ for all $p \in \Delta\Omega$, and

$$\bar{S}_p'(\theta) = \Psi_\theta(p) \ .$$

Observing that, for all $\theta > \theta_0$, $[\theta_0, \theta] \times \Omega$ is of finite $(\mathcal{L} \otimes \mu)$-measure, define G_f as

$$G_f(\theta) = \int_{[\theta_0, \theta] \times \Omega} G(t, \omega) f(\omega) \mathrm{d}\mathcal{L} \otimes \mu(t, \omega) \ ,$$

for every $f \in L^1(\mu)$. By application of Fubini's theorem,

$$G_f(\theta) = \int_{\theta_0}^{\theta} \Psi_t(f)\mathrm{d}t \ .$$

Consequently G_f can be written as a Lebesgue integral, implying that given some fixed $f \in L^1(\mu)$, G_f is differentiable almost everywhere, and $G'_f(\theta) = \Psi_\theta(f)$.

As $L^1(\mu)$ is separable, there exists a countable set $\mathcal{F}$ dense in $L^1(\mu)$. A countable union of measure zero sets remains of measure zero, so there exists a set $\mathcal{Z}$ of Lebesgue measure zero such that, for all $f \in \mathcal{F}$, and all $\theta \notin \mathcal{Z}$, G_f is differentiable at θ and $G'_f(\theta) = \Psi_\theta(f)$. Choose $\mathcal{Z}$ so as to include the extremities of interval Θ.

We can generalize to all $f \in L^1(\mu)$. Let $f \in L^1(\mu), T \notin \mathcal{Z}$, and $K > 0$ an upper bound of $|G|$. Let $\epsilon > 0$. We want to find a $\delta > 0$ such that, if t satisfies $|T - t| \leq \delta$, then

$$\left| \frac{G_f(T) - G_f(t)}{T - t} - \Psi_t(f) \right| \leq \epsilon \ ,$$

or equivalently,

$$\left| \frac{1}{T-t} \int_t^T \int_\Omega G(r,\omega) f(\omega)\mathrm{d}\mu(\omega)\mathrm{d}r - \int_\Omega G(T,\omega) f(\omega)\mathrm{d}\mu(\omega) \right| \leq \epsilon \ . \tag{8}$$

Let $\tilde{f} \in \mathcal{F}$ such that $\|f - \tilde{f}\| \leq \frac{\epsilon}{3K}$. Then,

$$\left| \frac{1}{T-t} \int_t^T \int_\Omega G(r,\omega)[f(\omega) - \tilde{f}(\omega)]\mathrm{d}\mu(\omega)\mathrm{d}r \right| \leq \frac{\epsilon}{3} \ .$$

Besides, $G_{\tilde{f}}$ is differentiable at T and so, for some $\delta > 0$, $|T - t| \leq \delta$ implies

$$\left| \frac{1}{T-t} \int_t^T \int_\Omega G(r,\omega) \tilde{f}(\omega)\mathrm{d}\mu(\omega)\mathrm{d}r - \int_\Omega G(t,\omega) \tilde{f}(\omega)\mathrm{d}\mu(\omega) \right| \leq \frac{\epsilon}{3} \ .$$

Finally,

$$\left| \int_\Omega G(t,\omega)[f(\omega) - \tilde{f}(\omega)]\mathrm{d}\mu(\omega) \right| \leq \frac{\epsilon}{3} \ .$$

Summing the three inequalities above yields (8). In particular, for all $\theta \notin \mathcal{Z}$, $\bar{S}_p$ is differentiable at θ for all $p \in \Delta\Omega$, and

$$\bar{S}'_p(\theta) = \Psi_\theta(p) \ .$$

STEP 2. Assume S is proper, and let $\theta \notin \mathcal{Z}$. If $p \in \{\Gamma = \theta\}$, $\Gamma(p) \notin \mathcal{Z}$ and so $\bar{S}_p(\Gamma(p))' = 0$, which yields $\{\Gamma = \theta\} \subset \ker \Psi_\theta$. By Theorem 1, there exists a continuous linear functional Φ_θ on $\mathcal{C}(\Omega)$ such that $\{\Gamma = \theta\} = \ker \Phi_\theta \cap \Delta\Omega$. As $\{\Gamma = \theta\}$ is nonempty, applying Lemma 4 yields $\ker \Phi_\theta = \langle\{\Gamma = \theta\}\rangle$ and, as $\{\Gamma = \theta\} \subset \ker \Psi_{\theta|\mathcal{C}(\Omega)}$ we have that $\ker \Phi_\theta \subseteq \ker \Psi_{\theta|\mathcal{C}(\Omega)}$. Consequently there exists a real number $\xi(\theta)$ such that $\Psi_{\theta|\mathcal{C}(\Omega)} = \xi(\theta)\Phi_\theta$ (Lemma 3.1 of Megginson, 1998). If $\Phi_\theta = 0$ choose $\xi(\theta) = 0$. By the Banach extension theorem, Φ_θ can be extended by continuity to the whole space $L^1(\mu)$, and by density of $\mathcal{C}(\Omega)$ in $L^1(\mu)$, we have that $\Psi_\theta = \xi(\theta)\Phi_\theta$ on $L^1(\mu)$. Let $\xi(\theta) = 0$ for all $\theta \in \mathcal{Z}$.

We can choose without loss $\|\Phi_\theta\| = 1$. In the proof of Theorem 1, we showed that Φ_θ can be chosen such that $\theta \mapsto \Phi_\theta(p)$ be Lebesgue measurable for all $p \in \Delta\Omega$. Since, from its definition, $\theta \mapsto \Psi_\theta(p)$ is Lebesgue measurable, writing $\xi(\theta) = \Psi_\theta/\Phi_\theta$ leads to Lebesgue measurability of ξ. Besides, noting that $\|G(\theta,\cdot)\|_\infty = \|\Psi_\theta\| = |\xi(\theta)|\|\Phi_\theta\| = |\xi(\theta)|$, boundedness of ξ follows from boundedness of G.

Therefore, for all $p \in \Delta\Omega$, and all θ,

$$\bar{S}_p(\theta) = \int_{\theta_0}^{\theta} \Psi_t(p)\mathrm{d}t = \int_{\theta_0}^{\theta} \xi(t)\Phi_t(p)\mathrm{d}t \ .$$

By Proposition 1, S is order sensitive. This implies $\xi \geq 0$. Indeed, suppose $\xi(\theta) < 0$ for some $\theta \notin \mathcal{Z}$. Take, for example, $p \in \{\Gamma > \theta\}$. Then,

$$\bar{S}_p'(\theta) = \xi(\theta)\Phi_\theta(p) < 0 \ ,$$

and $\bar{S}_p$ is not (weakly) increasing on $\{t < \Gamma(p)\}$, contradicting order sensitivity of S. Hence $\xi \geq 0$. Assume that, in addition, S is strictly proper. Take any $\theta_1 < \theta_2$ and $p \in \{\Gamma = \theta_2\}$. Then,

$$\begin{aligned} 0 < |\bar{S}_p(\theta_2) - \bar{S}_p(\theta_1)| &= \left| \int_{\theta_1}^{\theta_2} \int_\Omega \xi(t)\Phi_t(p)\mathrm{d}t \right| \ , \\ &\leq \|f\| \int_{\theta_1}^{\theta_2} \xi \ , \end{aligned}$$

implying $\int_{\theta_1}^{\theta_2} \xi > 0$.

B Proofs of Section 4

Through the proofs of this section, I shorten notation and often write, for a scoring rule $S : \Theta \times \Omega \mapsto \mathbb{R}$, $S(\theta)$ to denote the random variable $S(\theta, \cdot)$. For a subset of $\mathcal{S}$ of a vector space, I denote by $\dim \mathcal{V}$ the dimension of its linear span. A convex polyhedra in a convex subset $\mathcal{C}$ of a vector space is nondegenerate when it has the same dimension as $\mathcal{C}$.

The proofs make use of the following lemma.

Lemma 8. *If there exists a strictly proper scoring rule for (Θ, F), then, for all θ, $F(\theta)$ is a nondegenerate closed convex polyhedra of $\Delta\Omega$, and, when the intersection of two elements $F(\theta_1)$ and $F(\theta_2)$ is not empty, it is a degenerate closed convex polyhedron.*

Proof. The lemma is a direct consequence of Theorem 3, which asserts that when a strictly proper scoring rule exists, the $F(\theta)$'s form a power diagram of distributions.

□

B.1 Proof of Theorem 4

The proof uses the following lemma.

Lemma 9. *Let $\mathcal{E}$ be an n-dimensional Hilbert space with an inner product $\langle \cdot, \cdot \rangle$. Let $y_1, \ldots, y_m$ be m vectors that generate $\mathcal{E}$. Consider the two systems of inequalities*

$$\langle y_i, x \rangle \geq 0 \ , \qquad i \in \{1, \ldots, m\} \tag{9}$$

and

$$\langle y_i, x \rangle > 0 \ , \qquad i \in \{1, \ldots, m\} \ . \tag{10}$$

If both systems admit a nonempty set of solutions, then there exist vectors $s_1, \ldots, s_\ell$ of $\mathcal{E}$ such that the set of solutions of (9) is $\{\lambda_1 s_1 + \cdots + \lambda_\ell s_\ell, \lambda_1, \ldots, \lambda_\ell \geq 0\}$ while the set of solutions of (10) is $\{\lambda_1 s_1 + \cdots + \lambda_\ell s_\ell, \lambda_1, \ldots, \lambda_\ell > 0\}$.

Proof. As (9) is a homogeneous system of weak inequalities, its set of solutions is a cone. Let $\{s_1, \ldots, s_\ell\}$ be a set of directrices of the edges of this cone. As by assumption there exists a nonzero solution, this set is not empty. The parametric form of the solutions of (9) is given by the set $\{\sum_i \lambda_i s_i, \ \lambda_1, \ldots, \lambda_\ell \geq 0\}$ (Eremin, 2002). We shall see that the cone $\mathcal{C} = \{\sum_i \lambda_i s_i,\ \lambda_1, \ldots, \lambda_\ell > 0\}$ is the set of solutions of (10).

Part 1. This part shows that any element of $\mathcal{C}$ is solution of (10).

Each vector s_k of $\{s_1, \ldots, s_\ell\}$ is solution of a $(n-1)$-boundary system of the form

$$\begin{cases} \langle y_i, s_k \rangle = 0 \ , & i \notin I_k \ , \\ \langle y_i, s_k \rangle > 0 \ , & i \in I_k \ , \end{cases} \tag{11}$$

for I_k a subset of $\{1, \ldots, m\}$. Let x_0 be a solution of (10). Then x_0 is also solution of (9) and so $x_0 = \sum_i \lambda_i s_i$, with $\lambda_i \geq 0$ for all i. There cannot exist j with $\langle y_j, s_k \rangle = 0$ for all k, otherwise $\langle y_j, x_0 \rangle = 0$ and x_0 would not be solution of (10). Therefore $\cup_k I_k = \{1, \ldots, m\}$.

Let $\hat{x} \in \mathcal{C}$, with $\hat{x} = \sum_i \mu_i s_i$, with $\mu_i > 0$ for all i. Since $\cup_k I_k = \{1, \ldots, m\}$, for all j there exists k such that $\mu_k \langle y_j, s_k \rangle > 0$ and $\mu_k \langle y_i, s_k \rangle \geq 0$ for all $i \neq j$. By summation, for all i, $\langle y_i, \hat{x} \rangle > 0$, and so $\hat{x}$ is solution of (10).

Part 2. This part shows the converse, that any solution of (10) is in $\mathcal{C}$.

Let $\hat{x}$ be a solution of (10). Let $\mathcal{B}_0$ be the open ball of diameter δ centered on $\hat{x}$, and $\mathcal{B}_1$ the open ball of diameter $\frac{3}{4}\delta$ with the same center. If δ is chosen small enough, any vector of $\mathcal{B}_0$ is solution of (10) since its inequalities define an open set of $\mathcal{E}$.

For $\epsilon > 0$, let $t = \epsilon \sum_i s_i$, and let $\mathcal{B}_1' = \mathcal{B}_1 + t$ be the translated ball by t. If ϵ is chosen small enough, the open ball $\mathcal{B}_1'$ remains contained in $\mathcal{B}_0$. In such a case, $\hat{x}$, which also belongs to $\mathcal{B}_1'$, is the image of some $x_0 \in \mathcal{B}_1$. As x_0 is solution of (9), we can write $x_0 = \sum_i \lambda_i s_i$, with $\lambda_i \geq 0$ for all i, hence $\hat{x} = \sum \mu_i s_i$, with $\mu_i = \lambda_i + \epsilon > 0$ for all i. Therefore $\hat{x} \in \mathcal{C}$. This concludes the proof of the lemma. □

I now turn back to the proof of the main theorem. Denote by $\mathcal{S}$ the space of scoring rules, *i.e.*, the linear space of functions $S : \Theta \times \Omega \mapsto \mathbb{R}$, considered as a Hilbert space whose inner product is defined as $\langle S_1, S_2 \rangle = \sum_{\theta,\omega} S_1(\theta, \omega) S_2(\theta, \omega)$.

Part 1. Suppose that there exists a strictly proper scoring rule for the statistic (Θ, F). $S \in \mathcal{S}$ is proper if, and only if, for all $\theta, \hat{\theta} \in \Theta$,

$$\langle S(\theta), p \rangle = \langle S(\hat{\theta}), p \rangle \qquad \forall p \in F(\theta) \cap F(\hat{\theta}) \ , \tag{12}$$

$$\langle S(\theta), p \rangle \geq \langle S(\hat{\theta}), p \rangle \qquad \forall p \in F(\theta) \backslash F(\hat{\theta}) \ , \tag{13}$$

with the last inequality being strict if and only if S is strictly proper.

By Lemma 8, for all $\theta \in \Theta$, the level set $F(\theta)$ is a bounded convex polyhedron, and so is the convex hull of a set of vertices $\mathcal{V}_\theta$. I supplement the set of vertices $\mathcal{V}_\theta$ of each polyhedron $F(\theta)$ by vertices of the other polyhedra that belong to its boundary, in such a way that, for all $\theta, \hat{\theta} \in \Theta$, and all p belonging to both $F(\theta)$ and $\mathcal{V}_{\hat{\theta}}$, p also belong to $\mathcal{V}_\theta$. Let us write $\mathcal{V}_\theta$ as $\{p_1^\theta, \ldots, p_{\ell_\theta}^\theta\}$.

Let $S \in \mathcal{S}$ be proper (resp. strictly proper). Let $\theta, \hat{\theta} \in \Theta$. If $p \in \mathcal{V}_\theta \cap \mathcal{V}_{\hat{\theta}}$, then $p \in F(\theta) \cap F(\hat{\theta})$ and so by (12), $\langle S(\theta), p\rangle = \langle S(\hat{\theta}), p\rangle$. If $p \in \mathcal{V}_\theta \backslash \mathcal{V}_{\hat{\theta}}$, then $p \in F(\theta)$ and $p \notin F(\hat{\theta})$, since by construction of $\mathcal{V}_\theta$, $p \in F(\hat{\theta})$ and $p \in \mathcal{V}_\theta$ implies $p \in \mathcal{V}_{\hat{\theta}}$. So by (13), $\langle S(\theta), p\rangle \geq \langle S(\hat{\theta}), p\rangle$ (resp. $\langle S(\theta), p\rangle > \langle S(\hat{\theta}), p\rangle$).

We shall show the sufficiency of these two conditions. Assume that if $p \in \mathcal{V}_\theta \cap \mathcal{V}_{\hat{\theta}}$, then $\langle S(\theta), p\rangle = \langle S(\hat{\theta}), p\rangle$, and if $p \in \mathcal{V}_\theta \backslash \mathcal{V}_{\hat{\theta}}$, then $\langle S(\theta), p\rangle \geq \langle S(\hat{\theta}), p\rangle$ (resp. $\langle S(\theta), p\rangle > \langle S(\hat{\theta}), p\rangle$). Let $p \in F(\theta) \cap F(\hat{\theta})$. Then p is a linear combination of vectors in $\mathcal{V}_\theta$ and $\mathcal{V}_{\hat{\theta}}$, and since the equality $\langle S(\theta), q\rangle = \langle S(\hat{\theta}), q\rangle$ holds for all vectors q that belong to these two sets, by linearity $\langle S(\theta), p\rangle = \langle S(\hat{\theta}), p\rangle$. Now let $p \in F(\theta) \backslash F(\hat{\theta})$. Then $p = \sum_i \lambda_i p_i^\theta$ for some nonnegative scalars λ_i that sum to one. Since $p \notin F(\hat{\theta})$, there exists k such that $\lambda_k > 0$ and $p_k^\theta \notin F(\hat{\theta})$. Hence $p_k^\theta \in \mathcal{V}_\theta \backslash \mathcal{V}_{\hat{\theta}}$, and $\langle S(\theta), p_k^\theta\rangle \geq \langle S(\hat{\theta}), p_k^\theta\rangle$ (resp. $\langle S(\theta), p_k^\theta\rangle > \langle S(\hat{\theta}), p_k^\theta\rangle$). For $i \neq k$, we either have $p_i^\theta \in \mathcal{V}_\theta \cap \mathcal{V}_{\hat{\theta}}$ or $p_i^\theta \in \mathcal{V}_\theta \backslash \mathcal{V}_{\hat{\theta}}$, and so in both cases $\langle S(\theta), p\rangle \geq \langle S(\hat{\theta}), p\rangle$. Hence

$$\langle S(\theta), p\rangle = \sum_i \lambda_i \langle S(\theta), p_i^\theta\rangle \geq \sum_i \lambda_i \langle S(\hat{\theta}), p_i^\theta\rangle = \langle S(\hat{\theta}), p\rangle$$

with a strict inequality when S is strictly proper. Therefore, we have shown that a scoring rule S is proper if, and only if, S is solution of the following finite linear system in the space $\mathcal{S}$,

$$\begin{cases} \langle S(\theta) - S(\hat{\theta}), p\rangle = 0 \ , & \theta, \hat{\theta} \in \Theta, p \in \mathcal{V}_\theta \cap \mathcal{V}_{\hat{\theta}} \ , \\ \langle S(\theta) - S(\hat{\theta}), p\rangle \geq 0 \ , & \theta, \hat{\theta} \in \Theta, p \in \mathcal{V}_\theta \backslash \mathcal{V}_{\hat{\theta}} \ , \end{cases} \tag{14}$$

and S is strictly proper if, and only if, S is solution of the system

$$\begin{cases} \langle S(\theta) - S(\hat{\theta}), p\rangle = 0 \ , & \theta, \hat{\theta} \in \Theta, p \in \mathcal{V}_\theta \cap \mathcal{V}_{\hat{\theta}} \ , \\ \langle S(\theta) - S(\hat{\theta}), p\rangle > 0 \ , & \theta, \hat{\theta} \in \Theta, p \in \mathcal{V}_\theta \backslash \mathcal{V}_{\hat{\theta}} \ . \end{cases} \tag{15}$$

Part 2. Let $\mathcal{S}_0$ be the space of solutions of the finite system of equalities (in $\mathcal{S}$)

$$\langle S(\theta) - S(\hat{\theta}), p \rangle = 0 \ , \quad \theta, \hat{\theta} \in \Theta, p \in \mathcal{V}_\theta \cap \mathcal{V}_{\hat{\theta}}$$

corresponding to the first part of (14) and (15).

Step 1. Let $\mathcal{S}_0^\perp$ be the orthogonal complement of $\mathcal{S}_0$ in $\mathcal{S}$. Let $S \in \mathcal{S}_0$. Then, for any vector X of $\mathcal{S}$, $\langle X, S \rangle = \langle X^{\perp\perp}, S \rangle$, with $X^{\perp\perp} \in \mathcal{S}_0$ and where $X^{\perp\perp} + X^\perp$ is the decomposition of X according to the direct sum $\mathcal{S} = \mathcal{S}_0 \oplus \mathcal{S}_0^\perp$. Therefore, there exists vectors $Y_1, \dots, Y_m$ in $\mathcal{S}_0$ such that the solutions of (14) in $\mathcal{S}$ are exactly the solutions of the finite system of weak linear inequalities in $\mathcal{S}_0$

$$\langle Y_i, S \rangle \geq 0, \quad i = 1, \dots, m \tag{16}$$

and the solutions of (15) are the solutions of the finite system of strict linear inequalities in $\mathcal{S}_0$

$$\langle Y_i, S \rangle > 0, \quad i = 1, \dots, m \ . \tag{17}$$

Step 2. Let $\mathcal{K}$ be the kernel of (16) in $\mathcal{S}_0$, and $\mathcal{K}^\perp$ be its orthogonal complement in $\mathcal{S}_0$. For each Y_i, write $Y_i^{\perp\perp} + Y_i^\perp$ its decomposition according to the direct sum $\mathcal{S}_0 = \mathcal{K} \oplus \mathcal{K}^\perp$.

We can easily describe $\mathcal{K}$: $S \in \mathcal{K}$ if and only if $S \in \mathcal{S}_0$, and if, for all $\theta, \hat{\theta} \in \Theta$ and all $p \in \mathcal{V}_\theta \backslash \mathcal{V}_{\hat{\theta}}$, $\langle S(\theta) - S(\hat{\theta}), p \rangle = 0$. Since $(\mathcal{V}_\theta \cap \mathcal{V}_{\hat{\theta}}) \cup (\mathcal{V}_\theta \backslash \mathcal{V}_{\hat{\theta}}) = \mathcal{V}_\theta$, $\mathcal{K}$ is simply the solution of

$$\langle S(\theta) - S(\hat{\theta}), p \rangle = 0 \ , \quad \theta, \hat{\theta} \in \Theta, p \in \mathcal{V}_\theta \ . \tag{18}$$

Any S such that $S(\theta) = S(\hat{\theta})$ for all $\theta, \hat{\theta} \in \mathcal{S}$ is solution. By Lemma 8, $F(\theta)$ has dimension $|\Omega|$ for all θ, and so the linear span of $\mathcal{V}_\theta$ is $\mathbb{R}^\Omega$. Consequently, if S is solution of (18), then $\langle S(\theta) - S(\hat{\theta}), p \rangle = 0$ for all $\theta, \hat{\theta}$ and all $p \in \mathbb{R}^\Omega$, implying $S(\theta) = S(\hat{\theta})$. Hence $\mathcal{K} = \{S \in \mathcal{S} \mid S(\theta, \omega) = S(\hat{\theta}, \omega) \, \forall \theta \neq \hat{\theta}\}$.

Step 3. Let's consider the following two systems of inequalities in $\mathcal{K}^\perp$:

$$\langle Y_i^\perp, S \rangle \geq 0 \ , \quad i = 1, \dots, m \tag{19}$$

and

$$\langle Y_i^\perp, S \rangle > 0 \ , \quad i = 1, \dots, m \ . \tag{20}$$

If $S \in \mathcal{K}^\perp$, $\langle Y_i, S \rangle = \langle Y_i^\perp, S \rangle$, and the solutions of (16) (resp. (17)) are the elements of $\mathcal{K}$ added to the solutions of (19) (resp. (20)). The systems (19) and (20) have full rank in $\mathcal{K}^\perp$, and since by assumption there exists a strictly proper scoring rule, both admit at least one solution. By Lemma 9, there exist vectors $S_1, \ldots, S_\ell \in \mathcal{K}^\perp$ such that S is solution of (19) (resp. of (20)) if and only if S is a nonnegative (resp. strictly positive) linear combination of $S_1, \ldots, S_\ell$.

Therefore, S is solution of (14) (resp. of (15)) if, and only if, $S = \kappa + \sum_i \lambda_i S_i$, for $\kappa \in \mathcal{K}$ and $\lambda_1, \ldots, \lambda_\ell \geq 0$ (resp. $\lambda_1, \ldots, \lambda_\ell > 0$).

B.2 Proof of Theorem 5

If part. The construction of strictly order-sensitive scoring rules shall be done in Theorem 6 and Proposition 2.

Only if part. Let S be a strictly order-sensitive scoring rule.

STEP 1. This first step shows that for all i and $j > i+1$, if $p \in F(\theta_i)$ and $p \in F(\theta_j)$ then $p \in F(\theta_{i+1})$. Suppose by contradiction that there exist i and p with $p \in F(\theta_i)$, $p \notin F(\theta_{i+1})$, and $p \in F(\theta_j)$ for some $j > i+1$. By Lemma 8, $F(\theta_i)$ is a convex polyhedron of nonempty relative interior. Since $p \in F(\theta_i)$, there exists a sequence of vectors $\{p_k\}_{k\geq 1}$ of the relative interior of $F(\theta_i)$ that converges to p. By continuity $\lim_{k\to+\infty} S(\theta_i, p_k) \to S(\theta_i, p)$. Let $\delta_k = S(\theta_i, p_k) - S(\theta_{i+1}, p_k)$. Since p_k and p both belong to $F(\theta_i)$, but not to $F(\theta_{i+1})$, $\delta_k > 0$, and δ_k converges to $\delta = S(\theta_i, p) - S(\theta_{i+1}, p) > 0$. Therefore $\inf\{\delta_k\}_{k\geq 1} > 0$. Let $\epsilon = \inf\{\delta_k/2\}_{k\geq 1}$. By continuity, there exists K such that

$$|S(\theta_i, p) - S(\theta_i, p_K)| \leq \epsilon/2 \ ,$$

and

$$|S(\theta_j, p) - S(\theta_j, p_K)| \leq \epsilon/2 \ ,$$

so that, since θ_i and θ_j both contain p, $S(\theta_i, p) = S(\theta_j, p)$ and

$$|S(\theta_i, p_K) - S(\theta_j, p_K)| \leq \epsilon \ .$$

Hence, $S(\theta_j, p_K) > S(\theta_i, p_K) - \epsilon = S(\theta_{i+1}, p_K) + \delta_K - \epsilon > S(\theta_{i+1}, p_K)$. However, p_K is in the relative interior of $F(\theta_i)$, which means according to Lemma 8 that θ_i is the only true value of the statistic for p_K. But, since $i < i+1 < j$, and S is strictly order sensitive, we should have $S(\theta_{i+1}, p_K) > S(\theta_j, p_K)$. Contradiction.

STEP 2. Now let $1 \leq j \leq n-1$. Let $B_j = F(\theta_1) \cup \cdots \cup F(\theta_j)$, and $C_j = F(\theta_{j+1}) \cup \cdots \cup F(\theta_n)$. By Lemma 8, B_j and C_j are polyhedra of dimension $|\Omega|$ and nonempty relative interior, with $B_j \cup C_j = \Delta\Omega$. Let $i \leq j < j+1 \leq k$. If $p \in F(\theta_i)$ and $p \in F(\theta_k)$, an iterative application of the claim of step 1 above yields $p \in F(\theta_i), F(\theta_{i+1}), \ldots, F(\theta_k)$. In particular, $p \in F(\theta_j) \cap F(\theta_{j+1})$. Therefore $B_j \cap C_j = F(\theta_j) \cap F(\theta_{j+1})$. By Lemma 8, the dimension of $F(\theta_j) \cap F(\theta_{j+1})$ is at most $|\Omega| - 1$, so that there is a hyperplane of distributions $\mathcal{H}$ that contains $B_j \cap C_j$. Suppose that there exists a distribution p of $\mathcal{H}$ that does not belong to $B_j \cap C_j$. Since $B_j \cup C_j = \Delta\Omega$, $p \in B_j$ or $p \in C_j$. Suppose for example that $p \in B_j$. Then there exists a distribution q in the relative interior of C_j with $q \notin \mathcal{H}$. Note that the segment $]p, q]$ contains only elements of B_j or C_j. Since both sets are closed, the segment intersects $B_j \cap C_j$, which is impossible since $]p, q]$ does not intersect $\mathcal{H}$. So $B_j \cap C_j$ must be the full hyperplane of distributions $\mathcal{H}$: $\mathcal{H} = B_j \cap C_j = F(\theta_j) \cap F(\theta_{j+1})$. This concludes the proof.

B.3 Proof of Theorem 6

Part 1. Define

$$S(\theta_k, \omega) = \kappa(\omega) + \sum_{1 \leq i < k} \lambda_i \mathbf{n}_i(\omega) \ ,$$

with $\lambda_1, \ldots, \lambda_{n-1} \geq 0$, and $\kappa \in \mathbb{R}^\Omega$.

As $\mathbf{n}_k$ is oriented positively, $\langle \mathbf{n}_k, p \rangle \geq 0$ for all $p \in F(\theta_{k+1}), \ldots, F(\theta_n)$, and $\langle \mathbf{n}_k, p \rangle \leq 0$ for all $p \in F(\theta_1), \ldots, F(\theta_k)$. The inequalities are strict if $p \notin F(\theta_k) \cap F(\theta_{k+1})$.

Let $p \in F(\theta_k)$. If $j < k$,

$$\mathop{\mathsf{E}}_{\omega \sim p}[S(\theta_k, \omega)] - \mathop{\mathsf{E}}_{\omega \sim p}[S(\theta_j, \omega)] = \sum_{j \leq i < k} \lambda_i \langle \mathbf{n}_i, p \rangle \geq 0 \ ,$$

and, if $j > k$,

$$\mathop{\mathsf{E}}_{\omega \sim p}[S(\theta_k, \omega)] - \mathop{\mathsf{E}}_{\omega \sim p}[S(\theta_j, \omega)] = -\sum_{k \leq i < j} \lambda_i \langle \mathbf{n}_i, p \rangle \geq 0 \ .$$

Therefore S is a proper scoring rule. If, in addition, $\lambda_1, \ldots, \lambda_{n-1} > 0$, the inequalities become strict when $p \notin F(\theta_j)$, making S strictly proper.

Part 2. Now assume S is a proper scoring rule. Then, for all $p \in F(\theta_k) \cap F(\theta_{k+1})$,

$1 \le k < n$, $\langle S(\theta_k), p\rangle = \langle S(\theta_{k+1}), p\rangle$, and so $\langle S(\theta_{k+1}) - S(\theta_k), p\rangle = 0$. Theorem 5 says that $F(\theta_k) \cap F(\theta_{k+1})$ is a hyperplane of $\Delta\Omega$. Its linear span is a hyperplane $\mathcal{H}_k$ of $\mathbb{R}^\Omega$, thus $S(\theta_{k+1}) - S(\theta_k) = \lambda_k \mathbf{n}_k$, where $\mathbf{n}_k$ is a normal to $\mathcal{H}_k$ oriented positively.

Let $p \in F(\theta_{k+1})$, $p \notin F(\theta_k)$. As S is proper, $\langle S(\theta_{k+1}), p\rangle \ge \langle S(\theta_k), p\rangle$, so $\lambda_k \langle \mathbf{n}_k, p\rangle \ge 0$. Since $p \notin \mathcal{H}_k$ and $\mathbf{n}_k$ is positively oriented, $\langle \mathbf{n}_k, p\rangle > 0$, implying $\lambda_k \ge 0$ ($\lambda_k > 0$ if S is strictly proper).

Therefore

$$S(\theta_k) = S(\theta_1) + \sum_{1 \le i < k} (S(\theta_{i+1}) - S(\theta_i)) = \kappa + \sum_{1 \le i < k} \lambda_i \mathbf{n}_i \ ,$$

with $\kappa = S(\theta_1)$.

B.4 Proof of Proposition 2

Assume the statistic accepts a strictly order-sensitive scoring rule with respect to the order relation $\prec$, and let $\theta_1 \prec \cdots \prec \theta_n$ be the elements of the value set of the statistic. Let S be a proper scoring rule. Theorem 6 shows that S takes the form

$$S(\theta_k, \omega) = \kappa(\omega) + \sum_{1 \le i < k} \lambda_i \mathbf{n}_i(\omega) \ ,$$

with $\lambda_1, \ldots, \lambda_{n-1} \ge 0$. Let $p \in \Delta\Omega$. Since the normals are positively oriented, $\langle \mathbf{n}_k, p\rangle \ge 0$ if $p \in F(\theta_{k+1}), \ldots, F(\theta_n)$ and $\langle \mathbf{n}_k, p\rangle \le 0$ if $p \in F(\theta_1), \ldots, F(\theta_k)$, the inequalities being strict if $p \notin F(\theta_k) \cap F(\theta_{k+1})$. So, for all $\theta, \theta_k, \theta_j$, if $\theta_j \prec \theta_k \prec \theta$ and $p \in F(\theta)$, then

$$\mathop{\mathsf{E}}_{\omega \sim p}[S(\theta_k, \omega)] - \mathop{\mathsf{E}}_{\omega \sim p}[S(\theta_j, \omega)] = \sum_{j \le i < k} \lambda_i \langle \mathbf{n}_i, p\rangle \ge 0 \ .$$

Similarly, if $\theta \prec \theta_k \prec \theta_j$,

$$\mathop{\mathsf{E}}_{\omega \sim p}[S(\theta_k, \omega)] - \mathop{\mathsf{E}}_{\omega \sim p}[S(\theta_j, \omega)] = -\sum_{k \le i < j} \lambda_i \langle \mathbf{n}_i, p\rangle \ge 0 \ .$$

Hence S is order sensitive. If S is strictly proper, the λ_i's are strictly positive, making the above inequalities strict, and S becomes strictly order sensitive.

C Proofs of Section 5

C.1 Proof of Lemma 1

Assume without loss of generality $0 \le u_t \le 1$. Suppose

$$\lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} u_t = 0 \ ,$$

then, for all $\epsilon > 0$,

$$\lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} u_t \mathbb{1}\{u_t \ge \epsilon\} = 0 \ .$$

As

$$\frac{1}{T} \sum_{t=1}^{T} \mathbb{1}\{u_t \ge \epsilon\} \le \frac{1}{T} \sum_{t=1}^{T} \frac{u_t}{\epsilon} \mathbb{1}\{u_t \ge \epsilon\} \ ,$$

we get that

$$\lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{1}\{u_t \ge \epsilon\} = 0 \ .$$

To get the converse, suppose that $\frac{1}{T}\sum_{t=1}^{T} u_t$ does not converge to zero. Then, there exists $\epsilon > 0$ and a subsequence $(w_t)_{t \ge 1}$ such that, for all T, $\frac{1}{T}\sum_{t=1}^{T} w_t \ge \epsilon$. Hence,

$$\frac{1}{T} \sum_{t=1}^{T} w_t \mathbb{1}\left\{w_t \ge \frac{\epsilon}{2}\right\} \ge \frac{\epsilon}{2} \ ,$$

so $\frac{1}{T}\sum_{t=1}^{T} u_t \mathbb{1}\{u_t \ge \epsilon/2\}$ does not converge to zero. As $u_t \mathbb{1}\{u_t \ge \epsilon/2\} \le \mathbb{1}\{u_t \ge \epsilon/2\}$, $\frac{1}{T}\sum_{t=1}^{T} \mathbb{1}\{u_t \ge \epsilon/2\}$ does not converge to zero.

C.2 Proof of Lemma 2

For a set of distributions over states $\mathcal{P}$, let $\mathcal{P}^c$ be the complement of $\mathcal{P}$ in $\Delta\Omega$, and define $\mathcal{B}_\delta(\mathcal{P}) = \{q \in \Delta\Omega \mid \exists p \in \mathcal{P}, \|q - p\| < \delta\}$.

Fix $\theta, \hat{\theta} \in \Theta$. By contradiction, suppose that for some $\delta > 0$ and all $\epsilon > 0$, there exists some $p \in F(\theta)$ such that $p \in \mathcal{B}_\delta(F(\hat{\theta}))^c$ and $S(\hat{\theta}, p) \ge S(\theta, p) - \epsilon$. Choosing $\epsilon = 1/n$, we can generate a sequence of distributions $(p_n)_{n \ge 1}$ such that $p_n \in F(\theta) \cap \mathcal{B}_\delta(F(\hat{\theta}))^c$ and $S(\hat{\theta}, p) \ge S(\theta, p) - 1/n$. Observing that $F(\theta) \cap \mathcal{B}_\delta(F(\hat{\theta}))^c$ is compact, we can extract a subsequence that converges to some $p_\infty \in F(\theta) \cap \mathcal{B}_\delta(F(\hat{\theta}))^c$ and that satisfies, by continuity, $S(\hat{\theta}, p_\infty) \geq S(\theta, p_\infty)$. But as S is strictly proper and $p_\infty \in F(\theta)$, it must be the case that $p_\infty \in F(\hat{\theta})$, contradicting $p \in \mathcal{B}_\delta(F(\hat{\theta}))^c$.

The constant ϵ may depend on the choice of $\theta, \hat{\theta}$, however since there are finitely many such pairs, the result also holds uniformly.