Econometrica Supplementary Material

SUPPLEMENT TO "MONOTONE ADDITIVE STATISTICS"

(*Econometrica*, Vol. 92, No. 4, July 2024, 995–1031)

XIAOSHENG MU
Department of Economics, Princeton University

LUCIANO POMATTO
Division of the Humanities and Social Sciences, California Institute of Technology

PHILIPP STRACK
Department of Economics, Yale University

OMER TAMUZ
Division of the Humanities and Social Sciences, California Institute of Technology

APPENDIX D: PROOF OF THEOREM 2

THE PROOF IS CONSIDERABLY MORE COMPLEX THAN the proof of Theorem 1, so we break it into several steps below.

D.1. *Step 1: Catalytic Order on L_M*

We first establish a generalization of Theorem 6 to unbounded random variables. For two random variables X and Y with c.d.f. F and G respectively, we say that X *dominates Y in both tails* if there exists a positive number N with the property that

$$G(x) > F(x) \quad \text{for all } |x| \geq N.$$

In particular, X needs to be unbounded from above, and Y unbounded from below.

LEMMA D.1: *Suppose $X, Y \in L_M$ satisfy $K_a(X) > K_a(Y)$ for every $a \in \mathbb{R}$. Suppose further that X dominates Y in both tails. Then there exists an independent random variable $Z \in L_M$ such that $X + Z \geq_1 Y + Z$.*

PROOF: We will take Z to have a normal distribution, which does belong to L_M. Following the proof of Theorem 6, we let $\sigma(x) = G(x) - F(x)$, and seek to show that $[\sigma * h](y) \geq 0$ for every y when h is a Gaussian density with sufficiently large variance. By assumption, $\sigma(x)$ is strictly positive for $|x| \geq N$. Thus, there exists $\delta > 0$ such that $\int_{N+1}^{N+2} \sigma(x)\,\mathrm{d}x > \delta$, as well as $\int_{-N-2}^{-N-1} \sigma(x)\,\mathrm{d}x > \delta$. We fix $A > 0$ that satisfies $\mathrm{e}^A \geq \frac{4N}{\delta}$.

Similarly to (9), we have for $h(x) = \mathrm{e}^{-\frac{x^2}{2V}}$ that

$$\mathrm{e}^{\frac{y^2}{2V}} \int \sigma(x) h(y - x)\,\mathrm{d}x = \int_{-\infty}^{\infty} \sigma(x) \cdot \mathrm{e}^{\frac{y}{V} \cdot x} \cdot \mathrm{e}^{-\frac{x^2}{2V}}\,\mathrm{d}x. \tag{S1}$$

The variance V is to be determined below.

Xiaosheng Mu: xmu@princeton.edu
Luciano Pomatto: luciano@caltech.edu
Philipp Strack: philipp.strack@yale.edu
Omer Tamuz: tamuz@caltech.edu

 https://doi.org/10.3982/ECTA19967

We first show that the right-hand side is positive if $V \geq (N+2)^2$ and $\frac{y}{V} \geq A$. Indeed, since $\sigma(x) > 0$ for $|x| \geq N$, this integral is bounded from below by

$$\begin{aligned}
&\int_{-N}^{N} \sigma(x) \cdot \mathrm{e}^{\frac{y}{V} \cdot x} \cdot \mathrm{e}^{-\frac{x^2}{2V}} \, \mathrm{d}x + \int_{N+1}^{N+2} \sigma(x) \cdot \mathrm{e}^{\frac{y}{V} \cdot x} \cdot \mathrm{e}^{-\frac{x^2}{2V}} \, \mathrm{d}x \\
&\quad \geq -2N \cdot \mathrm{e}^{\frac{y}{V} \cdot N} + \delta \cdot \mathrm{e}^{\frac{y}{V} \cdot (N+1)} \cdot \mathrm{e}^{-\frac{(N+2)^2}{2V}} \\
&\quad = \mathrm{e}^{\frac{y}{V} \cdot N} \cdot \left(-2N + \delta \cdot \mathrm{e}^{\frac{y}{V}} \cdot \mathrm{e}^{-\frac{(N+2)^2}{2V}}\right) \\
&\quad > 0,
\end{aligned}$$

where the last inequality uses $\mathrm{e}^{\frac{y}{V}} \geq \mathrm{e}^{A} \geq \frac{4N}{\delta}$ and $\mathrm{e}^{-\frac{(N+2)^2}{2V}} \geq \mathrm{e}^{-\frac{1}{2}} > \frac{1}{2}$. By a symmetric argument, we can show that the right-hand side of (S1) is also positive when $\frac{y}{V} \leq -A$.

It remains to consider the case where $\frac{y}{V} \in [-A, A]$. Here we rewrite the integral on the right-hand side of (S1) as

$$\int_{-\infty}^{\infty} \sigma(x) \cdot \mathrm{e}^{\frac{y}{V} \cdot x} \cdot \mathrm{e}^{-\frac{x^2}{2V}} \, \mathrm{d}x = M_\sigma\left(\frac{y}{V}\right) - \int_{-\infty}^{\infty} \sigma(x) \cdot \mathrm{e}^{\frac{y}{V} \cdot x} \cdot \left(1 - \mathrm{e}^{-\frac{x^2}{2V}}\right) \mathrm{d}x,$$

where $M_\sigma(a) = \int_{-\infty}^{\infty} \sigma(x) \cdot \mathrm{e}^{ax} \, \mathrm{d}x = \frac{1}{a}\mathbb{E}[\mathrm{e}^{aX}] - \frac{1}{a}\mathbb{E}[\mathrm{e}^{aY}]$ is by assumption strictly positive for all a. By continuity, there exists some $\varepsilon > 0$ such that $M_{\sigma(a)} > \varepsilon$ for all $|a| \leq A$. So it only remains to show that when V is sufficiently large,

$$\int_{-\infty}^{\infty} \sigma(x) \cdot \mathrm{e}^{ax} \cdot \left(1 - \mathrm{e}^{-\frac{x^2}{2V}}\right) \mathrm{d}x < \varepsilon \quad \text{for all } |a| \leq A. \tag{S2}$$

To estimate this integral, note that $M_\sigma(A) = \int_{-\infty}^{\infty} \sigma(x) \cdot \mathrm{e}^{Ax} \, \mathrm{d}x$ is finite. Since $\sigma(x) > 0$ for $|x|$ sufficiently large, we deduce from the Monotone Convergence theorem that $\int_{-\infty}^{T} \sigma(x) \cdot \mathrm{e}^{Ax} \, \mathrm{d}x$ converges to $M_\sigma(A)$ as $T \to \infty$. In other words, $\int_{T}^{\infty} \sigma(x) \cdot \mathrm{e}^{Ax} \, \mathrm{d}x \to 0$. We can thus find a sufficiently large $T > N$ such that $\int_{T}^{\infty} \sigma(x) \cdot \mathrm{e}^{Ax} \, \mathrm{d}x < \frac{\varepsilon}{4}$, and likewise $\int_{-\infty}^{-T} \sigma(x) \cdot \mathrm{e}^{-Ax} \, \mathrm{d}x < \frac{\varepsilon}{4}$.

As $1 - \mathrm{e}^{-\frac{x^2}{2V}} \geq 0$ and $\mathrm{e}^{ax} \leq \mathrm{e}^{A|x|}$ when $|a| \leq A$, we deduce that

$$\int_{|x| \geq T} \sigma(x) \cdot \mathrm{e}^{ax} \cdot \left(1 - \mathrm{e}^{-\frac{x^2}{2V}}\right) \mathrm{d}x < \frac{\varepsilon}{2} \quad \text{for all } |a| \leq A.$$

Moreover, for this fixed T, we have $\mathrm{e}^{-\frac{T^2}{2V}} \to 1$ when V is large, and thus

$$\int_{|x| \leq T} \sigma(x) \cdot \mathrm{e}^{ax} \cdot \left(1 - \mathrm{e}^{-\frac{x^2}{2V}}\right) \mathrm{d}x < 2T\mathrm{e}^{AT}\left(1 - \mathrm{e}^{-\frac{T^2}{2V}}\right) < \frac{\varepsilon}{2} \quad \text{for all } |a| \leq A.$$

These estimates together imply that (S2) holds for sufficiently large V. This completes the proof. *Q.E.D.*

D.2. *Step 2: A Perturbation Argument*

With Lemma D.1, we know that if Φ is a monotone additive statistic defined on L_M, then $K_a(X) \geq K_a(Y)$ for all $a \in \mathbb{R}$ implies $\Phi(X) \geq \Phi(Y)$ *under the additional assumption* *that* X dominates Y in both tails (same proof as for Lemma 1). Below, we deduce the same result without this extra assumption. To make the argument simpler, assume X and Y are unbounded both from above and from below; otherwise, we can add to them an independent Gaussian random variable without changing either the assumption or the conclusion. In doing so, we can further assume X and Y admit probability density functions.

We first construct a heavy right-tailed random variable as follows:

LEMMA D.2: *For any $Y \in L_M$ that is unbounded from above and admits densities, there exists $Z \in L_M$ such that $Z \geq 0$ and $\frac{\mathbb{P}[Z>x]}{\mathbb{P}[Y>x]} \to \infty$ as $x \to \infty$.*

PROOF: For this result, it is without loss to assume $Y \geq 0$ because we can replace Y by $|Y|$ and only strengthen the conclusion. Let $g(x)$ be the probability density function of Y. We consider a random variable Z whose p.d.f. is given by $cxg(x)$ for all $x \geq 0$, where $c > 0$ is a normalizing constant to ensure $\int_{x\geq 0} cxg(x)\,\mathrm{d}x = 1$. Since the likelihood ratio between $Z = x$ and $Y = x$ is cx, it is easy to see that the ratio of tail probabilities also diverges. Thus, it only remains to check $Z \in L_M$. This is because

$$\mathbb{E}\big[\mathrm{e}^{aZ}\big] = c \int_{x\geq 0} xg(x)\mathrm{e}^{ax}\,\mathrm{d}x,$$

which is simply c times the derivative of $\mathbb{E}[\mathrm{e}^{aY}]$ with respect to a. It is well known that the moment generating function is smooth whenever it is finite. So this derivative is finite, and $Z \in L_M$. *Q.E.D.*

In the same way, we can construct heavy left-tailed distributions:

LEMMA D.3: *For any $X \in L_M$ that is unbounded from below and admits densities, there exists $W \in L_M$ such that $W \leq 0$ and $\frac{\mathbb{P}[W\leq x]}{\mathbb{P}[X\leq x]} \to \infty$ as $x \to -\infty$.*

With these technical lemmata, we now construct “perturbed” versions of any two random variables X and Y to achieve dominance in both tails. For any random variable $Z \in L_M$ and every $\varepsilon > 0$, let Z_ε be the random variable that equals Z with probability ε, and 0 with probability $1 - \varepsilon$. Note that Z_ε also belongs to L_M.

LEMMA D.4: *Given any two random variables $X, Y \in L_M$ that are unbounded on both sides and admit densities. Let $Z \geq 0$ and $W \leq 0$ be constructed from the above two lemmata. Then for every $\varepsilon > 0$, $X + Z_\varepsilon$ dominates $Y + W_\varepsilon$ in both tails.*

PROOF: For the right tail, we need $\mathbb{P}[X + Z_\varepsilon > x] > \mathbb{P}[Y + W_\varepsilon > x]$ for all $x \geq N$. Note that $W_\varepsilon \leq 0$, so $\mathbb{P}[Y + W_\varepsilon > x] \leq \mathbb{P}[Y > x]$. On the other hand,

$$\mathbb{P}[X + Z_\varepsilon > x] \geq \mathbb{P}[X \geq 0] \cdot \mathbb{P}[Z_\varepsilon > x] = \mathbb{P}[X \geq 0] \cdot \varepsilon \cdot \mathbb{P}[Z > x].$$

Since by assumption X is unbounded from above, the term $\mathbb{P}[X \geq 0] \cdot \varepsilon$ is a strictly positive constant that does not depend on x. Thus, for sufficiently large x, we have

$$\mathbb{P}[X \geq 0] \cdot \varepsilon \cdot \mathbb{P}[Z > x] > \mathbb{P}[Y > x]$$

by the construction of Z. This gives dominance in the right tail. The left tail is similar. *Q.E.D.*

D.3. *Step 3: Monotonicity w.r.t. K_a*

The next result generalizes the key Lemma 1 to our current setting:

LEMMA D.5: *Let $\Phi: L_M \to \mathbb{R}$ be a monotone additive statistic. If $K_a(X) \geq K_a(Y)$ for all $a \in \mathbb{R}$, then $\Phi(X) \geq \Phi(Y)$.*

PROOF: As discussed, we can without loss assume X, Y are unbounded on both sides, and admit densities. Let Z and W be constructed as above; then, for each $\varepsilon > 0$, $X + Z_\varepsilon$ dominates $Y + W_\varepsilon$ in both tails, and $K_a(X + Z_\varepsilon) > K_a(X) \geq K_a(Y) > K_a(Y + W_\varepsilon)$ for every $a \in \mathbb{R}$, where the inequalities are strict as Z, W are not identically zero.

Thus, the pair $X + Z_\varepsilon$ and $Y + W_\varepsilon$ satisfy the assumptions in Lemma D.1. We can then find an independent random variable $V \in L_M$ (depending on ε), such that

$$X + Z_\varepsilon + V \geq_1 Y + W_\varepsilon + V.$$

Monotonicity and additivity of Φ then imply $\Phi(X) + \Phi(Z_\varepsilon) \geq \Phi(Y) + \Phi(W_\varepsilon)$, after canceling out $\Phi(V)$. The desired result $\Phi(X) \geq \Phi(Y)$ follows from the lemma below, which shows that our perturbations only slightly affect the statistic value. *Q.E.D.*

LEMMA D.6: *For any $Z \in L_M$ with $Z \geq 0$, it holds that $\Phi(Z_\varepsilon) \to 0$ as $\varepsilon \to 0$. Similarly, $\Phi(W_\varepsilon) \to 0$ for any $W \in L_M$ with $W \leq 0$.*

PROOF: We focus on the case for Z_ε. Suppose for contradiction that $\Phi(Z_\varepsilon)$ does not converge to zero. Note that as ε decreases, Z_ε decreases in first-order stochastic dominance. So $\Phi(Z_\varepsilon) \geq 0$ also decreases, and non-convergence must imply there exists some $\delta > 0$ such that $\Phi(Z_\varepsilon) > \delta$ for every $\varepsilon > 0$. Let μ_ε be image measure of Z_ε. We now choose a sequence ε_n that decreases to zero very fast, and consider the measures

$$\nu_n = \mu_{\varepsilon_n}^{*n},$$

which is the nth convolution power of μ_{ε_n}. Thus, the sum of n i.i.d. copies of Z_{ε_n} is a random variable whose image measure is ν_n. We denote this sum by U_n.

For each n, we choose ε_n sufficiently small to satisfy two properties: (i) $\varepsilon_n \leq \frac{1}{n^2}$, and (ii) it holds that

$$\mathbb{E}\big[\mathrm{e}^{nU_n} - 1\big] \leq 2^{-n}.$$

This latter inequality can be achieved because $\mathbb{E}[\mathrm{e}^{nU_n}] = (\mathbb{E}[\mathrm{e}^{nZ_{\varepsilon_n}}])^n$, and as $\varepsilon_n \to 0$ we also have $\mathbb{E}[\mathrm{e}^{nZ_{\varepsilon_n}}] = 1 - \varepsilon_n + \varepsilon_n \mathbb{E}[\mathrm{e}^{nZ}] \to 1$ since $Z \in L_M$.

For these choices of ε_n and corresponding U_n, let $H_n(x)$ denote the c.d.f. of U_n, and define $H(x) = \inf_n H_n(x)$ for each $x \in \mathbb{R}$. Since $H_n(x) = 0$ for $x < 0$, the same is true for $H(x)$. Also note that each $H_n(x)$ is a non-decreasing and right-continuous function in x, and so is $H(x)$.

We claim that $\lim_{x\to\infty} H(x) = 1$. Indeed, recall that U_n is the n-fold sum of Z_{ε_n}, which has mass $1 - \varepsilon_n$ at zero. So U_n has mass at least $(1 - \varepsilon_n)^n \geq (1 - \frac{1}{n^2})^n \geq 1 - \frac{1}{n}$ at zero. In other words, $H_n(0) \geq 1 - \frac{1}{n}$. By considering the finitely many c.d.f.s $H_1(x), H_2(x), \ldots, H_{n-1}(x)$, we can find N such that $H_i(x) \geq 1 - \frac{1}{n}$ for every $i < n$ and $x \geq N$. Together with $H_i(x) \geq H_i(0) \geq 1 - \frac{1}{i} \geq 1 - \frac{1}{n}$ for $i \geq n$, we conclude that $H_i(x) \geq 1 - \frac{1}{n}$ whenever $x \geq N$, and so $H(x) \geq 1 - \frac{1}{n}$. Since n is arbitrary, the claim follows. The fact that $H_n(x) \geq 1 - \frac{1}{n}$ also shows that in the definition $H(x) = \inf_n H_n(x)$, the "inf" is actually achieved as the minimum.

These properties of $H(x)$ imply that it is the c.d.f. of some non-negative random variable U. We next show $U \in L_M$, that is, $\mathbb{E}[\mathrm{e}^{aU}] < \infty$ for every $a \in \mathbb{R}$. Since $U \geq 0$, we only need to consider $a \geq 0$. To do this, we take advantage of the following identity based on integration by parts:

$$\mathbb{E}\big[\mathrm{e}^{aU_n} - 1\big] = -\int_{x\geq 0} \big(\mathrm{e}^{ax} - 1\big)\,\mathrm{d}\big(1 - H_n(x)\big) = a \int_{x\geq 0} \mathrm{e}^{ax}\big(1 - H_n(x)\big)\,\mathrm{d}x.$$

Now recall that we chose U_n so that $\mathbb{E}[\mathrm{e}^{nU_n} - 1] \leq 2^{-n}$. So $\mathbb{E}[\mathrm{e}^{aU_n} - 1] \leq 2^{-n}$ for every positive integer $n \geq a$. It follows that the sum $\sum_{n=1}^{\infty} \mathbb{E}[\mathrm{e}^{aU_n} - 1]$ is finite for every $a \geq 0$. Using the above identity, we deduce that

$$a \int_{x\geq 0} \mathrm{e}^{ax} \sum_{n=1}^{\infty} \big(1 - H_n(x)\big)\,\mathrm{d}x < \infty,$$

where we have switched the order of summation and integration by the Monotone Convergence theorem. Since $H(x) = \min_n H_n(x)$, it holds that $1 - H(x) \leq \sum_{n=1}^{\infty}(1 - H_n(x))$ for every x. And thus

$$\mathbb{E}\big[\mathrm{e}^{aU} - 1\big] = a \int_{x\geq 0} \mathrm{e}^{ax}\big(1 - H(x)\big)\,\mathrm{d}x < \infty$$

also holds. This proves $U \in L_M$.

We are finally in a position to deduce a contradiction. Since by construction the c.d.f. of U is no larger than the c.d.f. of each U_n, we have $U \geq_1 U_n$ and $\Phi(U) \geq \Phi(U_n)$ by monotonicity of Φ. But $\Phi(U_n) = n\Phi(Z_{\varepsilon_n}) > n\delta$ by additivity, so this leads to $\Phi(U)$ being infinite. This contradiction proves the desired result. *Q.E.D.*

D.4. *Step 4: Functional Analysis*

To complete the proof of Theorem 2, we also need to modify the functional analysis step in our earlier proof of Theorem 1. One difficulty is that for an unbounded random variable X, $K_a(X)$ takes the value ∞ as $a \to \infty$. Thus, we can no longer think of $K_X(a) = K_a(X)$ as a real-valued continuous function on $\overline{\mathbb{R}}$.

We remedy this as follows. Note first that if Φ is a monotone additive statistic defined on L_M, then it is also monotone and additive when restricted to the smaller domain of bounded random variables. Thus, Theorem 1 gives a probability measure μ on $\mathbb{R} \cup \{\pm\infty\}$ such that

$$\Phi(X) = \int_{\overline{\mathbb{R}}} K_a(X)\,\mathrm{d}\mu(a)$$

for all $X \in L^{\infty}$. In what follows, μ is fixed. We just need to show that this representation also holds for $X \in L_M$.

As a first step, we show μ does not put any mass on $\pm\infty$. Indeed, if $\mu(\{\infty\}) = \varepsilon > 0$, then for any bounded random variable $X \geq 0$, the above integral gives $\Phi(X) \geq \varepsilon \cdot \max[X]$. Take any $Y \in L_M$ such that $Y \geq 0$ and Y is unbounded from above. Then monotonicity of Φ gives $\Phi(Y) \geq \Phi(\min\{Y, n\}) \geq \varepsilon \cdot n$ for each n. This contradicts $\Phi(Y)$ being finite. Similarly, we can rule out any mass at $-\infty$.

The next lemma gives a way to extend the representation to certain unbounded random variables.

LEMMA D.7: *Suppose* $Z \in L_M$ *is bounded from below by* 1 *and unbounded from above, while* $Y \in L_M$ *is bounded from below and satisfies* $\lim_{a\to\infty} \frac{K_a(Y)}{K_a(Z)} = 0$; *then*,

$$\Phi(Y) = \int_{(-\infty,\infty)} K_a(Y)\, \mathrm{d}\mu(a).$$

PROOF: Given the assumptions, $K_a(Z) \geq 1$ for all $a \in \mathbb{R}$, with $\lim_{a\to\infty} K_a(Z) = \infty$. Let L_M^Z be the collection of random variables $X \in L_M$ such that X is bounded from below, and $\lim_{a\to\infty} \frac{K_a(X)}{K_a(Z)}$ exists and is finite. L_M^Z includes all bounded X (in which case $\lim_{a\to\infty} \frac{K_a(X)}{K_a(Z)} = 0$), as well as Y and Z itself. L_M^Z is also closed under adding independent random variables.

Now, for each $X \in L_M^Z$, we can define

$$K_{X|Z}(a) = \frac{K_a(X)}{K_a(Z)},$$

which reduces to our previous definition of $K_X(a)$ when Z is the constant 1. This function $K_{X|Z}(a)$ extends by continuity to $a = -\infty$, where its value is $\frac{\min[X]}{\min[Z]}$, as well as to $a = \infty$ by definition of L_M^Z. Thus, $K_{X|Z}(\cdot)$ is a continuous function on $\overline{\mathbb{R}}$.

Since Φ induces an additive statistic when restricted to L_M^Z, and $K_{X|Z} + K_{Y|Z} = K_{X+Y|Z}$, we have an additive functional F defined on $\mathcal{L} = \{K_{X|Z} : X \in L_M^Z\}$, given by

$$F(K_{X|Z}) = \frac{\Phi(X)}{\Phi(Z)}.$$

Because $Z \geq 1$ implies $\Phi(Z) \geq 1$, F is well-defined, and $F(1) = 1$. By Lemma D.5, F is also monotone in the sense that $K_{X|Z}(a) \geq K_{Y|Z}(a)$ for each $a \in \mathbb{R}$ implies $F(K_{X|Z}) \geq F(K_{Y|Z})$.

Likewise, we can show F is 1-Lipschitz. Note that $K_{X|Z}(a) \leq K_{Y|Z}(a) + \frac{m}{n}$ is equivalent to $K_a(X) \leq K_a(Y) + \frac{m}{n} K_a(Z)$ and equivalent to $K_a(X^{*n}) \leq K_a(Y^{*n} + Z^{*m})$, where we use the notation X^{*n} to denote the sum of n i.i.d. copies of X. If this holds for all a, then by Lemma D.5, we also have $\Phi(X^{*n}) \leq \Phi(Y^{*n} + Z^{*m})$, and thus $\Phi(X) \leq \Phi(Y) + \frac{m}{n}\Phi(Z)$ by additivity. An approximation argument shows that for any *real* number $\varepsilon > 0$, $K_{X|Z}(a) \leq K_{Y|Z}(a) + \varepsilon$ for all a implies $\Phi(X) \leq \Phi(Y) + \varepsilon\Phi(Z)$. Thus, the functional F is 1-Lipschitz.

Given these properties, we can exactly follow the proof of Theorem 1 to extend the functional F to be a positive linear functional on the space of all continuous functions over $\overline{\mathbb{R}}$ (the majorization condition is again satisfied by constant functions, as $K_{Z|Z} = 1$). Therefore, by the Riesz Representation theorem, we obtain a probability measure μ_Z on $\overline{\mathbb{R}}$ such that, for all $X \in L_M^Z$,

$$\frac{\Phi(X)}{\Phi(Z)} = \int_{\overline{\mathbb{R}}} \frac{K_a(X)}{K_a(Z)}\, \mathrm{d}\mu_Z(a).$$

In particular, for any X bounded from below such that $\lim_{a\to\infty} \frac{K_a(X)}{K_a(Z)} = 0$, it holds that

$$\Phi(X) = \int_{[-\infty,\infty)} K_a(X) \cdot \frac{\Phi(Z)}{K_a(Z)}\, \mathrm{d}\mu_Z(a),$$

where we are able to exclude ∞ from the range of integration (this is useful below).

If we define the measure $\hat{\mu}_Z$ by $\frac{\mathrm{d}\hat{\mu}_Z}{\mathrm{d}\mu_Z}(a) = \frac{\Phi(Z)}{K_a(Z)} \leq \Phi(Z)$, then since $K_a(X)$ is finite for $a < \infty$, we have

$$\Phi(X) = \int_{[-\infty,\infty)} K_a(X)\,\mathrm{d}\hat{\mu}_Z(a).$$

This in particular holds for all bounded X, so plugging in $X = 1$ gives that $\hat{\mu}_Z$ is a probability measure. But now we have two probability measures μ and $\hat{\mu}_Z$ on $\overline{\mathbb{R}}$ that lead to the same integral representation for bounded random variables, so Lemma 5 implies that $\hat{\mu}_Z$ coincides with μ and is supported on the standard real line. Plugging in $X = Y$ in the above display then yields the desired result. *Q.E.D.*

The next lemma further extends the representation:

LEMMA D.8: *For every $X \in L_M$ that is bounded from below,*

$$\Phi(X) = \int_{(-\infty,\infty)} K_a(X)\,\mathrm{d}\mu(a).$$

PROOF: It suffices to consider X that is unbounded from above. Moreover, without loss we can assume $X \geq 0$, since we can add any constant to X. Given the previous lemma, we just need to construct $Z \geq 1$ such that $\lim_{a\to\infty} \frac{K_a(X)}{K_a(Z)} = 0$. Note that $\mathbb{E}[\mathrm{e}^{aX}]$ strictly increases in a for $a \geq 0$. This means we can uniquely define a sequence $a_1 < a_2 < \cdots$ by the equation $\mathbb{E}[\mathrm{e}^{a_n X}] = \mathrm{e}^n$. This sequence diverges as $n \to \infty$. We then choose any increasing sequence b_n such that $b_n > n$ and $a_n b_n > 2n^2$.

Consider the random variable Z that is equal to b_n with probability $\mathrm{e}^{-\frac{a_n b_n}{2}}$ for each n, and equal to 1 with remaining probability. To see that $Z \in L_M$, we have

$$\mathbb{E}\big[\mathrm{e}^{aZ}\big] \leq \mathrm{e}^a + \sum_{n=1}^{\infty} \mathrm{e}^{-\frac{a_n b_n}{2}} \cdot \mathrm{e}^{a b_n} = \mathrm{e}^a + \sum_{n=1}^{\infty} \mathrm{e}^{(a-\frac{a_n}{2})\cdot b_n}.$$

For any fixed a, $\frac{a_n}{2}$ is eventually greater than $a+1$. This, together with the fact that $b_n > n$, implies the above sum converges.

Moreover, for any $a \in [a_n, a_{n+1})$, we have

$$\mathbb{E}\big[\mathrm{e}^{aZ}\big] \geq \mathbb{E}\big[\mathrm{e}^{a_n Z}\big] \geq \mathbb{P}[Z = b_n] \cdot \mathrm{e}^{a_n b_n} \geq \mathrm{e}^{\frac{a_n b_n}{2}} > \mathrm{e}^{n^2},$$

whereas $\mathbb{E}[\mathrm{e}^{aX}] \leq \mathbb{E}[\mathrm{e}^{a_{n+1}X}] \leq \mathrm{e}^{n+1}$. Thus,

$$\frac{K_a(X)}{K_a(Z)} = \frac{\log \mathbb{E}\big[\mathrm{e}^{aX}\big]}{\log \mathbb{E}\big[\mathrm{e}^{aZ}\big]} \leq \frac{n+1}{n^2},$$

which converges to zero as a (and thus n) approaches infinity. *Q.E.D.*

D.5. *Step 5: Wrapping up*

By a symmetric argument, the representation $\Phi(X) = \int_{(-\infty,\infty)} K_a(X)\,\mathrm{d}\mu(a)$ also holds for all X bounded from above. In the remainder of the proof, we will use an approximation argument to generalize this to all $X \in L_M$. We first show a technical lemma:

LEMMA D.9: *The measure μ is supported on a compact interval of* $\mathbb{R}$.

PROOF: Suppose not, and without loss assume the support of μ is unbounded from above. We will construct a non-negative $Y \in L_M$ such that $\Phi(Y) = \infty$ according to the integral representation. Indeed, by assumption, we can find a sequence $2 < a_1 < a_2 < \cdots$ such that $a_n \to \infty$ and $\mu([a_n, \infty)) \geq \frac{1}{n}$ for all large n. Let Y be the random variable that equals n with probability $\mathrm{e}^{-\frac{a_n \cdot n}{2}}$ for each n, and equals 0 with remaining probability. Then, similarly to the above, we can show $Y \in L_M$. Moreover, $\mathbb{E}[\mathrm{e}^{a_n Y}] \geq \mathrm{e}^{\frac{a_n \cdot n}{2}}$, implying that $K_{a_n}(Y) \geq \frac{n}{2}$. Since $K_a(Y)$ is increasing in a, we deduce that for each n,

$$\int_{[a_n,\infty)} K_a(Y)\,\mathrm{d}\mu(a) \geq K_{a_n}(Y) \cdot \mu\big([a_n, \infty)\big) \geq \frac{n}{2} \cdot \frac{1}{n} = \frac{1}{2}.$$

The fact that this holds for $a_n \to \infty$ contradicts the assumption that $\Phi(Y) = \int_{(-\infty,\infty)} K_a(Y)\,\mathrm{d}\mu(a)$ is finite. *Q.E.D.*

Thus, we can take N sufficiently large so that μ is supported on $[-N, N]$. To finish the proof, consider any $X \in L_M$ that may be unbounded on both sides. For each positive integer n, let $X_n = \min\{X, n\}$ denote the truncation of X at n. Since $X \geq_1 X_n$, we have

$$\Phi(X) \geq \Phi(X_n) = \int_{[-N,N]} K_a(X_n)\,\mathrm{d}\mu(a).$$

Observe that for each $a \in [-N, N]$, $K_a(X_n)$ converges to $K_a(X)$ as $n \to \infty$. Moreover, the fact that $K_a(X_n)$ increases both in n and in a implies that for all a and all n,

$$\begin{aligned}\big|K_a(X_n)\big| &\leq \max\big\{\big|K_a(X_1)\big|, \big|K_a(X)\big|\big\} \\ &\leq \max\big\{\big|K_{-N}(X_1)\big|, \big|K_N(X_1)\big|, \big|K_{-N}(X)\big|, \big|K_N(X)\big|\big\}.\end{aligned}$$

As $K_a(X_n)$ is uniformly bounded, we can apply the Dominated Convergence theorem to deduce

$$\Phi(X) \geq \lim_{n\to\infty} \int_{[-N,N]} K_a(X_n)\,\mathrm{d}\mu(a) = \int_{[-N,N]} K_a(X)\,\mathrm{d}\mu(a).$$

On the other hand, if we truncate the left tail and consider $X^{-n} = \max\{X, -n\}$, then a symmetric argument shows

$$\Phi(X) \leq \lim_{n\to\infty} \int_{[-N,N]} K_a\big(X^{-n}\big)\,\mathrm{d}\mu(a) = \int_{[-N,N]} K_a(X)\,\mathrm{d}\mu(a).$$

Therefore, for all $X \in L_M$, it holds that

$$\Phi(X) = \int_{[-N,N]} K_a(X)\,\mathrm{d}\mu(a).$$

This completes the entire proof of Theorem 2.

APPENDIX E: OMITTED PROOFS FOR SECTION 4

E.1. *Proof of Proposition 5*

The result can be derived as a corollary of Proposition 6 which we prove below, but we also provide a direct proof here. We focus on the "only if" direction because the "if" direction follows immediately from the monotonicity of $K_a(X)$ in a. Suppose μ is not supported on $[-\infty, 0]$; we will show that the resulting monotone additive statistic Φ does not always exhibit risk aversion. Since μ has positive mass on $(0, \infty]$, we can find $\varepsilon > 0$ such that μ assigns mass at least ε to $(\varepsilon, \infty]$. Now consider a gamble X which is equal to 0 with probability $\frac{n-1}{n}$ and equal to n with probability $\frac{1}{n}$, for some large positive integer n. Then $\mathbb{E}[X] = 1$ and $K_a(X) \geq \min[X] = 0$ for every $a \in \overline{\mathbb{R}}$. Moreover, for $a \geq \varepsilon$, we have

$$K_a(X) \geq K_\varepsilon(X) = \frac{1}{\varepsilon}\log\left(\frac{n-1}{n} + \frac{1}{n}\mathrm{e}^{\varepsilon n}\right) \geq \frac{n}{2}$$

whenever n is sufficient large. Thus,

$$\Phi(X) = \int_{\overline{\mathbb{R}}} K_a(X)\,\mathrm{d}\mu(a) \geq \int_{[\varepsilon,\infty]} K_a(X)\,\mathrm{d}\mu(a) \geq \frac{n}{2}\varepsilon.$$

We thus have $\Phi(X) > 1 = \mathbb{E}[X]$ for all large n, showing that the preference represented by Φ sometimes exhibits risk-seeking.

Symmetrically, if μ is not supported on $[0, \infty]$, then Φ must sometimes exhibit risk aversion (by considering X equal to 0 with probability $\frac{1}{n}$ and equal to n with probability $\frac{n-1}{n}$). This completes the proof.

E.2. *Proof of Proposition 6*

We first show that conditions (i) and (ii) are necessary for $\int_{\overline{\mathbb{R}}} K_a(X)\,\mathrm{d}\mu_1(a) \leq \int_{\overline{\mathbb{R}}} K_a(Y)\,\mathrm{d}\mu_2(a)$ to hold for every X. This part of the argument closely follows the proof of Lemma 5. Specifically, by considering the same random variables $X_{n,b}$ as defined there, we have the key equation (11). Since the limit on the left-hand side is smaller for μ_1 than for μ_2, we conclude that for every $b > 0$, $\int_{[b,\infty]} \frac{a-b}{a}\,\mathrm{d}\mu_1(a)$ on the right-hand side must be smaller than the corresponding integral for μ_2. Thus, condition (i) holds, and an analogous argument shows condition (ii) also holds.

To complete the proof, it remains to show that when conditions (i) and (ii) are satisfied,

$$\int_{\overline{\mathbb{R}}} K_a(X)\,\mathrm{d}\mu_1(a) \leq \int_{\overline{\mathbb{R}}} K_a(X)\,\mathrm{d}\mu_2(a)$$

holds for every X. Since μ_1 and μ_2 are both probability measures, we can subtract $\mathbb{E}[X]$ from both sides and arrive at the equivalent inequality

$$\int_{\overline{\mathbb{R}}_{\neq 0}} \big(K_a(X) - \mathbb{E}[X]\big)\,\mathrm{d}\mu_1(a) \leq \int_{\overline{\mathbb{R}}_{\neq 0}} \big(K_a(X) - \mathbb{E}[X]\big)\,\mathrm{d}\mu_2(a). \tag{S3}$$

Note that we can exclude $a = 0$ from the range of integration because $K_a(X) = \mathbb{E}[X]$ there. Below, we show that condition (i) implies

$$\int_{(0,\infty]} \big(K_a(X) - \mathbb{E}[X]\big)\,\mathrm{d}\mu_1(a) \leq \int_{(0,\infty]} \big(K_a(X) - \mathbb{E}[X]\big)\,\mathrm{d}\mu_2(a). \tag{S4}$$

Similarly, condition (ii) gives the same inequality when the range of integration is $[-\infty, 0)$. Adding these two inequalities would yield the desired comparison in (S3).

To prove (S4), we let $L_X(a) = a \cdot K_a(X) = \log \mathbb{E}[\mathrm{e}^{aX}]$ be the cumulant generating function of X. It is well known that $L_X(a)$ is convex in a, with $L'_X(0) = \mathbb{E}[X]$ and $\lim_{a\to\infty} L'_X(a) = \max[X]$. Then the integral on the left-hand side of (S4) can be calculated as follows:

$$\begin{aligned}
&\int_{(0,\infty]} \big(K_a(X) - \mathbb{E}[X]\big)\,\mathrm{d}\mu_1(a) \\
&\quad = \int_{(0,\infty)} \big(K_a(X) - \mathbb{E}[X]\big)\,\mathrm{d}\mu_1(a) + \big(\max[X] - \mathbb{E}[X]\big) \cdot \mu_1\big(\{\infty\}\big) \\
&\quad = \int_{(0,\infty)} \big(L_X(a) - a\mathbb{E}[X]\big)\,\mathrm{d}\frac{\mu_1(a)}{a} + \big(\max[X] - \mathbb{E}[X]\big) \cdot \mu_1\big(\{\infty\}\big).
\end{aligned}$$

Note that since the function $g(a) = L_X(a) - a\mathbb{E}[X]$ satisfies $g(0) = g'(0) = 0$, it can be written as

$$g(a) = \int_0^a g'(t)\,\mathrm{d}t = \int_0^a \int_0^t g''(b)\,\mathrm{d}b\,\mathrm{d}t = \int_0^a g''(b) \cdot (a-b)\,\mathrm{d}b.$$

Plugging back to the previous identity, we obtain

$$\begin{aligned}
&\int_{(0,\infty]} \big(K_a(X) - \mathbb{E}[X]\big)\,\mathrm{d}\mu_1(a) \\
&\quad = \int_{(0,\infty)} \int_0^a L''_X(b) \cdot (a-b)\,\mathrm{d}b\,\mathrm{d}\frac{\mu_1(a)}{a} + \big(\max[X] - \mathbb{E}[X]\big) \cdot \mu_1\big(\{\infty\}\big) \\
&\quad = \int_0^\infty L''_X(b) \int_{[b,\infty)} (a-b)\,\mathrm{d}\frac{\mu_1(a)}{a}\,\mathrm{d}b + \big(L'_X(\infty) - L'_X(0)\big) \cdot \mu_1\big(\{\infty\}\big) \\
&\quad = \int_0^\infty L''_X(b) \int_{[b,\infty)} \frac{a-b}{a}\,\mathrm{d}\mu_1(a)\,\mathrm{d}b + \int_0^\infty L''_X(b) \cdot \mu_1\big(\{\infty\}\big)\,\mathrm{d}b \\
&\quad = \int_0^\infty L''_X(b) \int_{[b,\infty]} \frac{a-b}{a}\,\mathrm{d}\mu_1(a)\,\mathrm{d}b,
\end{aligned}$$

where the last step uses $\frac{a-b}{a} = 1$ when $a = \infty > b$.

The above identity also holds when μ_1 is replaced by μ_2. We then see that (S4) follows from condition (i) and $L''_X(b) \geq 0$ for all b. This completes the proof.

E.3. *Proof of Theorem 5*

The "if" direction is straightforward: if $\succeq_1$ and $\succeq_2$ are both represented by a monotone additive statistic Φ, then they satisfy responsiveness and continuity. In addition, combined choices are not stochastically dominated because if $X \succ_1 X'$ and $Y \succ_2 Y'$, then $\Phi(X) > \Phi(X')$ and $\Phi(Y) > \Phi(Y')$. Thus, $\Phi(X+Y) > \Phi(X'+Y')$ and $X'+Y'$ cannot stochastically dominate $X+Y$.

Turning to the "only if" direction, we suppose $\succeq_1$ and $\succeq_2$ satisfy the axioms. We first show that these preferences are the same. Suppose for the sake of contradiction that

$X \succeq_1 Y$ but $Y \succ_2 X$ for some X, Y. Then, by continuity, there exists $\varepsilon > 0$ such that $Y \succ_2 X + \varepsilon$. By responsiveness, we also have $X \succeq_1 Y \succ Y - \frac{\varepsilon}{2}$. Thus, $X \succ_1 Y - \frac{\varepsilon}{2}$, $Y \succ_2 X + \varepsilon$, but $X + Y$ is strictly stochastically dominated by $Y - \frac{\varepsilon}{2} + X + \varepsilon = X + Y + \frac{\varepsilon}{2}$, contradicting Axiom 7.

Henceforth, we denote both $\succeq_1$ and $\succeq_2$ by $\succeq$. We next show that for any X and any $\varepsilon > 0$, $\max[X] + \varepsilon \succ X \succ \min[X] - \varepsilon$. To see why, suppose for contradiction that X is weakly preferred to $\max[X] + \varepsilon$ (the other case can be handled similarly). Then we obtain a contradiction to Axiom 7 by observing that $X \succ \max[X] + \frac{\varepsilon}{2}$, $\frac{\varepsilon}{4} \succ 0$ but $X + \frac{\varepsilon}{4} <_1 \max[X] + \frac{\varepsilon}{2} + 0$.

Given these upper and lower bounds for X, we can define $\Phi(X) = \sup\{c \in \mathbb{R} : c \preceq X\}$, which is well-defined and finite. By definition of the supremum and responsiveness, for any $\varepsilon > 0$ it holds that $\Phi(X) - \varepsilon \prec X \prec \Phi(X) + \varepsilon$. Thus, by continuity, $\Phi(X) \sim X$ is the (unique) certainty equivalent of X.

It remains to show that Φ is a monotone additive statistic. For this, we show that $X \sim Y$ implies $X + Z \sim Y + Z$ for any independent Z. Suppose for contradiction that $X + Z \succ Y + Z$. Then, by continuity, we can find $\varepsilon > 0$ such that $X + Z \succ Y + Z + \varepsilon$. By responsiveness, it also holds that $Y + \frac{\varepsilon}{2} \succ Y \sim X$. But the sum $(X + Z) + (Y + \frac{\varepsilon}{2})$ is stochastically dominated by $(Y + Z + \varepsilon) + X$, contradicting Axiom 7.

Therefore, from $X \sim \Phi(X)$ and $Y \sim \Phi(Y)$ we can apply the preceding result twice to obtain $X + Y \sim \Phi(X) + Y \sim \Phi(X) + \Phi(Y)$ whenever X, Y are independent, so that $\Phi(X + Y) = \Phi(X) + \Phi(Y)$ is additive. Finally, we show Φ is monotone. Consider any $Y \geq_1 X$, and suppose for contradiction that $X \succ Y$. Then there exists $\varepsilon > 0$ such that $X \succ Y + \varepsilon$. This leads to a contradiction since $X \succ Y + \varepsilon$, $\frac{\varepsilon}{2} \succ 0$, but $X + \frac{\varepsilon}{2}$ is stochastically dominated by $Y + \varepsilon + 0$.

This completes the proof that both preferences $\succeq_1$ and $\succeq_2$ are represented by the same certainty equivalent $\Phi(X)$, which is a monotone additive statistic.

APPENDIX F: MONOTONE ADDITIVE STATISTICS AND THE INDEPENDENCE AXIOM

In this Appendix, we discuss the classic independence axiom and what it implies for preferences represented by monotone additive statistics.

AXIOM F.1—Independence: *For all X, Y, Z and all $\lambda \in (0, 1)$, $X \succeq Y$ implies $X_\lambda Z \succeq Y_\lambda Z$.*

PROPOSITION F.1: *Suppose a preference $\succeq$ is represented by a monotone additive statistic $\Phi(X) = \int_{\overline{\mathbb{R}}} K_a(X)\, \mathrm{d}\mu(a)$. Then $\succeq$ satisfies the independence axiom if and only if μ is a point mass at some $a \in \overline{\mathbb{R}}$.*

PROOF: The "if" direction is relatively straightforward. If $a = 0$, then $\Phi(X) = \mathbb{E}[X]$. In this case, $\mathbb{E}[X] \geq \mathbb{E}[Y]$ does imply

$$\mathbb{E}[X_\lambda Z] = \lambda\mathbb{E}[X] + (1 - \lambda)\mathbb{E}[Z] \geq \lambda\mathbb{E}[Y] + (1 - \lambda)\mathbb{E}[Z] = \mathbb{E}[Y_\lambda Z].$$

If $a > 0$, then $\Phi(X) \geq \Phi(Y)$ implies $\mathbb{E}[\mathrm{e}^{aX}] \geq \mathbb{E}[\mathrm{e}^{aY}]$ and thus

$$\lambda\mathbb{E}\big[\mathrm{e}^{aX}\big] + (1 - \lambda)\mathbb{E}\big[\mathrm{e}^{aZ}\big] \geq \lambda\mathbb{E}\big[\mathrm{e}^{aY}\big] + (1 - \lambda)\mathbb{E}\big[\mathrm{e}^{aZ}\big],$$

so that $\Phi(X_\lambda Z) \geq \Phi(Y_\lambda Z)$. A similar argument applies to the case of $a < 0$. Finally, it is easy to see that $\max[X] \geq \max[Y]$ implies $\max[X_\lambda Z] \geq \max[Y_\lambda Z]$ and the same holds for the minimum. So the above independence axiom holds for $a = \pm\infty$ as well.[1]

We turn to the "only if" direction of the result. By the independence axiom, whenever c is a constant we have $X \succeq c$ implies $X_\lambda c \succeq c$ and $c \succeq X$ implies $c \succeq X_\lambda c$. Therefore, $X \sim c$ implies $X_\lambda c \sim c$, which allows us to directly apply Lemma 7 from before. It remains to show that independence rules out $\Phi(X) = \beta K_{-a\beta}(X) + (1-\beta)K_{a(1-\beta)}(X)$ for some $\beta \in (0,1)$ and $a \in (0,\infty]$.

Suppose Φ takes the above form. If $a = \infty$, then $\Phi(X) = \beta \min[X] + (1-\beta)\max[X]$ for some $\beta \in (0,1)$. To see that it violates independence, choose X supported on 0 and $\frac{1}{1-\beta}$, and $Y = 1$ so that $\Phi(X) = \Phi(Y)$. But with Z being a sufficiently large constant, we see that $X_\lambda Z$ has the same maximum as $Y_\lambda Z$, but a strictly smaller minimum. Hence, $\Phi(X_\lambda Z) < \Phi(Y_\lambda Z)$, contradicting independence.

If instead $a \in (0,\infty)$, then we can do a similar construction by choosing X and Y such that $\Phi(X) > \Phi(Y)$ but $K_{-a\beta}(X) < K_{-a\beta}(Y)$. For example, let $Y = 1$, and let X be supported on $\{0, k\}$, with $\mathbb{P}[X = k] = \frac{1}{k}$. Then

$$K_b(X) = \frac{1}{b}\log \mathbb{E}\left[1 - \frac{1}{k} + \frac{e^{bk}}{k}\right].$$

For k tending to infinity, $K_b(X)$ tends to zero if $b < 0$, and to infinity if $b > 0$. Hence, for k large enough, X and Y will have the desired property.

Now let $Z = n$ where n is a large positive integer. Then

$$K_b(Y_\lambda n) = \frac{1}{b}\log \mathbb{E}\big[\lambda \mathbb{E}\big[e^{bY}\big] + (1-\lambda)e^{bn}\big],$$

$$K_b(X_\lambda n) = \frac{1}{b}\log \mathbb{E}\big[\lambda \mathbb{E}\big[e^{bX}\big] + (1-\lambda)e^{bn}\big],$$

and so

$$K_b(Y_\lambda n) - K_b(X_\lambda n) = \frac{1}{b}\log\left(\frac{\lambda \mathbb{E}\big[e^{bY}\big] + (1-\lambda)e^{bn}}{\lambda \mathbb{E}\big[e^{bX}\big] + (1-\lambda)e^{bn}}\right).$$

It easily follows that for fixed $\lambda \in (0,1)$ and b,

$$\lim_{n\to\infty} K_b(Y_\lambda n) - K_b(X_\lambda n) = 0 \quad \text{if } b > 0;$$

$$\lim_{n\to\infty} K_b(Y_\lambda n) - K_b(X_\lambda n) = K_b(Y) - K_b(X) \quad \text{if } b < 0.$$

Thus, as n tends to infinity,

$$\begin{aligned}
&\lim_n \Phi(Y_\lambda n) - \Phi(X_\lambda n) \\
&\quad = \lim_n \beta\big[K_{-a\beta}(Y_\lambda n) - K_{-a\beta}(X_\lambda n)\big] + (1-\beta)\big[K_{a(1-\beta)}(Y_\lambda n) - K_{a(1-\beta)}(X_\lambda n)\big] \\
&\quad = \beta\big[K_{-a\beta}(Y_\lambda n) - K_{-a\beta}(X_\lambda n)\big] > 0.
\end{aligned}$$

[1]Note, however, that $\Phi(X) = \max[X]$ or $\min[X]$ would violate a stronger form of independence that additionally requires $X \succ Y$ to imply $X_\lambda Z \succ Y_\lambda Z$ with strict preferences. This is related to the fact that these extreme monotone additive statistics do not satisfy mixture continuity.

Therefore, for n large enough, we have found X and Y such that $\Phi(X) > \Phi(Y)$ but $\Phi(X_\lambda n) < \Phi(Y_\lambda n)$. This implies $X \succ Y$ but $X_\lambda n \prec Y_\lambda n$, which contradicts the independence axiom and completes the proof of Proposition F.1. *Q.E.D.*

F.1. *Proof of Proposition 1*

We now prove Proposition 1 as a corollary of Proposition F.1. The first observation is that, under time invariance, strong stochastic dynamic consistency is equivalent to the following property of the preference $\succeq$:

AXIOM F.2—Strong Stochastic Stationarity: *For every pair of time lotteries* (x, T), (y, S) *and every* $D \in L^\infty_+$ *not necessarily independent, if* $(x, T_d) \succeq (y, S_d)$ *for almost every realization* d *of* D, *then* $(x, T + D) \succeq (y, S + D)$.

Indeed, suppose strong stochastic dynamic consistency is satisfied, and $(x, T_d) \succeq (y, S_d)$ holds for almost every realization d of D. Then, by time invariance, $(x, T_d) \succeq_{t+d} (y, S_d)$ also holds for almost every d. Strong stochastic dynamic consistency thus implies $(x, T + D) \succeq_t (y, S + D)$ and therefore strong stochastic stationarity. A similar argument shows that conversely, strong stochastic stationarity also implies strong stochastic dynamic consistency.

For the "only if" direction of Proposition 1, suppose that $\succeq$ is an MSTP that satisfies strong stochastic stationarity. Let $\succeq_*$ denote the preference over random times induced by $\succeq$ when fixing the payoff. That is, $T \succeq_* S$ if and only if $(x, T) \succeq (x, S)$ for any and every $x > 0$.

Fix any $X \succeq_* Y$ and any $Z \in L^\infty_+$, which can be considered as random times. For a given $\lambda \in (0, 1)$, choose D to be a random variable that is equal to either 0 or 1, with probability λ and $1 - \lambda$, respectively. Let $\tilde{X}$ be a random variable that, conditioned on $D = 0$, has the same distribution as $X + 1$, and conditioned on $D = 1$, has the same distribution as Z. Likewise, let $\tilde{Y}$ be a random variable that, conditioned on $D = 0$, has the same distribution as $Y + 1$, and conditioned on $D = 1$, has the same distribution as Z.

By construction, $\tilde{X}_D \succeq_* \tilde{Y}_D$ for every possible value of D, so by strong stochastic stationarity, $\tilde{X} + D \succeq_* \tilde{Y} + D$ must hold. But $\tilde{X} + D$ has the same distribution as $(X_\lambda Z) + 1$ while $\tilde{Y} + D$ has the same distribution as $(Y_\lambda Z) + 1$, so $(X_\lambda Z) + 1 \succeq_* (Y_\lambda Z) + 1$. Since this is an MSTP, we deduce $X_\lambda Z \succeq_* Y_\lambda Z$ as the independence axiom requires.

Note that even though $\succeq_*$ and the associated monotone additive statistic Φ are defined only for non-negative bounded random variables, it can be extended to all of L^∞ as shown in the proof of Proposition 7. Given additivity, it is easy to see that the extension preserves independence. So we can assume $\succeq_*$ and Φ satisfy independence on L^∞. This allows us to apply Proposition F.1 and deduce that Φ must have a point-mass mixing measure μ, which proves the "only if" direction of Proposition 1.

As for the "if" direction, we need to verify that an MSTP represented by $V(x, T) = u(x) \cdot \mathrm{e}^{-rK_a(T)}$ does satisfy strong stochastic stationarity. First consider $a = 0$, in which case the representation simplifies to $u(x) \cdot \mathrm{e}^{-\mathbb{E}[T]}$ with the normalization $r = 1$. If $(x, T_d) \succeq (y, S_d)$ for almost every d, then $u(x) \cdot \mathrm{e}^{-\mathbb{E}[T_d]} \geq u(y) \cdot \mathrm{e}^{-\mathbb{E}[S_d]}$, which can be rewritten as $\mathbb{E}[S_d] - \mathbb{E}[T_d] \geq \log(u(y)/u(x))$. Averaging across different realizations d, this implies $\mathbb{E}[S] - \mathbb{E}[T] \geq \log(u(y)/u(x))$, and thus $\mathbb{E}[S + D] - \mathbb{E}[T + D] \geq \log(u(y)/u(x))$. After rearranging, this yields $u(x) \cdot \mathrm{e}^{-\mathbb{E}[T+D]} \geq u(y) \cdot \mathrm{e}^{-\mathbb{E}[S+D]}$. So $(x, T + D) \succeq (y, S + D)$ as demanded by strong stochastic stationarity.

Next, consider $a > 0$. In this case, we normalize $r = a$ and adjust u accordingly, to arrive at an equivalent representation $V(x, T) = u(x)/\mathbb{E}[\mathrm{e}^{aT}]$. From $(x, T_d) \succeq (y, S_d)$, we obtain $u(x) \cdot \mathbb{E}[\mathrm{e}^{aS_d}] \geq u(y) \cdot \mathbb{E}[\mathrm{e}^{aT_d}]$ and thus

$$u(x) \cdot \mathbb{E}\big[\mathrm{e}^{a(S_d+d)}\big] \geq u(y) \cdot \mathbb{E}\big[\mathrm{e}^{a(T_d+d)}\big].$$

Averaging across different realizations d then yields $u(x) \cdot \mathbb{E}[\mathrm{e}^{a(S+D)}] \geq u(y) \cdot \mathbb{E}[\mathrm{e}^{a(T+D)}]$, which after rearranging gives the desired conclusion $V(x, T + D) \geq V(y, S + D)$.

If instead $a < 0$, then we normalize $r = -a$ and recover the usual EDU representation $V(x, T) = u(x) \cdot \mathbb{E}[\mathrm{e}^{aT}]$. Essentially the same argument as above applies to this case.

Finally, consider $a = \infty$, so that $V(x, T) = u(x) \cdot \mathrm{e}^{-\max[T]}$ after normalizing $r = 1$. In this case, $(x, T_d) \succeq (y, S_d)$ implies $\max[S_d] - \max[T_d] \geq \log(u(y)/u(x))$, and thus

$$\max[S_d + d] - \max[T_d + d] \geq \log\big(u(y)/u(x)\big).$$

Let $\alpha = \max[S + D]$ and $c = \log(u(y)/u(x))$ be constants. Then the above implies that for almost every realization d of D, $T_d + d \leq \alpha - c$. Thus, $T + D \leq \alpha - c$ almost surely, which gives $\max[S + D] - \max[T + D] \geq c$. This implies $V(x, T + D) \geq V(y, S + D)$ as desired.

A similar argument applies to the case of $a = -\infty$, completing the proof.

Co-editor Barton L. Lipman handled this manuscript.

Manuscript received 14 July, 2021; final version accepted 4 April, 2024; available online 9 April, 2024.