

Spectral Risk Sharing

Stephen J. Mildenhall

2026-04-18

This note presents a summary of the spectral risk sharing theorem from Jouini et al. (2008). The theorem describes how two agents with spectral risk cost functions optimally share risk. It provides a wonderful, intuitive, and constructive solution to the problem, involving a horizontal layering of the total loss, with each horizontal layer assumed by the agent with lowest cost of risk at the tail probability level. The proof can be extended to any finite number of agents. First, we provide an elementary proof. Then, we give a detailed sketch of the original proof. The original proof uses convex duality and is much more conceptual than our elementary one, but correspondingly more sophisticated and involved: it yields more for more effort. We follow the original argument closely but translate the notation to use cost functions rather than utilities and the actuarial loss sign convention rather than the asset payoff convention.

1 The Problem

Consider an economy with two agents $i = 0, 1$. Each agent evaluates losses using a spectral distortion risk measure. Specifically, for a risk $X \geq 0$ is a bounded loss, they evaluate the cost of holding X using

$$g_i(X) = \int_0^\infty g_i(S_X(x)) dx.$$

We use the same notation for the distortion and its associated functional and $S_X(x) = P(X > x)$, and we use the actuarial sign convention that larger X is worse.

Let X denote the total loss in the economy. The problem is to minimize the shared cost of bearing X between the two agents:

$$\text{minimize } \rho_0(X_0) + \rho_1(X_1) \quad \text{over } X_0 + X_1 = X,$$

among non-negative X_i . That is, we want the most economical way for the economy to share the risk X between the two actors. In this section we give an elementary proof of a constructive solution to the optimization problem. In general, the solution is not unique.

Proposition 1.1. *With the above notation, let*

$$\begin{aligned} A &= \{x \geq 0 : g_0(S_X(x)) \leq g_1(S_X(x))\}, \\ B &= \{x \geq 0 : g_0(S_X(x)) > g_1(S_X(x))\}, \end{aligned}$$

and define

$$X_0 = \int_0^X A(s) ds, \quad X_1 = \int_0^X B(s) ds.$$

Then (X_0, X_1) is a comonotone Pareto optimal allocation of X , and

$$g_0(X_0) + g_1(X_1) = (g_0 \wedge g_1)(X).$$

Proposition 1.1 is a weaker version of Proposition 3.1 from Jouini et al. (2008). The original version characterizes all comonotone allocations, but our proposition includes the most important parts—existence and flatness—of the original. The ambiguity in our solution lives in $\{g_0 = g_1\}$ where risk can be shared arbitrarily.

2 Elementary Proof

Setting

Let $X \in L_+^\infty$ be the aggregate loss. For $i = 0, 1$, let g_i be a nondecreasing concave distortion on $[0, 1]$, and use the same symbol for the induced Choquet functional

$$g_i(Y) = \int_0^\infty g_i(S_Y(x)) dx, \quad S_Y(x) = \mathbb{P}(Y > x).$$

An allocation of X is a pair (X_0, X_1) with

$$X_0 + X_1 = X.$$

We seek an allocation that is both comonotone and Pareto optimal. The candidate common pricing distortion is the pointwise minima

$$h := g_0 \wedge g_1 = \min\{g_0, g_1\}.$$

Step 1: A Universal Lower Bound

The function h is again a nondecreasing concave distortion. Hence its Choquet functional is subadditive. Therefore, for every allocation (Y_0, Y_1) of X ,

$$h(X) = h(Y_0 + Y_1) \leq h(Y_0) + h(Y_1).$$

Since $h \leq g_0$ and $h \leq g_1$ pointwise, monotonicity with respect to the distortion gives

$$h(Y_0) \leq g_0(Y_0), \quad h(Y_1) \leq g_1(Y_1).$$

Combining the two inequalities,

$$h(X) \leq g_0(Y_0) + g_1(Y_1)$$

for every allocation (Y_0, Y_1) of X . Therefore $h(X)$ is a universal lower bound for the total evaluation.

Step 2: Construct the Allocation by Assigning Layers

Define the layer sets as in Proposition 1.1. Clearly A and B partition $[0, \infty)$. Now define

$$X_0 = \int_0^X A(s) ds, \quad X_1 = \int_0^X B(s) ds.$$

Since $A(s) + B(s) = 1$ for all s ,

$$X_0 + X_1 = \int_0^X 1 ds = X.$$

Also, both X_0 and X_1 are increasing 1-Lipschitz functions of X , and so (X_0, X_1) is comonotone.

The economic meaning is transparent: at each loss level s , we give that marginal layer to the agent whose distortion is cheaper at the corresponding survival probability $S_X(s)$.

Step 3: The Key Layer Lemma

The key technical point is to compute the evaluation of a loss built by selecting layers of X .

Lemma 2.1. *Let $a : [0, \infty) \rightarrow [0, 1]$ be measurable, and define a random variable*

$$Y_a = \int_0^X a(s) ds.$$

Then for every concave distortion g , the functional

$$g(Y_a) = \int_0^\infty a(s) g(S_X(s)) ds.$$

Proof. We build up to general a from simple functions. First take $a = 1_{(u,v]}$. Then

$$Y_a = \int_0^X 1_{(u,v]}(s) ds = (X - u)^+ - (X - v)^+.$$

Equivalently,

$$(X - u)^+ = Y_a + (X - v)^+.$$

All three terms are increasing functions of X , hence comonotone. By comonotone additivity of the Choquet functional,

$$g((X - u)^+) = g(Y_a) + g((X - v)^+).$$

Therefore

$$g(Y_a) = g((X - u)^+) - g((X - v)^+).$$

Now

$$\begin{aligned} g((X - u)^+) &= \int_0^\infty g(\mathbb{P}((X - u)^+ > y)) dy \\ &= \int_0^\infty g(\mathbb{P}(X > u + y)) dy \\ &= \int_u^\infty g(S_X(s)) ds. \end{aligned}$$

Similarly for $(X - v)^+$, so subtracting,

$$g(Y_a) = \int_u^v g(S_X(s)) ds = \int_0^\infty 1_{(u,v]}(s) g(S_X(s)) ds.$$

The next step is a simple selector. If

$$a = \sum_{m=1}^n c_m 1_{(u_m, v_m]}, \quad 0 \leq c_m \leq 1,$$

then

$$Y_a = \sum_{m=1}^n c_m \int_0^X 1_{(u_m, v_m]}(s) ds.$$

Each summand is an increasing function of X , so the summands are comonotone. Using positive homogeneity and comonotone additivity,

$$\begin{aligned} g(Y_a) &= \sum_{m=1}^n c_m g\left(\int_0^X 1_{(u_m, v_m]}(s) ds\right) \\ &= \sum_{m=1}^n c_m \int_{u_m}^{v_m} g(S_X(s)) ds \\ &= \int_0^\infty a(s) g(S_X(s)) ds. \end{aligned}$$

Finally, for general measurable selectors, choose simple functions $a_n \uparrow a$. Then

$$Y_{a_n} \uparrow Y_a.$$

By monotone convergence in the survival representation,

$$g(Y_{a_n}) \uparrow g(Y_a),$$

and also

$$\int_0^\infty a_n(s) g(S_X(s)) ds \uparrow \int_0^\infty a(s) g(S_X(s)) ds.$$

Passing to the limit proves the lemma. □

Step 4: The Common Contact Distortion

Now apply Lemma 2.1 with the selector A . First, using $g = g_0$,

$$g_0(X_0) = \int_0^\infty A(s) g_0(S_X(s)) ds = \int_A g_0(S_X(s)) ds.$$

Next, using $g = h$,

$$h(X_0) = \int_0^\infty A(s) h(S_X(s)) ds.$$

But on A we have by definition

$$h(S_X(s)) = \min \{ g_0(S_X(s)), g_1(S_X(s)) \} = g_0(S_X(s)).$$

Hence

$$h(X_0) = g_0(X_0).$$

Exactly the same argument on B gives

$$g_1(X_1) = \int_B g_1(S_X(s)) ds, \quad h(X_1) = g_1(X_1),$$

because on B we have

$$h(S_X(s)) = g_1(S_X(s)).$$

So the distortion $h = g_0 \wedge g_1$ prices each allocated piece exactly as the corresponding agent does:

$$h(X_0) = g_0(X_0), \quad h(X_1) = g_1(X_1).$$

That is the key contact property. The minimum distortion is not just the right lower bound for the total problem; it also acts as a common pricing rule for the two pieces we construct.

Step 5: The Lower Bound Is Attained

Since X_0 and X_1 are comonotone, comonotone additivity of h gives

$$h(X) = h(X_0 + X_1) = h(X_0) + h(X_1).$$

Using the contact identities from Step 4,

$$h(X) = g_0(X_0) + g_1(X_1).$$

Now compare with the universal lower bound from Step 1. For every allocation (Y_0, Y_1) of X we have

$$h(X) \leq g_0(Y_0) + g_1(Y_1).$$

Hence

$$h(X) \leq \inf_{Y_0+Y_1=X} (g_0(Y_0) + g_1(Y_1)).$$

On the other hand, since (X_0, X_1) is itself an allocation of X ,

$$\inf_{Y_0+Y_1=X} (g_0(Y_0) + g_1(Y_1)) \leq g_0(X_0) + g_1(X_1) = h(X).$$

Therefore both inequalities are equalities and, in particular,

$$g_0(X_0) + g_1(X_1) = h(X) = \inf_{Y_0+Y_1=X} (g_0(Y_0) + g_1(Y_1)).$$

Step 6: Pareto Optimality

A Pareto improvement would lower at least one of g_0 or g_1 without raising the other, and therefore would strictly lower the sum

$$g_0(Y_0) + g_1(Y_1).$$

But our constructed allocation already minimizes that sum over all allocations. Therefore no Pareto improvement exists. Therefore, (X_0, X_1) is Pareto optimal and we have proved Proposition 1.1.

A Remark on Ambiguity

The only ambiguity lies on the tie set

$$T = \{ x \geq 0 : g_0(S_X(x)) = g_1(S_X(x)) \}.$$

Any non-negative measurable split of T between the two agents gives the same total value. So the construction is not unique there, but every such choice is optimal.

A Trivial Representation With Non-Trivial Implications

The representation

$$g(X) = \sup \{ h(X) : h(s) \leq g(s) \ \forall s \in [0, 1], \ h \text{ a distortion} \}$$

is trivial as a value identity, since we can always take $h = g$. Its real use is not the equality itself but that it tells us what it means for a smaller distortion $h < g$ to price a given loss X : h has to agree with g only on the layers of X that actually matter.

For a fixed loss X ,

$$g(X) = \int_0^\infty g(S_X(x)) dx.$$

So if $h \leq g$ yet

$$h(S_X(x)) = g(S_X(x)) \quad \text{for a.e. } x,$$

then automatically

$$h(X) = g(X).$$

Thus contact is only needed along the image of the survival profile $x \mapsto S_X(x)$, and only up to Lebesgue-null sets in x .

That is exactly where the jump issue disappears. If q_X has a flat spot, then S_X has a jump, and the corresponding problematic survival level is just a single value of $S_X(x)$. Changing a distortion at that single value does not affect $g(X)$ because the set of x where $S_X(x)$ equals that exact value has Lebesgue measure zero. The Choquet integral does not see isolated values of the distortion.

The non-trivial implication is that if $h \leq g$ prices X then

$$\{s : h(s) < g(s)\} \subset \{dq_X = 0\}.$$

The right-hand set is values of p where q_X is flat.

In our construction, the distortion $h = g_0 \wedge g_1$ prices X , X_0 and X_1 , $X = X_0 + X_1$ and all three variables are comonotone. Thus on $\{g_0 < g_1\}$ we have $h = g_1 < g_0$. Since h prices X_0 it follows q_{X_0} is flat. If q_X is also flat then q_{X_1} must be flat too since $q_X = q_{X_0} + q_{X_1}$. If q_X is not flat then X_1 must increase with X . Likewise, if $h = g_0 < g_1$ then q_{X_1} is flat. The set $\{g_0 = g_1\}$ is ambiguous. These relationships are summaries in Table 1. The last row uses the fact that $q_X = q_{X_0} + q_{X_1}$ by comonotonicity.

Table 1: Characteristics of optimal allocations, $h = g_0 \wedge g_1$.

Component	$h = g_0 < g_1$	$h = g_1 < g_0$
Since h prices X_0		q_{X_0} flat
Since h prices X_1	q_{X_1} flat	
On $\{dq_X > 0\}$	X_0 increases like X	X_1 increases like X
On $\{dq_X = 0\}$	X, X_1 flat $\Rightarrow q_{X_0}$ flat	X, X_0 flat $\Rightarrow q_{X_1}$ flat

3 The Original Conceptual Proof

In this section we describe the original, conceptual proof from Jouini et al. (2008).

We work on an atomless standard probability space $(\Omega, \mathcal{F}, \mathbb{P})$. All losses lie in L^∞ and so are bounded. For a loss X , write F_X for its distribution function and

$$q_X(u) := \inf \{x : F_X(x) \geq u\}, \quad 0 \leq u \leq 1,$$

for its quantile function. When $\mu \in (L^\infty)^*$ is a bounded finitely additive functional, the pairing with X is $\langle \mu, X \rangle$; if μ is countably additive with density $Z \in L^1$, then $\langle \mu, X \rangle = \mathbb{P}(ZX)$. The paper keeps the full dual $(L^\infty)^*$ in play because, in the comonotonic additive spectral case, subgradients need not live in L^1 when there is a jump at 0 in the distortion function.

Let $X \in L^\infty$, and let g_0, g_1 be comonotone law-invariant coherent risk measures. We write the same symbol g_i for the associated concave distortion on $[0, 1]$ given by

$$g_i(Y) = \int_0^1 q_Y(s) d\check{g}(s),$$

where $\check{g}(s) = 1 - g(1 - s)$ is the dual of g . When g_i is absolutely continuous, the risk measure is written as

$$g_i(Y) = \int_0^1 q_Y(s) g'_i(1 - s) ds,$$

so the supporting density Z matches by

$$q_Z(1 - s) = \check{g}_i(s) = g'_i(1 - s).$$

In the general case, the same formula is understood in the Stieltjes-Choquet sense, with a possible atom at 0 corresponding to a singular part. Define

$$(g_0 \square g_1)(X) := \inf \{ g_0(Y_0) + g_1(Y_1) : Y_0 + Y_1 = X \},$$

and the set of comonotone allocations

$$\mathbb{A}^\uparrow(X) := \{ (Y_0, Y_1) : Y_0 + Y_1 = X, Y_0 \uparrow X, Y_1 \uparrow X \}.$$

Theorem 3.1 is the loss-side translation of Jouini et al. (2008) Theorem 3.2 for parts (1) and (2), Theorem 3.1 together with Lemma 6.2 for part (3), and Proposition 3.1 for part (4).

Theorem 3.1. *With the above notation:*

1. An optimal allocation may be chosen in $\mathbb{A}^\uparrow(X)$.
2. There exists an allocation (X_0, X_1) attaining $(g_0 \square g_1)(X)$.
3. If $(X_0, X_1) \in \mathbb{A}^\uparrow(X)$ attains $(g_0 \square g_1)(X)$, then there exists a finitely additive probability measure m simultaneously pricing X_0 , X_1 , and X :

$$m(X_0) = g_0(X_0), \quad m(X_1) = g_1(X_1), \quad m(X) = (g_0 \square g_1)(X).$$

Equivalently,

$$m \in \partial g_0(X_0) \cap \partial g_1(X_1) = \partial(g_0 \square g_1)(X).$$

4. If $(X_0, X_1) \in \mathbb{A}^\uparrow(X)$, then the following are equivalent:

- (X_0, X_1) is Pareto optimal;
- for each $i = 0, 1$, the quantile q_{X_i} is flat on

$$\{g_i > g_{1-i}\} \cap \{dq_X > 0\}.$$

Moreover, in that case both quantiles are flat on

$$\{g_0 \neq g_1\} \cap \{dq_X = 0\}.$$

Plan of the Proof

We first reduce the infimum to increasing, hence comonotone, allocations. That step is the Landsberger-Meilijson rearrangement argument plus preservation of second-order stochastic dominance by law-invariant monetary risk measures. We then prove existence by compactness: a minimizing sequence in $\mathbb{A}^\uparrow(X)$ is written as $(f_n(X), X - f_n(X))$, the family (f_n) is uniformly bounded and equicontinuous, and Arzela-Ascoli yields a uniformly convergent subsequence. Next we describe dual pricing measures directly in distortion language: a finitely additive probability m determines a concave primitive g_m , the support set of a distortion risk measure is $\{m : g_m \leq g\}$, and m prices Y precisely when $g_m \leq g$ and q_Y is flat where the inequality is strict. The dual of the inf-convolution then comes from intersecting support sets, which produces the distortion $g_0 \wedge g_1$. The last step proves the flatness criterion: a common pricing measure for an optimal split must coincide with $g_0 \wedge g_1$ on the active part of dq_X , and flatness follows; conversely, if the flatness holds, the supporter associated with $g_0 \wedge g_1$ prices both pieces, so the split is optimal.

In summary:

1. we reduce the search to increasing allocations;
2. compactness gives an optimizer;
3. every increasing optimizer admits a common pricing measure;
4. the common pricing distortion is $g_0 \wedge g_1$ on the active part of q_X ;
5. flatness of the allocated quantiles on the regions where one agent is more risk averse than the other is equivalent to Pareto optimality.

Preliminaries

An allocation (Y_0, Y_1) is Pareto optimal when there is no other feasible split (Z_0, Z_1) with

$$g_i(Z_i) \leq g_i(Y_i), \quad i = 0, 1,$$

and at least one strict inequality. Because the risk measures are cash invariant, Pareto optimality is equivalent to minimizing the sum $g_0(Y_0) + g_1(Y_1)$ subject to $Y_0 + Y_1 = X$, which follows directly by contradiction.

For a finitely additive probability m with decomposition

$$m = Z \cdot P + m^s,$$

$Z \in L^1$ and m^s singular and purely finitely additive, define the associated primitive distortion by

$$g_m(0) := 0, \quad g_m(t) := \|m^s\|_{ba} + \int_0^t q_Z(1-u) du, \quad 0 < t \leq 1.$$

Thus g_m is increasing and concave, with a possible jump at 0. In the countably additive differentiable case,

$$g'_m(t) = q_Z(1-t).$$

Write

$$\mathcal{M}_g := \{m \in (L^\infty)_+^* : m(\Omega) = 1, g_m \leq g\}.$$

Then $m \in \mathcal{M}_g$ means that m is an admissible dual pricing measure for g . This follows from the representation

$$g(X) = \int_0^\infty g(S(x)) dx; \quad S(x) := \mathbf{P}(X > x)$$

for non-negative X and from which it is easy to see

$$g(X) = \sup_{h \leq g} h(X)$$

where the sup is over distortions pointwise dominated by g . Hence the support-set identity for coherent distortions is

$$g^*(m) = \chi_{\mathcal{M}_g}(m),$$

and therefore

$$g(Y) = \sup_{m \in \mathcal{M}_g} m(Y).$$

It is a general fact from convex analysis that

$$\begin{aligned} (f \square g)^*(x^*) &= \sup_x (\langle x^*, x \rangle - (f \square g)(x)) \\ &= \sup_x \left(\langle x^*, x \rangle - \inf_y (f(y) + g(x-y)) \right) \\ &= \sup_x \sup_y (\langle x^*, x \rangle - f(y) - g(x-y)) \\ &= \sup_{y,z} (\langle x^*, y+z \rangle - f(y) - g(z)) \\ &= \sup_{y,z} (\langle x^*, y \rangle - f(y) + \langle x^*, z \rangle - g(z)) \\ &= \sup_y (\langle x^*, y \rangle - f(y)) + \sup_z (\langle x^*, z \rangle - g(z)) \\ &= f^*(x^*) + g^*(x^*). \end{aligned}$$

where we change variables $z := x - y$, $x = y + z$ midway through. In the case of spectral measures, the simple form of the conjugate shows

$$\begin{aligned} (g_0 \square g_1)^* &= g_0^* + g_1^* \\ &= \chi_{\mathcal{M}_{g_0}} + \chi_{\mathcal{M}_{g_1}} \\ &= \chi_{\mathcal{M}_{g_0} \cap \mathcal{M}_{g_1}} \\ &= \chi_{\mathcal{M}_{g_0 \wedge g_1}} \end{aligned}$$

Thus $g_0 \square g_1$ is again a distortion risk measure, with distortion $g_0 \wedge g_1$. It is obviously law invariant and hence also Fatou, an important fact that is not true in general.

Finally, the subgradient criterion is:

$$m \in \partial g(Y) \iff g_m \leq g \text{ and } q_Y \text{ is flat on } \{g_m < g\}.$$

The proof is the usual integration-by-parts argument: if m prices Y , then

$$g(Y) - m(Y) = 0,$$

and after rewriting both terms through the same quantile pairing one obtains

$$\int_0^1 (g - g_m)(t) dq_Y(t) + (q_Y(0+) - q_Y(0))(g(0+) - g_m(0+)) = 0.$$

Since $g - g_m \geq 0$, equality forces q_Y to be flat wherever $g_m < g$.

Proof of (1): Reduction to Increasing Allocations

Take any feasible allocation (Y_0, Y_1) with $Y_0 + Y_1 = X$. By the Landsberger-Meilijson rearrangement theorem (Landsberger and Meilijson (1994) Dana and Meilijson (2003)), there exists a comonotone feasible allocation (\hat{Y}_0, \hat{Y}_1) such that each \hat{Y}_i dominates Y_i in the sense of second-order stochastic dominance. Law-invariant monetary risk measures preserve that order (Svindland (2014)), so

$$g_i(\hat{Y}_i) \leq g_i(Y_i), \quad i = 0, 1.$$

Hence every feasible allocation is weakly improved by a comonotone one. By Denneberg's lemma (Denneberg (1994)), the comonotone feasible allocations are exactly the allocations increasing with X , so

$$(g_0 \square g_1)(X) = \inf_{(Y_0, Y_1) \in \mathbb{A}^\uparrow(X)} (g_0(Y_0) + g_1(Y_1)).$$

Proof of (2): Existence

Choose a minimizing sequence (X_0^n, X_1^n) in $\mathbb{A}^\uparrow(X)$. Write

$$X_0^n = f_n(X), \quad X_1^n = X - f_n(X),$$

where each f_n is nondecreasing on the range

$$[a, b] := [\text{ess inf } X, \text{ess sup } X].$$

By shifting a constant from one coordinate to the other, we may assume

$$\text{ess inf } X_0^n = 0.$$

Then

$$0 \leq f_n \leq b - a,$$

and since both f_n and $\text{Id} - f_n$ are nondecreasing, every f_n is 1-Lipschitz. Hence

$$B := \{ f : [a, b] \rightarrow \mathbb{R} : |f| \leq b - a, f \text{ and } \text{Id} - f \text{ are nondecreasing} \}$$

is bounded and equicontinuous. The Arzela-Ascoli theorem (Royden and Fitzpatrick (2010)) yields a subsequence, still denoted f_n , and a limit $f \in B$ such that

$$f_n \rightarrow f \quad \text{uniformly on } [a, b].$$

Set

$$X_0 := f(X), \quad X_1 := X - f(X).$$

Then $(X_0, X_1) \in \mathbb{A}^\uparrow(X)$ and

$$X_0^n \rightarrow X_0, \quad X_1^n \rightarrow X_1 \quad \text{in } L^\infty.$$

Each g_i is L^∞ -continuous, so

$$g_0(X_0) + g_1(X_1) = \lim_{n \rightarrow \infty} (g_0(X_0^n) + g_1(X_1^n)) = (g_0 \square g_1)(X).$$

Thus an optimizer exists.

Proof of (3): Common Pricing Measure for an Increasing Optimizer

Let $(X_0, X_1) \in \mathbb{A}^\uparrow(X)$ attain $(g_0 \square g_1)(X)$. By duality for the convolution problem, attainment is equivalent to the existence of a common dual supporter. In risk notation that means there exists a finitely additive probability m such that

$$m \in \partial g_0(X_0) \cap \partial g_1(X_1).$$

Then

$$m(X_0) = g_0(X_0), \quad m(X_1) = g_1(X_1).$$

Adding the two identities and using $X = X_0 + X_1$ gives

$$m(X) = g_0(X_0) + g_1(X_1) = (g_0 \square g_1)(X),$$

so the same m prices both pieces and the aggregate problem.

Because m prices X for the inf-convolution and because $(g_0 \square g_1)^* = \chi_{\mathcal{M}_{g_0 \wedge g_1}}$, we also have

$$g_m \leq g_0 \wedge g_1.$$

A useful refinement comes from conditioning on $\sigma(X)$. If m prices a law-invariant risk measure at Y , then $\mathbb{P}(m \mid Y)$ prices it as well, the regular part of the conditioned measure increases with Y , and the singular part concentrates on the worst-loss set $\{Y = \text{ess sup } Y\}$. The little trick in that proof is the measure-preserving swap: if singular mass sits away from the worst states, we move that mass to a better set of the same probability, lower the pairing with Y , and contradict optimality. Applying that lemma with $Y = X$, we may and do assume that the pricing measure in part (3) increases with X .

Proof of (4): Optimality Implies Flatness

Assume $(X_0, X_1) \in \mathbb{A}^\uparrow(X)$ is Pareto optimal. By the equivalence between Pareto optimality and attainment of the inf-convolution, the pair attains $(g_0 \square g_1)(X)$, so part (3) gives a common pricing measure m . Since m prices each X_i , the subgradient criterion gives

$$q_{X_i} \text{ is flat on } \{g_m < g_i\}, \quad i = 0, 1.$$

We now identify g_m on the active part of the aggregate quantile. Because $(X_0, X_1) \in \mathbb{A}^\uparrow(X)$, the pair is comonotone, so

$$q_X = q_{X_0} + q_{X_1}.$$

Also,

$$(g_0 \square g_1)(X) = \int_0^1 q_X(s) d(g_0 \wedge g_1)(1-s),$$

while

$$m(X) = \int_0^1 q_X(s) dg_m(1-s).$$

Since these two quantities are equal and $g_m \leq g_0 \wedge g_1$, the same integration-by-parts argument as above yields

$$g_m = g_0 \wedge g_1 \quad \text{on } \{dq_X > 0\}.$$

Therefore, on $\{dq_X > 0\}$,

$$\{g_m < g_i\} = \{g_{1-i} < g_i\}.$$

Hence for each $i = 0, 1$,

$$q_{X_i} \text{ is flat on } \{g_i > g_{1-i}\} \cap \{dq_X > 0\}.$$

That proves the second claim.

For the additional statement on $\{dq_X = 0\}$, suppose at some point t we have $g_0(t) \neq g_1(t)$; say $g_0(t) > g_1(t)$. Then

$$g_m(t) \leq g_1(t) < g_0(t),$$

so q_{X_0} is flat there by the subgradient criterion. Since

$$q_X = q_{X_0} + q_{X_1}$$

and $dq_X = 0$ there, q_{X_1} must be flat there as well. Interchanging the indices gives flatness of both quantiles on

$$\{g_0 \neq g_1\} \cap \{dq_X = 0\}.$$

Proof of (4): Flatness Implies Optimality

Now assume $(X_0, X_1) \in \mathbb{A}^\uparrow(X)$ and that

$$q_{X_i} \text{ is flat on } \{g_i > g_{1-i}\} \cap \{dq_X > 0\}, \quad i = 0, 1.$$

Set

$$g_* := g_0 \wedge g_1.$$

Choose a uniform random variable U with

$$X = q_X(U).$$

Let m_* be the pricing measure associated with the distortion g_* , with any jump at 0 assigned as singular mass on the worst-loss set $\{X = \text{ess sup } X\}$. Then m_* increases with X , hence with each X_i because $(X_0, X_1) \in \mathbb{A}^\uparrow(X)$.

By construction,

$$g_{m_*} = g_* \leq g_i.$$

On $\{dq_X > 0\}$ the assumed flatness gives

$$q_{X_i} \text{ flat on } \{g_* < g_i\} = \{g_{1-i} < g_i\}.$$

On $\{dq_X = 0\}$ both quantiles are flat automatically wherever $g_0 \neq g_1$, because $q_X = q_{X_0} + q_{X_1}$ and both X_i increase with X . Therefore

$$q_{X_i} \text{ is flat on } \{g_* < g_i\}$$

for each $i = 0, 1$. The subgradient criterion yields

$$m_* \in \partial g_0(X_0) \cap \partial g_1(X_1).$$

Thus m_* prices both coordinates, so

$$g_0(X_0) + g_1(X_1) = m_*(X) = (g_0 \square g_1)(X).$$

Hence (X_0, X_1) attains the inf-convolution and is therefore Pareto optimal.

Extensions

Theorem 3.2. *An allocation (X_1, X_2) is Pareto optimal if and only if $g_1(X_1) + g_2(X_2) = (g_1 \square g_2)(X)$.*

Proof. By definition, (X_1, X_2) is Pareto optimal if there is no feasible allocation (Y_1, Y_2) such that $g_1(Y_1) \leq g_1(X_1)$ and $g_2(Y_2) \leq g_2(X_2)$ with at least one strict inequality.

Assume (X_1, X_2) is Pareto optimal. Define two sets in \mathbb{R}^2 :

$$B = \{(g_1(Y_1), g_2(Y_2)) : Y_1 + Y_2 = X\} + \mathbb{R}_+^2$$

$$C = \{(g_1(X_1), g_2(X_2))\} - (\mathbb{R}_+^2 \setminus \{0\})$$

Both B and C are convex. Pareto optimality implies $B \cap C = \emptyset$. By the Hahn-Banach Separation Theorem (Royden and Fitzpatrick (2010)), there exists a non-zero vector $\lambda = (\lambda_1, \lambda_2)$ separating B and C :

$$\lambda \cdot b \geq \lambda \cdot c \quad \forall b \in B, c \in C$$

Since B extends infinitely in the positive directions, we must have $\lambda_1, \lambda_2 \geq 0$. Now, use the translation invariance of the risk measures: $g_i(Y_i + k) = g_i(Y_i) + k$. By shifting any allocation $Y_1 \rightarrow Y_1 + k$ and $Y_2 \rightarrow Y_2 - k$ (which preserves $Y_1 + Y_2 = X$), the value in B shifts by $\lambda_1 k - \lambda_2 k$. For the lower bound to hold for all $k \in \mathbb{R}$, we must have $\lambda_1 = \lambda_2$. Dividing by λ_1 , we see that (X_1, X_2) minimizes the sum $g_1(Y_1) + g_2(Y_2)$ over all valid allocations. \square

A bounded finitely additive measure $\mu \in (\mathbb{L}^\infty)^*$ decomposes uniquely into a regular (countably additive) part $Zd\mathbb{P}$ and a purely singular part μ^s .

Lemma 3.1. *If $\mu \in \partial g(Y)$ for a law-invariant risk measure g , then $-\mu$ is comonotone with Y , and its singular part μ^s is strictly concentrated on $\{Y = \text{ess sup } Y\}$.*

Proof. Suppose μ^s is not concentrated on the essential supremum. There exists an $\epsilon > 0$ such that the set $\Omega \setminus \{Y \geq \text{ess sup } Y - 2\epsilon\}$ has μ^s -measure $\alpha > 0$.

Choose two sequences of sets: B_n in the low-value region (where $Y < \text{ess sup } Y - 2\epsilon$) such that $\mu^s(B_n) = \alpha$ and C_n in the high-value region (where $Y \geq \text{ess sup } Y - \epsilon$) such that $\mu^s(C_n) = 0$.

Ensure $\mathbb{P}(B_n) = \mathbb{P}(C_n) \leq 1/n$. Construct a measure-preserving bijection τ_n that simply swaps B_n and C_n , leaving the rest of Ω unchanged. Define a new measure $\nu_n = \mu \circ \tau_n^{-1}$.

Because Y is strictly higher on C_n than B_n , swapping the evaluation of the singular mass α from the lower values to the higher values strictly decreases the inner product:

$$\langle \nu_n, Y \rangle \leq \langle \mu, Y \rangle - \alpha\epsilon$$

Thus, $\langle \nu_n, Y \rangle < \langle \mu, Y \rangle$. However, because g is law-invariant, $g(Y) \geq \langle \nu_n, Y \rangle$. This means $g(Y) > \langle \mu, Y \rangle$, which contradicts $\mu \in \partial g(Y)$. Therefore, μ^s must be zero everywhere except the essential supremum. \square

Ingredients

The conceptual proof uses a smörgåsbord of ingredients.

1. Convex duality and Fenchel contact: conjugates, biconjugates, Fenchel inequality, and subgradients as contact functionals.
2. The inf-convolution dualizes to sum:

$$(g_0 \square g_1)^* = g_0^* + g_1^*.$$

This is the algebraic backbone of the Pareto problem. On the utility side this is Lemma 2.1.

3. Pareto optimality equals common supporting subgradient: efficient allocation \iff one common dual object supports both agents. This is Theorem 3.1.
4. Hahn–Banach separation, plus cash-invariance: the separation argument gives Pareto \implies weighted optimum; cash-invariance collapses the two multipliers to one. This is an important hidden simplification.
5. The dual penalty of a coherent risk measure is an indicator:

$$g^* = I_Q,$$

so the dual problem becomes set intersection rather than a penalized optimization.

6. Kusuoka’s quantile representation for law-invariant criteria: general law-invariant costs have a variational quantile representation, and comonotone ones collapse to a single distortion profile. This is Theorems 2.2 and 2.3.
7. The lower-envelope rule: in the comonotonic additive case the aggregate criterion is generated by the lower envelope

$$g = g_0 \wedge g_1.$$

On the utility side this is Lemma 2.2.

8. The conditional-expectation reduction onto $\sigma(X)$: any Pareto allocation can be replaced by

$$(\mathbb{P}[X_0 | X], \mathbb{P}[X_1 | X])$$

without loss. This is one of the big structural moves in Theorem 3.1(vi).

9. Landsberger’s monotone-improvement idea: once only the law matters, unordered sharing can be rearranged into monotone sharing. This is not stated as a formal theorem in this paper, but it is part of the background logic behind the passage to increasing allocations.
10. Denneberg’s lemma: increasing in X is the same as comonotone with X , hence with each other. This is what turns the allocation problem into a quantile problem.
11. Quantile additivity under comonotonicity: for comonotone allocations,

$$q_X = q_{X_0} + q_{X_1}.$$

This is what makes the layer-by-layer assignment possible. This sits behind Theorem 3.2 and Proposition 3.1.

12. Singular-vs-regular dual decomposition: the dual lives in $(L^\infty)^*$, not just L^1 , because a jump of the distortion at 0 creates an $\text{ess sup } X$ term in your loss convention. The regular/singular split is therefore not bookkeeping; it is essential.
13. The rearrangement-comonotone-coupling trick for the dual variable: Lemma 4.1 lets one replace a supporting dual object by one aligned with the loss, and also push it down by conditional expectation to $\sigma(X)$. This is one of the real magic tricks in the proof.
14. The explicit description of the subgradient in distortion language: Lemma 4.2 says, in effect,

$$\mu \in \partial g(X)$$

iff the dual profile lies below the distortion profile and the quantile is flat where the inequality is strict. This is the bridge from abstract duality to the concrete contract shape.

15. A flatness criterion: in the optimal allocation, agent i ’s quantile is flat where $g_i > g_{1-i}$ on the moving part of q_X . This is Proposition 3.1 itself.
16. The stop-loss synthesis: when the profiles cross in finitely many places, the flatness rule integrates to a finite sum of layers, hence a sum of stop-loss contracts or options. That is Example 3.1.
17. In general, and surprisingly, law invariance of g_i does not imply that of $g_0 \square g_1$, see Liu et al. (2020). This is important because it means freebie Fatou can fail. Jouini et al. (2008) state that, in their monetary, convex context, “clearly” $U_0 \square U_1$ is law-invariant. Monetary utilities enjoy strong continuity properties and in that framework the inf-convolution does remain law-invariant. Liu et al. (2020) show that if we strip away that extra structure and keep only bare law invariance, the conclusion can fail: the obstruction is that inf-convolution depends not only on the marginal laws of the summands, but on which decompositions of a total risk are feasible on the underlying probability space, and that feasibility need not itself be law-invariant.

References

- Dana, R.-A., & Meilijson, I. (2003). Modelling agents’ preferences in complete markets by second order stochastic dominance, (May), 1–27.
- Denneberg, D. (1994). *Non-additive measure and integral*. Dordrecht: Kluwer Academic.
- Jouini, E., Schachermayer, W., & Touzi, N. (2008). Optimal risk sharing for law invariant monetary utility functions. *Mathematical Finance*, 18(2), 269–292. <https://doi.org/10.1111/j.1467-9965.2007.00332.x>
- Landsberger, M., & Meilijson, I. (1994). Co-monotone allocations, Bickel-Lehmann dispersion and the Arrow-Pratt measure of risk aversion. *Annals of Operations Research*, 52(2), 97–106. <https://doi.org/10.1007/BF02033185>

- Liu, P., Wang, R., & Wei, L. (2020). Is the inf-convolution of law-invariant preferences law-invariant? *Insurance: Mathematics and Economics*, 91(December), 144–154. <https://doi.org/10.1016/j.insmatheco.2020.01.004>
- Royden, H., & Fitzpatrick, P. (2010). *Real Analysis* (Fourth.). Pearson.
- Svindland, G. (2014). Dilatation monotonicity and convex order. *Mathematics and Financial Economics*, 8(3), 241–247. <https://doi.org/10.1007/s11579-013-0112-y>