ON COMONOTONICITY OF PARETO OPTIMAL RISK SHARING

MICHAEL LUDKOVSKI[†] AND LUDGER RÜSCHENDORF[‡]

Abstract. We establish various extensions of the comonotone improvement result of Landsberger and Meilijson (1994) which are of interest for the risk sharing problem. As a consequence we obtain general results on the comonotonicity of Pareto-optimal risk allocations using risk measures consistent with the stochastic convex order.

1. Introduction

With the ongoing development of convex risk measures (see e.g. Föllmer and Schied (2002)) there has been renewed interest in the problem of optimal risk exchange between economic agents. A key step in studying the structure of Pareto-optimal risk allocations is the *comonotonicity* property. This result was originally obtained in Landsberger and Meilijson (1994), who provided an algorithm to construct a convex order $\leq_{cx}$-improvement of any non-comonotone allocation. Since all law-invariant convex risk measures are consistent with the convex order (see e.g. Bäuerle and Müller (2006)) it follows that a Pareto-optimal risk allocation is necessarily comonotone. The improvement result in Landsberger and Meilijson (1994) was only stated for discrete and bounded allocations; later Dana and Meilijson (2003) constructed an extension to general bounded risks. In this note we establish various extensions of this result to unbounded random variables, as well versions working for certain classes of consistent risk measures. This is significant from a practical point of view where risks are often modeled as unbounded random variables.

2. Preliminaries

Consider the collection of real-valued random variables $L^0(\mathbb{P})$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. As usual, we write L^1 (resp. L^∞_+) for the collection of all integrable (positive bounded) random variables on $(\Omega, \mathbb{P})$.

Definition 1. *Two random variables Y and $Z \in L^0(\mathbb{P})$ are said to be* comonotone *if*

$$(Y(\omega_1) - Y(\omega_2))(Z(\omega_1) - Z(\omega_2)) \geq 0, \tag{1}$$

Date: May 18, 2007.

Key words and phrases. comonotonicity, convex order, Pareto-optimality.

We thank Virginia Young for many useful discussions.

$\mathbb{P}(d\omega_1) \times \mathbb{P}(d\omega_2)$-*almost surely. In other words, Y and Z move together.*

An equivalent definition of comonotonicity is that there exist non-decreasing functions h_Y and h_Z such that $h_Y(x) + h_Z(x) = x$, $Y = h_Y(Y + Z)$, and $Z = h_Z(Y + Z)$ almost surely (Denneberg 1994).

We next recall the stochastic convex order:

Definition 2. *$Y \in L^1$ is said to* precede (*or be* preferred *to*) *$Z \in L^1$ in* convex order *if $\mathbb{E}[f(Y)] \leq \mathbb{E}[f(Z)]$ for all convex functions f for which the expectations exist. We write $Y \leq_{cx} Z$.*

Note that convex order is equivalent to ordering with respect to second stochastic dominance with equal means, see Rothschild and Stiglitz (1970, 1971).

3. Main Results

Let $X \in L^1$ and consider the collection of integrable allocations of X, namely $\mathcal{A}(X) \triangleq \{\mathbf{Y} := (Y_1, Y_2, \ldots, Y_n) \colon X = \sum_{i=1}^n Y_i,\ Y_i \in L^1\}$.

Definition 3. *An allocation $\mathbf{Y} \in \mathcal{A}(X)$ is called comonotone if Y_i, Y_j are comonotone for all $1 \leq i \leq j \leq n$.*

We have trivially that if $\mathbf{Y}$ is a comonotone allocation then Y_i and X are comonotone for each i.

Theorem 1. *Let $\mathbf{Y} \in \mathcal{A}(X)$ be a discrete allocation taking on a countable number of values and bounded from below. Then there exists a comonotone allocation $\bar{\mathbf{Y}} \in \mathcal{A}(\mathbf{X})$ such that $\bar{Y}_i \leq_{cx} Y_i$, $i = 1, 2, \ldots, n$.*

The idea is to apply a variant of the Landsberger and Meilijson (1994) algorithm and take the limit using the properties of the basic single-crossing type improvement (Diamond and Stiglitz 1974). Theorem 1 essentially appeared in Landsberger and Meilijson (1994), but proof was only given in the case when X and $\mathbf{Y}$ take on finitely many values. By now the result is part of the folklore, however no complete proof has been published and the extension is not entirely trivial.

Proof. For completeness, we provide full details for the case $n = 2$, whence $\mathbf{Y} = (Y_1, Y_2)$. Let $\mathbf{y} = \bigcup_{i=1}^n \mathbf{y}^i$ be the union of the supports of Y_i. Then $\mathbf{y}$ is an ordered countable set in $\mathbb{R}$, bounded from below by some $\underline{y}$. Consider the partition $\Omega = \cup_k C_k$ such that (Y_1, Y_2) are constant on each partition element C_k, and the C_k's are ordered according to the values of $Y_1 + Y_2$. Our goal is to have each of Y_1, Y_2 also ordered with respect to their values

$y_1^1, y_2^1, \dots, y_1^2, y_2^2, \dots$ on C_k. Let $p_k = \mathbb{P}(C_k) > 0$ be the weights of different pairs. If Y_1 and Y_2 are not comonotone, there must be a minimal index k such that without loss of generality $y_1^1 \le y_2^1 \le \dots \le y_k^1 \le y_{k+1}^1$ and $y_1^2 \le y_2^2 \le \dots \le y_k^2$, but $y_{k+1}^2 < y_k^2$. Let $1 \le j \le k$ be the minimal index such that $y_j^2 > y_{k+1}^2$, so that the $(k+1)$-st pair "violates" the comonotonicity with all pairs between j and k.

Construct the update $(\tilde{Y}_1, \tilde{Y}_2)$ which takes on the same values $\tilde{y}_i^1 \leftarrow y_i^1, \tilde{y}_i^2 \leftarrow y_i^2$ as (Y_1, Y_2) except

$$\begin{cases} \tilde{y}_{k+1}^1 & \leftarrow y_{k+1}^1 - (y_j^2 - y_{k+1}^2) \cdot (\sum_{i=j}^k p_i)/(\sum_{i=j}^{k+1} p_i), \\ \tilde{y}_{k+1}^2 & \leftarrow y_{k+1}^2 + (y_j^2 - y_{k+1}^2) \cdot (\sum_{i=j}^k p_i)/(\sum_{i=j}^{k+1} p_i), \\ \quad \tilde{y}_i^1 & \leftarrow y_i^1 + (y_j^2 - y_{k+1}^2) \cdot p_{k+1}/(\sum_{i=j}^{k+1} p_i), \qquad j \le i \le k, \\ \quad \tilde{y}_i^2 & \leftarrow y_i^2 - (y_j^2 - y_{k+1}^2) \cdot p_{k+1}/(\sum_{i=j}^{k+1} p_i), \qquad j \le i \le k. \end{cases} \tag{2}$$

This change preserves $\tilde{Y}_1 + \tilde{Y}_2 = Y_1 + Y_2$, maintains $\mathbb{E}\tilde{Y}_1 = \mathbb{E}Y_1$ and $\mathbb{E}\tilde{Y}_2 = \mathbb{E}Y_2$ and is an improvement of the single-crossing type (Diamond and Stiglitz 1974), also known as mean-preserving spread (mps). It follows that $\tilde{\mathbf{Y}}$ is a component-wise $\le_{cx}$ and $\le_{mps}$ improvement of $\mathbf{Y}$. Moreover, after the change, $\tilde{y}_j^2 = \tilde{y}_{k+1}^2$ and $\tilde{y}_j^1 \le \tilde{y}_{k+1}^1$ so that the j-th pair is now comonotone with the $k+1$-st pair. Therefore, the next violation index pair (k, j) will be larger (in the lexicographic order) than the current one.

Iterate this argument to obtain an improvement sequence $(\tilde{\mathbf{Y}}^{(m)})_{m=1,2,\dots}$ of allocations. We now claim that this sequence converges almost surely and in L^1 to some limit $\bar{\mathbf{Y}}$. Indeed, let (k_m, j_m) be the violation index of the m-th update. Then $k_m \to \infty$ and therefore $p_{k_m} \to 0$. On the other hand, on a fixed subset C_k, once $k_m > k$, the value $y_k^{i,(m)}$ can change by at most

$$\begin{aligned} |y_k^{i,(m+1)} - y_k^{i,(m)}| &\le \max_{i=1,2} |y_k^{i,(m)} - y_{k_m+1}^{i,(m)}| \cdot p_{k_m+1}/(\sum_{j=k}^{k_m+1} p_j) \\ &\le \max_i (y_k^{i,(m)} - \underline{y}) \cdot p_{k_m+1}/p_k \\ &\le \left((y_k^1 + y_k^2 - 2\underline{y})/p_k\right) \cdot p_{k_{m+1}}, \end{aligned}$$

since $y_k^{1,(m)} + y_k^{2,(m)} = y_k^1 + y_k^2$ for all m.

Since $p_{k_m} \to 0$, for any $\epsilon > 0$ there is M large enough such that $\sum_{j > k_M} p_j < \epsilon$ and the respective tail sum is then bounded by $\sum_{n=M}^\infty |y_k^{i,(n)} - y_k^{i,(n+1)}| \le ((y_k^1 + y_k^2 - 2\underline{y})/p_k)\epsilon$. Thus, $(Y^{i,(m)})$ converges almost surely.

Moreover, $(Y^{i,(m)})$ also converges in L^1. This is obvious if Y^i is bounded since $\sup Y^{i,(m+1)} \le Y^{i,(m)}$. Otherwise note that for a fixed threshold index k', because of the mean-preserving spread the *tail mass* is non-increasing,

$$\sum_{k\geq k'} p_k \cdot y_k^{i,(m)} \leq \sum_{k\geq k'} p_k \cdot y_k^i. \tag{3}$$

For a fixed level K, the tail expectation $\mathbb{E}[Y^{i,(m)} 1_{\{Y^{i,(m)}>K\}}]$ will increase only if a point (y_k^i, p_k) is moved to the right of K as a result of e.g. the third line of (2). The algorithm of (2) operates by sliding points towards their average with changes in distance proportional to the weights. Thus, to slide an initial point y_k^i with mass p_k to level K requires an "energy" of $(K - y_k^i)\cdot p_k$. To do so, at least as much energy should be removed from the right of K. However, total available energy beyond K is $\max_i \mathbb{E}[Y^i 1_{\{Y^i>K\}}]$. Thus, if $(K - y_k^i)p_k > \mathbb{E}[Y^i 1_{\{Y^i>K\}}]$ then the k-th point will never contribute to $\mathbb{E}[Y^{i,(m)} 1_{\{Y^{i,(m)}>K\}}]$. Let $\underline{k}_K = \min\left\{k : (K - y_k^i)p_k < \mathbb{E}[Y^i 1_{\{Y^i>K\}}]\right\}$, be the first index that can affect the above tail expectation. Combining with (3) we obtain the uniform bound

$$\mathbb{E}[Y^{i,(m)} 1_{\{Y^{i,(m)}>K\}}] \leq \sum_{k\geq \underline{k}_K} p_k y_k^{i,(m)} \leq \sum_{k\geq \underline{k}_K} p_k y_k^i.$$

Finally, as $K\to\infty$, $\underline{k}_K \to \infty$ and $\sum_{k\geq \underline{k}_K} p_k y_k^i = \mathbb{E}[Y^i 1_{\{Y^i>K\}}] \to 0$, establishing the uniform integrability of $(Y^{i,(m)})_m$.

The limiting $\bar{\mathbf{Y}}$ is comonotone, since there are no comonotonicity violation pairs left in the limit and by (Müller and Stoyan 2002, Theorem 1.5.9), $\bar{Y}_i \leq_{cx} Y_i$, $i = 1, 2, \dots, n$.

□

When the probability space Ω is non-atomic, a more direct argument is available by extending the construction in Dana and Meilijson (2003). Namely we have,

Theorem 2. *Let $\mathbf{Y} \in \mathcal{A}(X)$ be an allocation of $X \in L^1$. Suppose Ω is non-atomic. Then there exists a comonotone allocation $\bar{\mathbf{Y}} \in \mathcal{A}(X)$ such that $\bar{Y}_i \leq_{cx} Y_i$, $i = 1, 2, \dots, n$.*

Proof. Recall that Dana and Meilijson (2003) proved the $\leq_{cx}$-improvement result for arbitrary $X \in L^\infty_+$. Defining

$$Y_i^{(m)} \triangleq Y_i \cdot 1_{|Y_i|\leq m}, \quad Y^{(m)} \triangleq \sum_{i=1}^n Y_i^{(m)}, \tag{4}$$

then $Y^{(m)} \to X$, $Y_i^{(m)} \to Y_i$, almost surely and in $L^1(\mathbb{P})$. Further, by Dana and Meilijson (2003) there exists $\mathbf{Z}^{(m)}$ comonotone, such that

$$Z_i^{(m)} \leq_{cx} Y_i^{(m)} \quad \text{and} \quad \sum_{i=1}^n Z_i^{(m)} = Y^{(m)}. \tag{5}$$

Let $F_{i,m}$ be the distribution function of $Z_i^{(m)}$. Since $\mathbf{Z}^{(m)}$ is comonotone and Ω is non-atomic, it follows that there exists a $U \sim Unif(0,1)$ random variable such that $Z_i^{(m)} \stackrel{d}{=} F_{i,m}^{-1}(U)$. Note that for each $1 \leq i \leq n$, $(Z_i^{(m)})_m$ are tight since by the convex ordering

$$\mathbb{E}|Z_i^{(m)}| \leq \mathbb{E}|Y_i^{(m)}| \leq \mathbb{E}|Y_i| < \infty.$$

Therefore there exist a subsequence, again labelled $(m) \subset \mathbb{N}$ along which the distribution functions converge, $F_{i,m} \to F_i$. This implies that $F_{i,m}^{-1}(U) \xrightarrow{a.s.} F_i^{-1}(U) =: Z_i$, and moreover

$$\sum_{i=1}^n Z_i^{(m)} \stackrel{d}{=} \sum_{i=1}^n F_{i,m}^{-1}(U) \to \sum_{i=1}^n Z_i \quad \text{a.s. and in } L^1(\mathbb{P}).$$

On the other hand, we already had

$$\sum_{i=1}^n Z_i^{(m)} = \sum_{i=1}^n Y_i^{(m)} \to \sum_{i=1}^n Y_i \quad \text{a.s. and in } L^1(\mathbb{P}).$$

Thus we obtain

$$\sum_{i=1}^n Z_i \stackrel{d}{=} \sum_{i=1}^n Y_i, \quad Z_i \leq_{cx} Y_i \quad \text{and in fact} \quad (Z_i^{(m)}) \stackrel{d}{=} (F_{i,m}^{-1}(U)) \to (Z_i) \quad \text{a.s.}$$

In particular, the limit allocation $\mathbf{Z}$ is comonotone. Therefore, for some measure preserving random variable U' it holds

$$\sum_{i=1}^n Z_i \circ U' = \sum_{i=1}^n Y_i = X \quad \text{a.s.}$$

Obviously $\bar{Y}_i \triangleq Z_i \circ U'$ satisfy

$$(\bar{Y}_i) \stackrel{d}{=} (Z_i), \quad \sum_{i=1}^n \bar{Y}_i = X, \quad \text{and} \quad \bar{Y}_i \leq_{cx} Y_i.$$

□

Remark 1. As in Theorem 1, the $\leq_{cx}$ improvement in (5) is in fact a mean-preserving spread improvement (Rothschild and Stiglitz 1970). However, the presence of $\leq_{mps}$-improvements in the above algorithm alone is not enough to establish tightness of $(Z_i^{(m)})$.

As a counterexample, consider a non-integrable random variable X which has a tail of the order $\mathbb{P}(X > x) \sim cx^{-1/4}$ on the positive real line, as well as a point mass $\epsilon > 0$ at zero. Then there exists a sequence (X_n) of $\leq_{mps}$ improvements $X = X_1 \geq_{mps} X_2 \geq_{mps} \cdots$, such that (X_n) is not tight. For the construction, observe that on the interval $(n^4, (n+1)^4)$, X has mass $\sim \frac{1}{n^2}$. We slide this mass n^2 times one unit to the left in exchange for shifting the point mass in zero one unit to the right (by formally splitting it into n^2 point masses of size $\frac{\epsilon}{n^2}$ and applying the single-crossing $\leq_{mps}$-improvement). Thus, the interval mass is shifted to the interval $(n^4 - n^2, (n+1)^4 - n^2)$. Next consider the interval $((n+1)^4, (n+2)^4)$ or a slight adjustment and repeat the procedure. The point mass in zero slides one unit to the right at each step and is finally shifted to $+\infty$, contradicting tightness.

The integrability assumption on X can be dropped if the initial allocation $\mathbf{Y}$ is bounded from below. Indeed this is sufficient for $(Z_i^{(m)})$ above to be tight; thus, we obtain the following

Corollary 1. *Consider a non-atomic probability space Ω and let $\mathbf{Y}$ be an allocation of $X \in L^0(\mathbb{P})$ bounded from below, $Y_i \geq \underline{y} > -\infty$. Define $Y_i^{(m)}, Y^{(m)}$ as in (4). Then there exists a subsequence $(m) \subset \mathbb{N}$ of comonotone allocations $\mathbf{Z}^{(m)}$ of $Y^{(m)}$ such that $Z_i^{(m)} \geq \underline{y}$ for $1 \leq i \leq n$ and*

(1) $Z_i^{(m)} \leq_{cx} Y_i^{(m)}$,
(2) $Z_i^{(m)} \xrightarrow{a.s.} Z_i$, *where $\mathbf{Z}$ is a comonotone allocation of X.*

Proof. The proof uses the construction of Theorem 2. The only remaining step is to establish tightness of the sequence $(Z_i^{(m)})_m$ for each $1 \leq i \leq n$. This however is a consequence of the fact that the Landsberger-Meilijson algorithm described in the proof of Theorem 1 respects lower and upper bounds of the initial allocation $\mathbf{Y}$. Indeed in (2) we have $\operatorname{ess\,inf} \tilde{Y}_i \geq \operatorname{ess\,inf} Y_i$, $\operatorname{ess\,sup} \tilde{Y}_i \leq \operatorname{ess\,sup} Y_i$. Consequently, for $Z_i^{(m)}$ obtained in (5) from the Dana and Meilijson (2003) result based on same algorithm we also have $\underline{y} \leq \operatorname{ess\,inf} Z_i^{(m)}$. Moreover,

$$Z_i^{(m)} = Y^{(m)} - \sum_{j \neq i} Z_j^{(m)} \leq X - (n-1)\underline{y}.$$

Thus, we have uniform bounds on the lower and upper tails of the distribution of $Z_i^{(m)}$ and the tightness of $(Z_i^{(m)})_m$ follows. Note that without L^1-convergence we cannot claim $Z_i \leq_{cx} Y_i$.

□

4. Application to Risk Sharing

We now interpret X as the total exposure of the n agents, and use convex risk measures ρ_i to define the subjective valuation (preference) functional of the i-th agent. Thus, $\rho_i \colon L^0(\mathbb{P}) \to \mathbb{R} \cup \{+\infty\}$, are the numerical translation-invariant representations of risk-preferences of agents with $\rho_i(X) \geq \rho_i(Y)$ meaning that agent i finds X riskier than Y. We make the following assumptions on ρ_i:

Assumption 1. *ρ_i is consistent with the natural order of $L^0(\mathbb{P})$: if $\mathbb{P}(X \geq Y) = 1$ then $\rho_i(X) \geq \rho_i(Y)$.*

Assumption 2. *ρ_i is consistent with the convex order: for $X, Y \in L^1(\mathbb{P})$ if $X \leq_{cx} Y$, then $\rho_i(X) \leq \rho_i(Y)$.*

Assumption 3. *ρ_i is continuous with respect to a.s. convergence: if $Y^{(n)} \xrightarrow{a.s.} Y$ and $\sup_n \rho_i(Y^{(n)}) < \infty$ then $\rho_i(Y^{(n)}) \to \rho_i(Y)$. In particular, ρ_i is continuous at $-\infty$:*

$$\lim_{d \to -\infty} \rho_i[\max(Y, d)] = \rho_i(Y), \qquad \text{for all } Y \in L^0(\mathbb{P}).$$

Assumption 3 implies that given $Y \in L^0(\mathbb{P})$ and $\epsilon > 0$, there exists $\tilde{Y}$ taking on a countable number of values and bounded from below such that $|\rho_i(\tilde{Y}) - \rho_i(Y)| \leq \epsilon$. For instance, one may take $\tilde{Y} = \max(-d, 2^{-K}(\lceil 2^K Y \rceil))$ for d, K large enough.

Remark 2. One usually defines risk measures on the space $L^\infty(\mathbb{P})$, in which case an extensive theory is available, see e.g. Föllmer and Schied (2002). When working with $L^0(\mathbb{P})$, one must allow $\rho_i(X) = +\infty$, i.e. positions X that are completely unacceptable and contain infinite risk. We also do not directly require ρ_i to be convex; see (Burgert and Rüschendorf 2006, Prop 2.2b) for relationship between convexity of ρ and its consistency with $\leq_{cx}$.

The risk sharing problem (see Aase (2002) for a recent survey) consists in finding an *optimal* allocation $\mathbf{Y}^* \in \mathcal{A}(X)$, namely an allocation such that $\mathbf{Y}^*$ is *Pareto optimal*, that is, no agent can be made strictly better off (in the sense of ρ_i-risk reduction) without another agent being made strictly worse off. Formally, a Pareto optimal risk exchange is defined as follows:

Definition 4. $\mathbf{X}^* \in \mathcal{A}(X)$ *is called a* Pareto optimal *risk exchange or allocation if whenever there exists an allocation $\mathbf{Y} \in \mathcal{A}(X)$ such that $\rho_i(Y_i) \leq \rho_i(X_i^*)$ for all $i = 1, 2, \ldots, n$, then $\rho_i(Y_i) = \rho_i(X_i^*)$ for all $i = 1, 2, \ldots, n$.*

To avoid trivialities we further restrict $\mathcal{A}(X)$ such that admissible allocations carry finite risk for each agent and are integrable:

$$\mathcal{A}(X) \triangleq \{\mathbf{Y} := (Y_1, Y_2, \ldots, Y_n) : X = \sum_{i=1}^n Y_i,\ \rho_i(Y_i) < \infty, Y_i \in L^1\},$$

and assume that the new $\mathcal{A}(X)$ is still non-empty.

The improvement results in this note show that in most situations Pareto-optimal allocations are necessarily comonotone, and so one may immediately restrict the attention to comonotone allocations. The latter fact strongly simplifies the structure of the problem. For instance, in Jouini et al. (2005), Ludkovski and Young (2007) finding Pareto-optimal allocations is reduced to minimizing the linear function $f(\mathbf{Y}) := \sum_{i=1}^n \alpha_i \rho_i(Y_i)$ over possible Lagrange multipliers $\alpha_i \geq 0$. Knowing that the $\mathbf{Y}$'s of interest are comonotone allows one to work directly with the non-decreasing 1-Lipschitz functions h_i that satisfy $h_i(X) = Y_i, h_1(x) + \ldots + h_n(x) = x$.

Theorems 1 and 2 imply that once ρ_i satisfy Assumptions 1-3, then without loss of generality one can indeed perform this minimization just over the set $\mathcal{C}(X) \triangleq \{\mathbf{Y} \in \mathcal{A}(X) : \mathbf{Y} \text{ comonotone}\}$ of comonotone allocations:

$$\inf_{\mathbf{Y}\in\mathcal{A}(X)} f(\mathbf{Y}) = \inf_{\mathbf{Y}\in\mathcal{C}(X)} f(\mathbf{Y}). \tag{6}$$

When Ω is non-atomic, (6) follows immediately from Theorem 2 combined with Assumption 2. In general, one can use the improvement of Theorem 1 together with Assumption 3 to obtain an ϵ-improvement with respect the risk measures:

Proposition 1. *If $\mathbf{Y} \in \mathcal{A}(X)$, then for any $\epsilon > 0$ there is a comonotone allocation $\bar{\mathbf{Y}} \in \mathcal{C}(X)$ such that*

$$\rho_i(\bar{Y}_i) \le \rho_i(Y_i) + \epsilon, \tag{7}$$

for all $i = 1, 2, \ldots, n$.

Clearly Proposition 1 is equivalent to (6). The proof is given in the Appendix. Note that one cannot directly "take the limit" in the discretization of Theorem 1 since $\le_{cx}$ is not necessarily stable under a.s.-limits.

An example of a family satisfying the above assumptions are the *distortion risk measures* (alternatively known as the law-invariant, comonotone-additive coherent risk measures). Denote by S_Y the (decumulative) distribution function of Y, that is, $S_Y(t) = \mathbb{P}(Y > t)$, and by S_Y^{-1} the (pseudo-)inverse of S_Y, which is unique up to a countable set (Denneberg 1994). For concreteness, take $S_Y^{-1}(p) = \sup\{t : S_Y(t) > p\}$; the inverse S_Y^{-1} thus defined is right continuous. Let $g : [0,1] \to [0,1]$ be a non-decreasing, concave function such that $g(0) = 0$, $g(1) = 1$. Take

$$\begin{aligned} \rho(Y) = \rho_g(Y) &= \int Y \, d(g \circ \mathbb{P}) = \int_0^1 S_Y^{-1}(p)\, dg(p) \\ &= \int_{-\infty}^0 \left(g[S_Y(t)] - 1\right)\, dt + \int_0^\infty g[S_Y(t)]\, dt. \end{aligned} \tag{8}$$

The function g is called a distortion because it modifies, or distorts, the tail probability S_Y before calculating an expectation. Note that if $g(p) = p$, then $\rho_g(Y) = \mathbb{E}Y$. For this reason, ρ_g is sometimes referred to an expectation with respect to a distorted probability.

Recall that for concave g, ρ_g is monotone, continuous at $-\infty$ and consistent with $\le_{cx}$ (Wang and Young 1998). Moreover, ρ_g is continuous with respect to a.s.-convergence. For $0 < p < 1$, let $\delta(p,\epsilon) > 0$ be such that $g(p+\eta) < g(p) + \epsilon$ for all $0 < \eta < \delta(p,\epsilon)$ (since g is uniformly continuous away from zero). Now define $\tilde{Y} \ge Y$ such that $\tilde{Y}$ takes on countably many values $\mathbf{y} = \{y_k\}_{k=1}^\infty$, $y_1 > -\infty$ and

$$S_Y(t) \le S_{\tilde{Y}}(t) \le S_Y(t) + \delta(S_Y(t), \epsilon \wedge \epsilon t^{-2}).$$

Without loss of generality, assume also that $S_{\tilde{Y}}(y_k) = S_Y(y_k)$ for all k. Because the distortion function g is concave, it is uniformly Lipschitz on $(\eta, 1]$ for any $\eta > 0$. Therefore, $\mathbf{y}$ forms a discrete set, i.e. $|y_{k'} - y_k| > \epsilon_2 > 0$ for any $k' \neq k$. Finally,

$$\begin{aligned}
\rho_g(Y) \leq \rho_g(\tilde{Y}) &= \int_{-\infty}^{0} [g(S_{\tilde{Y}}(t)) - 1]\, dt + \int_{0}^{\infty} g(S_{\tilde{Y}}(t))\, dt \\
&\leq \int_{-\infty}^{0} \left[g(S_Y(t)) + (\epsilon \wedge \epsilon t^{-2}) - 1\right] dt + \int_{0}^{\infty} \left[g(S_Y(t)) + (\epsilon \wedge \epsilon t^{-2})\right] dt \\
&\leq \rho_g(Y) + 4\epsilon.
\end{aligned}$$

where the second line follows by definition of $\delta(\cdot, \cdot)$.

Finally, Corollary 1 allows us to extend Proposition 1 to risk measures on $L^0(\mathbb{P})$ as long as the allocations are lower-bounded.

Proposition 2. *Let Ω be a non-atomic probability space and ρ_i be risk measures on $L^0(\mathbb{P})$ satisfying Assumptions 1-3. Suppose $X = \sum_{i=1}^{n} Y_i$, such that $Y_i \geq \underline{y}$ and $\rho_i(Y_i) < \infty$ for $i = 1, 2, \ldots, n$. Then there exists a comonotone allocation $\mathbf{Z}$, $\sum_{i=1}^{n} Z_i = X$, such that $\rho_i(Z_i) \leq \rho_i(Y_i)$, $i = 1, 2, \ldots, n$.*

Proof. Define $Y^{(m)}$ as in (4). By Corollary 1, there is a comonotone $\leq_{cx}$-improvement sequence of allocations $\mathbf{Z}^{(m)}$ of $Y^{(m)}$, such that $Z_i^{(m)} \geq \underline{y}$ and $Z_i^{(m)} \xrightarrow{a.s.} Z_i$ with $\mathbf{Z}$ comonotone. By Assumptions 1-2, for m large enough we have $\rho_i(Z_i^{(m)}) \leq \rho_i(Y_i^{(m)}) \leq \rho_i(Y_i)$ since $Y_i^{(m)} \leq Y_i$ once $m > |\underline{y}|$. By Assumption 3, $\rho_i(Z_i^{(m)}) \to \rho_i(Z_i) \leq \rho_i(Y_i)$.

□

Appendix

Proof of Proposition 1. Consider a given integrable allocation $\mathbf{Y}$, $Y_i \in L^1$. Fix $\epsilon > 0$. Using Assumption 3, construct a discrete $\tilde{Y}_i \geq Y_i$ such that

$$\rho_i(Y_i) \leq \rho_i(\tilde{Y}_i) \leq \rho_i(Y_i) + \epsilon.$$

Using Theorem 1, there is an $\leq_{cx}$-improvement $\mathbf{Y}^\epsilon$ of $\tilde{\mathbf{Y}}$. Unfortunately, $\mathbf{Y}^\epsilon$ is not directly comparable with $\mathbf{Y}$ since the sum of the risks is now too big, i.e. greater than X. The remaining steps show that one can further minorize $\mathbf{Y}^\epsilon$ such that the allocation remains comonotone and the sum is simply X.

First, we construct a comonotone improvement $\mathbf{Y}^\mathbf{c}$ of $\mathbf{Y}^\epsilon$ such that $\sum_i Y_i^c$ is comonotone with X. Let $\mathbf{y}$ be the support of $\sum_i Y_i^\epsilon$. Order $\mathbf{y} = \{\ldots < y_k < y_{k+1} < \ldots\}$. Since $\mathbf{Y}^\epsilon$ is comonotone, each Y_i^ϵ must be constant on an event $\{\omega \colon \sum_i Y_i^\epsilon = y_k\}$, say $Y_i^\epsilon = b_k^i$.

Define $B_k \triangleq \{\omega \colon y_{k-1} < X(\omega) \leq y_k\}$. Clearly, $\cup_k B_k = \Omega$ is a measurable partition. Now set $\mathbf{Y_i^c}(\omega) = b_k^i$, whenever $\omega \in B_k$. Then $\mathbf{Y}^\mathbf{c}$ is still comonotone and $\sum_i Y_i^c \geq X$

by construction. Moreover, $\mathbf{Y^c}$ and X are comonotone: consider $\omega_1 \in B_{k_1}$ and suppose $X(\omega_1) > X(\omega_2)$. If $\omega_2 \in B_{k_2}$ then necessarily $k_2 \leq k_1$ and $\sum_i Y_i^c(\omega_1) = y_{k_1} \geq y_{k_2} = \sum_i Y_i^c(\omega_2)$. Finally, $\mathbf{Y^c} \leq \mathbf{Y}^\epsilon$, since by construction $\mathbf{Y^c}$ is the smallest random variable with support in $\mathbf{y}$ that dominates X. For later use we also assume without loss of generality that $y_k = \operatorname{ess\,sup}_{\omega \in B_k} X(\omega)$.

The final step improves the allocation to $\bar{\mathbf{Y}}$, such that $\sum_i \bar{Y}_i = X$. Let $X' = \sum_i Y_i^c$. Since X and X' are comonotone, there exists a random variable Z such that $X = h(Z), X' = h'(Z)$ for some non-decreasing functions h, h'. Since a collection of comonotone random variables is also comonotone with their sum, there exist continuous increasing functions $\tilde{f}_i$ such that $\sum_i \tilde{f}_i(x) = x$ and $Y_i^c = \tilde{f}_i(X')$. Set $f_i = \tilde{f}_i \circ h'$, so that $Y_i^c = f_i(Z)$. Recall that each Y_i^c is discrete, so that f_i is piecewise constant and non-decreasing. The last step of the proof shows that one can construct non-decreasing $\bar{f}_i$'s such that $\bar{f}_i \leq f_i$ and $\sum_i \bar{f}_i = h \leq h' = \sum_i f_i$, which is geometrically intuitive.

Partition $\mathbb{R}$ into disjoint intervals $C_k \triangleq (z_k, z_{k+1})$ such that $\sum_i f_i$ is constant on each interval. Then without loss of generality there is an increasing sequence (a_k) such that on each interval, $a_k \leq h(z) \leq a_{k+1}$, $z_k \leq z \leq z_{k+1}$. Take $r_k = \sup\{z \in C_k \colon h(z) \leq (a_k + a_{k+1})/2\}$ to be the lower "half" of each C_k. By construction, $\sum_i b_k^i = \sup_{z \in C_k} h(z) = a_{k+1}$. Set $d_k^i = (b_{k+1}^i - b_k^i) \geq 0$, so that $\sum_i d_i^k = a_{k+1} - a_k > 0$. Define $\bar{f}_i(z) = b_k^i - \eta_k^i$ if $a_k \leq z \leq r_k$ and $\bar{f}_i(z) = b_k^i$ if $r_k < z \leq a_{k+1}$. Observe that the new family $\bar{f}_i$ is still non-decreasing, $\sum_i f_i \geq \sum_i \bar{f}_i(z) \geq h$ and $\| \sum_i \bar{f}_i - h\|_\infty \leq \frac{1}{2} \| \sum_i f_i - h\|_\infty$. This construction is illustrated in Figure 1. By repeating the argument, we obtain a monotonically decreasing sequence of comonotone allocations that converges pointwise. With a slight abuse of notation, let $\bar{\mathbf{Y}} = \bar{f}(Z)$ be the limiting allocation. Then the sum of $\bar{f}_i$ is h (i.e. $\sum_i \bar{Y}_i = X$), and $\bar{\mathbf{Y}}$ is still comonotone since each $\bar{f}_i$ is non-decreasing.

To conclude the proof observe that $\bar{\mathbf{Y}}$ is a comonotone allocation of X, and

$$\rho_i(\bar{Y}_i) \leq \rho_i(Y_i^\epsilon) \leq \rho_i(\tilde{Y}_i) \leq \rho_i(Y_i) + 4\epsilon, \tag{9}$$

for all $i = 1, 2, \ldots, n$, matching (7). □

References

Aase, K. K. (2002), Perspectives of risk sharing, *Scandinavian Actuarial Journal*, 2: 73–128.

Bäuerle, N. and A. Müller (2006), Stochastic orders and risk measures: Consistency and bounds, *Insurance: Mathematics and Economics*, 38(1): 132–148.

Burgert, C. and L. Rüschendorf (2006), On the optimal risk allocation problem, *Statistics and Decisions*, 24: 153-171.

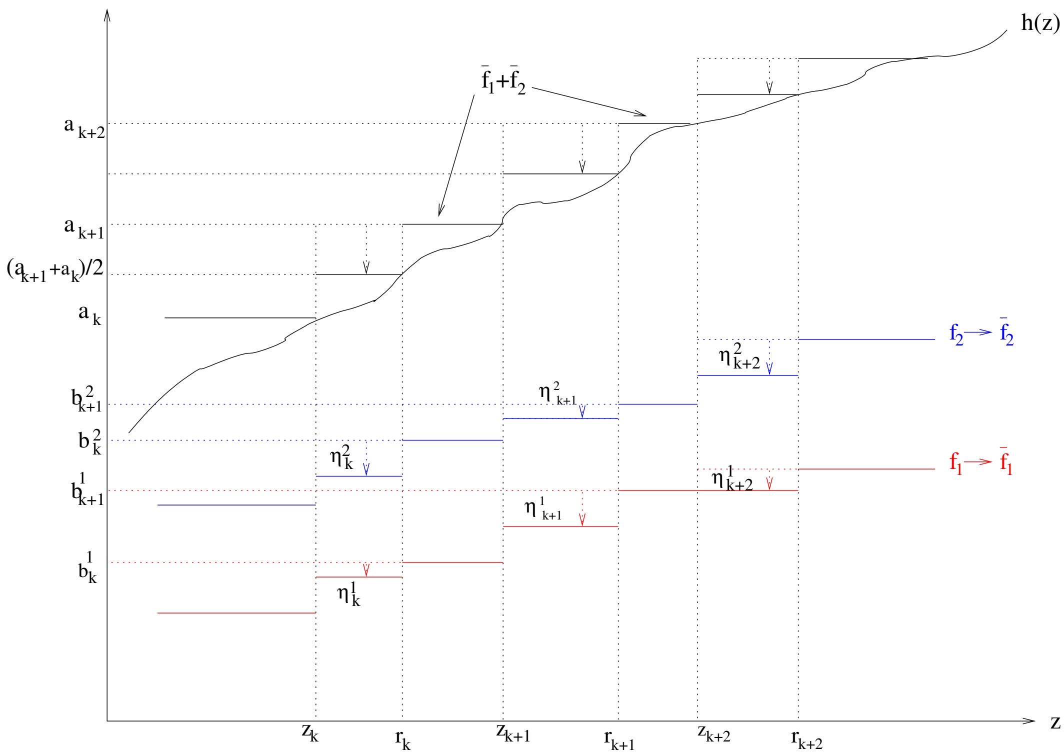

FIGURE 1. Third Step of Lemma 1 that improves the comonotone allocation described by (f_i) to $(\bar{f}_i)$ for the case $n = 2$.

Carlier, G. and R.-A. Dana (2003), Core of convex distortions of a probability, *Journal of Economic Theory*, 113:199–222.

Dana, R.-A. and I. Meilijson (2003), Modelling agents' preferences in complete markets by second order stochastic dominance, *Working Paper 0333*, Cahiers du CEREMADE.

Denneberg, D. (1994), *Non-Additive Measure and Integral*, Kluwer Academic Publishers, Dordrecht.

Diamond, P.A. and J.E. Stiglitz (1974), Increases in risk and in risk aversion, *Journal of Economic Theory*, 8(3):337–360.

Föllmer, H. and A. Schied (2004), *Stochastic Finance. An Introduction in Discrete Time*, 2nd Edition, Walter de Gruyter, Amsterdam.

Jouini, E., W. Schachermayer and N. Touzi (2005), Optimal risk sharing for law invariant monetary utility functions, *Working Paper*.

Landsberger, M. and I. Meilijson (1994), Co-monotone allocations, Bickel-Lehmann dispersion and the Arrow-Pratt measure of risk aversion, *Annals of Operations Research*, 52: 97-106.

Ludkovski, M. and V.R. Young (2007), Optimal risk sharing under distorted probabilities with transaction costs, *Working Paper*, University of Michigan.

Müller, A. and D. Stoyan (2002), *Comparison Methods for Stochastic Models and Risks*, Wiley Series in Probability and Statistics, New York.

Rothschild, M. and J. E. Stiglitz (1970), Increasing risk, I: a definition, *Journal of Economic Theory* 2: 225-243.

Rothschild, M. and J. E. Stiglitz (1971), Increasing risk, II: its economic consequences, *Journal of Economic Theory*, 3: 66-84.

Wang, S. S. and V. R. Young (1998), Ordering risks: Expected utility theory versus Yaari's dual theory of risk, *Insurance: Mathematics and Economics*, 22 (2): 145-161.

† Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109 USA

‡ Department of Mathematical Stochastics, University of Freiburg, D-79104 Freiburg, Germany

E-mail address: mludkov@umich.edu, ruschen@stochastik.uni-freiburg.de