THE CONVEXITY–CONES APPROACH TO COMPARATIVE RISK AND DOWNSIDE RISK

SALVATORE MODICA AND MARCO SCARSINI *

Introduction

What we call the convexity–cones approach consists of comparing two individuals' attitude towards (downside) risk by making them evaluate probability changes consisting of a joint increase in return —a first–order stochastic dominance shift— *and* in (downside) risk; that is, by confronting them with tradeoffs between risk and return. An individual is defined to be more (downside) risk averse than another if whenever he finds such a trade-off acceptable so does the other (that is, the other needs less return compensation for any given risk). Unlike in the classical approach to comparative risk aversion, here the test comparisons are always between two non-degenerate distributions; in the former they are bets against sure income (if the more risk averse accepts a bet so does the less risk averse).

For risk aversion (negative second derivative of the vNM utility function), Jewitt [7] has shown that the convexity–cones approach leads to the definition and characterization of Ross [12]. We report here that for downside risk aversion (positive third derivative) the same line is not fruitful, because of the non–decomposability of the dual of the relevant intersection of convexity cones.

Utility Space. We work with utilities in $C = C[0,1]$, the Banach space of continuous real functions on $[0,1]$ with supremum norm. Its dual space is $M = M[0,1]$, the space of Radon measures on $[0,1]$ (representable as functions of bounded variation on $[0,1]$, and including all distribution functions on $[0,1]$ as well as their differences); and the duality is $\mu(f) = \int f d\mu$ (cf. [3]).

Convexity Cones and their Duals. For $n \geq 0$, $C(1, x, x^2, \dots, x^n) \subseteq C$ will denote the cone of the $n+1$-st order convex functions on $[0,1]$. The complete definitions may be found in [1, 7, 9]; we will only need to know that $C(1)$ and $C(1, x)$ are the cones of non–decreasing and of convex functions respectively, and that $C(1, x, x^2)$ contains the functions with convex derivative. To get a general idea: in all of these cones the smooth functions are dense; and the smooth elements of $C(1, x, x^2, \dots, x^n)$ are the functions with positive $n+1$-st derivative ([9] ch.XI). We also let $C^-(1, x, x^2, \dots, x^n) = -C(1, x, x^2, \dots, x^n)$.

For a set of functions $K \subseteq C$, the set $K^* = \{\mu \in M : \mu(f) \geq 0 \forall f \in K\}$ is a convex cone, called the dual cone of K. Notice that all the convexity cones contain the constant functions; therefore any $\mu \in C(1, x, x^2, \dots, x^n)^*$ must have $\mu(1) = 0$, or $\int d\mu = 0$, and is thus interpretable as a probability change. Same applies to sums of duals of convexity cones. As to biduals, identifying the dual of M with C one defines $K^{**} = \{f \in C : \mu(f) \geq 0 \forall \mu \in K^*\}$ (it is the closed convex hull of K).

Risk

Ross–Jewitt. Recall that a probability change in the dual of a convexity cone is one which is favoured by all functions in the cone. A first–order stochastic

* March 2002. We gratefully acknowledge...

dominance shift is the type favoured by all non–decreasing functions, that is, is in $C(1)^*$; and an increase in risk is what the convex functions like, i.e. is in $C(1,x)^*$. For increasing functions, a trade–off between risk and return appears in the sum of the two, that is in a $\mu \in C(1)^* + C(1,x)^*$. Accordingly, the definition is: u is more risk averse than v if

$$\mu \in C(1)^* + C(1,x)^* \ \& \ \int u d\mu \geq 0 \ \Rightarrow \ \int v d\mu \geq 0. \tag{1}$$

Jewitt noticed that if in (1) μ is instead required to lie in the dual of a closed convex cone $\mathcal{C}$, then the assertion would become $v \in [\{u\} \cup \mathcal{C}]^{**}$; that is, $v(x) = a\, u(x) + w(x)$ for some $a \geq 0$ and $w \in \mathcal{C}$. He then pointed out that by a result of Amir and Ziegler [1], $C(1)^* + C(1,x)^*$ is indeed such a dual, precisely: $C(1)^* + C(1,x)^* = [C(1) \cap C(1,x)]^*$. Conclusion: *$u$ is more risk averse than v iff $v(x) = a\, u(x) + w(x)$ for some $a \geq 0$ and w increasing convex*. This is the characterization of Ross. For increasing utility functions, it is also equivalent, as is easily seen, to

$$\exists \lambda > 0 \ \ \forall x, y \ \ \frac{u''(x)}{v''(x)} > \lambda > \frac{u'(y)}{v'(y)}. \tag{2}$$

Asymmetry between Aversion and Attraction. The classical definition of u more risk averse than v, that is, "whenever u accepts a bet so does v", applies equally well to risk *loving* individuals; the property in quotes says that u is less risk attracted than v. This is not the case for the Ross-Jewitt definition; although the assertion (1) is still well defined, it yields no comparison of risk attitudes if u and v are increasing convex: if u is convex he will favour any μ as in (1), so the latter just says that also v is convex (increasing). As other side of the coin notice that the representation coming out of (1), $v(x) = a\, u(x) + w(x)$ with w increasing convex (may think of $u = x$ and $v = x + x^2$), gives no separation of derivative–ratios as in (2): computation yields that for increasing utilities, both ratios (of first and second derivatives) are above a.

To compare intensities of risk attraction in the Ross-Jewitt vein one clearly needs another trade-off, namely, increase in return versus *decrease* in risk: $\mu \in C(1)^* + C^-(1,x)^*$ (a decrease in risk is what the concave functions, i.e. those in $C^-(1,x)$, favour). As a consequence, for the resulting definition to yield a representation of the more risk lover v as $au + w$ with w convex one needs a separate result about duals of convexity cones, namely that the above sum be equal to the dual of $C(1) \cap C^-(1,x)$. And this holds (proof in last section):

Proposition 1. $C(1)^* + C^-(1,x)^* = [C(1) \cap C^-(1,x)]^*$.

One then defines v to be more risk attracted than u if

$$\mu \in C(1)^* + C^-(1,x)^* \ \& \ \int v d\mu \geq 0 \ \Rightarrow \ \int u d\mu \geq 0,$$

and the above proposition gives the correct characterization: *v is more risk attracted than u iff $v(x) = a\, u(x) + w(x)$ for some $a \geq 0$ and w <u>decreasing</u> convex* (to visualize may think of $u = x$, and $v = x^2$ up to $x = 1/2$ then straight). This is also equivalent to the right separation of derivative-ratios

$$\exists \lambda > 0 \ \ \forall x, y \ \ \frac{u''(x)}{v''(x)} < \lambda < \frac{u'(y)}{v'(y)}$$

(which is what you get when $v'' > 0$ if you back-of-the-enveolpe multiply the inequality $-u''/u' > -v''/v'$ by $-u'/v''$ and then separate).

Downside Risk

Identified by Menezes, Geiss and Tressler [10], an increase in downside risk is defined to be a mean–preserving spread coupled with a mean–preserving contraction occurring on its right, the pair leaving variance unchanged. In other words it is a mean-variance preserving transformation which shifts variability from the right to the left of a distribution (this is to be precise an 'elementary' change, a general one being given by a finite sum of such transormations). In [10] it is proved that all the functions with convex derivative *dislike* a probability change iff it is an increase in dowside risk; and accordingly, u is defined to be downside risk averse if it has convex derivative (if smooth: *positive* third derivative). In our terminology, the Menezes–Geiss–Tressler characterization says that a signed measure μ is an increase in downside risk according to their definition iff $\mu \in C^-(1,x,x^2)^*$ (what u dislikes is what $-u$ likes); and this will also be our definition. Of course the attraction counterpart is: μ is a decrease in downside risk if $\mu \in C(1,x,x^2)^*$.

Analogous to what we have seen in the previous section, the convexity–cones approach to comparative downside risk aversion, based on the risk–return tradeoff, is to define u to be more downside risk averse than v if

$$\mu \in C(1)^* + C^-(1,x,x^2)^* \;\&\; \int u d\mu \geq 0 \;\Rightarrow\; \int v d\mu \geq 0. \tag{3}$$

For risk attraction, the definition is: v is more downside risk attracted than u if

$$\mu \in C(1)^* + C(1,x,x^2)^* \;\&\; \int v d\mu \geq 0 \;\Rightarrow\; \int u d\mu \geq 0.$$

In both cases, one looks for a representation $v = au + w$ where w has concave derivative; and as before, in each case the representation hinges on the decomposability of the relevant cone. In the case of downside risk aversion for example, one would hope that $C(1)^* + C^-(1,x,x^2)^*$ be the dual of $C(1) \cap C^-(1,x,x^2)$; this would give $v = au + w$ with $a \geq 0$ and w increasing with concave derivative, and for increasing functions would also be equivalent to the 'right' separation of derivative ratios, namely

$$\exists \lambda > 0 \;\; \forall x, y \;\; \tfrac{u'''(x)}{v'''(x)} > \lambda > \tfrac{u'(y)}{v'(y)}.$$

Symmetrically, for risk attraction the result sought for is $C(1)^* + C(1,x,x^2)^* = [C(1) \cap C(1,x,x^2)]^*$, and the representation would be $v = au + w$ with w decreasing with concave derivative.

Unfortunately, while the latter result holds, the former, applying to the central case of downside risk aversion, does not. Indeed, it is proved in the next section that:

Proposition 2. $C(1)^* + C(1,x,x^2)^* = [C(1) \cap C(1,x,x^2)]^*$.

Proposition 3. $C(1)^* + C^-(1,x,x^2)^* \subsetneqq [C(1) \cap C^-(1,x,x^2)]^*$.

The content of the last proposition is that there are u more downside risk averse than v according to (3) which are not of the type $u = \alpha v + w$ where w is (decreasing) with convex derivative. In other words, the convex-derivative property of the (affine) transformation is stronger than u being more downside risk averse than v (cf. also Modica–Scarsini [11]). On the other hand, for *increasing* utility functions existence of the above transformation can be proved directly (cf. [11]).

Proofs

In all cases, the inclusion of the sum of duals in the dual of the intersection is a direct consequence of definitions; proofs are needed to show that a given μ in the dual of the intersection is decomposable in a sum. Most arguments in the sequel are applications of ideas of Amir–Ziegler [1] and Karlin–Studden [9]; some complications arise from dealing with non–consecutive cones, and from a non–differentiability problem in Proposition 1.

In each case the starting point is a useful characterization of the dual in terms of the extreme rays of the corresponding cone. And in each case, it can be checked by applying the approximation methods of [9] ch. XI that the smooth functions are dense in the relevant cone.

Basic notation: for $\mu \in M$, f a.e. continuous on $[0,1]$ and $t \in [0,1]$ let

$$P\mu(t) = \int_t^1 d\mu(x),\ Pf(t) = \int_t^1 f(x)dx, \text{ and } \quad Qf(t) = \int_0^t f(x)dx.$$

So we may write $P^2\mu(t) \equiv P(P\mu)(t) = \int_t^1 P\mu(x)dx$, and for $n > 2$ $P^n\mu(t) = \int_t^1 P^{n-1}\mu(x)dx$. Throughout, $\mathbf{1}_A$ will be the indicator of A.

Proof of Proposition 1. For $t \in [0,1]$ let $\tau(x;t)$ be the function of x which is equal to x up to t then constant:

$$\tau(x;t) = \mathbf{1}_{[0,t]}(x)\, x + \mathbf{1}_{(t,1]}(x)\, t, \quad x \in [0,1].$$

Lemma 1. $\mu \in [C(1) \cap C^-(1,x)]^*$ *iff*

$$\int_0^1 d\mu = 0 \quad \textit{and} \quad \int_0^1 \tau(x;t)d\mu(x) \geq 0,\ t \in [0,1]. \tag{4}$$

Proof of Lemma. We are talking about the cone of non–decreasing concave functions. So necessity is clear, since all constant functions belong to it (so $\int_0^1 d\mu = 0$ must hold), and so does $\tau(\cdot;t)$ for all t.

Sufficiency. Since $\tau(x;t) = \int_0^x \mathbf{1}_{[0,t]}(z)dz$, it is easily seen by interchanging order of integration that $\int_0^1 \tau(x;t)d\mu(x) = \int_0^t P\mu(x)dx = QP\mu(t)$. Assume μ satisfies the stated conditions; we want $\int \phi d\mu \geq 0$ for all $\phi \in C(1) \cap C^-(1,x)$, and by the denseness property mentioned before it suffices to take ϕ smooth. Then, noting that $P\mu(0) = \int_0^1 d\mu(x) = 0$,

$$\begin{aligned}\int_0^1 \phi(x)d\mu(x) &= -\int_0^1 \phi(x)dP\mu(x) = \int_0^1 \phi'(x)P\mu(x)d(x) \\ &= \int_0^1 \phi'(x)d(QP\mu)(x) = [\phi'(x)(QP\mu)(x)]_0^1 - \int_0^1 \phi''(x)(QP\mu)(x)dx.\end{aligned}$$

But $\phi' \geq 0$, $QP\mu \geq 0$, and $\phi'' \leq 0$, so this integral is non-negative. □

Now decomposition. We start with $\mu \in [C(1) \cap C^-(1,x)]^*$, and want $\mu = \mu_1 + \mu_2$ with $\mu_1 \in C(1)^*$ and $\mu_2 \in C^-(1,x)^*$. Characterizations of the latter duals are due to Karlin and Novikoff [8] (which is the starting point of this literature) and are reported in [1] (modulo a minus sign which is easily accounted for); for μ_1 and μ_2 they are

$$\begin{gathered} P\mu_1(0) = 0,\ P\mu_1(t) \geq 0\ \forall t \\ P\mu_2(0) = 0,\ P^2\mu_2(0) = 0, \text{ and } P^2\mu_2(t) \leq 0\ \forall t. \end{gathered} \tag{$\star$}$$

Since $\int_0^1 \tau(x;t)d\mu(x) = QP\mu(t) = P^2\mu(0) - P^2\mu(t)$, (4) may be written as

$$P\mu(0) = 0 \text{ and } P^2\mu(t) \leq A\ \forall t, \tag{4'}$$

where $A \equiv P^2\mu(0)$. Notice that $A = \int_0^1 x d\mu(x)$, so since the function x is increasing concave and $\mu \in [C(1) \cap C^-(1,x)]^*$, it must be $A \geq 0$. If $A = 0$, then by (4') $\mu \in C^-(1,x)^*$ (apply ($\star$) to μ), and decomposition is achieved with $\mu_1 \equiv 0$.

Take therefore μ satisfying (4') with $A > 0$. Given $\mu_2 = \mu - \mu_1$, the conditions ($\star$) are

$$P\mu_1(0) = P\mu(0) = 0,\ P^2\mu_1(0) = P^2\mu(0), \text{ and } P^2\mu_1(t) \geq P^2\mu(t)\ \forall t,$$

so finding μ_1 and μ_2 as wanted amounts to finding μ_1 satisfying:

$$P\mu_1(0) = 0,\ P\mu_1(t) \geq 0\,\forall t,\ \ P^2\mu_1(0) = A,\ P^2\mu_1(t) \geq P^2\mu(t)\,\forall t. \tag{5}$$

Now we follow Amir–Ziegler's lead (up to a point). If $F \in C$ is a smooth function such that $F(1) = DF(1) = 0$ (where D denotes derivative), then if μ_1 is defined by setting

$$\mu_1 = D^2F \tag{6}$$

(in the sense that $d\mu_1(x) = D^2F(x)dx$), it is easily verified that $P^2\mu_1 = F$ and $P\mu_1 = -DF$. Thus if μ_1 is defined by (6) for a smooth F with $F(1) = DF(1) = 0$, it will satisfy (5) iff F satisfies

$$F(0) = A,\ F(t) \geq P^2\mu(t)\,\forall t,\ DF(0) = 0,\ DF(t) \leq 0\,\forall t.$$

So we are looking for a smooth decreasing function dominating $P^2\mu$, equal to it at $t = 0, 1$ (to A and zero respectively), with zero derivative there; if such a function exists the proposition is proved.

If there is a left interval of $t = 1$ where $P^2\mu$ is negative such an F is found with no effort: construct it *backwardly* from $t = 1$ to $t = 0$ by making it smoothly reach height A within that interval, and from that point on have it constant $= A$ until t reaches zero.

So assume that in any left interval of $t = 1$ there are points where $P^2\mu$ is positive. Observe that $DP^2\mu(t) = -P\mu(t)$ if t is a continuity point of $P\mu$; in the case on hand one has to deal separately with the cases where $t = 1$ is/is not such a point.

Case $P\mu$ continuous at 1. Here the Amir–Ziegler argument still works: Given our current hypothesis, $DP^2\mu$ is zero at $t = 1$ and arbitrarily small (where it exists) in a sufficiently small left interval of it; so one can construct the wanted F, again *backwardly* from $t = 1$ to $t = 0$, by making it come out of 1 steep enough so that within this interval it stays above P^2 (still with $DF(1) = DP^2\mu(1) = 0$, but with DF negative and $< DP^2\mu$ in the interval), and also smoothly reaches height A; as before, from that point on we can have it constant $= A$ until t reaches zero, and we are done.

Case $P\mu$ not continuous at 1. In this case it may happen that $\lim_{t\to1^-} DP^2\mu(t) = -\lim_{t\to1^-} P\mu(t) < 0$, and then no twice differentiable function F with $F(1) = DF(1) = 0$ can stay above it just left of $t = 1$; so the Amir–Ziegler argument cannot be applied.

Recall that the the case where there is a left interval of $t = 1$ where $P^2\mu$ is negative has already been covered ($P\mu$ continuous or not). So we are dealing with the situation where $P\mu$ is not continuous at 1 and $\limsup_{t\to1^-} P\mu(t) > 0$.

In terms of $P\mu$ the conditions (4) or (4') characterizing our $\mu \in [C(1) \cap C^-(1, x)]^*$ with $A > 0$ may be written as (recall that $A = \int_0^1 P\mu(x)dx$)

$$P\mu(0) = P\mu(1) = 0,\ \int_0^1 P\mu(x)dx > 0,\ \text{and}\ \int_0^t P\mu(x)dx \geq 0\ \forall t. \qquad (7)$$

We are going to decompose this $P\mu$ into a sum, $P\mu = P_1 + P_2$, such that when we define μ_1 and μ_2 by

$$P_i(t) = \int_t^1 d\mu_i(x), \quad i = 1, 2$$

we obtain $\mu_1 \in C(1)^*$ and $\mu_2 \in C^-(1, x)^*$; since by construction $\mu = \mu_1 + \mu_2$, the proposition will be proved. To see what conditions we need on P_i notice that when μ_i is defined as above, one has

$$\forall n > 0\ \forall t \in [0, 1] \quad \int_t^1 (x - t)^n d\mu_i(x) = n \int_t^1 (x - t)^{n-1} P_i(x)dx. \qquad (8)$$

The respective conditions on μ_i are then seen to be (again cf. e.g. [1])

$$\begin{aligned} &P_i(t) = 0 \text{ for } t = 0, 1,\ i = 1, 2 \\ &P_1(t) \geq 0\ \forall t;\ \int_0^1 P_2(x)dx = 0,\ \int_0^t P_2(x)dx \geq 0\ \forall t. \end{aligned} \qquad (9)$$

Notice first that if $P\mu(t) \geq 0$ all t then $\mu \in C(1)^*$ and we are done; on the other hand it cannot be $P\mu \leq 0$ because it has positive intergal; so we will face a $P\mu$ oscillating between positive and negative (with $\limsup_{t\to1^-} P\mu(t) > 0$), with non-null positive and negative parts $(P\mu)^+, (P\mu)^-$. In this case also $\int_0^t P\mu(x)dx$

is oscillating, with $\lim_{t\to 1^-}\int_0^t P\mu(x)dx = \int_0^1 P\mu(x)dx > 0$. In the rest of the argument we write P for $P\mu$.

Since $P = P^+ - P^-$ and both P and P^- have positive integrals (the latter by right continuity), there exists $0 < \alpha < 1$ such that $\int_0^1 \big(\alpha P^+(x) - P^-(x)\big)dx = 0$. If for this α it is also $\int_0^t \big(\alpha P^+(x) - P^-(x)\big)dx \geq 0\,\forall t$, then we just set $P_1 = (1-\alpha)P^+$ and $P_2 = \alpha P^+ - P^-$, except maybe at the boundary points where we assign them vaule zero, and (9) holds. So assume that such a 'global' scaling down of P^+ cannot be done because the resulting P_2 would violate the condition $\int_0^t P_2(x)dx \geq 0\ \forall t$. The construction described below consists of scaling down P^+ 'locally'. We define $P_2 = P$ where the latter is negative; where P is positive, in some intervals P_2 will be again $= P$, in some others it will be a fraction of it (see below); and P_1 is defined as the difference $P - P_2$. Since by construction $P_2 \leq P$, it will then be $P_1 \geq 0$, and this is all we need of it (check (9)). So from now on we only speak of P_2; in fact only of P_2^+, because $P_2^- = P^-$.

If there are t such that $\int_0^t P(x)dx = 0$, take the largest of them, say t_0, define $P_2 = P$ on $[0, t_0]$, and begin the construction below from t_0. So in the following we assume wlog that $t_0 = 0$, i.e. that $\int_0^t P(x)dx > 0\,\forall t > 0$. Then for α sufficiently close to 1 it will still be $\int_0^t \big(\alpha P^+(x) - P^-(x)\big)dx > 0\,\forall t$. As α goes down, it will reach a value $\alpha_1 \in (0,1)$ where the above integral is zero for some t's. Let t_1 be the maximum of these t's (which exists because the integral is continuous), and define $P_2^+ = \alpha_1 P^+$ on $[0, t_1]$. Thus on this interval $\int_0^t P_2(x)dx \geq 0$, and $\int_0^{t_1} P_2(x)dx = 0$. If $t_1 = 1$, the proof ends (again by (9)). Otherwise, by construction of t_1 it is for all $t > t_1$

$$\int_{t_1}^t P(x)dx > \int_{t_1}^t \big(\alpha_1 P^+(x) - P^-(x)\big)dx = \int_0^t \big(\alpha_1 P^+(x) - P^-(x)\big)dx > 0.$$

Thus we can repeat the process just described for $[0,1]$ on the interval $[t_1, 1]$, and find α_2 and $t_2 > t_1$ such that defining $P_2^+ = \alpha_2 P^+$ on $(t_1, t_2]$ we have $\int_0^t P_2(x)dx \geq 0$ for $0 \leq t \leq t_2$, $\int_0^{t_1} P_2(x)dx = 0$; and $\int_{t_2}^t P(x)dx > 0$ all $t > t_2$. Iterating this process we get sequences (α_n) and (t_n) with the same properties we have just seen for $n = 1, 2$. Also, $t_n \to 1$ (as can be seen by contradiction); so P_2 is defined on all of $(0,1)$, and by construction $\int_0^t P_2(x)dx \geq 0\,\forall t$.

To complete the proof we must check that $\int_0^1 P_2(x)dx = \lim_{t\to 1}\int_0^t P_2(x)dx = 0$. Indeed for $\epsilon > 0$, there is t_ϵ such that for $t > t_\epsilon$ one has $\int_t^1 |P(x)|dx < \epsilon$; so take $t_n > t_\epsilon$, and for $t > t_n$ get $0 \leq \int_0^t P_2(x)dx = \int_{t_n}^t P_2(x)dx < \int_{t_n}^t |P(x)|dx < \epsilon$.

Proof of Proposition 2. Again start with $[C(1) \cap C(1, x, x^2)]^*$. The two characterizing families of extreme rays here are convex and concave increasing parts of parabolas; precisely, for $t \in [0,1]$ let

$$\phi_2(x;t) = \mathbf{1}_{[t,1]}(x)(x-t)^2/2 \quad\text{and}\quad \psi_2(x;t) = -\mathbf{1}_{[0,t]}(x)(x-t)^2/2,\ x \in [0,1].$$

Lemma 2. *$\mu \in [C(1) \cap C(1, x, x^2)]^*$ iff*

$$\int_0^1 d\mu(x) = 0, \quad \int_0^1 x d\mu(x) \geq 0,\ \textit{and}$$
$$\int_0^1 \phi_2(x;t)d\mu(x) \geq 0, \quad \int_0^1 \psi_2(x;t)d\mu(x) \geq 0, \quad t \in [0,1].$$

Proof of Lemma. Necessity is again direct. For sufficiency, observe that $\phi_2(x;t) = \mathbf{1}_{[t,1]}(x)\int_{x_1=t}^x \int_{x_2=t}^{x_1} dx_2 dx_1$ and $\psi_2(x;t) = -\mathbf{1}_{[0,t]}(x)\int_{x_1=x}^t \int_{x_2=x_1}^t dx_2 dx_1$; from this it easily follows that $\int_0^1 \phi_2(x;t)d\mu(x) = P^3\mu(t)$, and

$$\int_0^1 \psi_2(x;t)d\mu(x) = -\int_{x_2=0}^t \int_{x_1=0}^{x_2} \int_{x=0}^{x_1} d\mu(x)dx_1 dx_2 \equiv \Psi(t).$$

Now suppose μ satisfies the given conditions and integrate a smooth ϕ in the cone (want result non-negative); for this, observe that $\Psi'(t) = -\int_{x_1=0}^{t}\int_{x=0}^{x_1} d\mu(x)dx_1 = P^2\mu(0) - P^2\mu(t)$, last equality using $\int d\mu = 0$. Then obtain (using again $P\mu(0) = 0$)

$$\begin{aligned}
&\textstyle\int_0^1 \phi(x)d\mu(x) = -\int_0^1 \phi(x)dP\mu(x) = \int_0^1 \phi'(x)P\mu(x)dx\\
&= -\textstyle\int_0^1 \phi'(x)dP^2\mu(x) = \phi'(0)P^2\mu(0) + \int_0^1 \phi''(x)P^2\mu(x)dx\\
&= \phi'(0)P^2\mu(0) + \textstyle\int_0^t \phi''(x)[P^2\mu(0) - \Psi'(x)]dx - \int_t^1 \phi''(x)dP^3\mu(x)\\
&= \phi'(t)P^2\mu(0) - \textstyle\int_0^t \phi''(x)d\Psi(x) - \int_t^1 \phi''(x)dP^3\mu(x)\\
&= \phi'(t)P^2\mu(0) - \phi''(t)\big(\Psi(t) - P^3\mu(t)\big) + \textstyle\int_0^t \phi^{(3)}(x)\Psi(x)dx + \int_t^1 \phi^{(3)}(x)P^3\mu(x)dx.
\end{aligned}$$

By positivity and the endpoint conditions at $t = 0, 1$ of $P^3\mu$ and Ψ, there exists t such that $\Psi(t) - P^3\mu(t) = 0$; take t equal to this value in the last expression above, and you are left with only positive terms (recall that $P^2\mu(0) = \int_0^1 x d\mu(x) \geq 0$). □

Now decomposition: we start with $\mu \in [C(1)\cap C(1, x, x^2)]^*$, and want $\mu = \mu_1 + \mu_2$ with $\mu_1 \in C(1)^*$ and $\mu_2 \in C(1, x, x^2)^*$; again from [8] or [1], the conditions on μ_1 and μ_2 are then

$$P\mu_1(0) = 0, \quad P\mu_1(t) \geq 0,\ t \in (0, 1) \tag{10}$$

$$P\mu_2(0) = 0,\ P^2\mu_2(0) = 0,\ P^3\mu_2(0) = 0,\ P^3\mu_2(t) \geq 0,\ t \in (0, 1), \tag{11}$$

where recall that $P^2\mu_2(0) = \int x d\mu_2(x)$ and $P^3\mu_2(t) = 2^{-1}\int_t^1 (x - t)^2 d\mu_2(x)$. Let again $A = P^2\mu(0)$. Given $P\mu(0) = 0$, it easily checked that

$$\Psi(t) = P^3\mu(t) + At - P^3\mu(0). \tag{12}$$

From the lemma $A \geq 0$; suppose $A = 0$. Since the last two conditions on μ in the lemma assert non-negativity of $P^3\mu$ and Ψ on $[0, 1]$, and equation (12) (with $t = 1$) and $A = 0$ then implies $P^3\mu(0) \leq 0$, we may conclude that if $A = 0$ then also $P^3\mu(0) = 0$. But in this case $\mu \in C(1, x, x^2)^*$ (apply (11) to μ), and decomposition obtains with $\mu_1 \equiv 0$.

So assume $A > 0$. Then, since $(d/dt)P^3\mu(0) = -P^2\mu(0) < 0$ and $P^3\mu(t) \geq 0$ all t, it must be $P^3\mu(0) > 0$. Letting $B = P^3\mu(0)$, and using (12), the conditions on μ in lemma 2 can then be written as

$$\begin{gathered}
P\mu(0) = 0,\ P^2\mu(0) = A > 0,\ P^3\mu(0) = B > 0,\\
P^3\mu(t) \geq 0,\ P^3\mu(t) \geq B - At,\ t \in [0, 1].
\end{gathered} \tag{13}$$

On the other hand, using $\mu_2 = \mu - \mu_1$ and writing as in the previous proposition the conditions (11) in terms of μ and μ_1 we obtain conditions on μ_1 equivalent to (10) and (11), which in the present case are

$$\begin{gathered}
P\mu_1(0) = 0,\ P\mu_1(t) \geq 0\ \forall t,\ P^2\mu_1(0) = A,\\
P^3\mu_1(0) = B,\ P^2\mu_1(t) \leq P^3\mu(t)\ \forall t.
\end{gathered} \tag{14}$$

To sum up, we start with (13) and look for a μ_1 satisfying (14). Here adaptation of the Amir–Ziegler's line leads to the result. The starting point is again the fact that for any $f \in C$ differentiable j times (denoting $D^0 f = f$),

$$\text{if } D^i f(1) = 0,\ i = 0, 1, \ldots, j - 1,\ \text{then } P^j D^j f = (-1)^j f.$$

So if for an F which at $t = 1$ is zero with its first two derivatives we define

$$\mu_1 = -D^3 F, \tag{15}$$

then $P^3\mu_1 = F$, and also $P\mu_1 = D^2F$ and $P^2\mu_1 = -DF$. So if $F \in C$ has $F(1) = DF(1) = D^2F(1) = 0$, then μ_1 defined via (15) satisfies (14) iff F satisfies

$$F(0) = B,\ F(t) \le P^3\mu(t)\ \forall t,$$
$$DF(0) = -A,\ D^2F(0) = 0,\ D^2F(t) \ge 0\ \forall t. \tag{16}$$

Notice that at $t = 0$, F is required to be equal to $P^3\mu$ with its first two derivatives; and at $t = 1$ it is required to be equal to $P^3\mu$ with its first derivative. For the rest, F has to be convex and dominated by $P^3\mu$. If an F constant on a left neighbourhood of $t = 1$ and satisfying (16) exists, the proposition is proved (with μ_1 defined by (15)).

We already observed that $\Psi'(t) = P^2\mu(0) - P^2\mu(t)$; then $\Psi'(1) = A > 0$ which, together with $\Psi(t) \ge 0$ all t, implies $\Psi(1) > 0$; so from (12) letting $t = 1$ we get $A > B$. Hence there is $t_0 \in (0,1)$ such that $B - At_0 = 0$. Define

$$F_1(t) = \mathbf{1}_{[0,t_0]}(t)\big(B - At\big).$$

This F_1 satisfies all of (16) (convexity replacing $D^2F \ge 0$) except smoothness at t_0. To smooth it around t_0 (and end up with a function still below $P^3\mu$) we need to exclude that $P^3\mu(t_0) = 0$. But if this were the case, by smoothness of $P^3\mu$ we would have $P^3\mu(t) < B - At$ on a left neighbourhood of t_0, contradicting the last requirement of (13). Therefore we can smooth F_1 around t_0 (as done in Amir–Ziegler for example) to get the wanted F, and the proof is complete.

Proof of Proposition 3. To characterize the dual of $C(1) \cap C^-(1, x, x^2)$ define the following family of extreme rays for $t \in (0, 1]$ (parts of parabolas joined at t, increasing first convex then concave):

$$\begin{aligned}\pi_2(x;t) &= \mathbf{1}_{[0,t)}(x)\,\tfrac{1-t}{t}\tfrac{x^2}{2} + \mathbf{1}_{[t,1]}(x)\,\tfrac{-t+x(2-x)}{2}\\ &= \int_{x_1=0}^{x}\int_{x_2=0}^{x_1}\big[\tfrac{1-t}{t}\mathbf{1}_{[0,t)}(x_2) - \mathbf{1}_{[t,1]}(x_2)\big]\,dx_2\,dx_1.\end{aligned}$$

The last equality is elementarily checked, and it easily gives

$$\int_0^1 \pi_2(x;t)\,d\mu(x) = \tfrac{1}{t}\big[(1-t)B - P^3\mu(t)\big] \equiv \tfrac{1}{t}H(t), \tag{17}$$

where as before $B = P^3\mu(0)$. Incidentally, it will be again $A = P^2\mu(0)$.

Lemma 3. *$\mu \in \big[C(1) \cap C^-(1, x, x^2)\big]^*$ iff*

$$\int_0^1 d\mu(x) = 0,\ \int_0^1 x^2 d\mu(x) \ge 0,\ \int_0^1 -(x-1)^2 d\mu(x) \ge 0,\ \textit{and}$$
$$\int_0^1 \pi_2(x;t)\,d\mu(x) \ge 0,\ t \in (0,1).$$

Proof. Necessity is obvious by construction, and for sufficiency take a smooth $\phi \in C(1) \cap C^-(1, x, x^2)$ and integrate. The first equality below is obtained by using as usual $d\mu(x) = -dP\mu(x)$, $P\mu(x)dx = -dP^2\mu(x)$ and $P\mu(0) = 0$; then we use the fact that for the H defined in (17) it is $dH(t) = \big[P^2\mu(t) - B\big]dt$, and $H(0) = H(1) = 0$:

$$\begin{aligned}\int_0^1 \phi(x)\,d\mu(x) &= -\int_0^1 \phi'(x)\,dP^2\mu(x)\\ &= \big[-\phi'(x)P^2\mu(x)\big]_0^1 + \int_0^1 \phi''(x)\big[dH(x) + B\,dx\big]\\ &= \phi'(0)P^2\mu(0) + B\int_0^1 \phi''(x)\,dx + \big[\phi''(x)H(x)\big]_0^1 - \int_0^1 \phi^{(3)}(x)H(x)\,dx\\ &= \phi'(0)A + B[\phi'(1) - \phi'(0)] - \int_0^1 \phi^{(3)}H(x)\,dx\\ &= \phi'(1)B + \phi'(0)\big[A - B\big] - \int_0^1 \phi^{(3)}(x)H(x)\,dx.\end{aligned}$$

Now: $B = P^3\mu(0) = \int_0^1 \frac{x^2}{2}d\mu(x) \ge 0$ by hypothesis; and $A = P^2\mu(0) = \int_0^1 x\,d\mu(x)$, so using $\int_0^1 d\mu(x) = 0$ we get $A - B = \int_0^1 \frac{x(2-x)}{2}d\mu(x) = \int_0^1 \big(\frac{x(2-x)}{2} - \frac{1}{2}\big)d\mu(x) =$

$\int_0^1 -\frac{(x-1)^2}{2} d\mu(x) \geq 0$ by hypothesis; always by assumption, H, ϕ' and $\phi^{(3)}$ are also non-negative. Result then follows. [1] □

Now non–decomposability. We show that there exists $\mu \in \left[C(1) \cap C^-(1,x,x^2)\right]^*$ which admits no representation of the form $\mu = \mu_1 + \mu_2$ with $\mu_1 \in C(1)^*$ and $\mu_2 \in C^-(1,x,x^2)^*$. The conditions for $\mu_1 \in C(1)^*$ are still those in (10); those on μ_2 are found from (11) by observing that $\mu_2 \in C^-(1,x,x^2)^*$ iff $-\mu_2 \in C(1,x,x^2)^*$, hence they are

$$P\mu_2(0) = P^2\mu_2(0) = P^3\mu_2(0) = 0, \text{ and } P^3\mu_2(t) \leq 0,\ t \in (0,1).$$

As before using $\mu_2 = \mu - \mu_1$ and rewriting, we find that μ admits representation of the wanted type iff there exists μ_1 satisfying

$$\begin{aligned} &P\mu_1(0) = 0,\ P\mu_1(t) \geq 0\ \forall t \in (0,1), \\ &P^2\mu_1(0) = A,\ P^3\mu_1(0) = B, \text{ and } P^3\mu_1(t) \geq P^3\mu(t)\ \forall t \in (0,1). \end{aligned} \tag{18}$$

We now rewrite the conditions on μ in lemma 3. Given $\int_0^1 d\mu(x) = 0$, the condition $\int_0^1 -(x-1)^2 d\mu(x) \geq 0$ amounts to $A \geq B$; also, $\int_0^1 x^2 d\mu(x) = 2B$; finally, using (17) it is seen that the last condition in the lemma is $P^3\mu(t) \leq B(1-t)$ all $t \in (0,1)$, which obviously holds also for $t = 0, 1$. Hence we conclude that $\mu \in \left[C(1) \cap C^-(1,x,x^2)\right]^*$ iff

$$\int_0^1 d\mu(x) = 0,\ A \geq B \geq 0, \text{ and } P^3\mu(t) \leq B(1-t),\ t \in [0,1], \tag{19}$$

where recall $A = P^2\mu(0) = \int_0^1 x d\mu(x)$ and $B = P^3\mu(0) = \int_0^1 \frac{x^2}{2} d\mu(x)$.

If $A = 0$ then also $B = 0$, and decomposition does obtain trivially, with $\mu_1 \equiv 0$ (easy to check). We now give an example of μ satisfying (19), with $A > 0$ and $B = 0$, which cannot be decomposed. Recalling from page 6 that $P^3\mu(t) = \int_0^1 \phi_2(x;t) d\mu(x) = \int_t^1 \frac{(x-t)^2}{2} d\mu(x)$, conditions (19) in the present case read

$$\int_0^1 d\mu(x) = 0,\ \int_0^1 x d\mu(x) > 0, \int_0^1 x^2 d\mu(x) = 0, \text{ and } \int_t^1 (x-t)^2 d\mu(x) \leq 0\ \forall t. \tag{20}$$

Again we broadly follow the line Amir–Ziegler employed for their counterexample (to the decomposability of the dual of the intersection of four consecutive convexity cones), and define μ via a particular function P by setting

$$P(t) = \int_t^1 d\mu(x),\ t \in [0,1]. \tag{21}$$

Applying (8) to μ and P we see that μ defined by (21) satisfies (20) (and so is in the given dual) iff P is such that

$$P(0) = P(1) = 0,\ \int_0^1 P(x)dx > 0,\ \int_0^1 xP(x)dx = 0,\ \int_t^1 (x-t)P(x)dx \leq 0\ \forall t. \tag{22}$$

Define $R(t) = \int_t^1 (x-t)P(x)dx$. Then

$$R'(t) = -\int_t^1 P(x)dx = \int_0^t P(x)dx - A, \tag{23}$$

and if $\int_0^1 xP(x)dx = 0$ also $R(0) = 0$. P is defined as follows (a tent with basis $[0, 1/4]$, then 0 up to $t = 3/4$, then an upside–down tent with basis $[3/4, 1]$ and height 1/7 of the other):

$$P(t) = \mathbf{1}_{[0,\frac{1}{8}]}(t)\, t + \mathbf{1}_{(\frac{1}{8},\frac{1}{4}]}(t) \left(-t + \tfrac{1}{4}\right) + \mathbf{1}_{[\frac{3}{4},\frac{7}{8}]}(t)\, \tfrac{1}{7} \left(\tfrac{3}{4} - t\right) + \mathbf{1}_{(\frac{7}{8},1]}(t)\, \tfrac{1}{7} (x-1)\,.$$

This P is elementarily seen to satisfy all but the last of (22). For the latter, which is $R(t) \leq 0$ all t, integrate P from 0 to t, plot, and shift down by A; this function, which by (23) is $R'(t)$, increases on $[0, 1/4]$ from $-A$ to a positive value, then remains constant up to $t = 3/4$, then decreases to zero, which it reaches at

[1] Notice that although the condition $\int_0^1 x d\mu(x) \geq 0$ does not appear in the lemma, we have seen that it does hold: $\int_0^1 x d\mu(x) = A \geq B \geq 0$.

$t = 1$. But $\int_0^t R'(x)dx = R(t) - R(0) = R(t)$, so $R(t)$ decreases on $[0, 1/8]$, then increases; since $R(0) = R(1) = 0$, it must be $R(t) \leq 0$ all t.

The conclusion then is that the P defined above satisfies (22); thus the μ defined via P by (21) is in $\left[C(1) \cap C^-(1, x, x^2)\right]^*$. To finish the proof we shall show that it is not decomposable. Suppose it were, with $\mu = \mu_1 + \mu_2$, $\mu_1 \in C(1)^*$ and $\mu_2 \in C^-(1, x, x^2)^*$; then $\int_t^1 d\mu = \int_t^1 d\mu_1 + \int_t^1 d\mu_2 \equiv P_1 + P_2$; and if ν_i is defined by $d\nu_i(x) = P_i(x)dx$, $i = 1, 2$, it follows from (8) and the characterizations of the relevant duals (for the last time, cf. [1] or [9]) that

$$\nu_1 \in (C^+)^*, \ \nu_2 \in C^-(1, x)^*,$$

C^+ being the cone of positive functions. Therefore their densities P_1 and P_2 should in particular satisfy

$$P_1(t) \geq 0 \ \forall t, \ \int_0^1 P_2(x)dx = 0, \ \int_0^1 xP_2(x)dx = 0. \tag{24}$$

This and $P = P_1 + P_2$ give $\int_0^1 xP(x)dx = \int_0^1 xP_1(x)dx$; and the integral on the left is zero by (22), thus by non-negativity and right continuity of P_1 we conclude that (24) implies that the latter is identically zero. On the other hand $\int_0^1 P_2(x)dx = 0$ implies $\int_0^1 P_1(x)dx = \int_0^1 P(x)dx$, which is strictly positive by (22). A contradiction has been reached.

References

[1] Amir, D. and Z. Ziegler, (1968): Generalized Convexity Cones and their Duals, *Pacific Journal of Mathematics* **27**: 425–440

[2] Border, K.C. (1991): Functional Analytic Tools for Expected Utility Theory, in Aliprantis C.D. K.C. Border and W.A.J. Luxemburg editors, *Positive Operators, Riesz Spaces and Economics*, Springer–Verlag: 79–88

[3] Edwards, R.E. (1965): *Functional Analysis Theory and Applications*, Dover edition 1995

[4] Gollier, C. and Pratt, J.W. (1996): Risk Vulnerability and the Tempering Effect of Background Risk, *Econometrica* **64**: 1109–1123

[5] Hanson, D.L. and Menezes, C.F. (1971): On a Neglected Aspect of the Theory of Risk Aversion, *Western Economic Journal* **9**: 211–217

[6] Hicks, J. (1967): *Critical Essays in Monetary Theory*, Oxford University Press

[7] Jewitt, I. (1986): A Note on Comparative Statics and Stochastic Dominance, *Journal of Mathematical Economics* **15**: 249–254

[8] Karlin, S. and A. Novikoff, (1963): Generalized Convex Inequalities, *Pacific Journal of Mathematics* **13**: 1251–1279

[9] Karlin, S. and W.J. Studden, (1966): *Tchebycheff Systems with Applications in Analysis and Statistics*, Wiley Interscience

[10] Menezes C. Geiss C. and J. Tressler (1980): Increasing Downside Risk, *American Economic Review* **70**: 921–932

[11] Modica, S. and M. Scarsini (2002): Downside Risk Aversion in the Small and the Large, mimeo

[12] Ross, S.A. (1981): Some Stronger Measures of Risk Aversion in the Small and the Large with Applications, *Econometrica* **49**: 621–638

Salvatore Modica: Facoltà di Economia, Università di Palermo
E-mail address: `modica@unipa.it`

Marco Scarsini: Facoltà di Economia, Università di Pescara
E-mail address: `scarsini@sci.unich.it`