Caltech Division of the Humanities and Social Sciences

Pedantic Notes on the Risk Premium and Risk Aversion*

KC Border

May 1996
v. 2018.11.30::14.26

1 Preliminaries

The **risk premium** $\pi_u(w, Z)$ for an expected utility decision maker with Bernoulli utility function u is defined implicitly by the equation

$$u(w + \boldsymbol{E}\, Z - \pi_u(w, Z)) = \boldsymbol{E}\, u(w + Z). \qquad (\star)$$

It is the most that they would be willing to pay to completely insure against the risk Z to their initial wealth w.

The first obvious question is, how do we know that such π exists, and if it does exist how do we know that it is unique? Clearly, in order for equation $(\star)$ to make sense, we must place some restrictions on the random variable Z. Let us say that a random variable Z is **admissible for u at w** if $w + Z$ takes on values in the domain of u almost surely, both $\boldsymbol{E}\, Z$ and $\boldsymbol{E}\, u(w + Z)$ are finite, and $w + \boldsymbol{E}\, Z$ belongs to the domain of u. If the domain is an interval, this latter requirement follows from the others. Let us say that (w, Z) is an **admissible pair for u** if Z is admissible for u at w.

1 Lemma *Let D be an interval of the real line. Let u be a continuous strictly increasing real function on D. Suppose (w, Z) is an admissible pair for u. Then equation $(\star)$ is satisfied by a unique value $\pi_u(w, Z)$.*

Proof: Since u is continuous, the Intermediate Value Theorem (e.g., [2, Theorem 3.8, p. 144]) implies that the range of u is an interval. Moreover, the range contains the number $\boldsymbol{E}\, u(w+Z)$. It then follows that there is some π so that equation $(\star)$ is satisfied. If u is strictly increasing, then π must also be unique. ∎

So henceforth we shall require:

2 Assumption *Assume that the utility function is continuous, strictly increasing, and defined on an interval of $\boldsymbol{R}$.*

*I wrote these notes to reproduce some of the results in Pratt's [9] paper in my own style, since I was never facile with Landau's o and O notation. It also bothered me that Pratt assumed that the utility functions are "sufficiently regular to justify the proofs" (p. 123, line 2 from bottom). I wish my referees would let me make assumptions like that.

Recall that a concave function is automatically continuous on the interior of its domain, so if u is concave, strictly increasing, and defined on an open interval, then $\pi_u(w, Z)$ is unique.

A point worth noting is that subjecting the utility u to a positive affine transformation does not affect the validity of equation ($\star$), which is a good thing because it means that the risk premium is a property of the preference relation, not the particular Bernoulli utility function.

Another point worth mentioning is that if Z is admissible for u at w, then $Z - \boldsymbol{E}\, Z$ is admissible for u at $w + \boldsymbol{E}\, Z$, and

$$\pi_u(w, Z) = \pi_u(w + \boldsymbol{E}\, Z, Z - \boldsymbol{E}\, Z),$$

so without loss of generality we need only consider admissible pairs (w, Z) where $\boldsymbol{E}\, Z = 0$.

It is immediate that strictly increasing functions are one-to-one and so have inverses. The next easy lemma gives simple conditions for the continuity of the inverse.

3 Lemma *Let u be a strictly increasing real function on the domain D in $\boldsymbol{R}$. If D is either closed or an interval, then u^{-1} is continuous on the range of u.*

Proof: To see this, suppose $u(x_n) \to u(x)$. We wish to show that $x_n \to x$. By passing to a subsequence we may assume without loss of generality that either $u(x_n) \uparrow u(x)$ or $u(x_n) \downarrow u(x)$. Let's consider the case $u(x_n) \uparrow u(x)$. Since u is strictly increasing, the sequence x_n is increasing and bounded above by x. It thus has a limit y satisfying $x_1 \leqslant y \leqslant x$. If D is either closed or an interval, then $y \in D$. Now if $y < x$, then $u(x_n) \leqslant u(y) < u(x)$ for all n, which contradicts $u(x_n) \to u(x)$. Therefore $x_n \to y = x$, which shows the continuity of u^{-1}. The case of $u(x_n) \downarrow u(x)$ is disposed of similarly. ∎

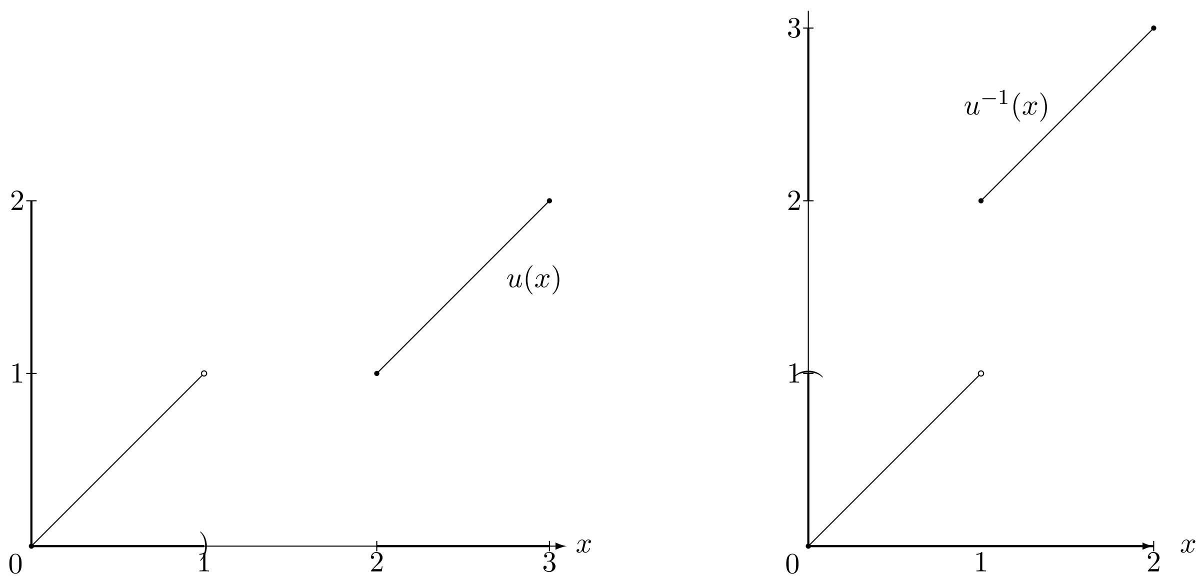


Figure 1. Discontinuity of the inverse.

4 Example (Continuous and discontinuous inverses) Note that the result above does not require continuity of u, but it can fail if the domain D is neither closed nor an interval. For instance, set $D = [0,1) \cup [2,3]$ and define u by $u(x) = x$ for x in $[0,1)$ and $u(x) = x-1$ for x in $[2,3]$. Then u is continuous, strictly increasing, and its range is the interval $[0,2]$. Nevertheless u^{-1} is not continuous, it jumps up at 1. See Figure 1.

Now flip the picture over and set $v = u^{-1}$. Then v is strictly increasing but not continuous, and its domain is the closed interval $[0,2]$. Observe that its inverse u is continuous on the range $[0,1) \cup [2,3]$ of v. □

We shall be interested in comparing different utility functions u and v defined on the same domain D. The next lemma will be extremely useful. It says that we can write either utility as a strictly increasing function of the other. Recall that a function is of class C^k on D if it is continuous on D and has a continuous k^{th}-order derivative on the interior of D.

5 Lemma *Let u and v be strictly increasing real functions defined on the common interval D. Then the function g defined on the range of v by*

$$g(x) = u(v^{-1}(x)).$$

is the unique function satisfying

$$u = g \circ v.$$

Furthermore g is strictly increasing.

If in addition, u is continuous, then g is continuous too.

If v is continuous, then the range of v, and hence the domain of g, is an interval.

If u and v are both C^k and v' is never zero, then g is C^k on its domain.

Proof: Since v is strictly increasing, it has an inverse v^{-1}, which is also strictly increasing, so g exists and is the unique function satisfying $u = g \circ v$. Clearly g is strictly increasing, being the composition of two strictly increasing functions.

Since v is strictly increasing, then by Lemma 3 v^{-1} is continuous on the range of v. So if u is continuous too, then $g(x) = u(v^{-1}(x))$ is continuous.

If v is continuous, as we remarked above, the Intermediate Value Theorem implies that its range is an interval.

If v is C^k and v' is never 0, the Inverse Function Theorem (see, e.g., [7, Theorem 1, p. 230]) implies that v^{-1} is also C^k. Thus g as the composition of two C^k functions is also C^k. ∎

2 Global results

The main theorem in the area is due to Pratt [9]. The next few theorems refine the published results.

6 Theorem *Let D be an interval of the real line. Let u and v be continuous strictly increasing real functions on D. Then the following statements are equivalent.*

(1) *For all w in D and random variables Z that are admissible for u and v at w and satisfy $\boldsymbol{E}\, Z = 0$,*

$$\pi_u(w, Z) \geqslant \pi_v(w, Z).$$

(2) *There exists a concave strictly increasing function g defined on the range of v satisfying*

$$u = g \circ v.$$

Proof: (1) $\implies$ (2): Lemma 5 gives the existence of the strictly increasing function $g = u \circ v^{-1}$, so the crux is to prove that g is concave. That is, for v_1, v_2 in the range of v, $g\big(\lambda v_1 + (1-\lambda)v_2\big) \geqslant \lambda g(v_1) + (1-\lambda)g(v_2)$ for $0 \leqslant \lambda \leqslant 1$. Note that Lemma 5 implies that the domain of g is an interval and hence a convex set, so g is actually defined for each $\lambda v_1 + (1-\lambda)v_2$.

To this end, let w_1, w_2 in D satisfy

$$v(w_1) = v_1 \quad \text{and} \quad v(w_2) = v_2,$$

and set

$$w = \lambda w_1 + (1-\lambda)w_2.$$

Define the random variable Z so that

$$Z = \begin{cases} (1-\lambda)(w_1 - w_2) & \text{with probability} \lambda \\ \lambda(w_2 - w_1) & \text{with probability} 1-\lambda. \end{cases}$$

Then $\boldsymbol{E}\, Z = 0$ and

$$w + Z = \begin{cases} w_1 & \text{with probability} \lambda \\ w_2 & \text{with probability} 1-\lambda. \end{cases}$$

Consequently, Z is admissible for u and v at w, and

$$\begin{aligned} u\big(w - \pi_u(w,Z)\big) &= \boldsymbol{E}\, u(w+Z) \\ &= \boldsymbol{E}\, g\big(v(w+Z)\big) \\ &= \lambda g(v_1) + (1-\lambda)g(v_2) \end{aligned}$$

and

$$\begin{aligned} u\big(w - \pi_v(w,Z)\big) &= g\Big(v\big(w - \pi_v(w,Z)\big)\Big) \\ &= g\big(\boldsymbol{E}\, v(w+Z)\big) \\ &= g\big(\lambda v_1 + (1-\lambda)v_2\big). \end{aligned}$$

But u is increasing and $\pi_u(x,Z) \geqslant \pi_v(x,Z)$, so $u\big(w - \pi_v(w,Z)\big) \geqslant u\big(w - \pi_u(w,Z)\big)$, so we conclude that

$$g\big(\lambda v_1 + (1-\lambda)v_2\big) \geqslant \lambda g(v_1) + (1-\lambda)g(v_2)$$

which proves that g is indeed concave.

(2) $\implies$ (1): This is clearly a job for Jensen's inequality. Let Z be admissible for u and v at w and satisfy $\boldsymbol{E}\, Z = 0$. Observe that

$$\begin{aligned} u\big(w - \pi_u(w,Z)\big) &= \boldsymbol{E}\, u(w+Z) \\ &= \boldsymbol{E}\, g\big(v(w+Z)\big) \\ &\leqslant g\big(\boldsymbol{E}\, v(w+Z)\big) \\ &= g\Big(v\big(w - \pi_v(w,Z)\big)\Big) \\ &= u\big(w - \pi_v(w,Z)\big), \end{aligned}$$

where the inequality follows from Jensen's inequality and the concavity of g, and the other equalities are either from the definition of the risk premium or the assumption that statement (2) holds.

Since u is strictly increasing it follows that $\pi_u(w, Z) \geqslant \pi_v(w, Z)$. ∎

The above proof can be trivially modified to show the following result.

7 Theorem *Let D be an interval of the real line. Let u and v be continuous strictly increasing real functions on D. Then the following statements are equivalent.*

1. *For all w in D and all* nondegenerate *random variables Z that are admissible for u and v at w and satisfy $\boldsymbol{E}\, Z = 0$,*
$$\pi_u(w, Z) > \pi_v(w, Z).$$

2. *There exists a* strictly *concave strictly increasing function g on the range of v satisfying*
$$u = g \circ v.$$

3 Local results

When u is twice differentiable, the **(Arrow–Pratt–de Finetti) coefficient of (local) risk aversion**[1] r_u is defined by
$$r_u(w) = -\frac{u''(w)}{u'(w)}.$$
Note that this coefficient is invariant under positive affine transformations of u, so it really is a property of the preferences.

8 Theorem *Let D be an interval of the real line. Let u and v be continuous strictly increasing functions on D that are twice differentiable with strictly positive derivatives everywhere on the interior of D. Then the following statements are equivalent.*

(1) *For all w in D and all random variables Z that are admissible for u and v at w and satisfy $\boldsymbol{E}\, Z = 0$,*
$$\pi_u(w, Z) \geqslant \pi_v(w, Z).$$

(2) *There exists a concave strictly increasing function g defined on the range of v satisfying*
$$u = g \circ v.$$

(3) *For all w in the interior of D,*
$$-\frac{u''(w)}{u'(w)} \geqslant -\frac{v''(w)}{v'(w)}.$$

[1] This is usually referred to as the Arrow–Pratt coefficient of risk aversion. Pratt [9, p. 123n] points out that the importance of this coefficient had been independently and earlier discovered by Arrow and by Schlaiffer, but he gives no references. This coefficient (or rather twice its inverse, $2u'/u''$) had been earlier discussed by de Finetti [4, p. 700] who essentially stated the relation used in the proof of (1) $\implies$ (3) below. See also the review article by Montesano [8].

Proof: The equivalence of (1) and (2) has already been established.

(1) $\implies$ (3): Fix w in the interior of D. Then there is an interval A containing zero such that $\varepsilon \in A$ implies $w \pm \varepsilon \in D$. For each ε in A let Z_ε be a random variable that takes on each of the values ε and $-\varepsilon$ with probability $\frac{1}{2}$. Then $\boldsymbol{E}\, Z_\varepsilon = 0$ and Z_ε is admissible for u and v at w. To simplify notation, define the real function p_u on A by $p_u(\varepsilon) = \pi_u(w, Z_\varepsilon)$, and similarly for p_v. Note that

$$p_u(0) = 0, \quad p_u(\varepsilon) = p_u(-\varepsilon),$$

and by equation $(\star)$,

$$u(w - p_u(\varepsilon)) = \tfrac{1}{2}u(w+\varepsilon) + \tfrac{1}{2}u(w-\varepsilon) \tag{1}$$

for all ε in A. Defining p_v in an analogous manner, we note that $p_u \geqslant p_v$ on A.

Note that (1) implies that the function p_u is twice differentiable on A. To see this, define the function $f\colon A \times (D - w) \to \boldsymbol{R}$ by

$$f(\varepsilon, p) = u(w - p) - \tfrac{1}{2}u(w+\varepsilon) - \tfrac{1}{2}u(w-\varepsilon) \tag{2}$$

and note that f is twice differentiable, $f(0,0) = 0$, and $\dfrac{\partial f(0,0)}{\partial p} = -u'(w) < 0$. Therefore by the Implicit Function Theorem (see, e.g., [7, Theorem 2, p. 235]) there is a unique twice differentiable function defined on a neighborhood of zero giving p as a function of ε to satisfy equation (2).[2] This function must be our p_u. (Naturally the same is true of p_v too.)

For the time being we shall drop the u subscript and simply write p for p_u. Since (1) holds everywhere in A, we may differentiate both sides to get

$$-u'(w - p(\varepsilon))p'(\varepsilon) = \tfrac{1}{2}u'(w+\varepsilon) - \tfrac{1}{2}u'(w-\varepsilon).$$

As a special case,

$$p'(0) = 0.$$

Differentiating a second time yields

$$u''(w - p(\varepsilon))\big(p'(\varepsilon)\big)^2 - p''(\varepsilon)u'(w - p(\varepsilon)) = \tfrac{1}{2}u''(w+\varepsilon) + \tfrac{1}{2}u''(w-\varepsilon).$$

Using $p(0) = p'(0) = 0$ we have

$$p''(0) = -\frac{u''(w)}{u'(w)} = r_u(w).$$

(So that's where that expression comes from.)

Having shown earlier that p is twice differentiable on A, for $\varepsilon > 0$ we can apply Taylor's Theorem [5, p. 290] to write

$$\begin{aligned} p(\varepsilon) &= p(0) + \varepsilon p'(0) + \frac{\varepsilon^2}{2}\big(p''(0) + R(\varepsilon)\big) \\ &= \frac{\varepsilon^2}{2}\big(p''(0) + R(\varepsilon)\big), \end{aligned} \tag{3}$$

[2]This conclusion may look a bit different from the statement of the theorem, which says that if u is k times continuously differentiable then p is k times continuously differentiable. What we have assumed is continuous once differentiability and everywhere twice differentiability (not necessarily continuously). But the formula for the derivative of p that is given below can be deduced from the continuous differentiability. But clearly the derivative is differentiable.

where $\lim_{\varepsilon\to 0} R(\varepsilon) = 0$.[3]

Now we have to start reattaching subscript us and vs. Using equation (3) and $p_u \geqslant p_v$, we conclude

$$p''_u(0) + R_u(\varepsilon) \geqslant p''_v(0) + R_v(\varepsilon)$$

for all $\varepsilon > 0$ in A. Letting $\varepsilon \downarrow 0$, we conclude that

$$p''_u(0) \geqslant p''_v(0).$$

That is,

$$-\frac{u''(w)}{u'(w)} \geqslant -\frac{v''(w)}{v'(w)}.$$

Since w is arbitrary, this completes the proof of the implication.

(3) $\implies$ (2): As above, Lemma 5 gives the existence of strictly increasing continuous g, so to prove statement (2) we need to show that g is concave. It follows from Lemma 5 that g is twice differentiable on the interior of its domain, so it suffices to prove that $g'' \leqslant 0$. I should put a reference here.

Since $u(w) = g(v(w))$ for all w, we may differentiate both sides to conclude

$$u'(w) = g'(v(w))v'(w) \tag{4}$$

and

$$u''(w) = g''(v(w))\big(v'(w)\big)^2 + g'(v(w))v''(w). \tag{5}$$

Dividing (5) by (4) yields

$$\frac{u''(w)}{u'(w)} = \frac{g''(v(w))}{g'(v(w))}v'(w) + \frac{v''(w)}{v'(w)}.$$

Thus $r_u \geqslant r_v$ implies $\dfrac{g''(v(w))}{g'(v(w))}v'(w) \leqslant 0$, but $u' > 0$ and $v' > 0$ imply $g'(v(w)) = \dfrac{u'(v(w))}{v'(w)} > 0$, so $g''(v(w)) \leqslant 0$. Thus g is concave. ∎

9 Remark (More nitpicking) Note that if D is open and u is concave and strictly increasing, then the assumption that $u' > 0$ is redundant, since $0 \in \partial u(w)$ implies that w is a global maximum of u, which cannot occur (since D is open and u is strictly increasing).

Again we may strengthen some of the inequalities.

10 Theorem *Let D be an interval of the real line. Let u and v be continuous strictly increasing functions on D that are twice differentiable with strictly positive derivatives everywhere on the interior of D. Consider following statements.*

(1) *For all w in D and all nondegenerate random variables Z that are admissible for u and v at w and satisfy $\boldsymbol{E}\, Z = 0$,*

$$\pi_u(w, Z) > \pi_v(w, Z).$$

[3]This form of Taylor's Theorem, given by Hardy [5], requires only twice differentiability at 0, not twice continuous differentiability on a neighborhood of 0.

(2) *There exists a strictly concave strictly increasing function g on the range of v satisfying*

$$u = g \circ v.$$

(3) *For all w in the interior of D,*

$$-\frac{u''(w)}{u'(w)} > -\frac{v''(w)}{v'(w)}.$$

Then (3) $\implies$ (2) $\implies$ (1).

I do not claim that (1) $\implies$ (3), since adapting the arguments in the proof of Theorem 8 only yield a weak inequality. Pratt himself makes the slightly weaker claim that (1) implies the following:

(3′): For all w in the interior of D,

$$-\frac{u''(w)}{u'(w)} \geqslant -\frac{v''(w)}{v'(w)},$$

and for every subinterval of D there is some w with

$$-\frac{u''(w)}{u'(w)} \geqslant -\frac{v''(w)}{v'(w)}.$$

Varian [12, pp. 182–183] makes the stronger claim, so I need to do some work here.

4 A more general interpretation

As an aside we mention another interpretation of (3). The variance of Z_ε is ε^2. So $\frac{p_u(\varepsilon)}{\varepsilon^2}$ is the fraction of the variance that someone with utility u would be willing to pay to insure against Z_ε. The limit of this fraction as $\varepsilon \to 0$ is then $\frac{1}{2}r_u(w)$. In fact, this generalizes to more general admissible random variables with variance $\varepsilon > 0$, but the only proof I can think of is subtle.

11 Theorem *Let D be an interval of the real line. Let u and v be continuous strictly increasing functions on D that are twice differentiable with strictly positive derivatives everywhere on the interior of D. Fix w in the interior of D, and let $\eta > 0$ satisfy $(w-\eta, w+\eta) \subset D$. For $0 < \varepsilon < \eta$, let X_ε be an admissible random variable for u at w satisfying $\boldsymbol{E}\, X_\varepsilon = 0$ and $\boldsymbol{E}\, X_\varepsilon^2 = \varepsilon^2$. Assume in addition that $\|X_\varepsilon\|_\infty \xrightarrow[\varepsilon\to 0]{} 0$. Then*

$$\lim_{\varepsilon \to 0} \frac{\pi_u(w, X_\varepsilon)}{\varepsilon^2} = \frac{1}{2} r_u(w).$$

Proof: To simplify notation, set $p(\varepsilon) = \pi_u(w, X_\varepsilon)$. Then by definition,

$$u(w - p(\varepsilon)) = \int_S u(w + X_\varepsilon(s))\, dP(s). \tag{6}$$

Using Taylor's Theorem write the left-hand side of (6) as

$$u(w - p(\varepsilon)) = u(w) - u'(w - \delta(\varepsilon))p(\varepsilon)$$

where for each ε, $\delta(\varepsilon) \in [0, p(\varepsilon)]$. Now us the stochastic version of Taylor's Theorem (Theorem 12 below) to write the right-hand side of (6) as

$$\begin{aligned}&\int_S u(w + X_\varepsilon(s))\,dP(s)\\ &\quad = \int_S \Big\{u(w) + u'(w)X_\varepsilon(s) + \tfrac{1}{2}u''(w + \xi_\varepsilon(s))X_\varepsilon^2(s)\Big\}\,dP(s),\end{aligned}$$

where each ξ_ε is a measurable function and for each s, $\xi_\varepsilon(s) \in [0, X_\varepsilon(s)]$. Dividing both terms by ε^2 and simplifying (using $\boldsymbol{E}\,X_\varepsilon = 0$), we can rewrite (6) as

$$-u'(w - \delta(\varepsilon))\frac{p(\varepsilon)}{\varepsilon^2} = \frac{1}{2}\int_S u''(w + \xi_\varepsilon(s))\frac{X_\varepsilon^2(s)}{\varepsilon^2}\,dP(s).$$

Now the left-hand side satisfies

$$-u'(w - \delta(\varepsilon))\frac{p(\varepsilon)}{\varepsilon^2} \xrightarrow[\varepsilon\to 0]{} -u'(w)\left(\lim_{\varepsilon\to 0}\frac{p(\varepsilon)}{\varepsilon^2}\right).$$

Since u'' is continuous, and $\|\xi_\varepsilon\|_\infty \leqslant \|X_\varepsilon\|_\infty \xrightarrow[\varepsilon\to 0]{}$, for each $\eta > 0$ there is some $\varepsilon_0 < 0$ such that for all $\varepsilon < \varepsilon_0$ we have $|u''(w + \xi_\varepsilon(s)) - u''(w)| < \eta$ for almost every s. So for such ε,

$$\left|\int_S u''(w + \xi_\varepsilon(s))\frac{X_\varepsilon^2(s)}{\varepsilon^2}\,dP(s) - \int_S u''(w)\frac{X_\varepsilon^2(s)}{\varepsilon^2}\,dP(s)\right| < \eta,$$

since $\int_S \frac{X_\varepsilon^2(s)}{\varepsilon^2}\,dP(s) = 1$. Therefore

$$\int_S u''(w + \xi_\varepsilon(s))\frac{X_\varepsilon^2(s)}{\varepsilon^2}\,dP(s) \xrightarrow[\varepsilon\to 0]{} u''(w). \tag{7}$$

Combining yields

$$\lim_{\varepsilon\to 0}\frac{p(\varepsilon)}{\varepsilon^2} = -\frac{1}{2}\frac{u''(w)}{u'(w)} = \frac{1}{2}r_u(w).$$

∎

The above proof of Theorem 11 relies on the following stochastic version of Taylor's Theorem.

12 Stochastic Taylor's Theorem *Let $h\colon [a,b] \to \boldsymbol{R}$ be continuous and possess a continuous n^{th}-order derivative on (a,b). Fix $c \in [a,b]$ and let X be a random variable on the probability space $(S, \mathcal{S}, P)$ such that $c + X \in [a,b]$ almost surely. Then there is a (measurable) random variable ξ satisfying $\xi(s) \in [0, X(s)]$ for all s (where $[0, X(s)]$ is the line segment joining 0 and $X(s)$, regardless of the sign of $X(s)$), and*

$$h(c + X(s)) = h(c) + \sum_{k=1}^{n-1}\frac{1}{k!}h^{(k)}(c)X^k(s) + \frac{1}{n!}h^{(n)}(c + \xi(s))X^n(s).$$

Proof: Taylor's Theorem asserts that there is such a $\xi(s)$ for each s, the trick is to show that there is a measurable version. To this end define the correspondence $\varphi\colon S \twoheadrightarrow \boldsymbol{R}$ by $\varphi(s) = [0, X(s)]$. It follows from [1, Theorem 18.5, p. 595] that φ is measurable and it clearly has compact values. Set $g(s) = h\big(c+X(s)\big) - h(c) - \sum_{k=1}^{n-1} \frac{1}{k!} h^{(k)}(c) X^k(s)$, $f(s,x) = \frac{1}{n!} h^{(n)}(c+x) X^n(s)$. Then g is measurable and f is a Carathéodory function. (See sections 4.10 and 18.1 of [1] for the definitions of Carathéodory functions and measurable correspondences.) By Filippov's Implicit Function Theorem [1, Theorem 18.17, p. 603] there is a measurable function ξ such that for all s, $\xi(s) \in \varphi(s)$ and $f\big(s, \xi(s)\big) = g(s)$, and we are done. ∎

5 Some odd examples

13 Example To see the need for the assumption $\|X_\varepsilon\|_\infty \to 0$ in Theorem 11, consider the following simple example, which shows how (7) can fail. Let

$$X_\varepsilon = \begin{cases} 1 & \text{with probability } \frac{\varepsilon}{\varepsilon+1} \\ -\varepsilon & \text{with probability } \frac{1}{\varepsilon+1} \end{cases}$$

Then $EX_\varepsilon = 0$ and $EX_\varepsilon^2 = \varepsilon$, so $\|X_\varepsilon\|_2 \to 0$, but $\|X_\varepsilon\|_\infty = 1$. The risk premium $p(\varepsilon)$ for X_ε is defined by

$$u\big((w - p(\varepsilon)\big) = \frac{\varepsilon}{\varepsilon+1} u(w+1) + \frac{1}{\varepsilon+1} u(w-\varepsilon). \tag{8}$$

Proceeding more along the lines of Pratt, we use Taylor's Theorem to write

$$u(w-p) = u(w) - p\big(u'(w) + R_1(p)\big)$$

and

$$u(w-\varepsilon) = u(w) - \varepsilon u'(w) + \frac{1}{2}\varepsilon^2\big(u''(w) + R_2(\varepsilon)\big),$$

and clearly

$$u(w+1) = u(w) + \frac{u(w+1) - u(w)}{u'(w)} u'(w).$$

Therefore (8) can be written

$$\begin{aligned} &u(w) - p(\varepsilon)\big(u'(w) + R_1(p(\varepsilon))\big) \\ &\quad= \frac{\varepsilon}{\varepsilon+1}\left(u(w) + \frac{u(w+1)-u(w)}{u'(w)} u'(w)\right) + \frac{1}{\varepsilon+1}\left(u(w) - \varepsilon u'(w) + \tfrac{1}{2}\varepsilon^2\big(u''(w) + R_2(\varepsilon)\big)\right). \end{aligned}$$

or regrouping,

$$\begin{aligned} &-p(\varepsilon)\big(u'(w) + R_1(p(\varepsilon))\big) \\ &\quad= \frac{\varepsilon}{\varepsilon+1}\left(\frac{u(w+1)-u(w)}{u'(w)} - 1\right) u'(w) + \frac{1}{2}\varepsilon^2\big(u''(w) + R_2(\varepsilon)\big). \end{aligned}$$

Dividing by $-\varepsilon\big(u'(w) + R_1(p(\varepsilon))\big)$ gives

$$\begin{aligned}\frac{p(\varepsilon)}{\operatorname{var} X_\varepsilon} &= \frac{p(\varepsilon)}{\varepsilon} \\ &= -\frac{1}{\varepsilon+1}\left(\frac{u(w+1)-u(w)}{u'(w)} - 1\right)\frac{u'(w)}{u'(w)+R_1(p(\varepsilon))} - \frac{1}{2}\frac{\varepsilon}{\varepsilon+1}\frac{u''(w)+R_2(\varepsilon)}{u'(w)+R_1(p(\varepsilon))} \\ &\longrightarrow 1 - \frac{u(w+1)-u(w)}{u'(w)} \quad \text{as}\varepsilon \to 0.\end{aligned}$$

This is not in general half the coefficient of risk aversion. If u is concave, then the supergradient inequality implies $1 - \frac{u(w+1)-u(w)}{u'(w)} > 0$.

To be more concrete, consider the quadratic utility

$$u(w) = -(a-w)^2$$

on the interval $[0, a]$. Then $u'(w) = 2(a-w)$ and $u''(w) = -2$, so the risk aversion coefficient is

$$r_u(w) = (a-w)$$

which is positive on $[0, a]$. Fix w and let $b = a - w$. Then (8) becomes

$$-(b+p)^2 = -\frac{\varepsilon}{\varepsilon+1}(b-1)^2 - \frac{1}{\varepsilon+1}(b+\varepsilon)^2$$

Expanding and simplifying yields $p^2 + 2bp - \varepsilon = 0$, so

$$p(\varepsilon) = \sqrt{b^2+\varepsilon} - b.$$

Now $p(\varepsilon)/\varepsilon \to p'(0) = 1/2b$ as $\varepsilon \to 0$, whereas $\frac{1}{2}r_u(a-b) = b/2$. Figure 2 shows plots of p and p/ε^2 for $b = 2$ (i.e., $w = a - 2$).

□

14 Example For a more dramatic example, let

$$X_\varepsilon = \begin{cases} 1/\varepsilon & \text{with probability } \varepsilon^3/(1+\varepsilon^3) \\ -\varepsilon^2 & \text{with probability } 1/(1+\varepsilon^3) \end{cases}$$

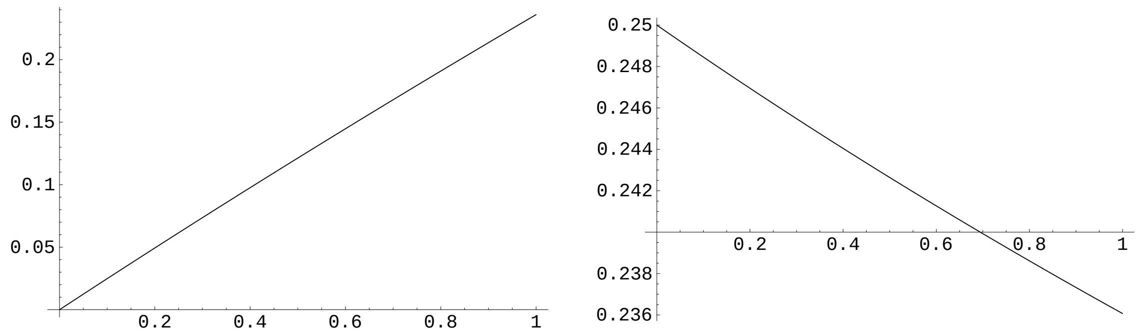
Then $EX_\varepsilon = 0$ and $EX_\varepsilon^2 = \varepsilon$, so $\|X_\varepsilon\|_2 \to 0$, but $\|X_\varepsilon\|_\infty = 1/\varepsilon$. The risk premium $p(\varepsilon)$ for X_ε is defined by

$$u((w - p(\varepsilon)) = \frac{\varepsilon^3}{1+\varepsilon^3}u(w + \tfrac{1}{\varepsilon}) + \frac{1}{1+\varepsilon^3}u(w-\varepsilon^2). \tag{9}$$

By Taylor's Theorem,

$$u(w-p) = u(w) - p\big(u'(w) + R_1(p)\big),$$

$$u(w-\varepsilon^2) = u(w) - \varepsilon^2 u'(w) + \frac{1}{2}\varepsilon^4\big(u''(w) + R_2(\varepsilon^2)\big),$$



Risk premium $p(\varepsilon)$. $p(\varepsilon)/\operatorname{var} X_\varepsilon$

Figure 2. The risk premium $p(\varepsilon)$ for the quadratic utility $u(w) = -(a-w)^2$ facing the risk X_ε of Example 13 at wealth level $w = a - 2$. (So $\frac{1}{2}r(w) = 1$.)

and also

$$u(w + 1/\varepsilon) = u(w) + \frac{u(w+1/\varepsilon) - u(w)}{u'(w)} u'(w).$$

Therefore (9) can be written

$$\begin{aligned} u(w) - p(\varepsilon)\big(u'(w) + R_1(p(\varepsilon))\big) \\ = \quad & \frac{\varepsilon^3}{1+\varepsilon^3}\left(u(w) + \frac{u(w+1/\varepsilon) - u(w)}{u'(w)} u'(w)\right) \\ & + \frac{1}{1+\varepsilon^3}\left(u(w) - \varepsilon^2 u'(w) + \frac{1}{2}\varepsilon^4\big(u''(w) + R_2(\varepsilon^2)\big)\right), \end{aligned}$$

or regrouping,

$$\begin{aligned} -p(\varepsilon)\big(u'(w) + R_1(p(\varepsilon))\big) \quad = \quad & \frac{\varepsilon^3}{1+\varepsilon^3}\left(\frac{u(w+1/\varepsilon) - u(w)}{u'(w)} - \frac{1}{\varepsilon}\right) u'(w) \\ & + \frac{1}{1+\varepsilon^3}\left(\frac{1}{2}\varepsilon^4\big(u''(w) + R_2(\varepsilon^2)\big)\right) \end{aligned}$$

Dividing by $-\varepsilon\big(u'(w) + R_1(p(\varepsilon))\big)$ gives

$$\begin{aligned} \frac{p(\varepsilon)}{\operatorname{var} X_\varepsilon} = \frac{p(\varepsilon)}{\varepsilon} \\ = \quad -\frac{\varepsilon^2}{1+\varepsilon^3}\left(\frac{u(w+1/\varepsilon) - u(w)}{u'(w)} - \frac{1}{\varepsilon}\right)\frac{u'(w)}{u'(w) + R_1(p(\varepsilon))} - \frac{1}{2}\frac{\varepsilon^3}{1+\varepsilon^3}\frac{u''(w) + R_2(\varepsilon)}{u'(w) + R_1(p(\varepsilon))}. \end{aligned}$$

Thus

$$\lim_{\varepsilon \to 0} p(\varepsilon)/\varepsilon = -\frac{1}{u'(w)} \lim_{\varepsilon \to 0} \varepsilon^2 u(w + 1/\varepsilon).$$

If u is bounded, then this limit is zero. If u is unbounded but concave, then by the supergradient inequality, $u(w) + u'(w)/\varepsilon \geqslant u(w + 1/\varepsilon)$, so

$$\lim_{\varepsilon \to 0} \varepsilon^2 u(w + 1/\varepsilon) \leqslant \lim_{\varepsilon \to 0} \varepsilon^2\big(u(w) + u'(w)/\varepsilon\big) = 0.$$

This makes sense. The risk premium will always be smaller than the maximum loss, and the maximum loss divided by the variance is ε, which tends to zero. □

6 Recovering the utility

We can recover the utility function from the coefficient of risk aversion.

15 Theorem *Let $r\colon D \to \boldsymbol{R}$ be a continuous function on an open interval D. Then there is a C^2 utility $u\colon D \to \boldsymbol{R}$ satisfying $u'(w) > 0$ for all $w \in D$ and*

$$-\frac{u''(w)}{u'(w)} = r(w) \quad \text{for all } w \in D.$$

The function u is unique up to positive affine transformation and is given by

$$u(w) = \alpha \int_a^w e^{-R(x)}\,dx + \beta$$

where $a \in D$, $\alpha > 0$,

$$R(x) = \int_a^x r(t)\,dt,$$

and β is an arbitrary real number.

Proof: Observe that we are seeking a solution to the linear differential equation

$$u''(x) - r(x)u'(x) = 0.$$

Letting $f(x)$ denote $u'(x)$, this becomes the first order equation

$$f'(x) + r(x)f(x) = 0.$$

It is well known, see Theorem 17 below, that for each $a \in D$ and $\alpha \in \boldsymbol{R}$, this equation has a unique solution satisfying $f(a) = \alpha$. It is given by

$$f(x) = \alpha e^{-R(x)}, \qquad \text{where } R(x) = \int_a^x r(t)\,dt.$$

(In Theorem 17, set $P = -r$ and $Q = 0$.) Any such solution is continuous, so by the same result (now with $P = 0$ and $Q = f$), the differential equation

$$u'(x) = f(x)$$

has a solution u given by

$$\begin{aligned} u(x) &= \int_a^x f(t)\,dt + \beta \\ &= \alpha \int_a^x e^{-R(t)}\,dt + \beta \end{aligned}$$

(as $P = 0$ implies $A = 0$ in Theorem 17). To guarantee that $u'(x) > 0$, we need only take $\alpha > 0$, as $e^{-R(x)} > 0$. Thus within the class of strictly increasing functions, given a, the utility u is unique up to the constants $\alpha > 0$ and β. ∎

16 Example Suppose the risk aversion function is constant:

$$r(x) = c$$

for all x. Then assuming $a = 0$ is in the domain, we have $R(x) = cx$, so

$$u(x) = \alpha \int_0^x e^{-ct}\,dt + \beta.$$

If $c = 0$, then $\int_0^x e^{-0t}\,dt = \int_0^x 1\,dt = x$

$$u(x) = \alpha x + \beta.$$

If $c \neq 0$, then $R(x) = cx$, and the primitive of $e^{-cx} = -\frac{1}{c}e^{-cx}$, so $\int_0^x e^{-ct}\,dt = -\frac{1}{c}e^{-cx} - (-\frac{1}{c}e^{-c0}) = -\frac{1}{c}(e^{-cx} - 1)$, so our formula gives

$$u(x) = -\frac{\alpha}{c}(e^{-cx} - 1) + \beta.$$

We can choose particularly aesthetic values of α and β. For $c = 0$, choose $\alpha = 1$ and $\beta = 0$, to get

$$u(x) = x \quad (c = 0).$$

For $c > 0$, set $\beta = 0$ and $\alpha = c > 0$ to get

$$u(x) = 1 - e^{-cx} \quad (c > 0).$$

For $c < 0$, set $\beta = 0$ and $\alpha = -c > 0$, to get

$$u(x) = e^{-cx} - 1 \quad (c < 0).$$

□

The following theorem is taken from Apostol [2, Theorem 8.3, p. 310]. See my notes [3] for an economic derivation and interpretation.

17 Theorem (First order linear differential equations) *Let P and Q be continuous functions on an open interval D. Let $a \in D$ and $b \in \boldsymbol{R}$. Then there is exactly one function f satisfying $f(a) = b$ and*

$$f'(x) + P(x)f(x) = Q(x).$$

It is given by

$$f(x) = be^{-A(x)} + e^{-A(x)} \int_z^x Q(x)e^{-A(x)}\,dx,$$

where

$$A(x) = \int_a^x P(x)\,dx.$$

7 Ross' generalization

Suppose initial wealth is random, or equivalently, you can only partially insure your wealth so there is still uncertainty after insurance. That is, after insurance the wealth is a random variable $\tilde{w}$, but without insurance it is $\tilde{w} + \tilde{z}$. It is natural to implicitly define the **generalized risk premium** $\hat{\pi}_u(\tilde{w}, \tilde{z})$ by the equation

$$\boldsymbol{E}\, u(\tilde{w} + \boldsymbol{E}\, \tilde{z} - \hat{\pi}_u(\tilde{w}, \tilde{z})) = \boldsymbol{E}\, u(\tilde{w} + \tilde{z}). \tag{$\star\star$}$$

But there are problems with this definition that are not present in the Pratt definition for the case of full insurance. One obvious problem is that if the domain of u is bounded below, for instance, if it is $\boldsymbol{R}_{++}$, and if the support of $\tilde{w}$ includes the lower bound then for any positive $\hat{\pi}$, the quantity $\tilde{w} - \hat{\pi}$ lies outside the domain of u with positive probability, so ($\star\star$) may have no solution.

A more subtle problem is that even for concave u, the risk premium $\hat{\pi}$ may actually be negative. That is, $\tilde{z}$ may already provide some insurance against fluctuations in $\tilde{w}$. A simple extreme example is where $\tilde{w}$ assumes the values $x \pm \varepsilon$ equally likely and $\tilde{z} = x - \tilde{w}$, so that $\tilde{w} + \tilde{z} = x$. Then $\hat{\pi}(\tilde{w}, \tilde{z}) < 0$ when u is concave. The problem here is that in this case $\tilde{z}$ does not represent a "risk." In order to be classified as a risk, we want $\tilde{w} + \tilde{z}$ to be riskier than $\tilde{z}$ in the sense of Rothschild and Stiglitz [11]. This amounts to the criterion that $\boldsymbol{E}(\tilde{z}|\tilde{w}) = 0$.

With these caveats in mind, let us say that a pair $(\tilde{w}, \tilde{z})$ of random variables is **admissible for u** if $\tilde{w}$ and $\tilde{w}+\tilde{z}$ take on values in the domain of u almost surely, $\boldsymbol{E}(\tilde{z}|\tilde{w}) = 0$ (which implies $\boldsymbol{E}\, \tilde{z} = 0$), and $\boldsymbol{E}\, \tilde{w}$, $\boldsymbol{E}\, u(\tilde{w})$, and $\boldsymbol{E}\, u(\tilde{w} + \tilde{z})$ are finite, and ($\star\star$) admits a solution $\hat{\pi}$.

Let us say that the utility function u is **strongly more risk averse** or **more risk averse in the sense of Ross** than v if u and v have a common domain and $\hat{\pi}_u(\tilde{w}, \tilde{z}) \geqslant \hat{\pi}_u(\tilde{w}, \tilde{z})$ for all admissible pairs $(\tilde{w}, \tilde{z})$.

Ross [10] provides the following characterization of strong risk aversion, similar to Pratt's theorem.

18 Theorem (Ross) *For u, v twice differentiable with $u', v' > 0$ and $u'', v'' < 0$, (strictly increasing and strictly concave), defined on a common nondegenerate interval D, the following are equivalent.*

1. *There exists $\lambda > 0$ and a decreasing concave function H satisfying $H' \leqslant 0$, $H'' \leqslant 0$ such that*
$$u = \lambda v + H.$$

2. *For all $\tilde{w}, \tilde{z}$ such that $\boldsymbol{E}(\tilde{z}|\tilde{w}) = 0$,*
$$\hat{\pi}_u(\tilde{w}, \tilde{z}) \geqslant \hat{\pi}_v(\tilde{w}, \tilde{z}).$$
That is, u is strongly more risk averse than v.

3. *There exists $\lambda > 0$ such that for all x, y in the interior of D,*
$$\frac{u''(x)}{v''(x)} \geqslant \lambda \geqslant \frac{u'(y)}{v'(y)}.$$

19 Remark There is an importance difference between this result and Pratt's result. If g is strictly concave and strictly increasing, then $g \circ v$ is always strictly more risk averse in Pratt's sense than v. Ross' theorem does not assert that if $\lambda > 0$ and H is a decreasing concave function, then $\lambda v + H$ is strongly more risk averse than v. Indeed, it may be that $\lambda v + H$ is decreasing. The theorem does say that if $\lambda v + H$ is strictly increasing and concave, then $\lambda v + H$ is strongly more risk averse than v. The Ross ordering thus admits maximal elements, which the Pratt ordering does not. In fact, if $v'(x) \to 0$ as $x \uparrow$, then v is maximally strongly risk averse. To see this, suppose $u = \lambda v + H$ is strictly increasing, where H is concave and nonincreasing. It follows that H is a constant. That is, u is a positive affine transformation of v, so that u and v are equivalent Bernoulli utilities.

Proof of Theorem 18: (1) $\implies$ (2): This part uses Jensen's inequality.

$$\begin{aligned}
\boldsymbol{E}\, u(\tilde w - \hat\pi_u) &= \boldsymbol{E}\, u(\tilde w + \tilde z) && \text{(by } (\star\star)\text{)}\\
&= \boldsymbol{E}\, \lambda v(\tilde w + \tilde z) + \boldsymbol{E}\, H(\tilde w + \tilde z) && \text{(by hypothesis)}\\
&= \boldsymbol{E}\, \lambda v(\tilde w + \tilde z) + \boldsymbol{E}\big[\boldsymbol{E}\big(H(\tilde w + \tilde z)|\tilde w\big)\big] \\
&\leqslant \boldsymbol{E}\, \lambda v(\tilde w + \tilde z) + \boldsymbol{E}\, H\big(\boldsymbol{E}(\tilde w + \tilde z|\tilde w)\big) && \text{(Jensen's inequality)}\\
&= \boldsymbol{E}\, \lambda v(\tilde w + \tilde z) + \boldsymbol{E}\, H(\tilde w) && \text{(since } \boldsymbol{E}(\tilde z|\tilde w) = 0\text{)}\\
&= \boldsymbol{E}\, \lambda v(\tilde w - \hat\pi_v) + \boldsymbol{E}\, H(\tilde w) && \text{(definition of } \hat\pi_v\text{)}\\
&\leqslant \boldsymbol{E}\, \lambda v(\tilde w - \hat\pi_v) + \boldsymbol{E}\, H(\tilde w - \hat\pi_v) && \text{(as } H \text{ is decreasing)}\\
&= \boldsymbol{E}\, u(\tilde w - \hat\pi_v) && \text{(by } (\star\star)\text{)}
\end{aligned}$$

Since u is increasing $\hat\pi_u \geqslant \hat\pi_v$.

(2) $\implies$ (3): Let $1 > p > 0$ and consider a random wealth $\tilde w$ with the following distribution

$$\tilde w = \begin{cases} y & \text{with probability } 1{-}p \\ x & \text{with probability } p \end{cases}$$

and a risk $\tilde z$ with the following distribution.

$$\tilde z = \begin{cases} 0 & \text{if } \tilde w = y \\ \varepsilon & \text{with probability } \frac12 \text{ if } \tilde w = x \\ -\varepsilon & \text{with probability } \frac12 \text{ if } \tilde w = x. \end{cases}$$

Then $\boldsymbol{E}\,(\tilde z|\tilde w) = 0$ and

$$\tilde w + \tilde z = \begin{cases} y & \text{with probability } 1{-}p \\ x+\varepsilon & \text{with probability } \frac{p}{2} \\ x-\varepsilon & \text{with probability } \frac{p}{2}. \end{cases}$$

Then $(\star\star)$ becomes

$$pu(x - \hat\pi_u) + (1{-}p)u(y - \hat\pi_u) = \tfrac12 pu(x+\varepsilon) + \tfrac12 pu(x-\varepsilon) + (1{-}p)u(y). \tag{10}$$

This implicitly defines $\hat\pi_u$ as a function of ε and p. Let us write $\tilde\pi(\varepsilon, p)$ for $\hat\pi_u(\tilde w, \tilde z)$. Note that

$$\tilde\pi(0, p) = 0$$

for any p. By the Implicit Function Theorem, since u is C^2, and $u' > 0$, we see that $\tilde{\pi}$ is C^2. We can use (10) to compute the partial derivatives $\frac{\partial \tilde{\pi}}{\partial \varepsilon}$ and $\frac{\partial^2 \tilde{\pi}}{\partial \varepsilon^2}$.

Differentiating (10) with respect to ε yields

$$-\left(pu'(x-\tilde{\pi}) + (1-p)u'(y-\tilde{\pi})\right)\frac{\partial \tilde{\pi}}{\partial \varepsilon} = \frac{p}{2}\left(u'(x+\varepsilon) - u'(x-\varepsilon)\right). \tag{11}$$

Evaluating this at $\varepsilon = 0$, since the right-hand side is zero and $u' > 0$, we have

$$\frac{\partial \tilde{\pi}(0,p)}{\partial \varepsilon} = 0.$$

Differentiating (11) with respect to ε yields

$$\begin{aligned} &\left(pu''(x-\tilde{\pi}) + (1-p)u''(y-\tilde{\pi})\right)\left(\frac{\partial \tilde{\pi}}{\partial \varepsilon}\right)^2 \\ &-\left(pu'(x-\tilde{\pi}) + (1-p)u'(y-\tilde{\pi})\right)\frac{\partial^2 \tilde{\pi}}{\partial \varepsilon^2} \quad = \quad \frac{1}{2}p\left(u''(x+\varepsilon) + u''(x-\varepsilon)\right). \end{aligned}$$

Evaluating at $\varepsilon = 0$, we have

$$\frac{\partial^2 \tilde{\pi}(0,p)}{\partial \varepsilon^2} = -\frac{pu''(x)}{pu'(x) + (1-p)u'(y)}. \tag{12}$$

By Taylor's theorem, for $\varepsilon \neq 0$,

$$\tilde{\pi}(\varepsilon,p) = \underbrace{\tilde{\pi}(0,p)}_{=0} + \underbrace{\frac{\partial \tilde{\pi}(0,p)}{\partial \varepsilon}}_{=0}\varepsilon + \frac{1}{2}\frac{\partial^2 \tilde{\pi}(\delta(\varepsilon),p)}{\partial \varepsilon^2}\varepsilon^2, \tag{13}$$

where $\delta(\varepsilon)$ is strictly between 0 and ε.

Reattaching subscripts u and v in (13) we see that $\tilde{\pi}_u(\varepsilon,p) \geqslant \tilde{\pi}_v(\varepsilon,p)$ implies

$$\frac{\partial^2 \tilde{\pi}_u(\delta_u(\varepsilon),p)}{\partial \varepsilon^2} \geqslant \frac{\partial^2 \tilde{\pi}_v(\delta_v(\varepsilon),p)}{\partial \varepsilon^2}.$$

Since each $\tilde{\pi}$ is a twice differentiable function of ε, letting $\varepsilon \to 0$ yields

$$\frac{\partial^2 \tilde{\pi}_u(0,p)}{\partial \varepsilon^2} \geqslant \frac{\partial^2 \tilde{\pi}_v(0,p)}{\partial \varepsilon^2}.$$

So by (12)

$$-\frac{pu''(x)}{pu'(x) + (1-p)u'(y)} \geqslant -\frac{pv''(x)}{pv'(x) + (1-p)v'(y)}.$$

Multiplying both sides by the positive quantity $-\frac{pu'(x) + (1-p)u'(y)}{pv''(x)}$ gives

$$\frac{u''(x)}{v''(x)} \geqslant \frac{pu'(x) + (1-p)u'(y)}{pv'(x) + (1-p)v'(y)}.$$

Since x and y are arbitrary, letting $p \downarrow 0$ gives the desired conclusion.

(3) $\implies$ (1): Set $H = u - \lambda v$, where λ is given by statement (3). Then $H' = u' - \lambda v'$, but $\lambda \geqslant \frac{u'}{v'}$ so $u' - \lambda v' \leqslant 0$. Therefore $H' \leqslant 0$.

Now $H'' = u'' - \lambda v''$. But also by statement (3) $\frac{u''}{v''} \geqslant \lambda$ so $u'' \leqslant \lambda v''$ as $v'' \leqslant 0$. Therefore $H'' \leqslant 0$. ∎

It follows from this theorem that even if u is more risk averse than v in the sense of Arrow and Pratt that u need not be strongly more risk averse than v. The following example due to Ross [10] drives this point home.

20 Example Consider the case where u and v have constant absolute risk aversion, and u is more risk averse than v. That is,

$$u(x) = -e^{-ax} \quad \text{and} \quad v(x) = -e^{-bx} \quad \text{with } a > b > 0.$$

For this case, (13) implies

$$\lim_{p \to 0} \lim_{\varepsilon \to 0} \frac{\tilde{\pi}_u}{\tilde{\pi}_v} = \frac{\frac{u''(y)}{u'(x)}}{\frac{v''(y)}{v'(x)}} = \frac{a}{b} e^{-(a-b)(x-y)}.$$

Thus, if ε and p are sufficiently small, and $x - y$ is large enough, then $\tilde{\pi}_u(\varepsilon, p) < \tilde{\pi}_v(\varepsilon, p)$ even though u is more risk averse in Pratt's sense. □

References

[1] C. D. Aliprantis and K. C. Border. 2006. *Infinite dimensional analysis: A hitchhiker's guide*, 3d. ed. Berlin: Springer–Verlag.

[2] T. M. Apostol. 1967. *Calculus*, 2d. ed., volume 1. Waltham, Massachusetts: Blaisdell.

[3] K. C. Border. 2016. The economics of first-order linear differential equations. http://people.hss.caltech.edu/~kcb/Notes/DiffEq.pdf

[4] B. de Finetti. 1952. Sulla preferibilità. *Giornale degli Economisti e Annali di Economia* 11(11/12):685–709. http://www.jstor.org/stable/23236169

[5] G. H. Hardy. 1952. *A course of pure mathematics*, 10th. ed. Cambridge: Cambridge University Press.

[6] M. J. Machina and W. S. Neilson. 1987. The Ross characterization of risk aversion: Strengthening and extension. *Econometrica* 55(5):1139–1149. http://www.jstor.org/stable/1911264

[7] J. E. Marsden. 1974. *Elementary classical analysis*. San Francisco: W. H. Freeman and Company.

[8] A. Montesano. 2009. De Finetti and the Arrow–Pratt measure of risk aversion. In M. C. Galavotti, ed., *Bruno de Finetti Radical Probabilist*, Texts in Philosophy, chapter 7, pages 115–127. College Publications. http://works.bepress.com/aldo_montesano/35/

[9] J. W. Pratt. 1964. Risk aversion in the small and in the large. *Econometrica* 32(1/2):122–136. http://www.jstor.org/stable/1913738

[10] S. A. Ross. 1981. Some stronger measures of risk aversion in the small and the large with applications. *Econometrica* 49(3):621–638. http://www.jstor.org/stable/1911515

[11] M. Rothschild and J. E. Stiglitz. 1970. Increasing risk I: A definition. *Journal of Economic Theory* 2:225–243. http://cowles.econ.yale.edu/P/cp/p03a/p0341a.pdf

[12] H. R. Varian. 1992. *Microeconomic analysis*, 3d. ed. New York: W. W. Norton & Co.