Corrective Taxation versus Liability

Steven Shavell

Online Appendix

This appendix contains the proof of the part of Proposition 4 that is not shown in the text, namely, *under the optimal joint tax and liability regime, the optimal tax* $t^{**} < E(y)$ *provided that injurer benefits* $b(x)$ *display decreasing absolute risk aversion.*

To establish this claim, observe that under the joint tax and liability regime, social welfare as a function of the tax t is

$$W_{TL}(t) = \int_0^m [b(x^*(t + py)) - x^*(t + py)y]f(y)dy \tag{1}$$

since it was shown already that $\lambda = 1$. Hence

$$W_{TL}'(t) = \int_0^m x^{*\prime}(t + py)[b'(x^*(t + py)) - y]f(y)dy, \tag{2}$$

where $x^{*\prime}(t + py)$ is the derivative of $x^*(t + py)$ with respect to t. It will be shown that

$$W_{TL}'(E(y)) = \int_0^m x^{*\prime}(E(y) + py)[b'(x^*(E(y) + py)) - y]f(y)dy < 0. \tag{3}$$

As I will note below, an essentially identical argument to what I am about to give will prove also that $W_{TL}'(t) < 0$ for any $t > E(y)$. Hence, it will follow that the optimal tax t^{**} must be less than $E(y)$.

Observe first that since the optimal tax under a tax only regime is $E(y)$ (from Proposition 1), $b(x^*(t)) - x^*(t)E(y)$ is maximized at $t = E(y)$. Therefore, $b'(x^*(E(y)))x^{*\prime}(E(y)) - x^{*\prime}(E(y))E(y) = 0$, which implies that $b'(x^*(E(y))) - E(y) = 0$. This is equivalent to

$$\int_0^m [b'(x^*(E(y)) - y]f(y)dy = 0. \tag{4}$$

It will now be shown that (4) implies

$$\int_0^m [b'(x^*(E(y) + py) - y]f(y)dy > 0. \tag{5}$$

To this end, rewrite (4) as

$$\int_0^{E(y)} [b'(x^*(E(y)) - y]f(y)dy + \int_{E(y)}^m [b'(x^*(E(y)) - y]f(y)dy = 0. \tag{6}$$

The first term in (6) is positive, since the integrand is positive for each $y < E(y)$. This claim about (6) is readily seen from Figure 1. In particular, in region A, an upward movement in the line $x^*(E(y))$ brings x closer to the optimum $x^*(y)$ at each y, and given the concavity of welfare $b(x) - xy$ in x, this change in x increases welfare.[1] The second term in (6) is negative, since the integrand is negative for each $y > E(y)$. The explanation is analogous to what was just stated; in regions B and C, an upward movement in the line $x^*(E(y))$ makes x more distant from $x^*(y)$ and thus lowers welfare at each y.

[1] That is, $b'(x) - y = 0$ at $x^*(y)$, and thus $b'(x) - y > 0$ for $x < x^*(y)$ since $b''(x) < 0$. Hence, $b'(x^*(E(y)) - y > 0$, since, for each y in A, $x^*(E(y)) < x^*(y)$. I will omit further explanations like this one that are easy to verify from concavity of $b(x) - xy$ in x.

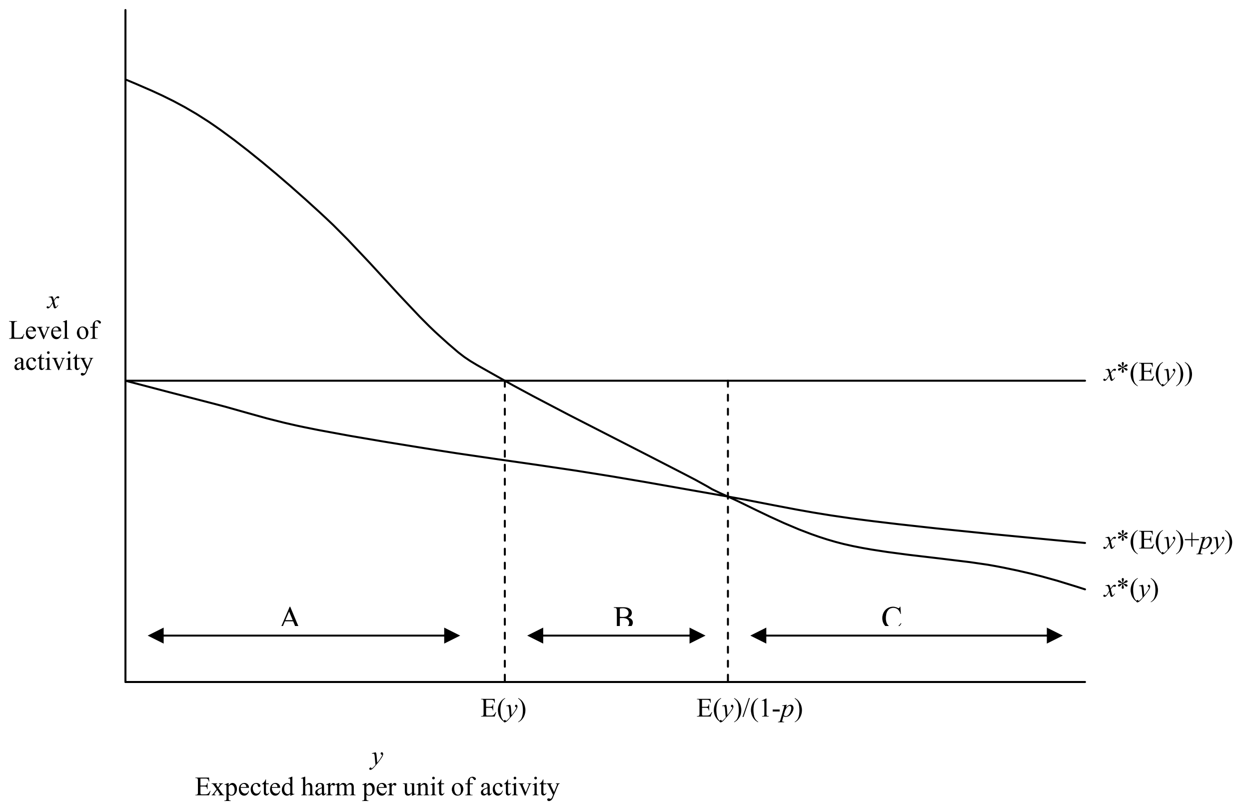
x
Level of
activity
x*(E(y))
x*(E(y)+py)
x*(y)
A
B
C
E(y)
E(y)/(1-p)
y
Expected harm per unit of activity

Figure 1

Next, observe that

$$\int_0^{E(y)}[b'(x^*(E(y)+py)-y]f(y)dy > \int_0^{E(y)}[b'(x^*(E(y))-y]f(y)dy. \quad (7)$$

This also can be seen from Figure 1. In region A, an increase in x from $x^*(E(y) + py)$ will increase welfare more than an increase in x from $x^*(E(y))$ since the former is more distant from $x^*(y)$ and welfare is concave in x. Similarly, we have that

$$\int_{E(y)}^{m}[b'(x^*(E(y)+py)-y]f(y)dy > \int_{E(y)}^{m}[b'(x^*(E(y))-y]f(y)dy. \quad (8)$$

To explain, in region B, an increase in x from $x^*(E(y) + py)$ will raise welfare since x will become closer to $x^*(y)$, whereas an increase in x from $x^*(E(y))$ will lower welfare since x will become farther from $x^*(y)$. In region C, an increase in x from $x^*(E(y) + py)$ will reduce welfare by less than an increase in x from $x^*(E(y))$ since the former is closer to $x^*(y)$. Hence, over both regions B and C, $b'(x^*(E(y) + py) - y > b'(x^*(E(y)) - y$, from which (8) follows. Finally, (7) and (8) imply (5).

I now show that (5) implies (3) given the assumption that b displays decreasing absolute risk aversion. Note first that the integrand of (3) equals the integrand of (5) multiplied by $x^{*\prime}(E(y) + py)$.

I first claim that $x^{*\prime}(E(y) + py) < 0$ and that it increases with y. To verify this, observe first that $x^{*\prime}(E(y) + py) = 1/b''(x^*(E(y) + py)) < 0$, for differentiation of $b'(x(t)) = t + py$ with respect to t gives $x'(t) = 1/b''(x(t))$. Second, note that $x^{*\prime}(E(y) + py)$ will increase with y if $b'''(x) > 0$. In particular, differentiation of $x'(t) = 1/b''(x(t))$ gives $x''(t) = -b'''(x(t))x'(t)/[b''(x(t))]^2$, so that the sign of $x''(t)$ equals the sign of $b'''(x(t))$. The assumption of decreasing absolute risk aversion implies that $b'''(x(t)) > 0$, for this assumption means that $-b''(x)/b'(x)$ decreases with x.

I now show that (3) holds. Recall that I demonstrated above that $[b'(x^*(E(y) + py) - y]$ is positive over regions A and B and negative over region C, and that $\int_0^{\infty}[b'(x^*(E(y) + py) - y]f(y)dy > 0$. It will follow that for any function $w(y)$ such that $w(y) > 0$ and $w'(y) < 0$, we must have

$$\int_0^{m} w(y)[b'(x^*(E(y)+py)-y]f(y)dy > 0. \quad (9)$$

To show (9), let w^* equal $w(E(y)/(1-p))$, namely, the value of w at the point between regions B and C. Then we have

$$\int_0^{m} w(y)[b'(x^*(E(y)+py)-y]f(y)dy > \int_0^{m} w^*[b'(x^*(E(y)+py)-y]f(y)dy, \quad (10)$$

since $w(y)[b'(x^*(E(y) + py) - y]f(y) > w^*[b'(x^*(E(y) + py) - y]f(y)$ for $y < E(y)/(1-p)$ (because for such y, $w(y) > w^*$ and $[b'(x^*(E(y) + py) - y]f(y) > 0$) as well as for $y > E(y)/(1-p)$ (because for such y, $w(y) < w^*$ and $[b'(x^*(E(y) + py) - y]f(y) < 0$). But

$$\int_0^{m} w^*[b'(x^*(E(y)+py)-y]f(y)dy = w^*\int_0^{m}[b'(x^*(E(y)+py)-y]f(y)dy > 0, \quad (11)$$

since w* > 0 and (5) holds. Hence, (9) is established. Now since $x^{*\prime}(E(y) + py) < 0$ and increases with y, we know that $-x^{*\prime}(E(y) + py) > 0$ and decreases with y. Thus, $-x^{*\prime}(E(y) + py)$ may play the role of $w(y)$, so that (9) implies

$$-\int_0^{m} x^{*\prime}(E(y)+py)[b'(x^*(E(y)+py)-y]f(y)dy > 0, \quad (12)$$

which is equivalent to (3).

Finally, the argument that has been given would apply essentially unchanged for any $t > E(y)$ and would show that $W_{TL}'(t) < 0$. The only difference would be that the graph of $x^*(t + py)$ would lie below that of $x^*(\mathrm{E}(y) + py)$ in Figure 1; but this would not affect the logic of any of the steps of the proof.