1 Introduction

In his 1986 paper, Esteban suggests that for many applications the size distribution of income may be usefully described by the *income share elasticity*, as an alternative to the conventional density representation. This notion is put forward as a convenient way to impose stylized-fact restrictions to be tested against the empirical evidence, and to provide criteria for identifying different classes of distributions.

On the other hand, in many economic applications the interesting feature to be studied is income dispersion, usually measured by indices of (first or second order) stochastic dominance. In this note we draw a link between the two, by providing sufficient conditions on the shape of the income share elasticity which support (first or second order) stochastic dominance – that is, such that a given shock to the income share elasticity has dispersion effects as measured by stochastic dominance.

The paper's main results are presented in the next section, while concluding remarks are gathered in section 3.

2 Income share elasticity and income dispersion

Income is distributed over some support (y_m, y_M), $y_M > y_m \geq 0$, according to the density $f(y,\theta) > 0$ for all $y \in (y_m, y_M)$, such that the distribution of income is defined by $F : (y_m, y_M) \times \mathbb{R} \to [0,1]$. The real parameter θ measures a shift of the distribution which may be thought of as an index of dispersion, to be made more precise below. Letting subscripts denote derivatives, Esteban's income share elasticity is defined as

$$\pi(y,\theta) = \lim_{h\to 0} \frac{d\log\left(\frac{1}{\mu}\int_y^{y+h} xf(x,\theta)dx\right)}{d\log y} = 1 + \frac{yf_y(y,\theta)}{f(y,\theta)} \tag{1}$$

and measures the relative marginal change in the share of income accruing to class y, brought about by a marginal increase in y. A one-to-one relationship exists between π and the conventional density representation of the size distribution of income. Esteban (1986, p.443) identifies three restrictions which seem well supported by empirical evidence, and can be formalized using (1): (*i*) the weak-weak Pareto Law, according to which π approaches some constant value $-\alpha < 0$ as y tends to infinity; (*ii*) the existence of at least one mode, which implies that $\pi(y,\theta) = 1$ has at least one solution over the support of the distribution; and (*iii*) that π exhibits a constant rate of decline, implying for all y that either $\pi_y(y,\theta)=0$, or the elasticity of $\pi_y(y,\theta)$ is a negative constant. A generalized three-parameter Gamma distribution satisfies all of these.[1]

As to stochastic dominance (SD), it is well known (e.g., Hirshleifer and Riley, 1992, ch.3) that a change in the parameter θ identifies a first or second order SD shift of the distribution whenever the following conditions are satisfied for all $y \in (y_m, y_M)$:

$$F_\theta(y,\theta) \leq 0 \tag{2}$$

for first order SD; and

$$\int_{y_m}^{y} F_\theta(x,\theta)dx \leq 0 \tag{3}$$

for second order SD; both inequalities hold strictly somewhere over the support of the distribution. It is a well known fact, widely used in economic applications, that the expected value of any increasing (increasing concave) function is increasing in θ whenever the latter measure first (second) order SD.

The following proposition establishes sufficient conditions on π, for θ to be a first-order SD parameter.

Proposition 1 *If $\pi_\theta(y,\theta) > 0$ for all $y \in (y_m, y_M)$, then θ is a first-order stochastic dominance parameter of $F(\cdot;\cdot)$, that is $F_\theta(y,\theta) \leq 0$ for all $y \in Y$.*

Proof. Let $\lambda(y,\theta) = f_\theta(y,\theta)/f(y,\theta)$, with $f(y,\theta) > 0$ for all $y \in (y_m, y_M)$. It is easily seen that $\pi_\theta(y,\theta) = y\lambda_y(y,\theta)$, so that $\pi_\theta(y,\theta) > 0$ for all y means that $\lambda(y,\theta)$ is monotonically increasing in y for any given θ. By definition $\int_{y_m}^{y_M} \lambda(y,\theta)f(y,\theta)dy = F_\theta(y_M,\theta) = 0$, which, $f(y,\theta)$ being positive and the overall integral nil, implies that $\lambda(y,\theta)$ takes both negative and positive values. Since $\lambda(y,\theta)$ is increasing in y, the smallest (negative) value of y identifies the minimum of λ, occurring at $y = y_m$ and, by the same token, $\lambda(y_M,\theta) > 0$ is a maximum for λ. Hence, there is a unique value $\widehat{y}$ of y such that $\lambda(\widehat{y},\theta) = 0$. Consider now the function $F_\theta(y,\theta) = \int_{y_m}^{y} f_\theta(x,\theta)dx$, the first derivative of which is $f_\theta(y,\theta) = \lambda(y,\theta)f(y,\theta)$. Since $\text{sign}\{f_\theta(y,\theta)\} = \text{sign}\{\lambda(y,\theta)\}$, $f_\theta(y,\theta)$ vanishes at $\widehat{y}$ which is the unique minimum for $F_\theta(y,\theta)$. As $\lambda(y,\theta)$ is negative (positive) for y close to y_m (y_M), so will be $f_\theta(y,\theta)$:

[1] Esteban points out that "the Pareto, Gamma and Normal density functions correspond to constant, linear and quadratic elasticities, respectively" (1986, p.442).

$F_\theta(y,\theta)$ points down (up) around y_m (y_M). As $F_\theta(y_m,\theta) = F_\theta(y_M,\theta) = 0$, $F_\theta(y,\theta)$ lies below the zero line: θ is then a first order SD parameter. ■

An immediate implication of Proposition 1 concerns Esteban's finding (1986, p.444) that the family of distributions obeying all three restrictions mentioned above[2] are such that

$$\pi(y,\cdot) = -\alpha + (\beta/\epsilon)y^{-\epsilon} \tag{4}$$

with $\alpha > 1$, and $\beta, \epsilon > 0$: a decrease (increase) in α (β) is a first order SD shock, and accordingly raises mean income. More generally, first order stochastic dominance clearly implies (inverse) Lorenz dominance. This can also be seen using directly Proposition 1:

Corollary *Consider two distributions $f(y,\theta_i)$, $i = 1,2$ such that $\pi_1 = \pi(y,\theta_1) > \pi_2 = \pi(y,\theta_2)$ for all y. Then $f(y,\theta_2)$ Lorenz-dominates $f(y,\theta_1)$.*

Proof. Define $\widehat{f}_i(y) = (1/\mu_i)yf(y,\theta_i) > 0$, which can be treated as a density. It is easily seen that the corresponding Esteban elasticity $\widehat{\pi}_i$ satisfies $\widehat{\pi}_i = 1 + \pi_i$. There follows that $\widehat{\pi}_1 > \widehat{\pi}_2$ and hence, by Proposition 1, $\widehat{F}_2(y) = \int_{y_m}^{y} \widehat{f}_2(x)dx > \widehat{F}_1 = \int_{y_m}^{y} \widehat{f}_1(x)dx$ for all y, which is equivalent to Lorenz dominance.[3] ■

The Corollary may be convenient whenever elasticities are involved in assessing Lorenz dominance. For example, it is well known that the distribution of post-tax income Lorenz dominates that of pre-tax income if taxation is progressive (e.g., Lambert, 2001, p.190). This result follows immediately if we identify $i = 1$ with pre-tax and $i = 2$ with post-tax income: it is then readily seen that $\widehat{\pi}_1 - \widehat{\pi}_2 = 1 - R$, where R is residual progression (which of course is less than one by the definition of progressive taxation).[4]

We now take up second order SD. The following proposition establishes conditions on π for θ to be a parameter of second order SD.

[2] With the proviso that the mode m is unique and satisfies $m = \left(\frac{\beta}{\epsilon}\frac{1}{1+\alpha}\right)^{1/\epsilon}$.

[3] An alternative proof, *not* involving first order stochastic dominance, may be based on Lambert's Lemma 8.1 (2001, p.200), according to which if $m(y)$ and $n(y)$ are two attributes of an income distribution, the m-concentration curve dominates the n-concentration curve iff the y-elasticity of m (e_m, say) lower than that of n (e_n). In our case, let the income distribution be $f(y,\theta_2)$, and define $m(y) = 1$ and $n(y) = f(y,\theta_1)/f(y,\theta_2) > 0$: then $e_m = 0$ and $e_n = \pi_1 - \pi_2$, so that $e_m < e_n$ is $\pi_1 > \pi_2$.

[4] More specifically, given the distribution $f(y)$ of pre-tax income, one has $\widehat{f}_1(y) = (1/\mu)yf(y)$, and $\widehat{f}_2(y) = [y - t(y)]f(y)/(\mu - T)$ for post-tax income, where $t(\cdot) > 0$ is the tax schedule and T is total tax liabilities. Then $\widehat{\pi}_1 = 1 + \pi$ and $\widehat{\pi}_2 = R + \pi$. Here, π is the Esteban elasticity of f and $R = y[1 - t'(y)]/[y - t(y)]$. See e.g. Lambert (2001, p.196-97).

Proposition 2 *Assume $\mu_\theta \geq 0$. Then, if $\pi_\theta(y,\theta)$ is monotonically decreasing in y and crosses zero at some $\widetilde{y} \in (y_m, y_M)$, θ is a second order stochastic dominance parameter.*

Proof. From Proposition 1, $\pi_\theta(y,\theta)$ monotonically decreasing in y implies that $y\lambda_y(y,\theta)$ is decreasing, i.e. $\lambda_y(y,\theta) + y\lambda_{yy}(y,\theta) < 0$. Since $\widetilde{y}\lambda_y(\widetilde{y},\theta) = 0 = \lambda_y(\widetilde{y},\theta)$ and $\lambda_y(\widetilde{y},\theta) + \widetilde{y}\lambda_{yy}(\widetilde{y},\theta) = \widetilde{y}\lambda_{yy}(\widetilde{y},\theta) < 0$ (all y being positive), $\widetilde{y}$ is the unique interior maximum of $\lambda(y,\theta)$. By definition $\int_{y_m}^{y_M} \lambda(y,\theta)f(y,\theta)dy = F_\theta(y_M,\theta) = 0$: hence, $\lambda(y,\theta)$ takes on both positive and negative values, which implies $\lambda(\widetilde{y},\theta) > 0$. As this is the unique turning point of λ, there are either one, or two interior values of y where $\lambda(\cdot,\theta) = 0$.

(a) Suppose there is only one such value, y_0 say. Then note that $y_0 < \widetilde{y}$. To see this, assume to the contrary $y_0 > \widetilde{y}$, so that $\lambda(y,\theta) > 0$ for $y < y_0$, and $\lambda(y,\theta) < 0$ for $y > y_0$; then $F_\theta(y,\theta) = \int_{y_m}^{y} f_\theta(x,\theta)dx > 0$ for all y: indeed, its first derivative is $f_\theta(y,\theta) = \lambda(y,\theta)f(y,\theta)$ and $f(y,\theta) > 0$ implies $\text{sign}\{f_\theta(y,\theta)\} = \text{sign}\{\lambda(y,\theta)\}$: $f_\theta(y,\theta)$ vanishes at y_0, which is the unique maximum for $F_\theta(y,\theta)$, with $f_\theta(y,\theta)$ positive (negative) for y lower (higher) than y_0. Thus $F_\theta(y,\theta)$ points up (down) around y_m (y_M) and, as $F_\theta(y_m,\theta) = F_\theta(y_M,\theta) = 0$, it lies above the zero line. But this contradicts the assumption $\mu_\theta \geq 0$, since it implies trivially $\mu_\theta < 0$. Hence, indeed $y_0 < \widetilde{y}$, and so $\lambda(y,\theta) < 0$ for $y < y_0$, and $\lambda(y,\theta) > 0$ for $y > y_0$. Now we can apply the same reasoning, and take the function $F_\theta(y,\theta)$. Its derivative $f_\theta(y,\theta)$ obeys $\text{sign}\{f_\theta(y,\theta)\} = \text{sign}\{\lambda(y,\theta)\}$ and vanishes at y_0, the unique minimum for $F_\theta(y,\theta)$. As $\lambda(y,\theta)$ is negative (positive) for y lower (higher) than y_0, so will be $f_\theta(y,\theta)$: $F_\theta(y,\theta)$ points down (up) around y_m (y_M). As $F_\theta(y_m,\theta) = F_\theta(y_M,\theta) = 0$, $F_\theta(y,\theta)$ lies below the zero line: θ is then a first (and hence second) order SD parameter.

(b) Consider now the case where there exist two values of y, $y_1 < y_2$ say, such that $\lambda(y_1,\theta) = \lambda(y_2,\theta) = 0$. Then (*i*) $\widetilde{y} \in (y_1,y_2)$; (*ii*) $\lambda(y,\theta) < 0$ for $y \notin (y_1,y_2)$ and $\lambda(y,\theta) > 0$ for $y \in (y_1,y_2)$. As $\text{sign}\{f_\theta(y,\theta)\} = \text{sign}\{\lambda(y,\theta)\}$, the same goes for $f_\theta(y,\theta)$: thus $F_\theta(y,\theta)$ is decreasing around y_m, and has a negative minimum in y_1; also, it is decreasing around y_M, as $F_\theta(y_M,\theta) = 0$ and y_2 identifies a positive maximum. This implies 'single crossing', as there is only one value $y^* \in (y_1,y_2)$ such that $F_\theta(y^*,\theta) = 0$. Define now $S(y,\theta) = \int_{y_m}^{y} F_\theta(x,\theta)dx$, so that obviously $S_y(y,\theta) = F_\theta(y,\theta)$. Since there is single crossing at y^* (where $S_y(y^*,\theta) = 0$) and $S_{yy}(y^*,\theta) = f_\theta(y^*,\theta) > 0$, y^*identifies the unique turning point, which is minimum. As $S(y_m,\theta) = 0$ and $\mu_\theta \geq 0$ implies $S(y_M,\theta) \leq 0$, $S(y,\theta)$ always lies below the zero axis and θ is an index of second order SD. If $\mu_\theta = 0 = S(y_M,\theta)$, it is a (inverse) mean preserving spread. ■

In analogy with Proposition 1, one implication of Proposition 2 is that it throws a bridge between the income share elasticity and Lorenz dominance. While obviously the discussion of Proposition 1 applies also to case (a) of the proof (which actually delivers first order SD), case (b) is connected with Shorrocks' generalized Lorenz dominance: as is well known, if the function $S(y, \theta)$ used in the proof does not change sign, generalized Lorenz curves never intersect (e.g., Lambert, 2001, p.55).[5]

3 Concluding remarks

The notion of income share elasticity can have useful economic applications, for example when dealing with the relationship between income distribution and the price elasticity market demand (Benassi *et al.*, 2002). In this note we have outlined the relationship between (first and second order) stochastic dominance, and the way income share elasticity depends on the distribution parameters; this also allows to see some related implications in terms of Lorenz dominance.

References

[1] Benassi C., A.Chirco and M.Scrimitore (2002): Income Concentration and Market Demand, *Oxford Economic Papers*, forthcoming.

[2] Esteban, J. (1986): Income Share Elasticity and the Size Distribution of Income, *International Economic Review*, **27**, 439-44.

[3] Hirshleifer J. and J.G.Riley (1992): *The Analytics of Uncertainty and Information*, Cambridge University Press, Cambridge.

[4] Lambert, P.J. (2001): *The Distribution and Redistribution of Income*, Manchester University Press, Manchester.

[5] This can be directly seen by defining the generalized Lorenz curve as $L(p, \theta) = \int_0^p y(p, \theta)dp$, where $y(p, \theta)$ satisfies $F(y, \theta) = p$ so that $dp = f(y, \theta)dy + F_\theta(y, \theta)d\theta$. By implicit differentiation, $y_\theta(p, \theta) = -F_\theta(y(p, \theta), \theta)/f(y(p, \theta), \theta)$ so that $L_\theta(p, \theta) = \int_0^p y_\theta(p, \theta)dp = -\int_0^p F_\theta(y(p, \theta), \theta)/f(y(p, \theta), \theta)dp = -\int_{y_m}^y F_\theta(y, \theta)dy = -S(y, \theta)$. As established above, the latter is positive in case (b) of the proof of Proposition 2.