Inequality and the multi-dimensional measurement of development: some remarks

François Bourguignon
Paris School of Economics

Oxford, OPHI, 14 June, 2010

Motivation

- Development essentially multi-dimensional
- Population heterogeneity with respect to development achievements
- Development measurement may be done explicitly on multi-dimensional basis (dart board for distinct population groups)
- Desirable to summarize the various dimensions and population heterogeneity into a single (or various) scalar(s)
- Most obvious route = 2-stage procedure used in multi-dimensional inequality measurement

Outline

1. Definition of 2-stage aggregation index and the 'decomposability' issue
2. An elementary aggregate decomposition
3. Sequential decomposition based on co-distribution
4. The FLS and the IA-HDI index
5. Conclusion

1. Definition of 2-stage aggregation index and 'decomposability'

- Individual attributes: y_i, z_i $i = 1, 2\ldots n$
- 1st stage: Aggregator ('utility') function at individual level: $u(y_i, z_i)$ with usual properties (increasing and concave)
- 2nd stage: Overall index (overall mean welfare):

$$W = \frac{1}{n}\sum_{i=1}^{n}\varphi\left[u(y_i, z_i)\right] \qquad or \qquad W = G\left\{\frac{1}{n}\sum_{i=1}^{n}\varphi\left[u(y_i, z_i)\right]\right\}$$

with φ () = individual 'welfare', φ'() > 0, φ"() < 0 and G'() > 0

- Decomposability: possible to express W as:

$$W = F(\bar{y}, \bar{z}; I_y; I_z; I_{yz}) \qquad ?$$

2. An elementary aggregate decomposition

- φ(u) = u; G'( ) =0
- Dalton multi-dimensional inequality measure

$$D_u = 1 - \frac{\sum_{i=1}^{n} u(y_i, z_i)}{nu(\bar{y}, \bar{z})}$$

$$nu(\bar{y}, \bar{z}) = Max \sum_{i=1}^{n} u(y_i, z_i) \ s.t. \sum_i y_i = n\bar{y}, \sum_i z_i = n\bar{z}$$

- Overall welfare: $W = u(\bar{y}, \bar{z}).(1 - D_u)$

- Equivalent specification with 'ede' (Multidimensional Atkinson and Tsui indices, Weymark, 2003)
- D_u incorporates implicitly I_y, I_z, I_{yz}. Issue is to make that relationship more explicit

The case of a linear utility function

- Consider the case φ(u) = u, $u_i = ay_i + bz_i$

Then of course: $$W = (a\bar{y} + b\bar{z})$$

- But with φ(u) # u then: $$W = \varphi(a\bar{y} + b\bar{z})(1 - D^{\varphi})$$

- D^{φ} is a measure of the inequality of u, and can be decomposed as in Shorrocks (1982)

$$W = \varphi(a\bar{y} + b\bar{z})(1 - D^{\varphi} s_y - D^{\varphi} s_z)$$

with $$s_y = \frac{a^2\sigma^2(y) + ab\,\mathrm{cov}(y,z)}{a^2\sigma^2(y) + 2ab\,\mathrm{cov}(y,z) + b^2\sigma^2(z)}, \quad s_z = 1 - s_y$$

Linear utility function (ct'd)

$$W = \varphi(a\bar{y} + b\bar{z})(1 - D^{\varphi} s_y - D^{\varphi} s_z)$$

$$s_y = \frac{a^2\sigma^2(y) + ab\,\mathrm{cov}(y,z)}{a^2\sigma^2(y) + 2ab\,\mathrm{cov}(y,z) + b^2\sigma^2(z)}, \quad s_z = 1 - s_y$$

- One can go further and decompose overall inequality into what is due to inequality in y (I_y), inequality in z (I_z) and the covariance between y and z (I_{yz})
- Note that this decomposition is highly non-linear
- Even with zero covariance between y and z, not possible to decompose W into a y- and a z-component.

The case of separability

$$u(y_i, z_i) = a(y_i) + b(z_i); \quad \varphi(u) = u; \quad G(x) \neq x$$

$$W = G\left\{\frac{1}{n}\sum_i [a(y_i) + b(z_i)]\right\} \Rightarrow W = G\left\{a(\bar{y}).(1 - D_y^a) + b(\bar{z}).(1 - D_z^b)\right\}$$

- Welfare aggregate index satisfies the general decomposition property
- Co-distribution of y and z does not matter, unlike in the preceding case
- Value of u_{yz} clearly important (as in Atkinson-Bourguignon)

3. Sequential decomposition based on co-distribution

- 'Inequality' of z conditionally on the distribution of y

$$D_u^{z|y} = 1 - \frac{\sum_{i=1}^{n} u(y_i, z_i)}{\sum_{i=1}^{n} u[y_i, z_i^*(y_i, \bar{z})]}$$

$$z_i^*(y_i, \bar{z}) = Arg\,Max_{z_i} \sum_{i=1}^{n} u(y_i, z_i) \ s.t. \ \sum_i z_i = n\bar{z}$$

- Properties of $z_i^*(y_i, \bar{z})$

$$\frac{\partial z_i^*}{\partial y_i} \geq 0 \ if \ \ u_{yz} > 0 \ ; \quad \frac{\partial z_i^*}{\partial y_i} \leq 0 \ if \ \ u_{yz} < 0$$

- Particular case: $z_i^*(y_i, \bar{z}) = y_i \frac{\bar{z}}{\bar{y}} \ \ if \ \ u(y, z) \ HOM \ \partial°1$

Sequential decomposition (ct'd)

- Defining "inequality of y after optimizing x"

$$D_{u*}^{y} = 1 - \frac{\sum_{i=1}^{n} u[y_i, z_i^*(y_i, \bar{z})]}{nu(\bar{y}, \bar{z})}$$

- Hence the sequential decomposition:

$$W = u(\bar{y}, \bar{z}).(1 - D_{u*}^{y}).(1 - D_{u}^{z|y})$$

- Remark: $D_{u}^{z|y}$ incorporates the inequality in z as well as the covariance with y

- This means that one dimension is given some priority

Sequential decomposition (end)

- Particular case: $u(y,z) = (y^{\alpha} + z^{\alpha})^{1/\alpha}$
- Then:

$$D_u^{z|y} = 1 - \frac{1}{n(1+\bar{z}/\bar{y})} \sum_i \left[1 + (\frac{y_i}{z_i})^{\alpha} \right]^{1/\alpha}$$

- This inequality index actually measures the deviation from the distribution of y!
- As $u(y_i, z^*_i)$ is now linear in y_i, the inequality index for y is now zero.
- The only that matters thus is the degree of non-proportionality of z and y.
- Things would be different with φ(u) # u!

4. The FLS and the IA-HDI index

- Focus here on the conceptual principles behind those indices not their empirical implementation

- In terms of the preceding framework, the FLS may be defined as follows:

$$u(y,z) = (y^{\alpha} + z^{\alpha})^{1/\alpha};\ \varphi(u) = u^{\alpha};\ G(x) = x^{1/\alpha};\ \alpha \leq 1$$

which is equivalent to :

$$u(y,z) = (y^{\alpha} + z^{\alpha});\ \varphi(u) = u;\ G(x) = x^{1/\alpha}$$

leading to:

$$FLS = [\frac{1}{n}\sum_{i}(y_i^{\alpha} + z_i^{\alpha})]^{1/\alpha}$$

FLS and IA-HDI: decomposition

$$FLS = [\frac{1}{n}\sum_i (y_i^\alpha + z_i^\alpha)]^{1/\alpha}$$

- According to earlier result on separability, this leads to the following decomposition

$$FLS = [(\bar{y}^\alpha (1 - D_\alpha^y) + \bar{z}^\alpha (1 - D_\alpha^z)]^{1/\alpha}$$

- Or using Atkinson rather than Dalton measures for the inequality of y and z:

$$FLS = \{[\bar{y}(1 - A_\alpha^y)]^\alpha + [\bar{z}(1 - A_\alpha^z)]^\alpha]\}^{1/\alpha}$$

- A dual decomposition is:

$$FLS = \{\sum_i u_i^\alpha\}^{1/\alpha} \ \ with\ u_i = (\frac{y_i + z_i}{2}).(1 - A(y_i, z_i))$$

FLS and IA-HDI : missing y-z correlation

- Separability implies that the co-distribution of y and z is ignored in these indices
- Yet this may be an important aspect of multi-dimensional inequality:
  - If y =income and z = health, co-distribution of y and z shows horizontal inequality w.r.t. health
  - This does not mean that health inequality does not matter per se (but probably not in an usual sense)
  - Issues are linked: health inequality very much affected by infant mortality + infant mortality higher in low income households

FLS and IA-HDI : the CES specification

- First level assumption in FLS = CES-α combination of individual atributes
- Second level assumptions in FLS, CES-α aggregation of individual utilities (equivalent to $\varphi(u) = u^{\alpha}$ and $G(x)=x^{1/\alpha}$)
- This double CES- α key for separability, decomposition formula and its dual
- First level assumption only –i.e. φ(u) = u and G(x)=x leads to rather different results

$$W = \frac{1}{n}\sum_{i=1}^{n}(y_i^{\alpha} + z_i^{\alpha})^{1/\alpha}$$

- As $u_{yz}>0$, optimal co-distribution of z = perfect positive correlation

FLS and IA-HDI : the two dimensions of 'inequality'

- Preceding (undesirable) result may be reversed by second-stage assumptions slightly different from FLS
- For instance, $\varphi(u) = u^{\beta}$ and $G(x)=x^{1/\beta}$ with $\beta < \alpha$ does not lead to separability and makes a negative correlation between y and z being optimal
- The distinction β vs. α makes very much sense.
  - α describes the way in which attributes combine to define individual utility, β describes the aversion of society to inequality
  - No reason for both to be the same!
  - Same problem as the confusion between risk aversion and intertemporal substitutability in consumer model (see also Shokaert).

5. Conclusion

- When co-distribution is not observed, FLS/IA-HDI is a clever and consistent way of dealing with multi-dimensionality
- Yet, partial information that may be available on co-distribution –e.g. infant mortality by income level- should be used
- Incomplete information makes the general decomposability issue especially relevant
- With complete information, always possible to correct for 'overall multi-dimensional inequality'
- Shares of attributes in total inequality very indicative: possible to go beyond the linearity case?