E C O N O M I C S   B U L L E T I N

Decomposition of Gini and the generalized entropy inequality measures

Stéphane Mussard
LAMETA, University of Montpellier I

Françoise Seyte
LAMETA, University of Montpellier I

Michel Terraza
LAMETA, University of Montpellier I

Abstract

In this article we provide an overview of the Gini decomposition and the generalized entropy inequality measures, a free access to their computation, an application on French wages, and a different way than Dagum to demonstrate that the Gini index is a more convenient measure than those issued from entropy: Theil, Hirschman−Herfindahl and Bourguignon.

Citation: Mussard, Stéphane, Françoise Seyte, and Michel Terraza, (2003) "Decomposition of Gini and the generalized entropy inequality measures." *Economics Bulletin*, Vol. 4, No. 7 pp. 1−6
Submitted: January 27, 2003. **Accepted:** January 27, 2003.
URL: http://www.economicsbulletin.com/2003/volume4/EB−03D30001A.pdf

1. Introduction

This paper analyses Dagum's (1997a, 1997b, 1998) articles about the Gini index decomposition and three particular cases of the entropy coefficient (Theil, Hirschman-Herfindahl and Bourguignon decompositions). The aim of this article is multiple. First, it facilitates the access to the computation of these decomposed measures and gives a theoretical overview (Section 2). Secondly, it presents an application on the French wages (Section 3). Finally, it gives a different approach to Dagum's (1998) assessment of the Gini ratio like a more complete measure, in studying the normative concepts of inequality measurements.

2. The Decompositions of Gini, Theil, Hirschman-Herfindahl and Bourguignon

On the web site: http://www.lameta.univ-montp1.fr/online/gini.html we propose a free program to estimate the four decompositions and all the needed directives to use it.

2.1 The Gini Decomposition

Let us consider a population P with n income units y_i (i = 1,…,n) where F(y), μ and G are respectively the cumulative distribution, the mean and the Gini index calculated on P, which is partitioned in k subpopulations P_j (j=1,…,k). The size and the income average of P_j are given by n_j and μ_j. The Gini index measured on P is:

$$G = \frac{\sum_{i=1}^{n}\sum_{r=1}^{n}|y_i - y_r|}{2n^2\mu} \ . \tag{1}$$

The Gini within the subpopulation P_j (within-group Gini) is given by:

$$G_{jj} = \frac{\sum_{i=1}^{nj}\sum_{r=1}^{nj}|y_i - y_r|}{2n_j^2\mu_j} \ , \tag{2}$$

and the between-group Gini (that calculates the inequalities between P_j and P_h) is:

$$G_{jh} = \frac{\sum_{i=1}^{nj}\sum_{r=1}^{nh}|y_{ji} - y_{hr}|}{\mu_j + \mu_h} \ . \tag{3}$$

All these ratios are included in the interval [0,1]. If they tend towards 1 the income repartition is unequal, and if they tend towards 0 the repartition is equal. Now, let us introduce two fundamental concepts. On one hand, the gross economic affluence (see Dagum 1997b) is expressed in the form:

$$d_{jh} = \int_0^\infty dF_j(y) \int_0^y (y - x) dF_h(x) \quad \forall\ \mu_j > \mu_h \ . \tag{4}$$

It is the expected income difference between the groups j and h such as: $y_{ji} > y_{hr}$ and $\mu_j > \mu_h$. On the other hand, the first order moment of transvariation is the expected income difference between P_j and P_h given that, $y_{ji} < y_{hr}$ and $\mu_j > \mu_h$:

$$p_{jh} = \int_0^\infty dF_h(y) \int_0^y (y - x)\ dF_j(x) \quad \forall\ \mu_j > \mu_h \ . \tag{5}$$

According to (4) and (5) we can introduce the relative economic affluence (Dagum 1980). It is a normalized index that indicates the "distance" between P_j and P_h:

$$D_{jh} = (d_{jh} - p_{jh}) / \Delta_{jh} = (d_{jh} - p_{jh}) / (d_{jh} + p_{jh}) \ . \tag{6}$$

Calculating $G_{jh}\times D_{jh}$, we proceed to the net measure of the between-group Gini. It symbolizes the inequalities derived from the non-overlap of the distributions j and h. The expression $G_{jh}\times(1-D_{jh})$ is the transvariation between P_j and P_h, which is the part of the inequality issued from the overlap of the distributions j and h.

If p_j and s_j are respectively the percentage of the individuals belonging to P_j and the income share of the subpopulation j, we have:

$$p_j = \frac{n_j}{n},\ \ s_j = \frac{n_j\ \mu_j}{n\ \mu}\ . \tag{7}$$

According to (3), (6) and (7) we can define the first component of the Gini decomposition. It is the net contribution of the between-group inequalities to the overall Gini measured on P:

$$G_b = \sum_{j=2}^{k}\sum_{h=1}^{j-1} G_{jh}\ D_{jh}(p_j\ s_h + p_h\ s_j)\ . \tag{8}$$

The second component is the contribution of the transvariation between the subpopulations to G:

$$G_t = \sum_{j=2}^{k}\sum_{h=1}^{j-1} G_{jh}\ (1 - D_{jh})(p_j\ s_h + p_h\ s_j)\ . \tag{9}$$

The third element is the contribution of the within-group inequalities to G:

$$G_w = \sum_{j=1}^{k} G_{jj}\ p_j\ s_j\ . \tag{10}$$

Finally, given (8), (9) and (10) the fundamental equation of the Gini decomposition in three components is:

$$G = G_w + G_b + G_t\ . \tag{11}$$

2.2. The Theil, Hirschman-Herfindahl and Bourguignon Decompositions

The Theil, Hirschman-Herfindahl (H-H) and Bourguignon indexes are three particular cases of the generalized entropy ratio given by:

$$I_\beta = \frac{1}{\beta(\beta+1)n}\sum_{j=1}^{k}\sum_{i=1}^{n_j}\frac{y_{ji}}{\mu}\left[\left(\frac{y_{ji}}{\mu}\right)^{\beta} - 1\right],\ \beta \text{ real.} \tag{12}$$

The generalized entropy index I_β can be decomposed in a within-group contribution $I_{\beta w}$ and a between-group contribution $I_{\beta b}$:

$$I_{\beta w} = \sum_{j=1}^{k}\frac{n_j\mu_j}{n\mu}\left(\frac{\mu_j}{\mu}\right)^{\beta} I_{\beta wj}\ , \tag{13}$$

$$I_{\beta b} = \frac{1}{\beta(\beta+1)}\sum_{j=1}^{k}\frac{n_j}{n}\ \frac{\mu_j}{\mu}\left[\left(\frac{\mu_j}{\mu}\right)^{\beta} - 1\right], \tag{14}$$

such as I_β is separable in two components,

$$I_\beta = I_{\beta w} + I_{\beta b}\ . \tag{15}$$

2.2.1. The Theil Decomposition

The Theil index T is the generalized entropy ratio when β tends towards 0:

$$T = \lim_{\beta\to 0} I_\beta = \frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n_j}\frac{y_{ji}}{\mu}\ log\ \frac{y_{ji}}{\mu}\ . \tag{16}$$

The between-group contribution T_b and the within-group contribution T_w are:

$$T_b = \lim_{\beta\to 0} I_\beta = \sum_{j=1}^{k}\frac{n_j}{n}\frac{\mu_j}{\mu}\ log\frac{\mu_j}{\mu}\ , \tag{17}$$

$$T_w = \lim_{\beta \to 0} I_\beta = \sum_{j=1}^{k} \frac{n_j \mu_j}{n\mu} \frac{1}{n_j} \sum_{i=1}^{n_j} \frac{y_{ji}}{\mu_j} \log \frac{y_{ji}}{\mu_j} , \quad (18)$$

such as,

$$T = T_w + T_b . \quad (19)$$

2.2.2. The Hirschman-Herfindahl Decomposition

The H-H index I_1 is the particular case of the generalized entropy when β tends towards 1:

$$I_1 = \lim_{\beta \to 1} I_\beta = \frac{1}{2n} \sum_{j=1}^{k} \sum_{i=1}^{n_j} \frac{y_{ji}}{\mu} \left(\frac{y_{ji}}{\mu} - 1 \right) . \quad (20)$$

The within-group contribution I_{1w} and the between-group contribution I_{1b} are:

$$I_{1w} = \lim_{\beta \to 1} I_{\beta w} = \frac{1}{2} \sum_{j=1}^{k} \frac{n_j \mu_j^2}{n\mu^2} \frac{Var\ y_j}{\mu_j^2} = \frac{1}{2} \sum_{j=1}^{k} \frac{n_j \mu_j^2}{n\mu^2} V^2(y_j) , \quad (21)$$

$$I_{1b} = \lim_{\beta \to 1} I_{\beta b} = \frac{1}{2} \sum_{j=1}^{k} \frac{n_j \mu_j}{n\mu} \left(\frac{\mu_j}{\mu} - 1 \right) = \frac{Var\ \mu_j}{2\mu^2} , \quad (22)$$

where *Var* and V^2 are respectively the variance and the coefficient of variation. Therefore, the breakdown of the Hirschman-Herfindahl index is:

$$I_1 = I_{1w} + I_{1b} . \quad (23)$$

2.2.3. The Bourguignon Decomposition

Bourguignon[1] (1979) presents a new coefficient B. Dagum (1997b) demonstrates that it is the limit of the entropy index when β tends towards -1:

$$B = \lim_{\beta \to -1} I_\beta = \log \mu - \log M_g . \quad (24)$$

In the same way than the two precedent ratios, the Bourguignon coefficient is separated in a within-group contribution B_w and a between-group contribution B_b:

$$B_w = \lim_{\beta \to -1} I_{\beta w} = \sum_{j=1}^{k} \frac{n_j}{n} (\log \mu_j - \log M_{gj}) , \quad (25)$$

$$B_b = \lim_{\beta \to -1} I_{\beta b} = \sum_{j=1}^{k} \frac{n_j}{n} \log \frac{\mu}{\mu_j} = \log \mu - \log M_{g\mu j} . \quad (26)$$

The expressions M_g, M_{gj} and $M_{g\mu j}$ are the geometric mean respectively measured on P, P_j and on the vector μ_j (j = 1, …,k). So, the breakdown of the Bourguignon index is:

$$B = B_w + B_b . \quad (27)$$

3. Application

Let us take a 1996-year wage sample of the south area of France. It represents 27,660 individuals ranked by sex (15,394 men and 12,266 women). The methods introduced above allow one to measure the components of the four decompositions. Furthermore, it is possible to know if the inequalities are generated by the wage gaps within the two groups or if the inequalities are engendered by the wage gaps between men and women.

Table 1 shows these results, in giving the percentage of each element in the global inequality. The three entropic indexes give the same contribution. Indeed, the

[1] We should attribute the paternity of the measure (24) to Hart (1970) p.80.

differences between the men and the women represent 2% of the global inequality and the contribution within the subpopulations represents 98% of the overall inequality, whereas the Gini index grants as much importance to the within-group element (50.9%) as to the between-group element (the net between-group component and the transvariation represent 49.1%). Only the Gini decomposition can provide the intensity of transvariation (35.6%), which is the part of the between-group disparities issued from the overlap between the two distributions.

Table 2 indicates that all the measures grant two times more wage gaps between the men than between the women. Nevertheless, the differences of results between the Gini and the entropic indexes are important. So, it is necessary to direct the choice of the users of inequality coefficients in examining the property they check.

4. Comment

The main concern of Dagum's article (1998), we want to comment, is the properties of the social choice theory that the four indexes integrate. Dagum chooses to discuss about the following principles:

- (A) the interpersonal utility comparisons;
- (B) the inequality aversion (the utility function U is concave: $U''<0$);
- (C) and an increasing utility function ($U'>0$).

In his paper, Dagum demonstrates that the Gini ratio satisfies with (A), (B) and (C) requirements and concludes we should retain the Gini coefficient as the principal measure, because the indexes issued from the entropy do not integrate the criteria of the interpersonal utility comparisons.

Nevertheless, under many conditions the interpersonal utility comparisons and the concavity of the utility function are incompatible. If the interpersonal utility comparisons are permitted, individuals have conscience of the crucial last dollar earned by the other individuals (for instance the poor persons). Indeed, persons have more satisfaction because they know that when their incomes rise they actively participate to the future redistribution in order to decrease inequalities and poverty. So, the growth of the utility function can not decrease when incomes increase.

Finally, we can doubt about the complementarity between the (A) and (B) principles. So, it is more convenient for income inequality indexes to satisfy only one of these two properties. Therefore, in a world where only the (A), (B) and (C) principles exist we should accept the coefficients derived from the generalized entropy index. However, because this world does not exist and because the between-group contributions (14), (17), (22) are obtained like a residual ($I_{\beta b} = I_{\beta} - I_{\beta w}$), it is preferable to use the Gini decomposition because the between-group index (G_{jh}) is specified and also because G_b and G_t can not be considered as residuals.

5. Conclusion

We have provided the way for the computation of the Gini decomposition and for the entropic indexes with their specifications. Then, we firstly see that in the south area of France the Gini index attributes as much importance to the contribution between groups as to the within-group component, whereas the Theil, H-H and Bourguignon coefficients show that the inequalities are generated within the groups (98%). The Gini and the three particular cases of the entropy coefficient indicate that the men are two times more concerned with the disparities than the women are. Nevertheless, these two types of measures are too distant, and in order to motivate the choice of users of these measures we show, in a different way than Dagum (1998), that the Gini decomposition is a better index. Even if Theil, H-H and Bourguignon indexes check the (B) and (C) properties which are more important than the combination of (A), (B), (C), we incite to privilege the Gini decomposition in particular because it is built on a better between-group specification.

References

Bourguignon, F. (1979) “Decomposable Inequality Measures” *Econometrica* **47**, 901-920.

Dagum, C. (1980) “Inequality Measures between Income Distributions with Applications” *Econometrica* **48**, 1791-1803.

Dagum, C. (1997a) “A New Approach to the Decomposition of the Gini Income Inequality Ratio” *Empirical Economics* **22(4)**, 515-531.

Dagum, C. (1997b) “Decomposition and Interpretation of Gini and the Generalized Entropy Inequality Measures” *Proceedings of the American Statistical Association, Business and Economic Statistics Section, 157th Meeting,* 200-205.

Dagum, C. (1998) “Fondements de bien-être social et décomposition des mesures d’inégalité dans la répartition du revenu” *Economie Appliquée*, 151-202.

Hart, P. E. (1970) “Entropy and Other Measures of Concentration” *Journal of the Royal Statistical Society* **A134**, 73-85.

Mussard, S., Seyte, F., and M. Terraza (2002) *Program for Dagum’s Gini Decomposition*, http://www.lameta.univ-montp1.fr/online/gini.html

Theil, H. (1967) *Economics and Information Theory*, North-Holland Publishing Company, Amsterdam.

Table 1: Contributions of each element of the four decompositions to the overall inequality

Indexes	*% of the within-group component*	*% of the between-group component*	*% of transvariation*
G	50.9	13.5	35.6
T	97.8	2.2	NA*
I_1	98.2	1.8	NA*
B	98.2	1.8	NA*

*NA: Non available for this type of index

Table 2: Contributions of the men and the women to the global inequality

Indexes	*Wage inequalities within the men group (%)*	*Wage inequalities within the women group (%)*
G	34.3	16.6
T	62.7	35.1
I_1	71	27.2
B	56.8	41.4