Consortium of Household Panels for European Socio-Economic Research

CHER

Extension de la décomposition de l'indice de Gini de Dagum : la multi-décomposition non paramétrique

by Stéphane Mussard, Maria Noel Pi Alperin, Françoise Seyte, Michel Terraza

Document n° 17 2007

Corresponding author:

Stéphane Mussard
CEPS/INSTEAD, Luxembourg;
GREDI, Université de Sherbrooke;
GEREM, Université de Perpignan

E-mail: SMussard@adm.usherbrooke.ca

Comparative Research on Household Panel Studies

This series presents the results of research projects based on the analysis of one or more household panel studies from the CHER micro database. Papers will cover the wide range of substantive topics and investigations of the particular problems of comparative research.

The series contains in papers no. 1 – xx, among other papers, the results of all of the work being carried out as part of the CHER project, which was funded by the European Commission under within the program "Improving the Human Research potential and the Socio-Economic Knowledge Base. CHER aims to develop instruments for analyzing, programming and stimulating socio-economic policies, and for comparative research on policy issues such as labor force participation, income distribution, unpaid work, poverty, household composition change, and problems of the elderly.

Coordination of the CHER project is provided by CEPS/INSTEAD, Differdange, Luxembourg.

The main partners are:
Joachim R. Frick, Deutsches Institut für Wirtschaftsforschung , DIW Berlin, Germany
Kimberly Fisher, Institute for Social & Economic Research, ISER Essex University, Colchester, UK

The distinguished partners are:

Alain Degenne, Institut du Longitudinal CNRS-LASMAS, Caen, France
Franco Peracchi, Centre for International Studies on Economic Growth, CEIS, Tor Vergata University, Rome, Italy
Sergi Jimenez, Department of Economics, Universidad Carlos III, UIIIM, Madrid, Spain
Zoltan Fabian, Social Research Informatics Center , TARKI Budapest, Hungary
Brunon Gorecki, Dept. of Economics, University of Warsaw, UWARS, Poland
Ruud Muffels, WORC/TISSER, Katholieke Universiteit Brabant, Tilburg, Netherlands
Rudy Marynissen, University of Antwerp, UIA, Antwerp, Belgium
John Kallas, National Centre for Social Research, EKKE, Athens, Greece
Erwin Zimmermann, The Swiss Household Panel, SHP, Neuchâtel, Switzerland

For more information about this series, or to submit papers for inclusion, contact:

Günther Schmaus
CEPS/INSTEAD
Anc. Bât. administratif ARBED
Rue E. Mark, Boîte postale 48
L- 4501 Differdange
Tel: +352 58 58 55-509
Fax: +352 58 55 88

e-mail:gunther.schmaus@ci.rech.lu
url: http://www.ceps.lu/cher/accueil.cfm

 ISBN 978-2-87987-425-8

Internal CEPS Id 05.07.0361-E

Extension de la décomposition de l'indice de Gini de Dagum : la multi-décomposition non paramétrique[1]

Stéphane Mussard [1]
Maria Noel Pi Alperin [2]
Françoise Seyte [2]
Michel Terraza [2]

1 CEPS/INSTEAD, Luxembourg ; GREDI, Université de Sherbrooke ; GEREM, Université de Perpignan

2 LAMETA, Université de Montpellier I, Avenue de la mer, BP 9606, 34054, Montpellier Cedex 1 ,Tel : 33(0)4.67.15.84.20, Fax : 33(0)4.67.15.84.67, Email : f-seyte@lameta.univ-montp1.fr

Résumé

En 1997, Dagum propose une nouvelle décomposition de l'indicateur de Gini en trois indices : les inégalités à l'intérieur des groupes, les inégalités moyennes entre les groupes, et les inégalités de transvariation entre les groupes. Nous développons cette décomposition et l'appliquons aux revenus Luxembourgeois de 2001.

Mots-clés: Décomposition en sources, Décomposition en sous-groupes, Gini.

Classification JEL: D63, D31.

I. INTRODUCTION

Les méthodes de décomposition des mesures d'inégalité considèrent que l'étude des inégalités ne peut pas être réalisée en supposant la population réduite à un agent représentatif. Les techniques de décomposition permettent donc de comprendre que les populations sont constituées d'agents hétérogènes et que des inégalités entre ces agents peuvent être recensées. La décomposition de l'indicateur de Gini fut longtemps considérée comme un outil peu attractif. En le comparant aux indices dérivés de l'entropie, comme l'indicateur de THEIL (1967), on s'aperçoit que le coefficient de Gini possède une structure singulière. En effet, ce dernier repose sur une propriété de décomposition différente de celle de l'entropie ou des mesures d'inégalités utilitaristes telles que la classe des mesures de ATKINSON-KOLM-SEN. Néanmoins les économistes continuent à s'intéresser à la décomposition de la mesure de GINI. La première décomposition peut être attribuée à SOLTOW (1960). Ensuite, BHATTACHARYA et MAHALANOBIS (1967) ont introduit une décomposition de l'indice de GINI afin d'étudier les disparités régionales en Inde. En 1969, RAO est le premier à étudier à la fois la décomposition de la mesure de GINI en sous-groupes et en sources de revenu. Sa décomposition en sous-groupes va beaucoup plus loin que celle de BHATTACHARYA et MAHALANOBIS (1967). Il propose une décomposition en deux composantes unies par une forme quadratique. Il va ainsi montrer la différence fondamentale qui existe avec les indices de l'entropie. Selon lui, l'indice de GINI se décompose en la somme d'un indice intragroupe et d'un indice intergroupe particulier. L'indice intragroupe représente les inégalités entre les individus appartenant à un même groupe. Les inégalités intergroupes particulières sont les inégalités entre chaque paire de groupe, alors que la définition d'un indice intergroupe repose sur les différences moyennes entre les groupes d'une population.

[1] Les données utilisées dans cet article proviennent de la version publique de la base de données CHER « Consortium of Household Panels for European Socio-Economic Research », et sont utilisées avec la permission du consortium CHER (représenté par le CEPS/INSTEAD au Luxembourg). Les auteurs remercient le consortium CHER ainsi que le CEPS/INSTEAD.

Cette particularité va exclure l'indicateur de GINI des mesures dites « décomposables ». En effet, certains auteurs vont démontrer que l'indice de GINI possède une troisième composante, un terme d'interaction encore nommé résidu [voir MUKHERJEE et SHORROCKS (1982)]. Cette période est de surcroît marquée par la montée des indices dérivés de l'entropie et des généralisations faites sur les mesures décomposables [voir par exemple BOURGUIGNON (1979), SHORROCKS (1980), COWELL (1980a, 1980b)]. Ces approches prouvent que l'entropie possède de meilleures propriétés de décomposabilité comme la cohérence en sous-groupes. Lorsque l'inégalité totale augmente, les composantes intergroupes et intragroupes augmentent aussi. Cette monotonie en sous-groupes n'est pas respectée par l'indicateur de Gini.

Les décompositions qui suivent vont s'intéresser à la troisième composante. SILBER (1989) reprend l'approche matricielle de PYATT (1976) et propose une nouvelle décomposition en sous-groupes et en sources de revenu. DAGUM (1997a) prouve, au même titre que SILBER (1989), que la troisième composante provient du chevauchement entre les distributions. Il inclut la distance économique comme un élément permettant de dissocier les inégalités intergroupes en deux sous-indices. Il va ainsi montrer que le terme d'interaction n'est en fait que l'intensité de transvariation définie au début du 20ème siècle par GINI (1916). Cette intensité permet de mesurer l'influence des inégalités qui proviennent des souspopulations en moyenne les plus pauvres.

Cet article repose sur la décomposition de DAGUM (1997), DAGUM et *alii* (2002), Mussard et *alii* (2002). Cette décomposition, comme celles qui les précèdent, reposent sur une seule partition de la population en sous-groupes. SALAS (2002) introduit l'idée selon laquelle les sous-groupes peuvent eux-mêmes être déterminés par des groupes encore plus petits. Il propose ainsi une décomposition en multi-niveaux. En reprenant cette idée, nous démontrons que l'indicateur de GINI peut être décomposé en multi-niveaux. Ensuite, nous extrapolons les travaux de MUSSARD (2004, 2006) qui réconcilie la décomposition en sources de revenu avec la décomposition en sous-groupes. Nous montrons que ce modèle de multi-décomposition peut être unifié à la décomposition en multi-niveaux pour former un seul modèle général. Nous montrons ensuite que ce modèle général de décomposition peut être étudié en insérant des modèles paramétriques linéaires afin d'augmenter et de tester les déterminants des inégalités intragroupes et intergroupes.

II. LA DECOMPOSITION DE L'INDICATEUR DE GINI EN SOUS-GROUPES DE DAGUM (1997a, 1997b)

Soit une population mère P, où prévalent n unités de revenu : x_i (i=1,…, n). La population P est partitionnée en k sous-groupes (souspopulations) P_j (j=1,…, k) où le $j^{\text{ème}}$ groupe est de taille n_j, de fonction de répartition $F_j(x)$ et de moyenne μ_j. On note $F(x)$ et μ, respectivement, la fonction de répartition et la moyenne mesurées sur P. Le revenu de l'individu i appartenant $j^{\text{ème}}$ groupe est : x_{ij}. On peut donc définir le vecteur des revenus en faisant apparaître les k sous-groupes de P : (1)

$$\left(x_1, \ldots, x_i, \ldots, x_n\right) = \left(\left(x_{11}, \ldots, x_{1n_1}\right), \ldots, \left(x_{j1}, \ldots, x_{jn_j}\right), \ldots, \left(x_{k1}, \ldots, x_{kn_k}\right)\right).$$

L'indicateur de GINI, défini à partir de la différence moyenne de GINI, est donné par :

$$G = \frac{\sum_{i=1}^{n}\sum_{r=1}^{n}\left|x_i - x_r\right|}{2n^2\mu}. \tag{2}$$

Définition 1 : *Le coefficient de GINI mesuré sur les revenus de la soupopulation P_j est :*

$$G_{jj} = \frac{\sum_{i=1}^{n_j}\sum_{r=1}^{n_j}\left|x_i - x_r\right|}{2n_j^2\mu_j}. \tag{3}$$

Définition 2 : *Le coefficient de GINI intergroupe [DAGUM (1987)] permet de quantifier les inégalités de revenu qui prévalent entre deux souspopulations. Par définition, il mesure la différence de* *revenu espérée entre un individu tiré au hasard de la souspopulation P_j et un individu tiré au hasard de la souspopulation P_h.*[2] *Il s'exprime sous la forme :*

$$G_{jh} = \frac{\sum_{i=1}^{n_j}\sum_{r=1}^{n_h}\left|x_{ij} - x_{rh}\right|}{(\mu_j + \mu_h)\, n_j\, n_h}\,,\ \forall\, j, h = 1,\ldots, k\ . \tag{4}$$

On constate que l'indicateur de GINI intragroupe [G_{jj}] est construit en regroupant les différences binaires de revenu inhérentes au groupe P_j. L'indice de GINI intergroupe [G_{jh}] est construit en prenant l'ensemble des différences binaires de revenu qui existent entre les groupes P_j et P_h. Par conséquent, en regroupant les différences de revenu qui existent à l'intérieur des k sous-groupes et les différences de revenu qui existent entre chaque paire de sous-groupes, on peut établir une première décomposition :

$$G = \frac{\sum_{j=1}^{k}\left(\sum_{i=1}^{n_j}\sum_{r=1}^{n_j}\left|x_{ij} - x_{rj}\right|\right)}{2\mu n^2} + \frac{2\sum_{j=2}^{k}\sum_{h=1}^{j-1}\left(\sum_{i=1}^{n_j}\sum_{r=1}^{n_h}\left|x_{ij} - x_{rh}\right|\right)}{2\mu n^2}, \tag{5}$$

$$= G_w + G_{gb}\ . \tag{6}$$

Les différences de revenu à l'intérieur des groupes fournissent l'indicateur de Gini intragroupe [G_w]. Il donne la contribution des inégalités qui existent à l'intérieur des groupes au coefficient de GINI global. La contribution brute des inégalités intergroupes [G_{gb}] au coefficient global représente les inégalités de revenu entre chaque paire de sous-groupe. On constate que la contribution brute des inégalités intergroupes est différente des inégalités intergroupes qui apparaissent dans la littérature, qui sont, les inégalités moyennes entre les groupes. L'expression G_{gb} procure donc une meilleure information puisqu'elle concerne, comme le souligne RAO (1969), les différences binaires de revenu entre les paires de groupes.
La décomposition peut s'écrire comme une moyenne pondérée des indices intragroupes [G_{jj}] et des indicateurs intergroupes [G_{jh}]. Pour cela, on introduit les pondérations suivantes :

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

Le poids p_j est le pourcentage d'individus appartenant au groupe j et s_j le pourcentage de revenu détenu par la sous-population j.
La décomposition du coefficient de GINI se réécrit ainsi par :

$$\begin{aligned} G &= \sum_{j=1}^{k} G_{jj}\ p_j\, s_j + \sum_{j=2}^{k}\sum_{h=1}^{j-1} G_{jh}\left(p_j\, s_h + p_h\, s_j\right) \\ &= G_w + G_{gb}\,. \end{aligned} \tag{8}$$

On constate que deux formulations équivalentes [(5) et (8)] permettent de définir la décomposition de l'indicateur de GINI en deux éléments. Néanmoins, le coefficient de GINI offre trois composantes [voir par exemple MUKHERJEE et SHORROCKS (1982) et SILBER (1989) parmi d'autres]. Pour ce faire, la richesse économique brute et le moment d'ordre 1 de transvariation sont introduits.

Lemme 1 : *La distance directionnelle brute d_{jh} est une moyenne pondérée des différences de revenu $x_{ji}-x_{hr}$ pour chaque revenu x_{ji} d'un membre de P_j supérieur au revenu x_{hr} d'un membre de P_h, étant donné que le groupe P_j est en moyenne plus riche que le groupe P_h [$\mu_j > \mu_h$] :*

$$d_{jh} = \int_0^{\infty} dF_j(x) \int_0^{x} (x-y)\, dF_h(y)\ . \tag{9}$$

Lemme 2 : *Le moment d'ordre 1 de transvariation p_{jh} entre la $j^{ème}$ et la $h^{ème}$ souspopulation (tel que $\mu_j > \mu_h$) est la moyenne pondérée des différences de revenu $x_{hr}-x_{ji}$ pour chaque revenu x_{hr} d'un membre de P_h plus important que le revenu x_{ji} d'un membre de P_j:*

$$p_{jh} = \int_0^{\infty} dF_h(x) \int_0^{x} (x-y)\, dF_j(y)\ . \tag{10}$$

---

[2] On retrouve un des aspects essentiels des travaux de PYATT (1976) en rapport avec la théorie des jeux et le gain espéré que chaque individu peut gagner en espérant.

L'expression transvariation [GINI (1916), DAGUM (1959, 1960, 1961)] vient du fait que les différences de revenu considérées sont de signes opposés à la différence des moyennes de leur sous-groupe correspondant.

La distance directionnelle relative brute et le moment d'ordre 1 de transvariation permettent de définir *la distance économique relative et directionelle* entre les souspopulations P_j et P_h :

$$D_{jh} = \frac{d_{jh} - p_{jh}}{d_{jh} + p_{jh}} \ . \tag{11}$$

D_{jh} est inclus dans l'intervalle fermé [0,1], et est un nombre sans dimension car d_{jh}, p_{jh} ont la dimension de la richesse. La richesse économique relative sépare les inégalités intergroupes en deux composantes :

- la contribution nette à l'inégalité totale des inégalités de revenu entre les souspopulations P_j et P_h, obtenue par le produit $G_{jh} \times D_{jh}$;
- et la contribution des transvariations de revenu entre Pj et P_h, obtenue par $G_{jh} \times (1 - D_{jh})$.

L'addition de ces deux produits mesure de manière brute les inégalités de revenu entre deux souspopulations P_j et P_h.

Il est donc possible de séparer l'indicateur de Gini en trois éléments. Il suffit de décomposer la contribution brute des inégalités intergroupes [G_{gb}] en deux éléments par l'intermédiaire de la distance D_{jh} : (12)

$$\begin{aligned} G_{gb} &= \sum_{j=2}^{k} \sum_{h=1}^{j-1} G_{jh} \left(p_j s_h + p_h s_j\right) \\ &= \sum_{j=2}^{k} \sum_{h=1}^{j-1} G_{jh} D_{jh} \left(p_j s_h + p_h s_j\right) + \sum_{j=2}^{k} \sum_{h=1}^{j-1} G_{jh} \left(1 - D_{jh}\right) \left(p_j s_h + p_h s_j\right) \\ &= G_{nb} + G_t \, . \end{aligned}$$

En multipliant la contribution des inégalités intergroupes par D_{jh} puis par 1-D_{jh} (soit au total par 1), on distingue les inégalités intergroupes nettes [G_{nb}] des inégalités intergroupes de transvariation [G_t]. La contribution des inégalités de revenu inhérentes à l'intensité de la transvariation,

$$G_t = \sum_{j=2}^{k} \sum_{h=1}^{j-1} G_{jh} \left(1 - D_{jh}\right) \left(p_j s_h + p_h s_j\right) , \tag{13}$$

permet de mesurer le poids des inégalités intergroupes issues du chevauchement entre les distributions. L'intensité de la transvariation permet de savoir si les inégalités sont générées par les hauts revenus des souspopulations (en moyenne) les plus pauvres. A contrario, les disparités provenant des revenus élevés des souspopulations riches (en moyenne) sont données par la contribution nette des inégalités intergroupes à l'inégalité totale :

$$G_{nb} = \sum_{j=2}^{k} \sum_{h=1}^{j-1} G_{jh} D_{jh} \left(p_j s_h + p_h s_j\right) . \tag{14}$$

Théorème [DAGUM (1997a)] : *Le coefficient de GINI mesuré sur une population P, partagée en k sous-groupes, peut être décomposée comme suit :*

$$G = G_w + G_{nb} + G_t . \tag{15}$$

III. MULTI-DECOMPOSITIONS NON PARAMETRIQUES

III. 1. La multi-décomposition

MUSSARD (2004, 2006) introduit un lien unissant la décomposition de l'indicateur de GINI en sous-groupes et l'indicateur de GINI en sources de revenu. Il s'agit d'insérer une décomposition en sources de revenu dans la décomposition en sous-groupes. Le point de départ concerne la formule selon laquelle :

$$\left|x_i - x_r\right| = x_i + x_r - 2 \ min \left\{x_i \ , \ x_r\right\}. \tag{16}$$

L'indice de GINI se réécrit alors :

$$G = \frac{\sum_{i=1}^{n}\sum_{r=1}^{n}\left(x_i + x_r - 2\ min\{x_i\ ,\ x_r\}\right)}{2\mu n^2}. \qquad (17)$$

Les revenus sont constitués d'une somme de q facteurs (q sources de revenu). Par exemple le revenu de la personne i est divisée de la manière suivante : $x_i = \sum_{m=1}^{q} x_i^m$.

(18)

L'indice de GINI global se réécrit donc :

$$G = \frac{\sum_{i=1}^{n}\sum_{r=1}^{n}\left(\sum_{m=1}^{q} x_i^m + \sum_{m=1}^{q} x_r^m - 2\ min\{x_i\ ,\ x_r\}\right)}{2\mu n^2}. \qquad (19)$$

On sépare ensuite les termes $2\ min\{x_i\ ,\ x_r\}$ en sources de revenu :

$$2\ min\{x_i\ ,\ x_r\} = \sum_{m=1}^{q} 2x_{Q,ir}^{*m}. \qquad (20)$$

Proposition 1 : *Une décomposition de l'indicateur de GINI en sources de revenu peut être définie par :*

$$G = \sum_{m=1}^{q}\left(\frac{\sum_{i=1}^{n}\sum_{r=1}^{n}\left(x_i^m + x_r^m - 2x_{ir}^{*m}\right)}{2\mu n^2}\right) = \sum_{m=1}^{q}\left(S^m\right), \qquad (21)$$

où S^m est la contribution du facteur m à l'inégalité totale.

Preuve : ***cf.*** **MUSSARD (2004, 2006).** ■

Corollaire 1 : *Si les q sources de revenu sont q réplications (c'est-à-dire q variables i.i.d.) alors,*

$$S^1 = \ldots = S^m = \ldots = S^q \Rightarrow G = qS^m. \qquad (22)$$

Preuve : ***cf.*** **MUSSARD (2004, 2006).** ■

On constate que l'inégalité d'une source particulière est définie de manière proportionnelle par rapport à la différence moyenne de GINI. On peut ensuite démontrer que les indices de GINI mesurés sur les k groupes [G_{jj}] et entre les paires de groupe j et h [G_{jh}] sont décomposables en facteurs :

$$G_{jj} = \sum_{m=1}^{q}\left(\frac{\sum_{i=1}^{n_j}\sum_{r=1}^{n_j}\left(x_{ij}^m + x_{rj}^m - 2x_{irj}^{*m}\right)}{2\mu_j n_j^2}\right) = \sum_{m=1}^{q}\left(S_{jj}^m\right), \qquad (23)$$

$$G_{jh} = \sum_{m=1}^{q}\left(\frac{\sum_{i=1}^{n_j}\sum_{r=1}^{n_h}\left(x_{ij}^m + x_{rh}^m - 2x_{irjh}^{*m}\right)}{\left(\mu_j + \mu_h\right) n_j\, n_h}\right) = \sum_{m=1}^{q}\left(S_{jh}^m\right). \qquad (24)$$

On peut donc obtenir la décomposition de l'indicateur de GINI en sous-groupes mêlée avec la décomposition en facteurs.

Proposition 2 : *L'indicateur de GINI est multi-décomposable en deux éléments. La multi-décomposition est parfaite* :

$$G = \sum_{j=1}^{k}\sum_{m=1}^{q}\left(S_{jj}^m\right) p_j\, s_j \qquad (G_w) \qquad (25)$$

$$+ \sum_{j=2}^{k}\sum_{h=1}^{j-1}\sum_{m=1}^{q} S_{jh}^m\left(p_j\, s_h + p_h\, s_j\right). \qquad (G_{gb})$$

Preuve : ***cf.*** **MUSSARD (2004, 2006).** ■

Cette décomposition bidimensionnelle permet de décomposer les inégalités à la fois en sources de revenu et en sous-groupes. De plus, la multi-décomposition est parfaite dans le sens où tous les termes (hors mis le dénominateur) sont à la fois décomposés en sous-groupes et en sources sans termes redondants, contrairement aux indices issus de l'entropie généralisée. En développant l'équation (25), tous les termes du numérateur sont décomposés :

(26)

$$G=\sum_{m=1}^{q}\left(\frac{\sum_{j=1}^{k}\left(\sum_{i=1}^{nj}\sum_{r=1}^{nj}\left(x_{ij}^{m}+x_{rj}^{m}-2x_{irj}^{*m}\right)\right)}{2\mu n^{2}}\right)+\sum_{m=1}^{q}\left(\frac{2\sum_{j=2}^{k}\sum_{h=1}^{j-1}\left(\sum_{i=1}^{nj}\sum_{r=1}^{nh}\left(x_{ij}^{m}+x_{rh}^{m}-2x_{irjh}^{*m}\right)\right)}{2\mu n^{2}}\right)$$

On démontre ainsi que la multi-décomposition en deux indices est exacte. On peut ensuite construire une multi-décomposition en trois indicateurs. On utilise pour cela la décomposition en sous-groupes de DAGUM (1997) en trois éléments. MUSSARD (2006) démontre qu'il est nécessaire de décomposer la distance économique D_{jh} en sources de revenu :

$$D_{jh}^{m}=\sum_{m=1}^{q}\left(\frac{\left(\sum_{x_{j,i}>x_{h,r}}\sum_{i=1}^{nj}\sum_{r=1}^{nh}\left(x_{j,i}^{m}+x_{h,r}^{m}-2x_{jh,ir}^{*m}\right)\right)-\left(\sum_{x_{j,i}<x_{h,r}}\sum_{i=1}^{nj}\sum_{r=1}^{nh}\left(x_{j,i}^{m}+x_{h,r}^{m}-2x_{jh,ir}^{*m}\right)\right)}{\sum_{i=1}^{nj}\sum_{r=1}^{nh}\left|x_{j,i}-x_{h,r}\right|}\right)$$

$$=\sum_{m=1}^{q}D_{jh}^{m}\ ,\forall\ \mu_j>\mu_h\ .\qquad(27)$$

On obtient alors :

$$G=\sum_{j=1}^{k}\sum_{m=1}^{q}\left(S_{jj}^{m}\right)p_j\,s_j \qquad (G_w) \qquad (28)$$

$$+\sum_{j=2}^{k}\sum_{h=1}^{j-1}\sum_{m=1}^{q}D_{jh}^{m}\left(p_j\,s_h+p_h\,s_j\right)G_{jh} \qquad (G_{nb})$$

$$+\sum_{j=2}^{k}\sum_{h=1}^{j-1}\sum_{m=1}^{q}\left(1-D_{jh}^{m}\right)\left(p_j\,s_h+p_h\,s_j\right)G_{jh}\ . \qquad (G_t)$$

Cependant, pour la transvariation [G_t], qui est fonction de ($1-D_{jh}$), il est impossible de mesurer la contribution de chaque source de revenu à G_t. Il faut donc transformer le terme $1-D_{jh}$. On pose alors :

$$P_{jh}=1-D_{jh}=\sum_{m=1}^{q}\frac{2\left(\sum_{x_{j,i}<x_{h,r}}\sum_{i=1}^{nj}\sum_{r=1}^{nh}\left(x_{j,i}^{m}+x_{h,r}^{m}-2x_{jh,ir}^{*m}\right)\right)}{\sum_{i=1}^{nj}\sum_{r=1}^{nh}\left|x_{j,i}-x_{h,r}\right|}=\sum_{m=1}^{q}P_{jh}^{m}\quad,\ \forall\ \mu_j>\mu_h$$

(29)

où P_{jh}^{m} est la contribution de la source m au ratio de chevauchement P_{jh}. A titre d'information, P_{jh} est la composante symétrique de D_{jh}. Il tend vers 1 lorsque les distributions j et h se chevauchent et vers 0 quand les distributions ne se chevauchent pas. Ce ratio de chevauchement, P_{jh}, décomposable en facteurs, permet de mesurer la contribution d'une source particulière à l'intensité de transvariation G_t :

$$G_t=\sum_{j=2}^{k}\sum_{h=1}^{j-1}\sum_{m=1}^{q}P_{jh}^{m}\left(p_j\,s_h+p_h\,s_j\right)G_{jh}\ . \qquad (30)$$

Proposition 3 : *L'indicateur de GINI est multi-décomposable en trois éléments. La multi-décomposition est parfaite* :

$$G = \sum_{j=1}^{k}\sum_{m=1}^{q} \left(S_{jj}^{m}\right) p_j\, s_j \qquad (G_w) \qquad (31)$$

$$+ \sum_{j=2}^{k}\sum_{h=1}^{j-1}\sum_{m=1}^{q} D_{jh}^{m}\left(p_j\, s_h + p_h\, s_j\right) G_{jh} \qquad (G_{nb})$$

$$+ \sum_{j=2}^{k}\sum_{h=1}^{j-1}\sum_{m=1}^{q} P_{jh}^{m}\left(p_j\, s_h + p_h\, s_j\right) G_{jh}\ . \qquad (G_t)$$

Preuve : ***cf.*** **MUSSARD (2004, 2006).** ■

Le modèle permet de connaître l'influence de chaque source de revenu sur le montant de l'indicateur intragroupe [G_w], sur la contribution nette intergroupe [G_{nb}], et sur l'intensité de transvariation entre les groupes [G_t]. Plus précisément, il est possible d'évaluer la contribution d'une source m à l'inégalité générée par un groupe j ; d'évaluer la contribution de la source m à l'inégalité nette entre deux groupes j et h ; et enfin d'évaluer la contribution d'une source m à l'intensité de transvariation entre deux groupes j et h.

III.2. La décomposition en multi-niveaux

COWELL (1985), WODON (1999) et SALAS (2002) ont introduit la technique de décomposition en multi-niveaux. La population est divisée en une partition K, comprenant k groupes P_j (j =1,…, k). Soit S une seconde partition. : chaque groupe P_j est composé de q_j sous-groupes $P_{j,w}$ et $P_{j,z.}$ (w, z =1,…, q_j).

Le résultat de Salas est le suivant :

$$I = I^{b,K} + I^{b,SK} + I^{w,S}\ , \qquad (32)$$

où I est l'indice général d'entropie, $I^{b,K}$ est l'inégalité intergroupe (moyenne) pour les groupes de la première partition K; $I^{b,SK}$ est la mesure intergroupe (moyenne) pour les sous-groupes de la deuxième partition S; et $I^{w,S}$ l'inégalité au sein des sous-groupes.

Pour construire un résultat similaire avec la décomposition de la mesure de Gini, on définit tout d'abord n_z et μ_z (n_w, μ_w) comme le nombre d'individus et la moyenne de revenus du z-ième sous-groupe (w-ième) de la partition S. L'inégalité à l'intérieur du z-ième sous-groupe du j-ième groupe est :

$$G_{j,zz} = \frac{\sum_{i=1}^{n_z}\sum_{r=1}^{n_z}\left|x_{iz} - x_{rz}\right|}{2 n_z^2 \mu_z}\ . \qquad (33)$$

Les inégalités entre les sous-groupes z et w de P_j sont données par :

$$G_{j,wz} = \frac{\sum_{i=1}^{n_w}\sum_{r=1}^{n_z}\left|x_{iw} - x_{rz}\right|}{\left(\mu_z + \mu_w\right) n_z\, n_w}\ . \qquad (34)$$

Les proportions de population et de revenu du sous-groupe z (dans P_j) sont :

$$p'_z = \frac{n_z}{n_j}\ ;\ s'_z = \frac{n_z\ \mu_z}{n_j\ \mu_j}\ . \qquad (35)$$

Proposition 5 : *La multi-décomposition la décomposition en multi-niveaux peuvent être combinées en une "super" multi-décomposition :*

$$G = \sum_{j=1}^{k}\left(\sum_{z=1}^{q_j}\sum_{m=1}^{q} S_{j,zz}^{m} p'_z\, s'_z\right) p_j s_j \qquad (G^{w,S})\ (36)$$

$$+ \sum_{j=1}^{k}\left(\sum_{z=2}^{q_j}\sum_{w=1}^{z-1}\sum_{m=1}^{q} D_{j,wz}^{m}\ G_{j,wz}\left(p'_w\, s'_z + p'_z\, s'_w\right)\right) p_j s_j \qquad (G^{nb,SK})$$

$$+ \sum_{j=1}^{k}\left(\sum_{z=2}^{q_j}\sum_{w=1}^{z-1}\sum_{m=1}^{q} P_{j,wz}^{m}\ G_{j,wz}\left(p'_w\, s'_z + p'_z\, s'_w\right)\right) p_j s_j \qquad (G^{t,SK})$$

$$+\sum_{j=2}^{k}\sum_{h=1}^{j-1}\sum_{m=1}^{q} D_{jh}^{m} G_{jh}\left(p_j s_h + p_h s_j\right) \qquad (G^{nb,K})$$

$$+\sum_{j=2}^{k}\sum_{h=1}^{j-1}\sum_{m=1}^{q} P_{jh}^{m} G_{jh}\left(p_j s_h + p_h s_j\right). \qquad (G^{t,K})$$

Preuve : ***cf.*** **MUSSARD et *alii* (2005).** ■

On obtient :

$G^{w,S}$ la contribution des inégalités intragroupes pour les sous-groupes de S appartenant à K ;

$G^{nb,SK}$ la contribution nette des inégalités entre les sous-populations de S appartenant à K (les inégalités nettes intergroupes de deuxième ordre) ;

$G^{tb,SK}$ la contribution des inégalités de transvariation entre les sous-populations de S appartenant à K (les inégalités intergroupes de transvariation de deuxième ordre) ;

$G^{nb,K}$ la contribution nette des inégalités entre les groupes de la première partition (les inégalités nettes intergroupes de premier ordre) ;

$G^{tb,K}$ la contribution des inégalités de transvariation des groupes de la partition K (les inégalités de transvariation intergroupes de premier ordre)

IV. UNE APPLICATION AUX MENAGES LUXEMBOURGEOIS

Nous utilisons les données de l'enquête Panel Socio-Economique *Liewen zu Lëtzebuerg* (PSELL-2) sur les conditions de vie et de travail au Luxembourg obtenues *via* le « Consortium of Household Panels for European Socio-Economic Research » (CHER) . L'enquête PSELL-2 est menée chaque année depuis 1994 auprès d'un échantillon représentatif des ménages et des individus qui les composent, résidant sur le territoire luxembourgeois. Nous utilisons ici les données de la vague 2001. Les ménages sont regroupés selon une première partition : les personnes vivant en ville (Urban) représentant près de 68% de la population et possédant 66% du revenu global, et celles vivant à la campagne (Rural) près de 32% de la population et possédant uniquement 34% du revenu global (*cf.* Tableau 1). La population est décomposée en une seconde partition : les ménages de nationalité Luxembourgeoise, et les ménages de nationalité étrangère. Ainsi, nous pouvons connaître les caractéristiques des sous-groupes, autrement dit, les caractéristiques des ménages vivant en ville de nationalité Luxembourgeoise (Nat Urban), celles des ménages vivant en ville de nationalité étrangère, celles des ménages vivant à la campagne de nationalité Luxembourgeoise (Nat Rural), et celles des ménages vivant à la campagne de nationalité étrangère (Not Nat Rural).

Tableau 1.

	Urban **(1)**	Rural **(2)**	Nat Urban **(3)**	Not Nat Urban **(4)**	Nat Rural **(5)**	Not Nat Rural **(6)**
p_i	0,6776	0,3224	0,6502	0,3498	0,7257	0,2743
s_i	0,6590	0,3410	0,6590	0,3410	0,7369	0,2631

Les indices de GINI pour les quatre sous-groupes sont les suivants. Il existe de fortes inégalités de revenu entre les ménages vivant à la campagne de nationalité Luxembourgeoise (G = 0,312). Ensuite, les inégalités concernent les ménages de nationalité Luxembourgeoise vivant en ville (G = 0,307). Puis, nous trouvons les ménages de nationalité étrangère vivant en ville (G = 0,304) et ceux qui vivent à la campagne (G = 0,289).[3] Ces indicateurs donnent des informations intéressantes mais ne permettent pas de comprendre la contribution de chaque sous-groupe au montant de l'inégalité globale, ni de leurs interactions. L'inégalité globale correspond à un indice de GINI, G = 0,307. C'est cet indice que nous proposons de

[3] Les indices de Gini sont légèrement surestimés. Ceci provient de la non prise en compte des pondérations associées à chaque ménage.

décomposer selon la multi-décomposition (36). Pour cela, nous disposons de 10 sources de revenu : les salaires, les revenus des artisans et commerçants, les revenus immobiliers, les transferts privés, l'assurance chômage, transferts dus aux handicaps et à la santé, transferts liés à la famille, autres transferts, pensions, revenus d'autres sources.
Les résultats de la décomposition de l'indice de GINI total sont présentés dans les tableaux 2 et 3. On constate que le sous-groupe 3 (Nat Urban) est celui où les inégalités sont les plus importantes, que ce soit au niveau des salaires, des revenus liés à sa propre profession ou bien des revenus immobiliers (*cf.* $G^{w3,S}$ Tableau 2). C'est ce même sous-groupe qui permet de diminuer les inégalités globales grâce aux transferts de revenu privés et à l'assurance chômage.
Par ailleurs, ce sont les inégalités intergroupes, entre les ménages vivant en ville, qui sont les plus importantes. En particulier, ce sont les inégalités entre les ménages de nationalité Luxembourgeoise et étrangères qui sont les plus élevées en matière de salaire et des revenus ($G^{nb34,SK}$: revenus des artisans et commerçants et revenus immobiliers). C'est aussi entre les ménages de ces mêmes sous-groupes que les transferts et les pensions permettent de réduire fortement les inégalités.
Concernant les inégalités intergroupes de transvariation, celles qui permettent d'expliquer que certains ménages issus de groupes en moyenne moins aisés (Not Nat Rural) parviennent à créer des écart de revenu avec les ménages appartenant à des groupes aisés (Nat Rural), celles-ci sont parfois négatives comme pour les salaires (*cf.* $G^{t56,SK}$). Ceci signifie que les sous-groupes les moins riches n'arrivent pas à créer des écarts de revenus avec les plus riches, diminuant ainsi l'inégalité globale. Pour d'autres sources de revenu comme les revenus imobiliers, la transvariation est positive, signifiant que les ménages des sous-groupes les plus pauvres commencent à rejoindre certains niveaux de revenus des sous-groupes les plus riches.

Tableau 2

	Salaires	**Revenus des artisans et commercants**	**Revenus immobiliers**	**Transferts privés**	**Assurance chômage**
$G^{w3,S}$	**0,0420**	0,0041	0,0080	-0,0001	-0,0002
$G^{w4,S}$	0,0142	0,0008	0,0006	0,0000	0,0000
$G^{w5,S}$	0,0118	0,0016	0,0011	0,0000	-0,0002
$G^{w6,S}$	0,0019	0,0002	0,0001	0,0000	0,0000
$G^{nb12,K}$	**0,0836**	0,0072	0,0126	-0,0003	-0,0016
$G^{t12,K}$	0,0165	0,0034	-0,0020	0,0000	0,0008
$G^{nb34,SK}$	**0,0686**	0,0035	0,0009	-0,0003	-0,0002
$G^{t34,SK}$	-0,0202	0,0003	0,0046	0,0002	0,0001
$G^{nb56,SK}$	0,0126	0,0009	0,0001	-0,0001	-0,0001
$G^{t56,SK}$	-0,0031	0,0002	0,0005	0,0000	0,0000

Le tableau 3 expose le détail des différents types de transfert et leus rôles dans l'inégalité globale. Les contributions peuvent être négatives. Par exemple, les trasnferts pour les personnes malades et handicapées permettent de diminuer les inégalités entre les résidents et les non résidents, que ce soit en zone urbaine ou rurale.

Tableau 3

	Trasnferts dus aux handicaps et à la santé	**Transferts liés à la famille**	**Autres transfers**	**Pensions**	**Revenus d'autres sources**

$G^{w3,S}$	0,0002	0,0061	0,0003	-0,0017	0,0000
$G^{w4,S}$	0,0001	0,0011	0,0001	-0,0007	0,0000
$G^{w5,S}$	0,0001	0,0036	0,0003	0,0000	0,0000
$G^{w6,S}$	0,0000	0,0001	0,0000	0,0000	0,0000
$G^{nb12,K}$	0,0008	0,0086	0,0004	0,0068	0,0000
$G^{t12,K}$	0,0000	0,0091	0,0010	-0,0102	0,0000
$G^{nb34,SK}$	-0,0007	0,0069	0,0003	-0,0210	0,0000
$G^{t34,SK}$	0,0011	-0,0015	0,0000	0,0193	0,0000
$G^{nb56,SK}$	-0,0001	0,0014	0,0001	-0,0032	0,0000
$G^{t56,SK}$	0,0001	0,0003	0,0001	0,0031	0,0000

De la même manière, les pensions réduisent les inégalités entre les sous-groupes. Au contrarire, les transferts liés à la famille ne permettent pas de réduire les inégalités intergroupes.

IV CONCLUSION

L'indice de GINI appartient à la famille des indices décomposables basés sur les comparaisons par paire. Ce principe est fondamental. Il n'est pas systématiquement vérifié par les mesures d'inégalité de l'entropie généralisée. Même si l'entropie généralisée est décomposable en sous-groupes comme en multi niveaux, elle ne satisfait pas la propriété de multi décomposition en sous-groupe et source de revenu. La décomposition de l'indice de GINI met en évidence un troisième terme qui correspond non pas à un terme résiduel mais à l'intensité de transvariation. La multi décomposition de GINI permet d'apprécier l'intensité de transvariation entre les groupes de chaque partition, la contribution nette des inégalités entre les sous-groupes de chaque partition et la contribution des inégalités intragroupes pour tous les sous-groupes. La multi décomposition non paramétrique met l'accent sur les composantes de l'inégalité et constitue donc une aide à la prise décision de politiques socio-économiques de redistribution.

BIBLIOGRAPHIE

BOURGUIGNON F. (1979), « Decomposable Inequality Measures »,*Econometrica*, Vol. 47 : p. 901-920.

COWELL F. A. (1980a), « Generalized entropy and the Measurement of Distributional Change »,*European Economics Review*, Vol. 13 : p. 147-159.

COWELL F. A. (1980b), « On the Structure of Additive Inequality Measures »,*Review of Economics Studies*, Vol. 47 : p. 521-531.

DAGUM C. (1959), *Transvariazione fra più di due distribuzioni*, *In : Gini, C.(ed.) Memorie di metodologia statistica*, Vol II, Libreria Goliardica, ROMA.

DAGUM C. (1960), « Teoria de la transvariacion, sus aplicaciones a la economia »,*Metron*, Vol. XX : p. 1-206.

DAGUM C. (1961), « Transvariacion en la hipotesis de varaibles aleatorias normales multidimensionales »,*Proceedings of the International Statistical Institute*, Vol. 38, Book 4, Tokyo, p. 473-486.

DAGUM C. (1980), « Inequality Measures Between Income Distributions with Applications », *Econometrica*, Vol. 48 (7) : p. 1791-1803.

DAGUM C. (1987), « Measuring the Economic Affluence Between Populations of Income Receivers »,*Journal of Business and Economic Statistics*, Vol. 5(1) : p. 5-12.

DAGUM C. (1997a), « A New Approach to the Decomposition of the Gini Income Inequality Ratio », *Empirical Economics*, Vol. 22(4) : p.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,* p. 200-205.

DAGUM C., MUSSARD S., SEYTE F. et TERRAZA M. (2002), *Programme pour la décomposition de l'indice de Gini de C. Dagum*, disponible sur <http://www.lameta.univ-montp1.fr/online/gini.html>.

DEUTSCH J. et SILBER J. (2002), « Earnings Functions and the Measurement of the Determinants of Wage Dispersion : Extending Oaxaca's Approach », *Séminaire général du LAMETA*.

GINI C. (1916), « Il concetto di transvariazione e le sue prime applicazioni »,*Giornale degli Economisti e Rivista di Statistica*, In : Gini C (ed.) (1959), p.21-44.

MORDUCH J. et SICULAR T. (2002), « Rethinking Inequality Decomposition, with Evidence from Rural China », *Economic Journal*, Vol. 112(476) : p. 93-106.

MUKHERJEE D. et A. SHORROCKS (1982), "A Decomposition Analysis of the Trend in UK Income Inequality",*The Economic Journal*, Vol. 92: p. 886-902.

MUSSARD S. (2004), *Décompositions multidimensionnelles du rapport moyen de Gini. Applications aux revenus italiens de 1989 et 2000*, Thèse, University of Montpellier I, France.

Mussard S. (2006), « Une réconciliation entre la décomposition en sous-groupes et la décomposition en sources de revenu de l'indice de Gini - La multi-décomposition de l'indicateur de Gini »,*Annales d'Economie et de Statistiques*, Vol. 81 : p. 1-25.

MUSSARD S., PI-ALPERIN M.N., SEYTE F. et TERRAZA M. (2005), « Extension of Dagum's Gini Decomposition »,*Colloque international en mémoire de Gini et Lorenz*, Sienne, Italie.

MUSSARD S., SEYTE F. et TERRAZA M. (2003), « Decomposition of Gini and the Generalized Entropy Inequality Measures »,*Economics Bulletin*, Vol.4 n°7 : p. 1-6.

PYATT G. (1976), « On the Interpretation and Disaggregation of Gini Coefficients », *Economic Journal*, Vol. 86 : p. 243-25.

RAO V.M. (1969), « Two Decompositions of Concentration Ratio »,*Journal of the Royal Statistical Society*, Séries A 132 : p. 418-425.

SALAS R. (2002), « Multilevel Interterritorial Convergence and Additive Multidimensional Inequality Decompostion »,*Social Choice and Welfare*, Vol. 19 : p. 207-218.

SHORROCKS A. F. (1980), « The Class of Additively Decomposable Inequality Measures », *Econometrica*, Vol. 48 : p. 613-625.

SILBER J. (1989), « Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality »,*Review of Economics and Statistics*, Vol. 71 : p. 107-115.

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

YITZHAKI S. (1979), « Relative Deprivation and the Gini Coefficient »*Quarterly Journal of Economics*, Vol. 93 : p. 321-324.