

# Décompositions des mesures d'inégalité : le cas des coefficients de Gini et d'entropie \*

Stéphane Mussard \*\*

*LAMETA, Université Montpellier I*

Michel Terraza \*\*\*

*LAMETA, Université Montpellier I*

## 1 Introduction

La mesure des inégalités du revenu est un domaine important de l'économie publique. L'intérêt pour ce type de mesure provient de la théorie de la répartition du revenu qui distingue la répartition fonctionnelle du revenu et la répartition personnelle du revenu.

La répartition fonctionnelle du revenu étudie la formation des prix, le rôle des facteurs de production dans la production totale et l'allocation du revenu aux propriétaires de chaque catégorie de facteurs. Elle permet donc de déterminer la part du revenu national imputée à chaque catégorie de facteurs de production.

La répartition personnelle du revenu analyse la répartition du revenu total entre les unités économiques (famille, ménage, individus, etc.). Elle met

---

\* Cette recherche est issu d'un projet dont les premiers résultats furent présentés à l'institut mondial pour la recherche en économie du développement de l'Université des Nations Unies (WIDER 2003). Les auteurs remercient les Professeurs Camilo Dagum, Jacques Silber et Henri Caussinus pour leurs commentaires sur une version antérieure à ce papier. La version finale de ce papier s'acheva lorsque Stéphane Mussard finissait ses recherches de post-doctorat au CEPS/INSTEAD Luxembourg. L'auteur remercie le CEPS/INSTEAD, Philippe Van Kerm, et le laboratoire GREDI de l'Université de Sherbrooke dans lequel il est chercheur associé.

\*\* CEPS/INSTEAD Luxembourg, GREDI Université de Sherbrooke, LAMETA Université de Montpellier I. Adresse : LAMETA, UFR Sciences économiques, Université de Montpellier I, Avenue de la Mer, CS 79806, 34960 Montpellier cedex 2, France, E-mail : smussard@adm.usherbrooke.ca.

\*\*\* LAMETA, Université Montpellier I, E-mail : mterraza@lameta.univ-montp1.fr

aussi l'accent sur les sources de revenu de chaque entité telles que le salaire, les primes, le profit, les taxes, etc., et repose sur la propriété privée des ressources initiales.

Les mesures d'inégalité sélectionnées dépendent des axiomes qui sont retenus pour les définir. En 1967, Theil s'interroge sur la notion de décomposabilité en sous-populations (sous-groupes) des mesures d'inégalité (par exemple hommes et femmes, catégories socioprofessionnelles, régions, groupes d'âge, etc.). La méthode permet de considérer que l'étude des inégalités ne peut pas être réalisée en supposant que la population est réduite à un agent représentatif. Les techniques de décomposition en sous-populations permettent donc de comprendre que les populations sont constituées d'agents hétérogènes et que des inégalités entre ces agents peuvent exister. Les mesures issues de l'entropie, tout comme celles appartenant à la classe des indices de Atkinson-Kolm-Sen (cf. Blackorby, Bossert et Donaldson (1999)), reposent sur un axiome de décomposition additive semblable à l'axiome d'indépendance qui caractérise l'espérance d'utilité de von Neumann et Morgenstern. Rappelons que cet axiome (repris dans l'axiomatique de Savage) aboutit aux paradoxes d'Allais et d'Ellsberg, lorsqu'on s'intéresse à la théorie de la décision en univers incertain. Une propriété moins exigeante que celle de décomposition additive serait donc bienvenue pour caractériser les indices d'inégalité.

De manière générale, la finalité de la décomposition en sous-groupes est d'expliquer les inégalités de revenu par le degré d'implication des différents groupes composant la population mère. Est-ce qu'un groupe participe plus qu'un autre à l'explication des inégalités ? Est-ce que les inégalités proviennent des différences de revenu à l'intérieur des groupes ou des différences de revenu entre les groupes ?

La deuxième technique de décomposition est celle de la décomposition en sources de revenu. Elle permet d'atteindre l'objectif suivant. Elle identifie les sources de rémunération qui expliquent la totalité des inégalités de revenu.

Pour mieux comprendre l'objectif de ces deux méthodes, considérons le tableau 1.

Sources → Inégalités ↓	Salaires	Primes	Taxes	Total
Inégalités entre hommes	×	...	×	30%
Inégalités entre femmes	×	...	×	20%
Inégalités entre hommes et femmes	×	...	×	50%
Total	80%	10%	10%	100%

**Tableau 1 :** Les possibilités de décomposition

Lorsque la population mère est partitionnée en deux groupes, hommes et femmes, la décomposition en sous-groupes met en évidence la participation des hommes (30%), des femmes (20%) à l'inégalité totale, et la contribution des inégalités entre les hommes et les femmes (50%) à l'inégalité totale. De manière analogue, la décomposition en sources de revenu indique la contribution des inégalités salariales (80%), des primes (10%), et des taxes (10%) à l'inégalité totale.

Lorsqu'un revenu s'observe sur une population dont les unités sont regroupées selon des critères naturels tels que le sexe, l'espace géographique ou la catégorie socioprofessionnelle et que par ailleurs ce revenu provient de l'agrégation d'éléments de diverses sources, il est intéressant de pouvoir mesurer l'inégalité totale ( $I$ ), celle des groupes ( $I_j$ ), ou celle des sources ( $I^h$ ). A l'heure actuelle, aucune mesure d'inégalité n'a pu être réécrite comme une moyenne pondérée des inégalités pour chaque groupe et chaque source ( $I_j^h$ ) et des inégalités entre chaque paire de groupe pour chaque source ( $I_{jh}^h$ ).

L'intérêt de ce type de décomposition de la structure des inégalités est d'établir un lien entre l'analyse statistique des inégalités et le rôle des politiques économiques de redistribution. L'utilisation de ces décompositions permettrait, par exemple, de tester l'impact du changement des politiques fiscales sur les inégalités de revenus inhérentes aux groupes et aux sources de revenu en s'intéressant aux contributions du tableau 1. Elle permettrait de répondre à la question suivante : est-ce que les taxes engendrées par une nouvelle politique fiscale vont entraîner de moindres inégalités intergroupes (entre les hommes et les femmes) ou au contraire vont-elles augmenter les inégalités intragroupes (au sein du groupe masculin ou du groupe féminin) ?

Dans le cadre de leur réflexion, les économistes ont étudié deux types de mesures d'inégalité dites régulières pouvant être classées dans deux ensembles non disjoints : les mesures construites en considérant le quotient d'une mesure de dispersion par une mesure de tendance centrale, comme la mesure de dispersion de Gini divisée par la moyenne et celles qui sont construites en considérant une divergence entre deux fonctions de répartition, comme l'entropie généralisée de Shorrocks (1980). L'analyse de ces deux types d'indicateurs suggère qu'une mesure  $I$  peut être décomposée en indicateurs secondaires  $C_j, C^h, C_j^h, C_{jh}^h$ , autrement dit en contributions permettant de calculer la participation à l'inégalité totale de chaque groupe, de chaque source et de chaque croisement de population et de source.

L'objet de cet article est d'exposer les techniques de décomposition en sous-groupes et en sources de revenu existantes avant d'étudier les méthodes synthétiques qui combinent les deux décompositions. Par exemple, Mussard (2004a, 2004b, 2006) montre, pour le cas particulier de la mesure de Gini, que ces mesures fournissent une explication aux inégalités qui s'exercent à l'intérieur des groupes et entre les groupes par les changements de la structure du revenu. La littérature concernant les mesures d'inégalités décomposables a

donc conduit aux solutions soulevées par le problème du no-bridge, c'est-à-dire par l'absence de lien entre les deux techniques de décomposition. Elle a en effet suggéré la construction et la généralisation de techniques synthétiques de type multi-décomposition, qui permettent d'espérer que le lien unissant les politiques économiques et les inégalités soit davantage précisé.

En retenant tout d'abord une approche descriptive, nous mettons en évidence les enjeux sous-jacents aux problèmes de la mesure des inégalités (Section 2). Nous orientons ensuite la recherche vers l'examen de la propriété de décomposition additive en sous-groupes permettant de concevoir une mesure d'entropie généralisée et les éléments de réflexion critiques pouvant légitimer la mesure de Gini et le recours à une autre conception de décomposition en sous-groupes (Section 3). Nous abordons ensuite (Section 4) trois méthodes de décomposition en sources de revenu dont l'une basée sur la valeur de Shapley permet de généraliser la mesure de la contribution de chaque source de revenu au montant de l'inégalité totale. Nous étudions (Section 5) l'adoption d'un nouveau type de décomposition, combinant la décomposition en sous-groupes et la décomposition en sources de revenu, qui découle naturellement des enseignements des deux sections précédentes. Cette technique est illustrée sur les revenus italiens de 2000 (Section 6).

## 2 La construction des indices d'inégalité

Les travaux de Dalton (1920), Kolm (1976a, 1976b), Shorrocks (1980) et Chakravarty (1999)<sup>1</sup> ont conduit les chercheurs à admettre que les mesures d'inégalité doivent respecter quatre propriétés essentielles. Soit  $I$  un indice d'inégalité mesuré sur une population de taille  $n$ , d'unités statistiques  $1, \dots, n$  et défini par  $I : \mathbb{R}_+^n \rightarrow \mathbb{R}_+, n \in \mathbb{N}^*$ .

**Axiome 2.1. Le principe de transfert de Pigou-Dalton (PD).** Soit une distribution de revenu  $x'$  (transposée de  $x$ ) obtenue à partir d'une distribution  $y'$  (transposée de  $y$ ) où l'on effectue un transfert de revenu progressif<sup>2</sup>, autrement dit d'une personne possédant un certain niveau de revenu vers une personne dont le niveau de revenu est plus faible :

$$I(x) < I(y) \tag{PD}$$

**Axiome 2.2 Le principe de population de Dalton (PP).** Pour une distribution  $x'$  concaténée<sup>3</sup>  $k$  fois, la mesure d'inégalité reste inchangée :

<sup>1</sup> Les travaux axiomatiques de Chakravarty sont antérieurs à 1999, mais son article de 1999 (cf. Handbook of Income Inequality Measurement) est une revue plus complète de la littérature.

<sup>2</sup>  $x$  est obtenu à partir de  $y$  par un seul transfert progressif s'il existe  $i$  et  $r$  tels que  $x_i < y_i \leq y_r < x_r$ ,  $y_i - x_i = x_r - y_r$ ,  $y_s = x_s$ , pour tout  $s \neq i, r$  impliquant  $I(x) < I(y)$ .

<sup>3</sup> Soit un vecteur  $x' = \{x_1, x_2, \dots, x_n\}$ . La concaténation d'ordre  $k$  consiste à répliquer  $k$  fois le vecteur  $x'$  :

$$x'^{(k)} = \underbrace{\{x_1, \dots, x_1\}}_{k \text{ fois}}, \dots, \underbrace{\{x_n, \dots, x_n\}}_{k \text{ fois}}$$

$$I(x^{(k)}) = I(x), k \geq 2. \quad (\text{PP})$$

Ce critère permet de comparer des indices mesurés sur des distributions de revenu dont les tailles sont différentes.

**Axiome 2.3 L'anonymat ou la symétrie (SM).** Elle désigne la propriété d'un indice  $I(\cdot)$  qui est une fonction symétrique de ses arguments. Par exemple, si  $x'$  est obtenue à partir de  $y'$  par une permutation quelconque des revenus, alors :

$$I(x) = I(y). \quad (\text{SM})$$

**Axiome 2.4 La normalisation (NM).** Il s'agit de la condition selon laquelle les indices d'inégalité prennent la valeur 0 pour des distributions égalitaires (autrement dit lorsque les individus possèdent tous le même revenu) :

$$I(k1_n) = 0. \quad (\text{NM})$$

où  $1_n$  est le vecteur de dimension  $n$  dont les éléments sont tous égaux à la valeur 1.

Les axiomes PD, PP, SM et NM définissent l'ensemble des indices réguliers.

**Axiome 5. L'invariance relative (IR).** La propriété d'invariance relative est l'homogénéité de degré zéro d'une fonction :

$$I(\lambda x) = I(x), \lambda \in \mathbb{R}_+^*. \quad (\text{IR})$$

Cette propriété permet de s'affranchir des unités de natures différentes<sup>4</sup>.

Le socle axiomatique constitué des cinq propriétés précédentes a permis aux chercheurs de trouver de manière analytique certaines classes de mesures d'inégalité comme celle de Atkinson-Kolm-Sen ou bien celle de Kolm-Pollak (cf. Blackorby, Bossert et Donaldson (1999)). Par la suite, d'autres auteurs vont s'intéresser aux notions de décomposition en sous-groupes et à la manière dont il est possible d'intégrer ces propriétés aux indices de mesure d'inégalité.

### 3 La décomposition en sous-groupes

La réflexion concernant la décomposition des mesures d'inégalité en sous-groupes, abordée par Theil (1967), fut axiomatisée et développée par Shorrocks (1980, 1984, 1988). Nous analysons dans un premier temps les résultats

<sup>4</sup> Il existe un débat au sein de la communauté des économistes concernant les axiomes d'invariance. Certains privilégient IR alors que d'autres utilisent IA : l'invariance absolue, autrement dit l'homogénéité de degré 1 (cf. l'indice de Kolm (1976)). Il en est ainsi lorsqu'on recourt aux fonctions distance (Hölder, distance transportée, etc.) pour évaluer les inégalités entre deux distributions de revenu. Certains auteurs, comme Bossert et Pfingsten (1990) utilisent des mesures d'inégalité avec un concept d'invariance intermédiaire afin que IR et IA deviennent des cas particuliers. Par conséquent, l'emploi de l'axiome IR est discutable. Les axiomes PD, PP, SM, et NM sont unanimement reconnus, ils définissent les mesures régulières.

auxquels Shorrocks est parvenu avant de proposer dans un deuxième temps quelques éléments de réflexion critiques basés sur la décomposition de la mesure de Gini.

### 3.1 L'entropie et la décomposition additive

Shorrocks (1980) reprend l'idée de Theil (1967) selon laquelle il est possible de recourir à la seconde loi de l'entropie, qui mesure le désordre d'un système thermodynamique, afin de mesurer les inégalités : plus l'entropie est importante, plus les inégalités de revenus le sont. Cette loi lui permet de s'intéresser à la notion de décomposabilité en sous-groupes. Shorrocks (1980, 1984, 1988) définit en effet dans un premier temps le concept de décomposabilité additive puis celui de décomposition cohérente en sous-populations.

Soit une population mère  $P$  de taille  $n$  dont la moyenne des revenus est notée  $m$ .  $P$  est divisée en  $k$  groupes  $P_j$  de moyenne  $m_j$  ( $j = 1, \dots, k$ ). Soit  $f_j$  la fréquence des individus appartenant au groupe  $P_j$  et  $s_j$  la part de revenu de  $P_j$  dans la masse totale des revenus :

$$f_j = \frac{n_j}{n}, s_j = \frac{n_j m_j}{n m}. \tag{1}$$

**Définition 3.1 (Shorrocks (1980))** *La décomposabilité additive (DA). L'inégalité totale est la somme de deux composantes, une inégalité intragroupe  $I_w$  et une inégalité intergroupe  $I_b$  :*

$$I = I_w + I_b. \tag{DA}$$

$$\text{où } I_w = \sum_{j=1}^k \omega_j(m_j, n_j) I_j(\tilde{x}_j)$$

$$\text{et } I_b = I(\underbrace{m_1, \dots, m_1}_{n_1 \text{ fois}} \dots \underbrace{m_k, \dots, m_k}_{n_k \text{ fois}}).$$

$I_j(\tilde{x}_j)$  représente l'indice d'inégalité mesuré sur la distribution de revenu  $\tilde{x}_j$  du groupe  $P_j$ ,  $\omega_j$  le poids associé à l'indicateur  $I_j$ ,  $I_w$  la contribution des inégalités intragroupes à l'inégalité totale et  $I_b$  la contribution des inégalités intergroupes à l'inégalité totale. L'inégalité intragroupe  $I_w$  est une moyenne pondérée des inégalités à l'intérieur des groupes. L'inégalité intergroupe  $I_b$  représente les inégalités entre les revenus moyens  $m_j$  des  $k$  sous-populations. Elle est construite en supposant que chaque individu possède le revenu moyen de son groupe correspondant.

En 1979, Bourguignon avait défini la classe des mesures agrégatives pour laquelle la quantité  $\sum_{j=1}^k \omega_j(m_j, n_j) I_j(\tilde{x}_j)$  n'est pas systématiquement égale à  $I_w$ . Les mesures additivement décomposables constituent donc un cas particulier des mesures agrégatives. Il avait ensuite démontré, en se

basant sur un principe de décomposition particulier, celui pour lequel  $\omega_j(m_p, n_j) = \frac{n_j m_j}{\sum_j n_j m_j}$ , que la différentiabilité, la symétrie et l'homogénéité de degré zéro permettaient de déterminer une mesure unique, celle de Theil (1967, cf. équation (4) *infra*), qui respecte par ailleurs le principe de transfert de Pigou-Dalton.

Sans imposer une structure particulière aux poids  $\omega_j(m_p, n_j)$ , et contrairement à Bourguignon (1979), Shorrocks (1980) s'appuie sur un principe de décomposabilité additive plus large (DA) afin de démontrer que le coefficient de Theil n'est pas l'unique mesure associée à un principe de décomposition.

**Théorème 3.2 (Shorrocks (1980))** *L'indice d'inégalité I (possédant des dérivées secondes continues) satisfait les axiomes IR, SM, NM et DA si et seulement si (ssi) I est un multiple positif de l'entropie généralisée  $S_c$  définie par,*

$$S_c = \sum_{i=1}^n \frac{(x_i/m)^c - 1}{nc(c-1)}, \quad c \in \mathbb{R} - 0 - 1 \quad (2)$$

$$S_0 = \frac{1}{n} \sum_{i=1}^n \log \frac{m}{x_i} \quad (3)$$

$$S_1 = \frac{1}{n} \sum_{i=1}^n \frac{x_i}{m} \log \frac{x_i}{m} \quad (4)$$

où  $x_i$  représente le revenu de l'individu  $i$ .<sup>5</sup>

Ces mesures dépendent du paramètre  $c$ , autrement dit, de l'importance que les analystes accordent aux différentes parties de la distribution des revenus. En effet, lorsque  $c$  diminue et tend vers  $-\infty$ , l'indice accorde toute son importance aux bas revenus et aux transferts de revenus qui s'exercent dans cette partie. Inversement, pour  $c \geq 2$ , l'indice pondère plus fortement les hauts revenus.

Les mesures intragroupes ( $S_{cw}$ ) et intergroupes ( $S_{cb}$ ) de l'entropie généralisée s'expriment par (cf. Annexe A) :

$$S_{cw} = \sum_{j=1}^k f_j \left( \frac{m_j}{m} \right)^c S_c(\tilde{x}_j) \quad (5)$$

<sup>5</sup> La construction de la mesure de Theil par Bourguignon (1979) ne nécessitait pas l'utilisation des axiomes PD et PP. Seul, l'axiome de transfert de Dalton (1920) était vérifié *a posteriori*. Shorrocks (1980) va lui aussi vérifier, *a posteriori*, pour la mesure  $S_0$ , le respect des axiomes PD et PP. La dérivabilité permet d'un point de vue technique, tout comme la différentiabilité supposée dans les théorèmes suivants, de recourir au théorème d'Euler avant de déterminer par intégration la forme de la décomposition recherchée.

$$S_{cb} = \sum_{j=1}^k f_j \frac{1}{c(c-1)} \left( \left( \frac{m_j}{m} \right)^c - 1 \right) \tag{6}$$

où  $S_c(\tilde{x}_j)$  est l'entropie mesurée sur le vecteur  $\tilde{x}_j$  de la sous-population  $P_j$  ; la somme des deux composantes donne la mesure globale de l'entropie généralisée :

$$S_c = S_{cw} + S_{cb}. \tag{7}$$

On peut donc mesurer la contribution des inégalités qui s'exercent à l'intérieur des groupes ( $S_{cw}/S_c$ ) et la contribution des inégalités entre les moyennes des groupes ( $S_{cb}/S_c$ ).

Shorrocks (1984, 1988) s'intéresse alors à la condition minimale qui permet d'obtenir une décomposition cohérente, procurant ainsi une famille d'indicateurs décomposables en sous-groupes beaucoup plus large. On considère à nouveau que les inégalités entre les groupes  $I_b$  sont uniquement dépendantes des moyennes des revenus de chaque groupe  $P_j$  à taille fixée. La condition minimale de cohérence respectée par l'entropie généralisée, encore appelée monotonie en sous-groupes, conçoit que si l'inégalité augmente au sein de chaque groupe, alors l'inégalité totale augmente, à condition que les  $f_j$  et les  $p_j$  ne se modifient pas, autrement dit, que l'inégalité intergroupe  $I_b$  reste invariante. En définitive, l'inégalité totale augmente (diminue) car la hausse (baisse) de l'inégalité intragroupe l'emporte sur la variation nulle de l'inégalité intergroupe.

**Définition 3.3 (Shorrocks (1984, 1988))** *La décomposition cohérente en sous-populations. Une mesure d'inégalité  $I$  satisfait la propriété de décomposition cohérente en sous-groupes si :*

$$I = g(I_1, \dots, I_p, \dots, I_k; f_1, \dots, f_p, \dots, f_k; s_1, \dots, s_p, \dots, s_k), \tag{8}$$

où  $g$  est une fonction croissante de ses  $k$  premiers arguments.

Supposons que les revenus du groupe  $P_j$  se modifient, alors que les revenus moyens et les tailles de chaque sous-population restent constants. Si les inégalités augmentent (diminuent) au sein du groupe  $P_j$ , alors l'inégalité totale augmente (diminue)<sup>6</sup>. Shorrocks (1984, 1988) montre ainsi que l'emploi de mesures non cohérentes, comme la mesure de Gini, peut conduire selon lui à des situations non désirables pour lesquelles les inégalités s'accroissent à l'intérieur de chaque groupe  $P_j$  alors que l'inégalité totale diminue.

Les travaux de Shorrocks (1980, 1984, 1988) conduisent donc à légitimer l'emploi de la mesure de l'entropie généralisée qui est un indicateur régulier, homogène de degré zéro et additivement décomposable donc véri-

<sup>6</sup> Cet exemple montre les lacunes liées à la notion de cohérence. En effet, en considérant un transfert progressif entre des individus appartenant à des groupes différents  $P_j$  et  $P_h$ , rien ne garantit la monotonie suggérée par la cohérence en sous-groupes. Il faut pour cela un transfert ne modifiant pas la part de revenu détenue par chaque groupe

fiant le principe de décomposition cohérente<sup>7</sup>. L'entropie généralisée de Shorrocks (1980) provient de la famille de divergence des puissances de Cressie et Read (1984). Il s'agit donc d'un cas particulier de  $\phi$ -divergence telle qu'elle est définie par Csiszár (1967). Les cas où  $c = 0$  ou  $c = 1$  permettent d'obtenir une décomposition cohérente des mesures construites sur la divergence de deux fonctions de répartition. En reprenant une remarque émise par Theil (1967), Shorrocks (1980) montre d'ailleurs que seuls les cas où  $c = 0$  et  $c = 1$  procurent une somme de coefficients de pondération  $\omega_j(m_j, n_j) = \frac{n_j \left(\frac{m_j}{m}\right)^c}{n}$  égale à l'unité. Ceci implique que la quantité  $1 - \omega_j(m_j, n_j)$  n'est pas proportionnelle à l'indice intergroupe  $I_b$ . Par conséquent, les mesures  $S_0$  et  $S_1$  ont l'avantage de posséder des coefficients de pondération intragroupes  $\omega_j(m_j, n_j)$  indépendants de la mesure intergroupe  $I_b$ . L'entropie généralisée n'est pas toujours construite sur la divergence de deux fonctions de répartition. On peut en effet démontrer, pour  $c = 2$ , que l'on obtient la décomposition additive d'une mesure de dispersion divisée par une mesure de tendance centrale : il s'agit du coefficient de variation au carré  $\frac{\sigma^2}{m^2}$  (cf. Dagum (1997b)). Une autre mesure de ce type, le coefficient de Gini, respecte aussi cette propriété mais dans un cadre d'analyse très particulier, celui du non-chevauchement entre les distributions de revenus des groupes de la population. L'existence de zones de non-chevauchement entre distributions deux à deux provient du fait que parmi les individus du groupe le plus pauvre en moyenne, aucun ne possède un revenu supérieur à un quelconque individu appartenant au groupe le plus riche en moyenne (cf. par exemple Pyatt (1976), Silber (1989), Lerman et Yitzhaki (1991) et Dagum (1997a, 1997b)).

**Théorème 3.4 (Ebert (1988))** *L'indice  $I$ , défini sur l'ensemble des distributions classées par ordre décroissant  $x_1 \geq \dots \geq x_n$  avec  $r_i$  le rang de l'individu  $i$  dans cette distribution, satisfait l'invariance IR, PD, PP, NM, la différentiabilité et la décomposabilité additive DA pour des distributions qui ne se chevauchent pas ssi  $I$  est l'indice de Gini  $G$ , ou l'entropie généralisée  $S_c$ . L'indice de Gini s'écrit :*

$$G = \sum_{i=1}^n \left( \frac{2r_i - 1}{n^2} \right) \left( \frac{x_i}{m} - 1 \right). \quad (9)$$

<sup>7</sup> Les travaux de Cowell (1980a, 1980b) permettent de trouver la même famille d'entropie généralisée. L'auteur définit l'axiome de séparabilité additive (SA) selon lequel :  $I(F(x)) = \psi(m(F(x)), \int \pi(x) d(F(x)))$ , où  $F(x)$  est la fonction de répartition de support  $\chi$ , où  $\mathfrak{F}$  est l'ensemble de toutes les distributions de revenu possibles tel que  $I : \mathfrak{F} \rightarrow \mathbb{R}$ ,  $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $\pi : \chi \rightarrow \mathbb{R}$  et  $m(F) := \int x dF(x)$ . Une mesure d'inégalité  $I$  satisfait PD, IR, SA ssi  $I$  est équivalente de manière ordinale à  $S_c(F) = \frac{1}{c-1} \int \left[ \left( \frac{x}{m(F)} \right)^c - 1 \right] dF(x)$ .

Ebert (1988) nuance les propos de Shorrocks en montrant que l'indicateur de Gini est aussi une mesure additivement décomposable. La décomposabilité additive est atteinte lorsque aucune distribution de revenus  $P_j$  et  $P_h$  (pour tout  $j \neq h$ ) ne se chevauche<sup>8</sup>. Le résultat de Ebert (1988) est donc restrictif. Il repose de plus sur l'hypothèse selon laquelle la mesure de Gini s'exprime en fonction des revenus ordonnés. Les revenus  $x_i$  sont en effet dépendants du rang des individus  $r_i$ , la mesure  $G$  ne garantit donc plus l'anonymat des individus (l'axiome de symétrie). Certains auteurs (cf. par exemple Pyatt (1976), Silber (1989), Lerman et Yitzhaki (1991) et Dagum (1997a, 1997b)) refusent par ailleurs d'utiliser la méthode de décomposition cohérente, jugée trop restrictive. Ils vont ainsi démontrer que la mesure de Gini peut être décomposée d'une manière particulière, permettant la construction d'un nouveau type de décomposition en sous-groupes et garantissant l'anonymat.

### 3.2 Mesure de Gini et éléments de réflexion critiques

Lorsqu'on utilise la différence moyenne de Gini (1912),

$$GMD = \sum_{i=1}^n \sum_{r=1}^n \frac{|x_i - x_r|}{n^2},$$

il devient possible de récrire le coefficient de Gini (9) en respectant l'axiome SM :

$$G = \sum_{i=1}^n \sum_{r=1}^n \frac{|x_i - x_r|}{2n^2m}. \tag{10}$$

Cette expression permet de décomposer la mesure de Gini en une composante intragroupe  $G_w$  et une composante intergroupe « brute »  $G_{gb}$ , autorisant la distinction entre  $G_{gb}$  et  $I_b$  (cf. Dagum (1997a)) :

$$G = \underbrace{\sum_{j=1}^k G_{jj} f_j s_j}_{G_w} + \underbrace{\sum_{j=2}^k \sum_{h=1}^{j-1} G_{jh} (f_j s_h + f_h p_j)}_{G_{gb}}, \tag{11}$$

où  $G_{jj}$  est la mesure de Gini mesurée sur le groupe  $P_j$ , et où  $G_{jh}$  est la mesure de Gini mesurant les inégalités entre les groupes  $P_j$  et  $P_h$ . Contrairement à l'entropie généralisée, la mesure brute intergroupe est néces-

<sup>8</sup> En pratique, mises à part les études concernant les revenus d'individus pauvres et non-pauvres, il n'est pas rare d'observer des chevauchements entre les distributions. En général, la plus riche des femmes a un revenu beaucoup plus élevé que le moins riche des hommes.

<sup>9</sup> La mesure de Gini entre deux groupes (cf. Dagum (1987)) est construite en calculant la somme des différences binaires de revenu entre individus de groupes différents, puis en normalisant cette somme. Formellement,  $G_{jh} = \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} \frac{|x_{ij} - x_{rh}|}{n_j n_h (n_j + n_h)}$ , où  $x_{ij}$  et  $x_{rh}$  sont respectivement les revenus des individus  $i$  et  $r$  des groupes  $P_j$  et  $P_h$ , avec  $G_{jh} \in [0, 1)$ .

sairement positive et cette positivité peut exister même si les groupes ont une même moyenne. Outre l'avantage que les inégalités entre les hommes et les femmes ne peuvent pas être négatives, ce type de décomposition permet de s'intéresser à l'influence du chevauchement s'exerçant entre les distributions de revenu. La mesure de Gini décomposée fait en effet apparaître un troisième terme qui évalue les inégalités provenant de la zone de chevauchement entre les distributions des différents groupes étudiés. Soit  $D_{jh}$  la distance économique directionnelle (Dagum (1997a)) permettant de mesurer la proportion de différences binaires de revenus calculée à partir de la zone de non-chevauchement entre les distributions de revenus des groupes  $P_j$  et  $P_h$ <sup>10</sup>. La décomposition du coefficient de Gini en trois composantes devient donc :

$$G = \underbrace{\sum_{j=1}^k G_{jj} f_j s_j}_{G_w} + \underbrace{\sum_{j=2}^k \sum_{h=1}^{j-1} G_{jh} D_{jh} (f_j s_h + f_h s_j)}_{G_{nb}} + \underbrace{\sum_{j=2}^k \sum_{h=1}^{j-1} G_{jh} (1 - D_{jh}) (f_j s_h + f_h s_j)}_{G_t}. \quad (12)$$

Par conséquent, il est possible de mesurer la contribution des inégalités à l'intérieur des groupes  $G_w/G$ , la contribution nette des inégalités entre les groupes  $G_{nb}/G$  et la contribution de l'intensité de transvariation entre les groupes  $G_t/G$ . L'indice  $G_{nb}$  mesure les inégalités générées par les revenus des groupes les plus riches en moyenne creusant des écarts avec ceux des groupes les moins riches. Notons qu'au même titre que l'entropie généralisée, dès lors que les groupes ont la même moyenne, les inégalités nettes entre les groupes sont nulles (ceci est possible puisque  $m_j = m_h \Rightarrow D_{jh} = 0$ ). Le troisième terme  $G_t$  mesure au contraire les inégalités provenant des revenus élevés des groupes en moyenne les moins riches qui créent des écarts de revenu avec les individus des groupes en moyenne les plus riches<sup>11</sup>.

La mesure de l'entropie est souvent préférée à celle de Gini en raison des propriétés de monotonie et de décomposition additive. Il est cependant légitime de penser que l'évolution des inégalités totales suit une logique bien plus complexe que la monotonie ne laisse supposer *a priori*. Un relâchement

<sup>10</sup> Pour construire  $D_{jh}$ , on mesure tout d'abord la distance directionnelle brute  $d_{jh}$  qui est une moyenne pondérée des différences de revenu  $x_{ij} - x_{rh}$  pour chaque revenu  $x_j$  d'un membre de  $P_j$  supérieur au revenu  $x_{rh}$  d'un membre de  $P_h$  ( $x_j > x_{rh}$  tel que  $m_j > m_h$ ) :  $d_{jh} = \int_0^{\infty} dF_j(x) \int_0^x (x-y) dF_h(y)$ , où  $F_j$  et  $F_h$  sont respectivement les fonctions de répartition des groupes  $P_j$  et  $P_h$ . on mesure ensuite le moment d'ordre 1 de transvariation  $p_{jh} : p_{jh} = \int_0^{\infty} dF_h(x) \int_0^x (x-y) dF_j(y)$  qui est la moyenne pondérée des différences binaires de revenu  $x_{rh} - x_{ij}$  ( $x_{ij} < x_{rh}$  tel que  $m_j > m_h$ ). Ainsi :  $D_{jh} = \frac{d_{jh} - p_{jh}}{d_{jh} + p_{jh}}$ .

<sup>11</sup> Avant que Dagum (1997a, 1997b) ne démontre que ce troisième terme  $G_t$  provient de la transvariation (Gini (1916), Dagum (1959, 1960 et 1961)), on a longtemps cru que l'indicateur de Gini n'était pas décomposable en sous-groupes.  $G_t$  fut nommé en premier lieu terme d'interaction par Mookherjee et Shorrocks (1982).

de la condition de monotonie permettrait d'envisager le cas selon lequel l'inégalité totale diminue (augmente) même si les inégalités à l'intérieur des groupes augmentent (diminuent). Il suffit pour cela que le montant des inégalités intergroupes l'emporte sur celui des inégalités intragroupes.

Illustrons ce propos à l'aide de l'exemple de Cowell (1988, p. 310) pour l'indice de Gini. A la période T1, la distribution de revenu globale est donnée par  $x' = (\tilde{x}'_1, \tilde{x}'_2)$  où  $\tilde{x}'_1 = (1, 6, 12)$  est la distribution du groupe 1, et  $\tilde{x}'_2 = (6, 6, 9)$  celle du groupe 2. En T2, la distribution devient  $y' = (\tilde{y}'_1, \tilde{y}'_2)$  où  $\tilde{y}'_1 = (3, 3, 13)$  et  $\tilde{y}'_2 = (6, 7, 8)$ . Les moyennes et les tailles des sous-populations sont constantes (les proportions  $s_j$  et  $f_j$  le sont aussi).

Entre T1 et T2, l'indice de Gini de chaque sous-groupe diminue :  $0,385 \rightarrow 0,351$  pour le groupe 1 ;  $0,095 \rightarrow 0,063$  pour le groupe 2. Au contraire, l'indice global augmente :  $0,266 \rightarrow 0,275$ , car l'intensité de transvariation  $G_t$  augmente ( $0,125 \rightarrow 0,150$ ) alors que l'indice de Gini net intergroupe  $G_{nb}$  reste constant ( $0,025$ ) entre T1 et T2. Le troisième terme mesure la dépréciation (ou l'exclusion) ressentie par les sous-groupes les mieux lotis (ceux dont la moyenne des revenus est élevée). Ces individus peuvent en effet se sentir exclus du fait que le seul haut revenu du groupe le moins bien loti (le revenu de valeur 12 du groupe 1) crée davantage d'inégalités en T2 qu'en T1. En T2, ce revenu engendre les trois différences suivantes :  $(13-6)+(13-7)-(13-8) = 18$ , alors qu'en T1 le montant de ces trois différences était plus faible :  $(12-6)-(12-6)+(12-9) = 15$ . C'est cette intensité de transvariation, ce sentiment d'exclusion qui émane des groupes relativement riches, qui explique la hausse de l'inégalité totale.

D'après Sen (1973), cette exclusion est le fruit de comparaisons par paires. Ces dernières sont permises dès lors qu'un individu a la possibilité de comparer son revenu (ou sa satisfaction, son utilité) à ceux des autres membres de la population, qu'ils appartiennent à son groupe ou non. Il évoque ce propos en ces termes :

*"[I]n any pair-wise comparison the man with the lower income can be thought to be suffering from some depression on finding his income to be lower. Let this depression be proportional to the difference in income. The sum total of all such depressions in all possible pair-wise comparisons takes us to the Gini coefficient."*

En 1976, Pyatt reprend l'idée de Sen afin de construire une nouvelle décomposition de la mesure de Gini en relation avec la théorie des jeux. L'indicateur de Gini est associé à un jeu où chaque convive a la possibilité de garder son revenu ou de prendre celui d'une autre personne tirée au hasard dans la population. Cette idée reposant sur le concept de comparaisons par paires fut reprise en 1998 par Dagum qui montre que les indices dérivés de l'entropie ne satisfont pas cette propriété. Selon l'auteur, l'importance de ce principe est liée au fait que les agents se livrent sans cesse à des comparaisons entre la satisfaction qu'ils retirent d'une situation (par exemple un bien de consommation) et celle que les autres individus obtiennent.

La mesure de Gini montre que la question de la décomposition des inégalités en sous-groupes est plus complexe que la propriété de décomposition additive. Cette complexité pourrait être mise en évidence par l'utilisation d'un concept de décomposition en sous-groupes basé sur les comparaisons par paires.

**Définition 3.5** *La décomposition en sous-groupes basée sur les comparaisons par paires.* Une mesure d'inégalité satisfait la propriété de décomposabilité basée sur les comparaisons par paires si la mesure intragroupe rassemble toutes les comparaisons binaires qui existent à l'intérieur des groupes et si la mesure intergroupe inclut toutes les comparaisons binaires possibles entre toutes les paires de groupes.

Cette propriété ouvre des perspectives dans la conception de l'inégalité intergroupe. Elle ne relève plus des différences moyennes qui existent entre les groupes mais des inégalités entre les revenus des individus qui appartiennent à des groupes différents : l'inégalité intergroupe brute. Comme le souligne Dagum (1997a), les indicateurs de mesure d'inégalité dont les composantes intergroupes sont uniquement fonction des moyennes des sous-groupes (comme toutes celles qui proviennent de la règle de décomposition cohérente) ne peuvent pas tenir compte des phénomènes d'asymétrie et de variance. Au même titre que l'analyse de la variance, tout se passe comme si les sous-groupes étaient normalement et également distribués, de même variance, et statistiquement indépendants. Comme elles sont construites en remplaçant les revenus des individus par la moyenne des revenus de leur groupe correspondant, ces mesures ignorent les variances et les asymétries qui prévalent dans chaque groupe. La décomposition par paire paraît donc heuristique. Pour formaliser une telle propriété, il est possible de recourir à la famille des indicateurs basés sur les paires de Kolm (1999) :

$$I(x) = F \left( \sum_{i < r} h(\phi(x_i, x_r)) \right). \quad (13)$$

où  $F(0) = h(0) = 0$ ,  $F$  est croissante,  $\phi$  est symétrique,  $\phi(\xi, \xi) = 0$ , et où  $\phi > 0$  dans les autres cas. Remarquons dans cet exemple qu'il est possible de regrouper les paires  $h(\phi(x_i, x_r))$  provenant d'un même groupe et celles qui sont issues de deux groupes différents. En supposant que  $F$  autorise le calcul des contributions qu'exercent les deux types de comparaisons par paires au montant global de l'inégalité, on obtient bien une nouvelle forme de décomposabilité en sous-groupes<sup>12</sup>.

Une autre propriété peut caractériser la famille des mesures basées sur les paires. En effet, ces mesures permettent de capturer des transferts de revenu intergroupes qui préservent la moyenne. Soit les distributions de revenu

<sup>12</sup> La fonction  $F$  ne permet pas systématiquement de mesurer la contribution des inégalités intragroupes ou intergroupes. Par exemple, une fonction puissance ne permettrait pas forcément de regrouper les paires intragroupes et intergroupes afin de déterminer  $I_w/I$  ou  $I_b/I$ .

$\tilde{z}_1 = (0, 2, 4)$  et  $\tilde{z}_2 = (0, 0, 6)$ . Supposons un transfert d'un Euro du troisième individu de la distribution  $\tilde{z}_2$  au premier de  $\tilde{z}_1$  et un transfert d'un Euro du troisième individu de  $\tilde{z}_1$  au premier de  $\tilde{z}_2$ , tels que les moyennes restent constantes. On obtient :  $\tilde{l}'_1 = (1, 2, 3)$  et  $\tilde{l}'_2 = (1, 0, 5)$ . Les mesures cohérentes et additivement décomposables fournissent une mesure intergroupe nulle. Elles ne permettent pas d'apprécier la manière dont les politiques de redistribution basées sur des transferts entre deux groupes qui préservent la moyenne peuvent diminuer les inégalités. Les mesures intergroupes basées sur les paires peuvent, au contraire, capter une variation d'inégalités, plus précisément une baisse d'inégalités s'il s'agit par exemple de la mesure de Gini brute intergroupe  $G'_{yb}$ .

En définitive, deux mesures régulières, homogènes de degré zéro, et décomposables en sous-groupes ont retenu l'attention des économistes. De nombreuses applications ont été menées afin d'expliquer l'évolution de la structure par groupe des inégalités de revenus<sup>13</sup> sans qu'une réponse suffisante ne soit fournie sur la différence qui prévaut entre les mesures additivement décomposables (l'entropie) et les mesures qui ne satisfont pas complètement ce principe (Gini). Néanmoins, les travaux de Gini (1916), Dagum (1959, 1960, 1961, 1997a, 1997b, 1998), Sen (1973), Pyatt (1976), et Kolm (1999) suggèrent implicitement que la différence entre l'entropie généralisée et la mesure de Gini fait naître une forme de décomposabilité basée sur les comparaisons par paires. Cette dernière peut réconcilier les deux mesures puisque, en effet, certains indices additivement décomposables peuvent être décomposés par paire.<sup>14</sup>

#### 4 La décomposition en sources de revenu

La décomposition des mesures d'inégalité en sources de revenu (facteurs) permet de comprendre qu'une source particulière (le revenu du travail ou le revenu du capital par exemple) puisse être à l'origine des inégalités dans la répartition des revenus. Les sources de revenu procurent des liens étroits avec la théorie économique. En 1978, Fei *et alii* montrent par exemple l'existence de relations entre les facteurs et la croissance économique, ainsi que la possibilité d'analyser la distance qui sépare les familles riches des familles pauvres par la détermination du poids des facteurs dans l'inégalité globale.

<sup>13</sup> cf. par exemple Jenkins (1995) pour l'entropie et Dagum (1997a, 1997b) pour la mesure de Gini.

<sup>14</sup> cf. Chameni (2006) pour le coefficient de variation au carré qui permet de construire une mesure intragroupe, une mesure nette intergroupe et une mesure de transvariation intergroupe, toutes analogues aux mesures décomposées de Gini.

Ces méthodes se sont révélées intéressantes dans la mesure où les formes analytiques à déterminer ne constituent pas un exercice simple. En effet, que ce soit pour l'entropie généralisée ou la mesure de Gini, la mesure globale est inférieure ou égale à la moyenne pondérée des mesures par source :

$$I(x) \leq \sum_{\ell} \rho^{\ell} I(\tilde{x}^{\ell}). \quad (14)$$

avec  $m^{\ell}$  la moyenne arithmétique de la source  $\ell$ ,  $I(\tilde{x}^{\ell}) =: I^{\ell}$  l'indice d'inégalité mesuré sur le vecteur de la source de revenu  $\ell$ , et  $\rho^{\ell} := \frac{m^{\ell}}{m}$ . Les économistes ont donc cherché à déterminer la pondération qui pouvait leur permettre d'atteindre l'égalité dans l'équation (14) ci-dessus. Cette recherche, conduite autour des indices de dispersion divisés par une valeur de tendance centrale, s'est effectuée de deux manières différentes : soit en effectuant la somme d'une combinaison linéaire des mesures obtenues pour chaque source soit en rajoutant à cette même combinaison linéaire un terme faisant intervenir les covariances des sources de revenu.

La paternité de la première approche peut être attribuée à Rao (1969) qui démontre que l'indice de concentration de Gini est décomposable en facteurs<sup>15</sup>. Supposons que les revenus de chaque individu soient constitués de  $q$  facteurs (sources de revenu, dépenses de consommation, etc.) :

$$x_i = \sum_{\ell=1}^q x_i^{\ell} \geq 0, \quad x_i^{\ell} \in \mathbb{R}, \quad (15)$$

où  $x_i^{\ell}$  est la  $\ell$ -ième source de revenu détenue par l'individu  $i$ . La décomposition en facteurs s'exprime alors par :

$$G = \sum_{\ell=1}^q \rho^{\ell} \bar{G}^{\ell}, \quad (16)$$

où  $\rho^{\ell} = m^{\ell}/m$  représente la part moyenne du facteur  $\ell$  dans le revenu total et  $\bar{G}^{\ell}$  un pseudo-indice de Gini mesuré sur le vecteur de la source  $\ell$ . Ce dernier mesuré sur la source  $\ell$  peut être négatif. Il se calcule à partir de l'équation (9) où les éléments  $x_i^{\ell}$  du vecteur  $\tilde{x}^{\ell}$  de la source  $\ell$  ne sont pas classés par ordre décroissant mais placés au même rang que celui du revenu global correspondant  $x_i$ . La contribution de chaque source à l'inégalité totale  $G$  est aussi représentée par une pseudo-inégalité de facteur pondérée par la part de cette source dans le revenu moyen. Rao (1969) remarque qu'il s'agit d'une décomposition non exacte puisque l'indice de Gini de la source  $\ell$ ,  $G^{\ell}$ , est ignoré. Fei *et alii* (1978) puis Fields (1979) justifient cependant

<sup>15</sup> L'indice de Gini étant un cas particulier de l'indice de concentration (cf. Pyatt *et alii* (1980)), la décomposition est donc aussi valable pour le premier.

son intérêt par la relation qui existe entre  $G^{\ell}$  et  $\bar{G}^{\ell}$  :  $\bar{G}^{\ell} = \Omega^{\ell} G^{\ell}$ , où  $\Omega^{\ell}$  représente un coefficient de rang mesurant la corrélation entre la source  $\ell$  et le revenu global<sup>16</sup>. La décomposition exacte du coefficient de Gini s'écrit alors :

$$G = \sum_{\ell=1}^q \rho^{\ell} \Omega^{\ell} G^{\ell}. \tag{16}$$

Cette équation a donné lieu à de nombreuses applications car elle permet des interprétations économiques de type compensation d'inégalités. La contribution positive des inégalités de revenus du capital à l'inégalité totale peut par exemple être compensée par la contribution négative des transferts. Ce type d'interprétation peut aussi s'effectuer en utilisant les mesures qui incorporent les covariances entre les sources. La variance, appartenant à la famille de l'entropie généralisée à  $1/m^2$  près, caractérise cette deuxième approche :

$$\sigma^2 = \sum_{\ell=1}^q \sigma^2(\tilde{x}^{\ell}) + 2 \sum_{\ell=2}^q \sum_{j=1}^{\ell-1} cov(\tilde{x}^{\ell}, \tilde{x}^j). \tag{18}$$

On retrouve l'effet principal associé à chaque source de revenu et leurs interactions deux à deux. Afin de réunir l'interprétation économique des décompositions en facteurs et la logique mathématique qui peut permettre de légitimer le recours à une mesure précise, Shorrocks (1982) introduit six hypothèses nécessaires et suffisantes, l'auteur refusant la dénomination axiome. Elles permettent de démontrer analytiquement que la variance et le coefficient de variation au carré sont les seules mesures décomposables en facteurs. Lerman et Yitzhaki (1985) ne retiennent pas ces hypothèses jugées trop restrictives. Ils montrent que seule la décomposition de la mesure de Gini (22) reste désirable puisqu'elle fait apparaître trois indices : la mesure d'inégalité pour chaque facteur, la part moyenne du facteur dans le revenu global, et un coefficient de corrélation entre le facteur et le revenu global. En effet, la variance ne fait intervenir au mieux que deux de ces indices voire un seul si elle s'exprime par :  $\sigma^2 = \sum_{\ell=1}^q cov(\tilde{x}^{\ell}, x)$ .

Afin de surmonter les difficultés auxquelles les économistes se heurtent lorsqu'ils proposent un arbitrage entre la logique économique, basée sur le calcul de la contribution d'un facteur à l'inégalité totale, et la logique mathématique, permettant de légitimer une mesure, Auvray et Trannoy (1992) proposent une généralisation de la technique de décomposition en facteurs par l'utilisation de la valeur de Shapley. Cette idée est reprise par Chantreuil

<sup>16</sup> Lerman et Yitzhaki (1985) montrent que  $\Omega^{\ell}$  est proche du coefficient de corrélation de Pearson. Cependant, son dénominateur est différent. Au lieu d'être basé sur le produit des écarts-types, il mesure la covariance entre le facteur  $\ell$  et son rang. Si  $\Omega^{\ell} = 1$ , la source  $\ell$  est une fonction croissante du revenu total. Si  $\Omega^{\ell} = -1$ , le facteur  $\ell$  est une fonction décroissante du revenu total. Enfin, si  $\Omega^{\ell} = 0$ , la source de revenu  $\ell$  est également et identiquement distribuée, donc sa contribution à l'inégalité totale est nulle.

et Trannoy (1999), Sastre et Trannoy (2002) et par Shorrocks (1999). Ce dernier retrouve à ce propos son résultat de 1982 concernant les mesures de la variance et du coefficient de variation au carré<sup>17</sup>.

La valeur de Shapley (1953), concept issu de la géométrie des simplexes, est utilisée en théorie des jeux coopératifs où l'on cherche à déterminer les modalités de partage d'une somme (un bénéfice, un coût, etc.).

On suppose qu'une mesure d'inégalité  $I$  est fonction de  $q$  facteurs  $\ell$  tels que :  $\ell \in \mathcal{L} = \{1, 2, \dots, q\}$ . L'algorithme de Shapley permet de mesurer la contribution de chaque facteur  $\ell$  à la mesure de l'inégalité totale  $I$  en supposant que ce facteur peut être éliminé lorsqu'on mesure  $I$ . Cet indicateur  $I$  dépend donc d'un ensemble mouvant  $\mathcal{R}$ . Il admet ainsi plusieurs formes fonctionnelles  $F(\mathcal{R})$  définies par  $F : \{\mathcal{R} : \mathcal{R} \subseteq \mathcal{L}\} \rightarrow \mathbb{R}$ .

Si on considère que le facteur  $\ell$  ( $\ell \in \mathcal{R}$ ) a été supprimé de l'ensemble des variables, le nombre de facteurs restants est égal à  $r$ , le cardinal de  $\mathcal{R}$ . En mesurant la différence entre l'indicateur d'inégalité global (sur l'ensemble de toutes les variables) et l'indicateur après la perte d'une variable  $\ell$ , on définit un impact marginal particulier associé à la variable  $\ell$ . L'opération est réitérée en supprimant  $\ell$  avec d'autres sources de revenu. La mesure de  $I$  permet alors de déterminer un nouvel impact marginal. On répertorie ainsi au total  $q!$  impacts marginaux pour chaque source de revenu.

Pour mesurer la contribution  $C^\ell$  de chaque facteur  $\ell \in \mathcal{L}$  à l'inégalité totale, on calcule la moyenne pondérée :

$$C^\ell(\mathcal{L}; F) = \sum_{r=0}^{q-1} \sum_{\mathcal{R} \subseteq \mathcal{L} \setminus \{\ell\}} \frac{(q-1-r)!r!}{q!} \Delta_\ell F(\mathcal{R}). \quad (19)$$

où

$$\Delta_\ell F(\mathcal{R}) := F(\mathcal{R} \cup \{\ell\}) - F(\mathcal{R}) \quad (20)$$

est un des impacts marginaux de la variable  $\ell$ . Chantreuil et Trannoy (1999) montrent que la valeur Shapley est une règle de décomposition cohérente car la somme de toutes les contributions procure la mesure totale de l'inégalité :

$$I = \sum_{\ell=1}^q C^\ell. \quad (21)$$

Ainsi, au lieu d'exprimer une relation de décomposition dans laquelle on retrouve les mesures d'inégalité associées à chaque source, la valeur de Shapley privilégie des mesures secondaires, autrement dit des indices de

<sup>17</sup> Shorrocks (1999) démontre que la valeur de Shapley peut aussi être appliquée dans d'autres domaines, comme celui des mesures de pauvreté.

contribution, s'appliquant à toutes les mesures d'inégalité<sup>18</sup>. Elle permet de réconcilier les propriétés mathématiques (que l'on trouve notamment dans Shorrocks (1982)) avec certaines hypothèses de l'analyse économique, comme la négativité de la contribution d'un facteur.

## 5 Une solution intermédiaire au problème du no-bridge : la multi-décomposabilité

L'utilisation exclusive des méthodes de décomposition en sous-groupes et en sources de revenu est-elle inévitable ? De nombreuses recherches, comme celle de Jenkins (1995) sur l'entropie, laissent penser que les ensembles associés aux mesures décomposables en sous-groupes et aux mesures décomposables en facteurs sont disjoints. Pourtant, les techniques de décomposition précédentes laissent supposer que les formes analytiques des deux types de décomposition peuvent être rapprochées pour mesurer, d'une manière plus complète, les inégalités. Pour cela, on utilise simultanément, les deux méthodes de décomposition en se basant soit sur la décomposition en sous-groupes par paires soit sur la valeur de Shapley.

Lorsqu'on s'intéresse à la mesure de l'entropie généralisée ou unidimensionnelle  $S_c$  de Shorrocks (1980), on constate qu'elle ne permet pas d'atteindre la propriété de multi-décomposabilité qui combine les deux méthodes de décomposition en identifiant les couples « source  $\ell$  / intragroupe » et « source  $\ell$  / intergroupe » dont la somme fournit l'indicateur global (cf. Annexe A). En effet, des termes redondants (comme des termes multiplicatifs entre deux sources de revenu) ou des termes non décomposables (comme le logarithme d'une somme) ne permettent pas la mesure de la contribution d'une source particulière au montant des inégalités intragroupes et intergroupes. Les possibilités de généralisation offertes par la valeur de Shapley permettent de tempérer ce résultat en appliquant l'algorithme de Shapley (24) à chaque composante  $I_w$  et  $I_b$  séparément, tout en respectant la règle de cohérence. Formellement :

$$I = \sum_{\ell=1}^q C_w^{\ell} + \sum_{\ell=1}^q C_b^{\ell}. \quad (22)$$

où  $C_w^{\ell}$  est la contribution de la source  $\ell$  à l'inégalité intragroupe et  $C_b^{\ell}$  la contribution de la source  $\ell$  à l'inégalité intergroupe. Les éléments du ta-

<sup>18</sup> Une décomposition exacte fait intervenir les mesures d'inégalité pour chaque source. Elle est attrayante puisque la base axiomatique existante pour la mesure globale s'applique aussi à la mesure de chaque source. Néanmoins, dans ce cadre de décomposition, si la logique économique est privilégiée par le calcul des contributions, rien ne prouve que les combinaisons linéaires de mesures par source, par exemple  $\rho^{\ell} \Omega^{\ell} C^{\ell}$  pour la mesure de Gini, respecte les axiomes de base.

bleau 1 peuvent être dans ce cas déterminés et toutes les mesures décomposables en deux éléments, comme les mesures cohérentes en sous-groupes, sont multi-décomposables<sup>19</sup>.

Ces multi-décompositions n'échappent pas cependant aux critiques énoncées précédemment concernant les mesures intergroupes déduites de la décomposition cohérente. Il en est ainsi car les restrictions imposées par cette propriété, impliquant notamment des sous-populations dont les caractéristiques obéissent à des lois normales de même variance et statiquement indépendantes, peuvent inciter à ne pas adopter ce type de multi-décomposition. Une version modernisée de l'entropie généralisée, l'entropie généralisée multidimensionnelle de Tsui (1999), fournit une multi-décomposabilité exacte (cf. Annexe B), mais ne permet pas, non plus, de contourner les difficultés liées à la structure de l'inégalité intergroupe.

Pour dépasser ces difficultés, il est possible de recourir à la famille des mesures d'inégalité décomposables en sous-populations par les comparaisons par paires (18). Elle réconcilie le cas particulier de l'entropie généralisée pour lequel  $c = 2$  et la mesure de Gini. Imposons une forme fonctionnelle à cette famille de telle sorte qu'elle permette de retrouver facilement ces deux indicateurs dont nous savons qu'ils sont réguliers et homogènes de degré zéro :

$$g(x) = \sum_{i=1}^n \sum_{r=1}^n \frac{|x_i - x_r|^\alpha}{f(\cdot)}, \forall f(\cdot) \neq 0, \alpha \geq 1. \quad (23)$$

Au numérateur apparaissent les inégalités de revenu entre chaque paire d'individus dépendant d'un paramètre  $\alpha$ <sup>20</sup>. Le dénominateur a pour rôle de normaliser l'indice (cf. par exemple Ebert (1999)). Pour des formes diverses de  $f$ , on obtient :

<sup>19</sup> Notons que la multi-décomposition peut aussi s'obtenir en appliquant une première fois la valeur de Shapley afin d'obtenir les composantes  $I_w$  et  $I_b$ , puis en appliquant une seconde fois la valeur de Shapley à chaque indice  $I_w$  et  $I_b$  afin d'obtenir les contributions de chaque source de revenu  $C'_w$  et  $C'_b$ . L'article d'Israëli (2007) montre, dans le cadre d'une régression linéaire, comment le  $R^2$  se décompose en variables primaires puis en variables secondaires en appliquant séquentiellement la valeur de Shapley. Deux procédures permettent d'atteindre ce résultat. La première procédure est la valeur Nested-Shapley de Chantreuil et Trannoy (1999) qui permet d'obtenir les contributions des variables secondaires indépendamment des variables primaires. La seconde procédure est la valeur d'Owen (reprise par Shorrocks (1999) dans le cadre des inégalités et de la pauvreté) pour laquelle les contributions des variables secondaires sont dépendantes des contributions des variables primaires. Cf. aussi Mussard et Peypoch (2006) pour le cas d'une multi-décomposition par la valeur de Shapley appliquée à l'indice de productivité de Malmquist.

<sup>20</sup> Contrairement au paramètre  $c$  de l'entropie généralisée,  $\alpha$  ne permet pas de se focaliser sur une partie de la distribution. Il peut néanmoins être considéré comme un paramètre d'aversion à l'inégalité, plus  $\alpha$  augmente et plus les inégalités s'amplifient.

$$g(x) = \begin{cases} G(x) & \text{si } f(x) = 2n^2m \text{ et si } \alpha = 1 \\ GMR(x) & \text{si } f(x_i, x_r) = n^2 \max\{x_i, x_r\} \text{ et si } \alpha = 1 \\ GMD(x) & \text{si } f(x) = n^2 \text{ et si } \alpha = 1 \\ S_2(x) & \text{si } f(x) = 2n^2m^2 \text{ et si } \alpha = 2 \end{cases} \quad (24)$$

où  $GMR$  est le rapport moyen de Gini (cf. Mussard (2004a)) et  $S_2$  le coefficient de variation au carré. Lorsque deux individus  $i$  et  $r$  sont tirés au hasard de la population globale  $P$  (avec remise), la différence moyenne de Gini mesure la différence de revenu espérée entre ces deux individus, alors que le rapport moyen de Gini<sup>21</sup> permet d'évaluer la part de revenu espérée du plus riche (en pourcentage) qui est détenue par le plus pauvre.

Le point commun à ces quatre mesures est la décomposabilité en sous-groupes. Lorsque la population mère  $P$  est partitionnée en  $k$  sous-groupes ( $j = 1, \dots, k$ ), on mesure la composante intragroupe  $g(x)_{iw}$  avec les comparaisons par paires à l'intérieur des groupes, et la mesure intergroupe brute  $g(x)_{gb}$  avec les comparaisons par paires entre chaque paire de groupes :

$$g(x) = \sum_{i=1}^n \sum_{r=1}^n \frac{|x_i - x_r|^\alpha}{f(\cdot)} = \underbrace{\sum_{j=1}^k \sum_{i=1}^{n_j} \sum_{r=1}^{n_j} \frac{|x_{ij} - x_{rj}|^\alpha}{f(\cdot)}}_{g(x)_{iw}} + 2 \underbrace{\sum_{j=2}^k \sum_{h=1}^{j-1} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} \frac{|x_{ij} - x_{rh}|^\alpha}{f(\cdot)}}_{g(x)_{gb}} \quad (25)$$

où  $x_{ij}$  est le revenu de l'individu  $i$  appartenant au groupe  $j$ <sup>22</sup>. Comme l'indique le tableau 1, ce type de ratio permet, par exemple, de mesurer les inégalités provenant du groupe des hommes, du groupe des femmes, et les inégalités entre les hommes et les femmes.

Pour combiner la décomposition en sous-groupes et la décomposition en sources de revenu, il faut mettre en évidence les sources de rémunération. Chaque individu possède au total  $q$  sources de revenu dont la somme fournit

<sup>21</sup> La mesure  $GMR$  satisfait NM, PP, SM, IR, une version relative de l'axiome de Pigou-Dalton et une version plus faible du principe de transfert issu du quasi-ordre différentiel relatif. La mesure  $GMD$  satisfait NM, SM, et PD.

<sup>22</sup> Pour les quatre mesures, il est possible de faire apparaître les indices pour chaque groupe  $g_j$  et les indices pour chaque paire de groupe  $g_{jh}$ . Il suffit pour cela de trouver la fonction de normalisation de  $g_j$  notée  $f(\cdot)_j$ , de spécifier un indice entre deux groupes  $g_{jh}$  et sa fonction de normalisation  $f(\cdot)_{jh}$ . La décomposition par paire s'exprime par :  $g(x) = \sum_{j=1}^k \frac{f(\cdot)_j}{f(\cdot)} g_j + \sum_{j=2}^k \sum_{h=1}^{j-1} \frac{f(\cdot)_{jh}}{f(\cdot)} g_{jh}$ . Par exemple, pour  $(GMD)_x$ , on peut avoir :  $f(\cdot)_j = n_j^2$ ,  $GMD_{jh} := \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} \frac{|x_{ij} - x_{rh}|}{n_j n_h}$  et  $f(\cdot)_{jh} = n_j n_h$ . De même, pour  $S_2$  (cf. Chameni (2006)) :  $f(\cdot)_j = 2n_j^2 m_j^2$ ,  $S_{2jh} := \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} \frac{|x_{ij} - x_{rh}|^2}{n_j n_h (m_j^2 + m_h^2)}$  et  $f(\cdot)_{jh} = n_j n_h (m_j^2 + m_h^2)$ . La décomposition ainsi obtenue est différente de celle de Shorrocks (1980) et évite les critiques liées à la structure de la mesure intergroupe.

le revenu global. Rappelons que le revenu de l'individu  $i$  se décompose selon :  $x_i = \sum_{\ell=1}^q x_i^\ell$ . On obtient alors :

$$g(x) = \sum_{j=1}^k \sum_{i=1}^{n_j} \sum_{r=1}^{n_j} \frac{|\sum_{\ell=1}^q x_{ij}^\ell - \sum_{\ell=1}^q x_{rj}^\ell|^\alpha}{f(\cdot)} + 2 \sum_{j=2}^k \sum_{h=1}^{j-1} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} \frac{|\sum_{\ell=1}^q x_{ij}^\ell - \sum_{\ell=1}^q x_{rh}^\ell|^\alpha}{f(\cdot)}. \quad (26)$$

Deux solutions peuvent alors être retenues en fonction des valeurs de  $\alpha$ . Pour  $\alpha \geq 2$ , on peut mesurer le poids de chaque source à l'inégalité totale ainsi que les interactions entre les sources. Les termes multiplicatifs entre les sources ne permettent pas, comme pour la décomposition de la variance (18), d'obtenir une interprétation économique claire en termes de corrélation croisée entre les sources. Pour  $\alpha = 1$ , on peut calculer, sans termes redondants, la contribution de chaque source à l'inégalité totale. En utilisant la formule  $|a - b| = a + b - 2\min\{a, b\}$ , on a :<sup>23</sup>

$$g(x) = \underbrace{\sum_{j=1}^k \sum_{i=1}^{n_j} \sum_{r=1}^{n_j} \frac{\sum_{\ell=1}^q x_{ij}^\ell + \sum_{\ell=1}^q x_{rj}^\ell - 2\min\{\sum_{\ell=1}^q x_{ij}^\ell, \sum_{\ell=1}^q x_{rj}^\ell\}}{f(\cdot)}}_{g(x)_w} + 2 \underbrace{\sum_{j=2}^k \sum_{h=1}^{j-1} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} \frac{\sum_{\ell=1}^q x_{ij}^\ell + \sum_{\ell=1}^q x_{rh}^\ell - 2\min\{\sum_{\ell=1}^q x_{ij}^\ell, \sum_{\ell=1}^q x_{rh}^\ell\}}{f(\cdot)}}_{g(x)_{gb}}. \quad (27)$$

Pour les applications empiriques, lorsque l'équation (27) est retenue, on peut déterminer le minimum de deux sommes. Les propriétés d'associativité et de commutativité de l'addition permettent ainsi de regrouper toutes les sources de revenu qui sont de même nature. On peut alors isoler le facteur salaire et mesurer sa contribution aux inégalités intragroupes (par exemple aux inégalités masculines et aux inégalités féminines) et sa contribution aux inégalités intergroupes brutes (par exemple aux inégalités entre hommes et femmes). En isolant l'effet de toutes les sources de rémunération, on obtient au total une mesure d'inégalité multi-décomposable. Elle autorise le calcul de tous les couples générateurs d'inégalités tels que « source  $\ell$  / groupe  $j$  » et « source  $\ell$  / entre les groupes  $j$  et  $h$  »<sup>24</sup>. L'équation (27) est

<sup>23</sup> D'autres formules équivalentes peuvent être utilisées afin d'obtenir le même résultat comme  $|a - b| = \max\{a, b\} - \min\{a, b\}$ .

<sup>24</sup> Cette solution (27) n'aborde pas les possibilités de décomposition en trois éléments. On pourrait en effet décomposer la mesure brute intergroupe  $g(x)_{gb}$  afin d'obtenir la mesure des inégalités moyennes entre les groupes et l'intensité de transvariation. Pour cela, il suffirait de regrouper par paquets les différences de revenu positives entre  $P_j$  et  $P_h$  issues du groupe le plus riche en moyenne, et celles qui proviennent du groupe le plus pauvre en moyenne. On obtiendrait alors des indicateurs multi-décomposables munis de trois composantes.

une solution intermédiaire au problème du no-bridge pour les mesures de Gini, *GMR* et *GMD*. En effet, une solution de multi-décomposition de premier rang consisterait à décomposer la mesure d'inégalité globale en une moyenne pondérée d'indices pour chaque groupe et chaque source ( $I_j^f$ ) et une moyenne pondérée d'indices pour chaque paire de groupe et chaque source ( $I_{jh}^f$ ). Il existe donc autant de problèmes ouverts que de mesures d'inégalités<sup>25</sup>.

## 6 Les revenus italiens en 2000

Afin de comprendre l'intérêt économique des multi-décompositions, étudions les résultats qu'il est possible d'obtenir lorsque les solutions intermédiaires de multi-décomposition sont envisagées. Utilisons, à titre d'exemple, la mesure *GMR* afin d'analyser les disparités de revenus italiennes. L'Italie est partitionnée en trois régions : Nord, Centre, Sud, où le Sud inclut aussi les îles telles que la Sicile. L'échantillon retenu est issu de la Banque d'Italie : Survey on Household's Income and Wealth et concerne l'année 2000. Il comprend 14254 individus et ses caractéristiques sont données dans le Tableau 2.

L'analyse est menée sur sept sources de revenu : (A) les salaires nets ; (B) les primes et autres avantages divers ; (C) pensions et arriérés ; (D) transferts (aide à la scolarité et pensions alimentaires) ; (E) revenus nets provenant de sa propre entreprise (entrepreneurs) ; (F) revenus immobiliers ; (G) revenus financiers. Les statistiques du tableau 2 sont annuelles. Le revenu annuel total de l'individu correspond à son revenu disponible net puisque les transferts sont comptabilisés.

Comme le montre le Tableau 3, la multi-décomposition fournit les contributions des sources à l'inégalité totale, les contributions des groupes, les contributions des inégalités entre les paires de groupes et les contributions des inégalités entre paires de groupes pour chaque source<sup>26</sup>.

<sup>25</sup> Pour la mesure de Gini, ce problème ouvert ne concerne que la partie intergroupe. La partie intragroupe peut en effet se réécrire comme une moyenne pondérée d'indicateurs de Gini pour chaque groupe et chaque source  $G_{jj}^f$ . D'après (11),  $G_w = \sum_{j=1}^k G_{jj} f_j s_j$ . D'après Fei et alii (1978) :  $G = \sum_{i=1}^q \rho^i \Omega^i G^i$ . Comme  $G_{jj}$  est construit de la même manière que  $G$ , on obtient  $G_{ij} = \sum_{i=1}^q \rho_i^f \Omega_j^f G_{ij}^f$ . Par conséquent :

$$G_w = \sum_{j=1}^k \sum_{i=1}^q f_j s_j \rho_j^f \Omega_j^f G_{ij}^f.$$

<sup>26</sup> cf. Annexe C pour le développement de la multi-décomposition de la mesure *GMR* et Mussard (2004b, 2004c) pour une application concernant la mesure de Gini.

Groupes $j \rightarrow$ Statistiques $\downarrow$	Nord	Centre	Sud-Iles	Total
Individus $n_j$	6540	3035	4679	14254
Moyenne (Euro) $m_j$	17325,65	15224,08	12643,93	15341,36
Ecart-type (Euro)	16052,85	14011,90	11159,05	14321,76
Moyenne source A : $m_j^A$	6643,38	5741,56	5043,19	5809,38
Moyenne source B : $m_j^B$	86,08	32,43	31,47	49,99
Moyenne source C : $m_j^C$	3823,35	3614,73	3198,97	3545,68
Moyenne source D : $m_j^D$	96,55	84,05	124,12	101,57
Moyenne source E : $m_j^E$	2552,93	2194,27	1888,27	2211,82
Moyenne source F : $m_j^F$	3407,48	3099,90	2150,06	2885,81
Moyenne source G : $m_j^G$	715,88	457,13	207,85	460,29

**Tableau 2** : Description des revenus italiens par région en 2000\*\*

\*\* Source : Banque d'Italie Survey on Household's Income and Wealth 2000

En 2000, les salaires expliquent 43,34% de l'inégalité totale (cf. le total de la colonne Source A). L'analyse de la structure des inégalités indique la tendance qu'elles ont à s'exercer entre les groupes. Les inégalités entre le Nord et le Sud (entre la région la plus industrialisée et la plus agricole) expliquent 30,96% de l'inégalité totale (cf. le total de la ligne Inégalités Nord / Sud). En combinant les deux plus fortes contributions marginales en facteurs et en sous-groupes, on vérifie que le couple Salaire / Nord-Sud possède la plus forte contribution à l'inégalité totale avec 14,03% (cf. Tableau 3). L'indépendance des lois marginales peut cependant conduire à la propriété selon laquelle la multi-décomposition ne procure pas forcément le même résultat que les croisements des décompositions marginales. Par exemple, considérons les inégalités de la région Centre qui s'élèvent à 4,38% de l'inégalité totale. Les transferts représentent le facteur le plus faible puisqu'il permet de diminuer les inégalités globales de l'ordre de 0,23%. Pourtant le croisement de ces deux contributions ne fournit pas le couple le plus faible, car les transferts de la région Centre expliquent -0,02% de l'inégalité totale alors que les transferts entre le Nord et le Sud permettent de diminuer les inégalités de 0,08%.

La technique de multi-décomposition évite dans certains cas les erreurs de politiques économiques qui consistent à interpréter uniquement les combinaisons des résultats marginaux donnés par la décomposition en sous-groupes et la décomposition en sources de revenu. La multi-décomposition détermine de manière plus claire et plus précise les couples source / groupe qui expliquent l'inégalité totale.

Sources $\ell \rightarrow$ Inégalités $\downarrow$	Source A	Source B	Source C	Source D	Source E	Source F	Source G	Total Groupes
Inégalités Nord	0,0393 [8,08]	0,0007 [0,14]	0,0120 [2,47]	-0,0002 [-0,04]	0,0172 [3,54]	0,0254 [5,23]	0,0050 [1,03]	0,0994 [20,45]
Inégalités Centre	0,0086 [1,77]	0,0001 [0,02]	0,0026 [0,53]	-0,0001 [-0,02]	0,0034 [0,7]	0,0059 [1,21]	0,0008 [0,16]	<b>0,0213</b> <b>[4,38]</b>
Inégalités Sud	0,0262 [5,39]	0,0002 [0,04]	0,0094 [1,93]	0 [0]	0,0087 [1,79]	0,0084 [1,73]	0,0009 [0,19]	0,0538 [11,07]
Inégalités Nord/Centre	0,0373 [7,67]	0,0005 [0,1]	0,0114 [2,35]	-0,0002 [-0,04]	0,0151 [3,11]	0,0242 [4,98]	0,0041 [0,84]	0,0924 [19,01]
Inégalités Nord/Sud	0,0682 [14,03]	0,001 [0,21]	0,0218 [4,48]	-0,0004 [-0,08]	0,024 [4,94]	0,0305 [6,27]	0,0054 [1,11]	0,1505 [30,96]
Inégalités Centre/Sud	0,0311 [6,4]	0,0002 [0,04]	0,0102 [2,1]	-0,0002 [-0,04]	0,0107 [2,2]	0,0147 [3,02]	0,002 [0,41]	0,0687 [14,13]
Total Sources	0,2107 [43,34]	0,0027 [0,56]	0,0674 [13,87]	<b>-0,0011</b> <b>[-0,23]</b>	0,0791 [16,27]	0,1091 [22,44]	0,0182 [3,74]	<b>GMR <math>\equiv</math></b> <b>0,4861</b>

**Tableau 3 : Multi-décomposition de GMR : Italie 2000\*\***

\*\* Source : Banque d'Italie « Survey on Household's Income and Wealth 2000 »

\*\*[.] : Contribution en pourcentage de l'inégalité totale

La technique permet enfin de réaliser en laboratoire des tests et des simulations de politiques économiques. Par exemple, même si globalement les transferts diminuent les inégalités globales (-0,23%), on s'aperçoit que les transferts entre les individus du Sud ne permettent pas de diminuer les inégalités (0% de l'inégalité totale). Ainsi, la multi-décomposabilité identifie et prévoit les combinaisons qui ne diminuent que trop faiblement ou qui tendent à accroître les inégalités totales.

## 7 Conclusion

Les mesures de l'entropie font office de référence dans le domaine des mesures d'inégalité décomposables. L'entropie unidimensionnelle généralisée  $S_c$  (cf. Shorrocks (1980) et Cowell (1980a, 1980b)) a longtemps été retenue car elle respecte le principe de décomposition additive. Certains auteurs, comme Dagum (1997a, 1997b), considèrent pourtant qu'elle ne conduit pas à la construction de bonnes mesures décomposables en raison des doutes concernant la propriété de décomposition cohérente. L'entropie généralisée unidimensionnelle, pourtant multi-décomposable par la valeur de Shapley, ne possède pas de composantes intergroupes  $I_b$  acceptables. Ces dernières sont, en effet, construites en supposant que les observations sont indépen-

dantes, que les groupes ont la même variance et sont normalement distribués (cf. Dagum (1997a)).

La classe des mesures de Gini  $g(x)$  est aussi une famille d'indices multi-décomposables qui lient la décomposition en sous-groupes et la décomposition en sources de revenu. Même si le revenu est uniquement déterminé par les dimensions que représentent les sources de rémunération, les mesures de Gini intergroupes (brutes)  $g(x)_{g,b}$  peuvent être considérées comme plus satisfaisantes puisqu'elles prennent en compte les phénomènes de variance et d'asymétrie entre les groupes<sup>27</sup>.

Les résultats de multi-décomposition intermédiaires suggérés par la classe de mesures  $g(x)$  permettent de manière générale de tester l'effet des politiques socio-économiques de redistribution en mesurant l'impact des transferts (et d'autres facteurs) aux inégalités intragroupes et intergroupes. Il devient possible de révéler, par exemple, les contresens des croisements de décompositions marginales en sous-groupes et en sources de revenu. La multi-décomposabilité offre par ailleurs des perspectives de recherche intéressantes. Il en est ainsi de l'étude de l'inférence statistique de toutes les composantes « source  $\ell$  / groupe  $j$  » et « source  $\ell$  / entre les groupes  $j$  et  $h$  » pour l'analyse de la significativité des variations des contributions d'inégalités au cours du temps et de la recherche d'une forme analytique qui permettrait l'écriture de l'inégalité comme une moyenne pondérée des inégalités pour chaque groupe et chaque source  $I'_j$  et des inégalités entre chaque paire de groupe pour chaque source  $I'_{jh}$ .

## Références

- Auvray C. et A.Trannoy (1992). Décomposition par source de l'inégalité des revenus à l'aide de la Valeur Shapley, *Journées de Microéconomie Appliquée*, Sfax, Tunisie.
- Blackorby C., W.Bossert, et D.Donaldson (1999). Income Inequality Measurement : The Normative Approach, in : Silber J. (éds.), *Handbook of Income Inequality Measurement*, Kluwer Academic Publishers, pp. 133-161.
- Bossert W. et A.Pfingsten (1990). Intermediate Inequality : Concepts, Indices and Welfare Implication, *Mathematical Social Science*, vol. 19, pp. 117-134.
- Bourguignon F. (1979). Decomposable Inequality Measures, *Econometrica*, vol. 47, pp. 901-920.

<sup>27</sup> Il existe aussi des possibilités (cf. Mussard (2006)) d'intégrer plusieurs dimensions dans l'analyse. La condition nécessaire étant que le revenu soit régressé sur plusieurs dimensions (éducation, santé, etc.) selon un modèle linéaire.

- Chakravarty S.R. (1999). Measuring Inequality : The Axiomatic Approach, in Silber J. (éds.), *Handbook of Income Inequality Measurement*, Kluwer Academic Publishers, pp. 163-186.
- Chameni C. (2006). A Note on the Decomposition of the Coefficient of Variation Squared: Comparing Entropy and Dagum's Methods, *Economics Bulletin*, vol. 4(8), pp. 1-8.
- Chantreuil F. et A.Trannoy (1999). Inequality Decomposition Values : The Trade-off Between Marginality and Consistency, *Document de travail n°9924*, THEMA.
- Csiszár I. (1967). Information Type Measures of Difference of Probability Distributions and Indirect Observations, *Studia Scientiarum Mathematicarum Hungarica*, vol. 2, pp. 299-318.
- Cowell F.A. (1980a). Generalized entropy and the Measurement of Distributional Change, *European Economic Review*, vol. 13, pp. 147-159.
- Cowell F.A. (1980b). On the Structure of Additive Inequality Measures, *Review of Economics Studies*, vol. 47, pp. 521-531.
- Cowell F.A. (1988). Inequality Decomposition : Three Bad Measures, *Bulletin of Economic Research*, vol. 40(4), pp. 309-312.
- Cressie N. et T.R.C.Read (1984). Multinomial Goodness-of-fit Tests, *Journal of the Royal Statistical Society, Séries B*, vol. 46, pp. 440-464.
- Dagum C. (1959). Transvariazione fra più di due distribuzioni, in Gini C. (éds.), *Memorie di metodologia statistica*, vol. II, Libreria Goliardica : Roma.
- Dagum C. (1960). Teoria de la transvariacion, sus aplicaciones a la economia, *Metron*, vol. XX, pp. 1-206.
- Dagum C. (1961). Transvariacion en la hipotesis de variables aleatorias normales multidimensionales, *Proceedings of the International Statistical Institute*, vol. 38(4), pp. 473-486.
- Dagum C. (1987). Measuring the Economic Affluence Between Populations of Income Receivers, *Journal of Business and Economic Statistics*, vol. 5(1), pp. 5-12.
- Dagum C. (1997a). A New Approach to the Decomposition of the Gini Income Inequality Ratio, *Empirical Economics*, vol. 22(4), pp. 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*, pp. 200-205.
- Dagum C. (1998). Fondements de bien-être social et décomposition des mesures d'inégalité dans la répartition du revenu, *Economie Appliquée*, vol. L1, n°4, pp. 151-202.
- Dalton H. (1920). The Measurement of Inequality of Incomes, *Economic Journal*, vol. 30, pp. 348-361.

- Ebert U. (1988). On the Decomposition of Inequality : Partitions into Non-overlapping Subgroups, in Eichorn W. (éds.), *Measurement In Economics*, New-York : Physica Verlag, pp. 399-412.
- Ebert U. (1999). Dual Decomposable Inequality Measures, *Canadian Journal of Economics*, vol. 32, n°1, pp. 234-46.
- Fei J. C. H., G.Ranis et S.W.Y.Kuo (1978). Growth and the Family Distribution of Income by Factor Components, *Quarterly Journal of Economics*, vol. 92, pp. 17-53.
- Fields G. (1979). Income Inequality in Urban Columbia : A decomposition Analysis, *Review of Income and Wealth*, vol. 25, n°3, pp. 327-341.
- Gini C. (1912). *Variabilità e mutabilità*, Memori di Metodologia Statistica I, Variabilità e Concentrazione, Libreria Eredi Virgilio Veschi : Rome, pp. 211-382.
- Gini C. (1916). Il concetto di transvariazione e le sue prime applicazioni, *Giornale degli Economisti e Rivista di Statistica*, in Gini C. (éds.) 1959, pp. 21-44.
- Isreali O. (2007). A Shapley-based Decomposition of the R-Square of a Linear Regression, *Journal of Economic Inequality*, vol. 5(2), pp. 199-212.
- Jenkins S. P. (1995). Accounting for Inequality Trends : Decomposition Analyses for the UK, 1971-1986, *Economica*, vol. 62, pp. 29-63.
- Kolm S-C. (1976a). Unequal Inequalities I, *Journal of Economic Theory*, vol. 12, pp. 416-442.
- Kolm S-C. (1976b). Unequal Inequalities II, *Journal of Economic Theory*, vol. 13, pp. 82-111.
- Kolm S-C. (1999). Rational Foundations of Income Inequality Measurement, in Silber J. (éds.), *Handbook of Income Inequality Measurement*, Kluwer Academic Publishers, pp. 19-94.
- Lerman R. et S.Yitzhaki (1985). Income Inequalities Effects by Income Source : A New Approach and Applications to United States, *Review of Economics and Statistics*, vol. 67, pp. 151-156.
- Lerman R. et S.Yitzhaki (1991). Income Stratification and Income Inequality, *Review of Income and Wealth*, vol. 37, n°3, pp. 313-329.
- Mookherjee D. et A.Shorrocks (1982). A Decomposition Analysis of the Trend in UK Income Inequality, *Economic Journal*, vol. 92, pp. 886-902.
- Mussard S. (2004a). *Décompositions multidimensionnelles du rapport moyen de Gini. Applications aux revenus italiens de 1989 et 2000*, Thèse, Université de Montpellier I.
- Mussard S. (2004b). The bidimensional Decomposition of the Gini Ratio. A Case Study: Italy, *Applied Economics Letters*, vol. 11, pp. 503-505.
- Mussard S. (2006). Une nouvelle décomposition de la mesure de Gini en sources de revenu, et la décomposition en sous-populations : une réconciliation, *Annales d'Economie et de Statistique*, vol. 81, pp. 1-25.

- Mussard S. et N.Peypoch (2006). On Multi-decomposition of the Aggregate Malmquist Productivity Index, *Economics Letters*, vol. 91, pp. 436-443.
- Pyatt G. (1976). On the Interpretation and Disaggregation of Gini Coefficients, *Economic Journal*, vol. 86, pp. 243-25.
- Pyatt G., Chen C.-N. et J.Fei (1980). The Distribution of Income by Factor Component, *Quarterly Journal of Economics*, vol. 95, n°3, pp. 451-473.
- Rao V.M. (1969). Two Decompositions of Concentration Ratio, *Journal of the Royal Statistical Society, Série A 132*, pp. 418-425.
- Sastre M. et A.Trannoy (2002). Shapley Inequality Decomposition by Factor Components: Some Methodological Issues, in Moyes P. Seidl C. et A.F.Shorrocks (éds.), *Journal of Economics*, vol. 9, supplément, pp. 51-90.
- Sen A. K. (1973). *On Economic Inequality*, Clarendon Press : Oxford.
- Shapley L. (1953). A Value for n-Person Games, in Kuhn H. W. et A.W.Tucker (éds.), *Contributions to the Theory of Games*, vol. 2, Princeton University Press.
- Shorrocks A. F. (1980). The Class of Additively Decomposable Inequality Measures, *Econometrica*, vol. 48, pp. 613-625.
- Shorrocks A. F. (1982). Inequality Decomposition by Factor Component, *Econometrica*, vol. 50, pp. 193-211.
- Shorrocks A. F. (1984). Inequality Decomposition by Factor Components and by Population Subgroups, *Econometrica*, vol. 53, pp. 1369-1386.
- Shorrocks A. F. (1988). Aggregation Issues in Inequality Measurement, in Heichhorn W. (éds.), *Measurement in Economics*, New York : Physica-Verlag, pp. 429-452.
- Shorrocks A. F. (1999). Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value, *Mimeo*, Université d'Essex.
- Silber J. (1989). Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality, *Review of Economics and Statistics*, vol. 71, pp. 107-115.
- Theil H. (1967). *Economics and Information Theory*, North-Holland Publishing Company : Amsterdam.
- Tsui K. (1999). Multidimensional Inequality and Multidimensional Generalized Entropy Measures : An axiomatic Derivation, *Social Choice and Welfare*, vol. 16, pp. 145-157.

## Annexe A : L'entropie généralisée n'est pas multi-décomposable

D'après la propriété DA :

$I_w = \sum_{j=1}^k \omega_j(m_j, n_j) I_j(\tilde{x}_j)$ . Si  $\omega_j(m_j, n_j) = f_j \left(\frac{m_j}{m}\right)^c$ , l'entropie généralisée intragroupe s'écrit :

$$S_{cw} = \sum_{j=1}^k f_j \left(\frac{m_j}{m}\right)^c S_c(\tilde{x}_j).$$

Toujours d'après DA :  $I_b = I(\underbrace{m_1, \dots, m_1}_{n_1 \text{ fois}}, \dots, \underbrace{m_k, \dots, m_k}_{n_k \text{ fois}})$ . En substituant dans  $S_c$  les revenus des individus  $x_i$  par la moyenne de leur groupe correspondant, on a :

$$S_{cb} = \sum_{j=1}^k f_j \frac{1}{c(c-1)} \left( \left(\frac{m_j}{m}\right)^c - 1 \right).$$

Etant donné que les sources de revenus sont liées par une relation additive  $x_i = \sum_{\ell=1}^q x_i^\ell$ , les moyennes des sources le sont aussi. La moyenne des revenus du groupe  $P_j$  se décompose en effet en la somme des moyennes de chaque facteur :  $m_j = \sum_{\ell=1}^q m_j^\ell$ . Par conséquent, l'inégalité intragroupe devient :

$$S_{cw} = \sum_{j=1}^k f_j \left( \frac{\sum_{\ell=1}^q m_j^\ell}{m} \right)^c S_c(\tilde{x}_j).$$

Le terme élevé à la puissance  $c$  ne permet pas de mesurer la contribution d'une source particulière au montant des inégalités intragroupes. La même remarque peut être faite concernant la mesure intergroupe :

$$S_{cb} = \sum_{j=1}^k f_j \frac{1}{c(c-1)} \left( \left( \frac{\sum_{\ell=1}^q m_j^\ell}{m} \right)^c - 1 \right).$$

L'entropie unidimensionnelle n'est donc pas multi-décomposable sans recourir à l'application de l'algorithme de Shapley sur chaque composante  $S_{cw}$  et  $S_{cb}$ .

## Annexe B : La multi-décomposition d'un cas particulier de l'entropie multidimensionnelle

Soit  $\Pi$  l'ensemble des matrices de taille  $n \times q$ . Supposons que l'on observe  $q$  caractéristiques (en colonnes) correspondant à  $n$  individus (en lignes) : par exemple le niveau d'éducation, de santé et de revenu de  $n$  individus. Pour  $X \in \Pi$ , une mesure d'inégalité multidimensionnelle est définie par

$I^n : \Pi \rightarrow \mathbb{R}, X \mapsto I^n(X), n \geq 2$ . Cette définition, dépassant le problème unidimensionnel de la mesure des inégalités, a permis à Tsui (1999) de proposer une mesure d'entropie généralisée multidimensionnelle. Un cas particulier de celle-ci, dont la forme analytique ressemble à celle de  $S_0$ , est donné par :

$$I^n(X) = \frac{1}{n} \sum_{i=1}^n \sum_{\ell=1}^q \delta^\ell \log \left( \frac{x_i^\ell}{m^\ell} \right) \cdot \delta^\ell > 0,$$

où  $I^n(X)$  est obtenue à une transformation quelconque près. On peut trouver dans Tsui (1999, p. 154), des informations complémentaires et en particulier l'obtention des paramètres  $\delta^\ell$ . A la différence de l'entropie unidimensionnelle,  $I^n$  est définie sur un ensemble de matrices dont les colonnes peuvent être caractérisées aussi bien par des dimensions que par des sources de revenu. Ainsi, si  $I^n$  était multi-décomposable, il serait possible de mesurer l'influence et la contribution de chaque source sur la valeur des inégalités intragroupes et celles des inégalités intergroupes afin de déterminer les valeurs manquantes du Tableau 1. On a dans ce cas :

$$I^n(X) = \sum_{j=1}^k f_j \frac{1}{n_j} \sum_{i=1}^{n_j} \sum_{\ell=1}^q \delta^\ell \log \left( \frac{x_{ij}^\ell}{m_j^\ell} \right) + \sum_{j=1}^k f_j \sum_{\ell=1}^q \delta^\ell \log \left( \frac{m_j^\ell}{m^\ell} \right).$$

Contrairement à l'entropie généralisée unidimensionnelle, ce cas particulier de l'entropie multidimensionnelle de Tsui (1999) est multi-décomposable. Il n'échappe cependant pas aux critiques concernant la spécification de la mesure intergroupe. Elle repose en effet sur des inégalités qui concernent les moyennes de chaque groupe.

## Annexe C : La multi-décomposition de la mesure *GMR*

Soit  $GMR_{jj}$  et  $GMR_{jh}$  les indices mesurant les inégalités au sein du groupe  $j$  et entre les groupes  $j$  et  $h$  :

$$GMR_j = \sum_{i=1}^{n_j} \sum_{r=1}^{n_j} \frac{|x_{ij} - x_{rj}|}{n_j^2 \max\{x_{ij}, x_{rj}\}}, \quad GMR_{jh} = \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} \frac{|x_{ij} - x_{rj}|}{n_j n_h \max\{x_{ij}, x_{rj}\}}.$$

Soit  $z_j = \frac{n_j^2}{n}$  et  $z_{jh} = \frac{n_j n_h}{n^2}$  les pondérations associées à  $GMR_{jj}$  et  $GMR_{jh}$ . La décomposition en sous-groupes s'écrit :

$$GMR(x) = \sum_{j=1}^k z_j GMR_{jj} + \sum_{j=2}^k \sum_{h=1}^{j-1} z_{jh} GMR_{jh}.$$

Soit l'opérateur  $x_{j,ir}^{\ell}$  donnant la  $\ell$ -ième source de rémunération issue du minimum entre les revenus  $x_{ij}$  et  $x_{rj}$ . Soit l'opérateur  $x_{jh,ir}^{\ell}$  donnant la  $\ell$ -ième source de rémunération issue du minimum entre les revenus  $x_{ij}$  et  $x_{rj}$ . En mesurant la contribution de chaque source aux indices  $GMR_{jj}$  et  $GMR_{jh}$ , on obtient la multi-décomposition de  $GMR$  :

$$GMR(x) = \underbrace{\sum_{\ell=1}^q \sum_{j=1}^k \sum_{i=1}^{n_j} \sum_{r=1}^{n_j} z_j \frac{x_{ij}^{\ell} + x_{rj}^{\ell} - 2x_{j,ir}^{\ell}}{n_j^2 \max\{x_{ij}, x_{rj}\}}}_{GMR_w} + 2 \underbrace{\sum_{\ell=1}^q \sum_{j=2}^k \sum_{h=1}^{j-1} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} z_{jh} \frac{x_{ij}^{\ell} + x_{rh}^{\ell} - 2x_{jh,ir}^{\ell}}{n_j n_h \max\{x_{ij}, x_{rh}\}}}_{GMR_{gb}}$$

où  $GMR_w$  est la mesure d'inégalité intragroupe et  $GMR_{gb}$  la mesure d'inégalité intergroupe (brute). En mesurant la contribution  $C_j^{\ell}$  de la source  $\ell$  aux inégalités du groupe  $j$  et la contribution  $C_{jh}^{\ell}$  de la source  $\ell$  aux inégalités entre les groupes  $j$  et  $h$ , les pourcentages du Tableau 3 concernant l'Italie peuvent donc être identifiés par :  $100 \frac{C_j^{\ell}}{GMR}$  et  $100 \frac{C_{jh}^{\ell}}{GMR}$ .