
Région et Développement

n° 47-2018

www.regionetdeveloppement.org

Les inégalités d'éducation : une nouvelle base mondiale (1950-2015)

Wail BENAABDELAALI* et Maurice CATIN**

Résumé - Une nouvelle base mondiale sur les inégalités d'éducation est proposée ici, s'inspirant de la base élaborée par Benaabdelaali et al. (2012) et s'appuyant sur la version 2016 de la base de Barro et Lee (2013). Elle couvre la période 1950-2015 pour un ensemble de 146 pays. Elle donne une estimation des inégalités d'éducation de la population par groupe d'âge de 5 ans et par tranche d'âges agrégés, en considérant 7 niveaux d'éducation. Elle présente, outre l'indice de Gini, plusieurs indices de la famille d'entropie généralisée et de la famille d'Atkinson. Elle offre des données plus appropriées en corrigeant l'erreur de pondération utilisée par Barro et Lee dans le calcul de la moyenne d'années de scolarisation, qui peut amener un biais important pour de nombreux pays. A partir de cette base de données un panorama de la réduction des inégalités d'éducation au niveau mondial est schématiquement présenté. Dans les pays en développement, la réduction des inégalités d'éducation est essentiellement liée à la baisse de la population n'ayant aucune scolarisation. Néanmoins, l'écart entre les pays à faible niveau de développement humain dans certaines régions du monde et ceux à développement humain élevé reste important.

Classification JEL

D63, I21, J24, O15

Mots-clés

Inégalités d'éducation
Indice de Theil
Indice de Gini
Base de données mondiale

La base mondiale des inégalités d'éducation 1950-2015 présentée ici est disponible au lien :
<http://lead.univ-tln.fr/bdd.php>

* Instance Nationale de l'Evaluation, Conseil Supérieur de l'Education, de la Formation et de la Recherche Scientifique, Rabat, Maroc ; wbenabdelaali@csefrs.ma

** LEAD, Université de Toulon ; maurice.catin@univ-tln.fr

1. INTRODUCTION

Alors qu'elles ont été largement étudiées dans le champ de la sociologie, les inégalités éducatives ont, pendant longtemps, été peu explorées dans le champ de l'économie de l'éducation. Au départ, comme le souligne Dessus (2001), la principale raison a été le manque de données comparables au niveau international : "The measure of dispersion of qualifications is more delicate to build. Ideally the good measure would be a Gini index of education, for which the years of education of each individual would replace the traditional measure of income. To our best knowledge, this measure is unfortunately unavailable for international comparisons".

La construction d'une base mondiale sur les inégalités éducatives requiert l'existence de données homogènes sur la répartition de l'éducation pour l'ensemble des pays considérés. En ce sens, le niveau de scolarité maximum atteint semble être la seule information statistique qui paraît disponible à l'échelle mondiale. Johnes (2006) rapporte "The distribution of education is not a straightforward thing to measure. It requires information not only about the proportions of the current cohort of young people who leave education after completing various levels, but also similar information for every generation still alive. For many countries, this extent of detail in data which are quasi historic in nature has become available only very recently". De même, Meschi et Scervini (2014) notent que "the effects of the distribution of educational attainment have not been so widely studied... empirical evidence on this topic remains scarce, possibly owing to the lack of data comparing educational inequality across countries".

La première base sur les inégalités d'éducation a été développée par Lopez et al. (1998) qui ont calculé un indice de Gini de l'éducation pour un échantillon de 12 pays¹ entre 1970 et 1995. Par la suite, les auteurs l'ont étendu pour 20 pays. En utilisant aussi l'indice de Gini, Thomas et al. (2001) ont été les premiers à construire une base sur les inégalités d'éducation pour un panel de 85 pays pour la période 1960-1990. Thomas et al. (2003) ont par la suite élargi leur base jusqu'à 140 pays de 1960 à 2000 en rajoutant l'indice de Theil. De même, en utilisant l'indice de Gini de l'éducation, Castelló et Doménech (2002), Checchi (2004) et Lim et Tang (2008) ont construit une base de données comprenant respectivement 108 pays de 1960 à 2000 (935 observations), 117 pays de 1960 à 1995 (848 observations) et 99 pays de 1960 à 2000 (891 observations).

Si les inégalités éducatives sont presque toutes mesurées à partir de la répartition de la population selon le niveau d'éducation maximum atteint, les sources mobilisées pour la construction de ces bases diffèrent. Ainsi, Lopez et al. (1998, 1999), Thomas et al. (2001) et Checchi (2004) se sont appuyés sur la base de Barro et Lee (1996). Castelló et Doménech (2002), Thomas et al. (2003) et Lim et Tang (2008) ont utilisé celle de Barro et Lee (2001). Araujo et al. (2004) et Meschi et Scervini (2014) se sont appuyés sur des bases d'enquêtes ménages. D'autres auteurs comme Castelló (2010a), Benaabdelaali et al. (2012), Morrisson et Murin (2013), Cuaresma et al. (2013), Földvári et Van Leeuwen (2014), Sauer et Zagler (2014) ont utilisé des bases plus récentes comme celles de Cohen et Soto (2007), Lutz et al. (2007) ou Barro et Lee (2013).

Plusieurs auteurs ont essayé de compiler des bases de données présentant la répartition de la population selon le niveau d'éducation maximum atteint. L'une des premières tentatives a été faite par Psacharopoulos et Arriagada (1986) qui, à partir de plusieurs recensements nationaux, ont rassemblé des données sur la composition de la population âgée de plus de 25 ans selon six niveaux d'éducation. Depuis,

¹ Brésil, Chili, Chine, Colombie, Inde, Corée, Mexique, Malaisie, Pérou, Philippines, Thaïlande et Venezuela.

de nombreux progrès ont été accomplis. Ainsi, la base de Barro et Lee (2013) présente un ensemble de données relatives à 146 pays par groupe d'âge et par genre de 1950 à 2010 sur des intervalles de cinq ans et qui donne à cet égard le pourcentage de la population – par tranche d'âge – ayant atteint, de façon achevée ou non, chacun des trois cycles d'enseignement (primaire, secondaire et supérieur). Barro et Lee ont utilisé exclusivement des données de recensements ventilées par groupe d'âge, ainsi que de nouvelles estimations des taux de mortalité et des taux d'achèvement différenciés selon l'âge et le niveau d'éducation, améliorant les données présentées dans les versions antérieures (Barro et Lee 1993, 1996, et 2001).

Il faut noter que la construction de bases internationales sur les inégalités éducatives a permis depuis l'émergence de nombreux travaux. Une littérature s'est développée pour explorer d'éventuelles relations entre les inégalités éducatives et diverses variables d'intérêt, notamment l'expansion de la scolarisation (Lim et Tang, 2008; Morrisson et Murtin, 2010, 2013; Meschi et Scervini, 2014), les inégalités de revenu (Checchi, 2004; Földvári et Van Leeuwen, 2014), la croissance économique (Castelló et Doménech, 2002; Castelló, 2010), la santé (Galea et al., 2007; Lê et al., 2010; Hori, 2011), la corruption (Patrawart, 2010), la cohésion sociale (Green et al., 2006; Vert, 2011), la migration (McKenzie et Rapoport, 2007), la démocratie (Castelló, 2008), la volatilité macroéconomique (Checchi et García-Peñalosa, 2004), voire même la rébellion populaire (Bartusevičius, 2014)...

Nous proposons dans ce qui suit une nouvelle base mondiale sur les inégalités éducatives. Elle couvre 146 pays observés tout au long de 65 années (1950-2015), et propose une estimation de l'inégalité d'éducation par tranches d'âge de 5 ans et pour les tranches d'âge agrégées. Nous partons de la base élaborée par Benaabdelali et al. (2012) qui présentait l'indice de Gini de l'éducation pour la période 1950-2010. A la différence, la nouvelle base de données construite ici présente, sur la période 1950-2015, outre l'indice de Gini, plusieurs indices de la famille d'entropie généralisée et de la famille d'Atkinson. De plus, la moyenne d'années de scolarisation est recalculée en corrigeant l'erreur de pondération de Barro et Lee (2013) qui nous est apparue conduire à un biais pour un grand nombre de pays (section 2). De là, un panorama des inégalités au niveau mondial et de leur évolution est schématiquement présenté à partir de cette base en section 3.

2. PROPOSITION D'UNE NOUVELLE BASE MONDIALE SUR LES INÉGALITÉS ÉDUCATIVES

Pour construire notre base mondiale sur les inégalités d'éducation, nous nous sommes appuyés sur la dernière base de Barro et Lee (2013) (version 2.1 Feb, 2016) qui, comme on l'a dit précédemment, offre des estimations plus précises et actualisées sur la répartition de l'éducation selon le niveau maximum atteint pour la période 1950-2010 et présente des avantages spécifiques par rapport aux autres bases mondiales. Pour l'année 2015, nous avons reconstitué nous-même la répartition de l'éducation selon le niveau maximum atteint par groupes d'âge quinquennaux en mobilisant la ventilation de la population par tranches d'âge des différents pays de la base à partir de la révision de 2017 des estimations de population des Nations Unies (World Population Prospect) selon la démarche adoptée par Barro et Lee (2013) (voir annexe A de Barro et Lee, 2013).

Outre les années retenues jusqu'à 2015, la base mondiale sur les inégalités d'éducation que nous proposons apporte une amélioration par rapport aux bases existantes, et ce sur plusieurs plans.

- Nous avons considéré sept niveaux d'éducation. Les bases de Castelló et Doménech (2002), Checchi (2004), Castelló (2010b), Cuaresma et al. (2013) et Van Leeuwen et Földvári (2014) ont toutes été construites en s'appuyant uniquement sur

quatre niveaux d'éducation. Ces auteurs ne font pas de distinction entre ceux qui ont effectivement achevé un cycle d'enseignement et ceux qui ne l'ont pas achevé puisqu'ils utilisent des données agrégées par cycle. Dans ce cas, l'inégalité estimée concerne uniquement l'inégalité inter-cycles alors que l'inégalité intra-cycles est de ce fait ignorée.

- Comme nous exploitons des données ventilées par tranches d'âge, nous prenons en compte la variation dans les durées des cycles d'enseignement. Il faut souligner qu'à l'exception de Benaabdelaali et al. (2012), l'ensemble des bases mondiales élaborées jusqu'à présent pour quantifier les inégalités éducatives ont toutes utilisé des durées des cycles d'enseignement fixes dans le temps. Tenir compte de ces changements nécessite au préalable de disposer de données désagrégées par tranches d'âge. La population âgée de 15 ans et plus est constituée d'individus, appartenant à différentes classes d'âge, n'ayant pas forcément passé au cours du temps par le même système d'éducation. En effet, durant la période 1950-2015, des restructurations du système éducatif ont été entreprises dans la plupart des pays impliquant des changements dans la durée totale de la scolarité aux cycles primaire, collégial et secondaire. Nous avons recensé 99 pays sur les 146 de la base de Barro et Lee (2013) qui ont connu des changements dans ces durées, dont 31 pays ont eu plus de trois changements. Bien entendu, ne pas tenir compte de ces changements risque de modifier significativement les estimations des indices d'inégalité. Ceci impacte aussi bien l'étendue que la répartition du nombre d'années d'études effectuées par les individus de ces différentes tranches d'âge.

- Finalement, la base de données que nous proposons est la première à offrir plusieurs indices d'inégalité de l'éducation ventilés en tranches d'âge de 5 ans, par genre, pour 146 pays, pour des groupes agrégés de pays et/ou d'intervalles d'âge, tout en tenant compte de la variation et de l'hétérogénéité au sein des groupes agrégés. Elle présente aussi pour les différents pays deux types de décomposition de quatre indices d'inégalité de l'éducation de 1950 à 2015 (Gini, indices d'entropie généralisée GE(0.5), de Theil GE(1) et d'Atkinson A(1)). Une première décomposition concerne deux groupes (les illettrés et les non illettrés) et se décompose en une inégalité intra-éduqués et une inégalité inter-éduqués-illettrés ; une seconde concerne 13 tranches d'âge, [15-19] à [75 et plus] et se décompose en une inégalité intragénérationnelle et une inégalité intergénérationnelle. Cette base est accessible en ligne sur le site du LEAD (<http://lead.univ-tln.fr/bdd.php>).

Nous présentons dans ce qui suit la méthodologie adoptée pour construire les différents indices d'inégalité en se basant sur la répartition de la population selon le niveau d'éducation maximum atteint.

2.1. L'erreur de pondération dans la base de Barro et Lee (2013)

Etant donné que les durées des cycles d'enseignement ne sont pas explicitement disponibles dans la base de Barro et Lee (2013), nous les avons calculées en utilisant la démarche adoptée par Benaabdelaali et al. (2012) (voir section 2.2). Elles sont appelées de manière générale :

$p_{c,i,t}^a$ la proportion de la population de la tranche d'âge "a" du pays "c" ayant atteint le niveau d'éducation "i" à la date "t" ;

$y_{c,i,t}^a$ est le nombre d'années de scolarisation cumulées de la tranche d'âge "a" du pays "c" pour atteindre le niveau d'éducation "j" à la date "t".

Afin de calculer un indice d'inégalité de l'éducation pour des grands groupes d'âge (15 ans et plus, 25 ans et plus, 15 à 65 ans, ...), la plupart des auteurs, même ceux qui ont utilisé la base de Barro et Lee (2013) comme Sauer et Zagler (2014), ne prennent pas en considération la variation par tranches d'âge des durées des cycles d'enseignement. Par exemple pour calculer l'indice de Theil de l'éducation pour la

population âgée de 15 ans et plus, ils supposent des « $y_{c,j,t}^a$ » invariants à la fois par tranche d'âge et dans le temps ($y_{c,j,t}^{15+} = y_{c,j}^a = y_{c,j} \forall a \in [1 - 13], \forall t \in [1950 - 2010]$).

Toutefois, même s'il s'agit du même pays "c" à la même date "t", les durées des cycles d'enseignement peuvent être différentes d'une tranche d'âge à l'autre. De manière générale, sur les 146 pays considérés, 99 pays ont procédé à des changements dans les durées des cycles d'enseignement.

Une première approche serait alors de calculer une moyenne des « $y_{c,j,t}^a$ » de manière à tenir compte de ces éventuelles variations par rapport aux différents groupes d'âge. Ainsi, comme $p_{c,j,t}^{15+} = \sum_{a=1}^A l_{c,t}^a p_{c,j,t}^a$, on pourrait a priori penser que $y_{c,j,t}^{15+}$ serait de même égale à $\sum_{a=1}^A l_{c,t}^a y_{c,j,t}^a$.

Or on peut montrer que Barro et Lee (2013) n'ont pas utilisé la pondération adéquate dans le calcul des durées $y_{c,j,t}^{15+}$. Après avoir déduit les durées $y_{c,j,t}^a$ à partir de leur base de données², il apparaît bien que nos moyennes d'années de scolarisation par tranches d'âge quinquennales ($MAS_{c,t}^a$) correspondent à celles calculées par Barro et Lee (2013). Toutefois, lorsqu'il s'agit de calculer la moyenne d'années de scolarisation de la tranche d'âge agrégée de 15 ans et plus, en appliquant la formule utilisée dans leur article $MAS_{c,t}^{15+} = \sum_{a=1}^A l_{c,t}^a MAS_{c,t}^a$ (la moyenne des moyennes d'années de scolarisation par tranches d'âge quinquennales) ainsi que leurs propres données (répartition de la population selon les niveaux d'éducation) nous ne retrouvons pas les mêmes résultats.

Il s'avère que, pour la population âgée de 15 ans et plus, les auteurs ont utilisé une autre formule tout aussi valable qui est égale à la somme des produits entre les durées et les proportions $MAS_{c,t}^{15+} = \sum_{j=1}^n p_{c,j,t}^{15+} y_{c,j,t}^{15+}$. Auquel cas, le problème proviendrait des durées agrégées $y_{c,j,t}^{15+}$. Il semble que les auteurs se soient appuyés sur une pondération qui correspondrait à celle utilisée pour les $p_{c,j,t}^{15+}$ ($p_{c,j,t}^{15+} = \sum_{a=1}^A l_{c,t}^a p_{c,j,t}^a$). Afin de le vérifier, nous avons recalculé les durées agrégées en utilisant cette dernière pondération tel que $\tilde{y}_{c,j,t}^{15+} = \sum_{a=1}^A l_{c,t}^a y_{c,j,t}^a$. Nous retrouvons alors bien les mêmes valeurs des moyennes d'années de scolarisation des 15 ans et plus que celles des auteurs³.

En partant des formules de la moyenne d'années de scolarisation de la tranche d'âge "a" ($MAS_{c,t}^a = \sum_{j=1}^n p_{c,j,t}^a y_{c,j,t}^a$) et de celle de la population âgée de 15 ans et plus ($MAS_{c,t}^{15+} = \sum_{a=1}^A l_{c,t}^a MAS_{c,t}^a = \sum_{j=1}^n p_{c,j,t}^{15+} y_{c,j,t}^{15+}$), on peut démontrer que l'agrégation des « $y_{c,j,t}^a$ » pour l'ensemble de la population âgée de 15 ans et plus doit s'écrire comme suit :

$$y_{c,j,t}^{15+} = \sum_{a=1}^A \left[\frac{l_{c,t}^a p_{c,j,t}^a}{\sum_{b=1}^A l_{c,t}^b p_{c,j,t}^b} \right] y_{c,j,t}^a$$

Certes, l'erreur de pondération faite par Barro et Lee (2013) peut, pour certains pays, ne pas avoir d'impact si les $y_{c,j,t}^a$ restent les mêmes pour l'ensemble des tranches d'âge. Toutefois comme les systèmes éducatifs de la majorité des pays ont subi des aménagements quant aux durées des cycles d'enseignement depuis 65 ans,

² La démarche de calcul des $y_{c,j,t}^a$ est présentée dans la section 2.2.

³ On a refait les calculs avec les versions précédentes de la base de Barro et Lee {1.0, 1.1, 1.2, 1.3, 2.0 et 2.1} et on obtient les mêmes résultats.

cette erreur peut changer significativement les résultats et ceci peut être illustré à travers l'exemple qui suit.

Tableau 1. Répartition de la population par tranche d'âge selon le niveau d'éducation maximum atteint pour le cas de la Bulgarie en 1990

Groupe d'âge	Pop.	Poids du groupe d'âge (l^a)	Proportion de la population selon le niveau d'éducation maximum atteint (en %) ($p_{c,j,t}^a$) {c=120,t=1990}						
			Non Scol. (p_1^a)	Prim. inachevé (p_2^a)	Prim. achevé (p_3^a)	Sec. inachevé (p_4^a)	Sec. achevé (p_5^a)	Sup. inachevé (p_6^a)	Sup. achevé (p_7^a)
15-19	641	9.24	1.90	6.40	63.50	22.76	5.24	0.09	0.01
20-24	585	8.43	1.40	2.80	17.30	14.97	57.13	4.26	2.04
25-29	580	8.36	1.40	2.70	17.40	12.35	46.35	6.33	13.37
30-34	606	8.74	1.40	2.70	17.40	13.59	45.11	6.33	13.37
35-39	625	9.01	1.40	3.40	25.70	14.34	34.06	6.72	14.18
40-44	638	9.20	1.40	3.40	25.70	16.46	31.94	6.72	14.18
45-49	527	7.60	2.70	6.80	37.90	16.70	19.00	5.37	11.33
50-54	513	7.40	2.70	6.80	37.90	18.73	16.97	5.37	11.33
55-59	560	8.07	5.60	16.30	42.80	14.26	9.64	3.63	7.67
60-64	529	7.63	5.60	16.30	42.80	14.86	9.04	2.75	8.55
65-69	474	6.83	12.00	32.30	36.70	8.55	3.75	1.61	4.99
70-74	240	3.46	12.00	32.30	36.70	8.95	3.35	1.09	5.51
75 et +	418	6.03	12.00	32.30	36.70	9.45	2.85	1.25	5.35
15 et +	6936	-	4.03	9.13	35.00	18.44	20.19	4.96	8.11

Source : Base de données de Barro et Lee (2013) (v. 2.1, Feb. 2016).

Tableau 2. Répartition de la population par tranche d'âge selon le niveau d'éducation maximum atteint pour le cas de la Bulgarie en 1990

Groupe d'âge	Moyenne d'années de scolarisation (l^a)	Nombre d'années cumulées de scolarisation ($y_{c,j,t}^a$) {c=120, t=1990}						
		Non Scol. (y_1^a)	Prim. inachevé (y_2^a)	Prim. achevé (y_3^a)	Sec. inachevé (y_4^a)	Sec. achevé (y_5^a)	Sup. inachevé (y_6^a)	Sup. achevé (y_7^a)
15-19	5.08	0.0	2.0	4.0	8.0	11.0	13.0	15.0
20-24	10.06	0.0	4.0	8.0	9.5	11.0	13.0	15.0
25-29	10.60	0.0	4.0	8.0	9.5	11.0	13.0	15.0
30-34	10.58	0.0	4.0	8.0	9.5	11.0	13.0	15.0
35-39	10.30	0.0	4.0	8.0	9.5	11.0	13.0	15.0
40-44	8.84	0.0	2.0	4.0	7.5	11.0	13.0	15.0
45-49	8.88	0.0	2.0	4.0	10.0	14.0	16.0	18.0
50-54	8.80	0.0	2.0	4.0	10.0	14.0	16.0	18.0
55-59	6.78	0.0	2.0	4.0	10.0	14.0	16.0	18.0
60-64	6.77	0.0	2.0	4.0	10.0	14.0	16.0	18.0
65-69	4.65	0.0	2.0	4.0	10.0	14.0	16.0	18.0
70-74	4.64	0.0	2.0	4.0	10.0	14.0	16.0	18.0
75 et +	4.62	0.0	2.0	4.0	10.0	14.0	16.0	18.0
15 et + Auteurs	7.96	0.0	2.2	4.8	9.3	11.6	14.1	16.3
Barro et Lee	8.42	0.0	2.7	5.4	9.4	12.4	14.4	16.4
Surestimation	0.46	0.0	0.5	0.6	0.1	0.8	0.3	0.1

Source : Calcul des auteurs à partir de la base de données de Barro et Lee (2013) (v. 2.1, Feb. 2016).

Les tableaux n°1 et n°2 présentent le cas de la Bulgarie en 1990 {c=120, t=1990}. Le tableau n°1 fournit la répartition de la population de la Bulgarie en 1990 par tranche d'âge selon le niveau d'éducation maximum atteint « $p_{c,j,t}^a$ » ainsi que les poids de chaque tranche d'âge dans la population âgée de 15 ans et plus « $l_{c,t}^a$ ». Le tableau n°2 présente les nombres d'années cumulées de scolarisation par tranche d'âge et selon chaque niveau d'éducation « $y_{c,j,t}^a$ » ainsi que la moyenne d'années de scolarisation par tranche d'âge. La différence entre les deux pondérations est ainsi présentée dans le tableau n°2.

Prenons par exemple le cas du primaire achevé $j=3$, $y_{c,3,t}^{15+}$. Selon Barro et Lee (2013), sa valeur est égale à $\sum_{a=1}^A l_{c,t}^a y_{c,3,t}^a = 5.4$ alors qu'elle doit plutôt être égale à $\sum_{a=1}^A \left[\frac{l_{c,t}^a p_{c,3,t}^a}{\sum_{b=1}^A l_{c,t}^b p_{c,3,t}^b} \right] y_{c,3,t}^a = 4.8$. Il en va de même du calcul pour les autres $y_{c,j,t}^{15+}$. La moyenne d'années de scolarisation de la population âgée de 15 ans et plus peut passer ainsi de 7.96 à 8.42 (différence de 0.46). Cette différence est relativement importante, puisqu'une moyenne d'années de scolarisation n'augmente en moyenne que d'une année de scolarisation tous les dix ans. D'où l'ampleur du biais dû à l'utilisation d'une pondération non adéquate.

2.2. Calcul des durées des cycles d'enseignement

Nous présentons dans cette section la démarche adoptée pour reconstituer les durées des cycles d'enseignement « $y_{c,j,t}^a$ » pour l'ensemble des 146 pays de 1950 à 2015. Celles-ci n'étant pas explicitement présentées dans la base de données de Barro et Lee (2013), nous les avons calculées – et mises à disposition dans notre base de données en ligne.

Nous avons adopté la même approche que Benaabdelaali et al. (2012). Nous fusionnons les deux bases de la population totale {MF} et féminine {F} considérées par Barro et Lee (2013) et utilisons les moyennes d'années de scolarisation par tranches d'âge de chaque cycle. Nous avons ainsi construit un ensemble de 24674 systèmes d'équations à sept inconnues chacun, ce qui nous a permis de retrouver « $y_{c,j,t}^a$ » qui correspond au nombre nécessaire d'années cumulées de scolarisation pour atteindre le niveau d'éducation "j" par la tranche d'âge "a" du pays "c" à la date "t" tel que $j=1...7$, $a=1...13$, $c=1...146$ et $t=1950, 1955...2010$.

Nous nous sommes basés sur les formules des moyennes d'années de scolarisation par cycle. Toutefois, il faut d'abord disposer des proportions d'individus qui n'ont pas achevé les trois cycles d'enseignement. En effet, la base de Barro et Lee (2013) ne les présente pas explicitement. Nous les avons calculées en faisant la différence entre la proportion de ceux qui ont atteint le cycle d'enseignement {notée "lp" "ls" "lh" pour les cycles primaire, secondaire et supérieur respectivement} en question et ceux qui l'ont terminé {notée "lpc" "lsc" "lhc" pour les cycles primaire, secondaire et supérieur respectivement}.

Ainsi, la moyenne d'années de scolarisation du cycle primaire peut être définie comme étant la somme du nombre d'années de scolarisation correspondant aux individus qui n'ont pas achevé le cycle primaire « $y_{c,2,t}^a - y_{c,1,t}^a$ » pondéré par la proportion de ces derniers « $p_{c,2,t}^a$ » plus le nombre d'années de scolarisation correspondant aux individus qui ont achevé le cycle primaire « $y_{c,3,t}^a - y_{c,1,t}^a$ » pondéré par la proportion des individus ayant au moins complété ce même niveau « $\sum_{j=3}^7 p_{c,j,t}^a$ », d'où la formule ci-après :

$$\underset{\text{Cycle primaire inachevé}}{aysp_{c,t}^a} = \underset{\text{Cycle primaire inachevé}}{p_{c,2,t}^a \left[y_{c,2,t}^a - y_{c,1,t}^a \right]} + \underset{\text{Cycle primaire achevé}}{\left(\sum_{j=3}^7 p_{c,j,t}^a \right) \left[y_{c,3,t}^a - y_{c,1,t}^a \right]} \quad (1)$$

Les moyennes d'années de scolarisation des cycles secondaire et supérieur peuvent être obtenues en suivant le même raisonnement :

$$ayss_{c,t}^a = \underbrace{p_{c,4,t}^a [y_{c,4,t}^a - y_{c,3,t}^a]}_{\text{Cycle secondaire inachevé}} + \underbrace{\left(\sum_{j=5}^7 p_{c,j,t}^a \right) [y_{c,5,t}^a - y_{c,3,t}^a]}_{\text{Cycle secondaire achevé}} \quad (2)$$

$$ayst_{c,t}^a = \underbrace{p_{c,6,t}^a [y_{c,6,t}^a - y_{c,5,t}^a]}_{\text{Cycle supérieur inachevé}} + \underbrace{p_{c,7,t}^a [y_{c,7,t}^a - y_{c,5,t}^a]}_{\text{Cycle supérieur achevé}} \quad (3)$$

L'objectif est de calculer les « $y_{c,j,t}^a$ » pour chaque quadruplet (a,c,t,j) où "a" est la tranche d'âge, "c" le pays, "t" la date et "j" le niveau d'éducation. Ceci revient à retrouver pour chaque triplet (a,c,t) les sept inconnues :

$y_{c,1,t}^a; y_{c,2,t}^a; y_{c,3,t}^a; y_{c,4,t}^a; y_{c,5,t}^a; y_{c,6,t}^a; y_{c,7,t}^a$.

On note ainsi :

$aysp_{(mf)_t}^a$ la moyenne d'années de scolarisation du cycle primaire de la tranche d'âge « a » de la population totale à la date « t ».

$ayss_{(mf)_t}^a$ la moyenne d'années de scolarisation du cycle secondaire de la tranche d'âge « a » de la population totale à la date « t ».

$ayst_{(mf)_t}^a$ la moyenne d'années de scolarisation du cycle supérieur de la tranche d'âge « a » de la population totale à la date « t ».

$aysp_{(f)_t}^a$ la moyenne d'années de scolarisation du cycle primaire de la tranche d'âge « a » de la population féminine à la date « t ».

$ayss_{(f)_t}^a$ la moyenne d'années de scolarisation du cycle secondaire de la tranche d'âge « a » de la population féminine à la date « t ».

$ayst_{(f)_t}^a$ la moyenne d'années de scolarisation du cycle supérieur de la tranche d'âge « a » de la population féminine à la date « t ».

$p_{(mf)_{j,t}}^a$ la proportion de la population totale appartenant à la tranche d'âge « a » ayant atteint le niveau d'éducation « j » à la date « t ».

$p_{(f)_{j,t}}^a$ la proportion de la population féminine appartenant à la tranche d'âge « a » ayant atteint le niveau d'éducation « j » à la date « t ».

Etant donné que le nombre d'années de scolarisation cumulées $y_{c,j,t}^a$ est le même pour les hommes que pour les femmes $y_{(mf)_{c,j,t}}^a = y_{(m)_{c,j,t}}^a = y_{(f)_{c,j,t}}^a$, nous avons reproduit les trois équations relatives à la moyenne d'années de scolarisation des trois cycles pour la population totale et féminine ($3 \times 2 = 6$ équations). En outre, la dotation éducative de la population non scolarisée est nulle ($y_{c,1,t}^a = 0$). Au total, nous pouvons obtenir pour chaque triplet (a,c,t) un système de sept équations et de sept inconnues :

$$aysp_{(mf)_{c,t}}^a = p_{(mf)_{c,2,t}}^a [y_{c,2,t}^a - y_{c,1,t}^a] + \left(\sum_{j=3}^7 p_{(mf)_{c,j,t}}^a \right) [y_{c,3,t}^a - y_{c,1,t}^a]$$

$$aysp_{(f)_{c,t}}^a = p_{(f)_{c,2,t}}^a [y_{c,2,t}^a - y_{c,1,t}^a] + \left(\sum_{j=3}^7 p_{(f)_{c,j,t}}^a \right) [y_{c,3,t}^a - y_{c,1,t}^a]$$

$$ayss_{(mf)_{c,t}}^a = p_{(mf)_{c,4,t}}^a [y_{c,4,t}^a - y_{c,3,t}^a] + \left(\sum_{j=5}^7 p_{(mf)_{c,j,t}}^a \right) [y_{c,5,t}^a - y_{c,3,t}^a]$$

$$ayss_{(f)_{c,t}}^a = p_{(f)_{c,4,t}}^a [y_{c,4,t}^a - y_{c,3,t}^a] + \left(\sum_{j=5}^7 p_{(f)_{c,j,t}}^a \right) [y_{c,5,t}^a - y_{c,3,t}^a]$$

$$ayst_{(mf)_{c,t}}^a = p_{(mf)_{c,6,t}}^a [y_{c,6,t}^a - y_{c,5,t}^a] + p_{(mf)_{c,7,t}}^a [y_{c,7,t}^a - y_{c,5,t}^a]$$

$$ayst_{(f)_{c,t}}^a = P_{(f)_{c,6,t}}^a [y_{c,6,t}^a - y_{c,5,t}^a] + P_{(f)_{c,7,t}}^a [y_{c,7,t}^a - y_{c,5,t}^a]$$

$$y_{c,1,t}^a = 0$$

On peut ainsi résoudre ce système d'équations pour chaque triplet (a,c,t) et les sept inconnues $\{y_{c,1,t}^a; y_{c,2,t}^a; y_{c,3,t}^a; y_{c,4,t}^a; y_{c,5,t}^a; y_{c,6,t}^a; y_{c,7,t}^a\}$ sont retrouvées comme suit :

$$y_{c,1,t}^a = 0$$

$$y_{c,2,t}^a = \frac{\left(\sum_{j=3}^7 P_{(f)_{c,j,t}}^a\right) (aysp_{(mf)_{c,t}}^a) - \left(\sum_{j=5}^7 P_{(mf)_{c,j,t}}^a\right) (aysp_{(f)_{c,t}}^a)}{P_{(mf)_{c,2,t}}^a \left(\sum_{j=3}^7 P_{(f)_{c,j,t}}^a\right) - P_{(f)_{c,2,t}}^a \left(\sum_{j=3}^7 P_{(mf)_{c,j,t}}^a\right)}$$

$$y_{c,3,t}^a = \frac{P_{(mf)_{c,2,t}}^a \left(aysp_{(f)_{c,t}}^a\right) - P_{(f)_{c,2,t}}^a \left(aysp_{(mf)_{c,t}}^a\right)}{P_{(mf)_{c,2,t}}^a \left(\sum_{j=3}^7 P_{(f)_{c,j,t}}^a\right) - P_{(f)_{c,2,t}}^a \left(\sum_{j=3}^7 P_{(mf)_{c,j,t}}^a\right)}$$

$$y_{c,4,t}^a = y_{c,3,t}^a + \frac{\left(\sum_{j=5}^7 P_{(f)_{c,j,t}}^a\right) (ayss_{(mf)_{c,t}}^a) - \left(\sum_{j=5}^7 P_{(mf)_{c,j,t}}^a\right) (ayss_{(f)_{c,t}}^a)}{P_{(mf)_{c,4,t}}^a \left(\sum_{j=5}^7 P_{(f)_{c,j,t}}^a\right) - P_{(f)_{c,4,t}}^a \left(\sum_{j=5}^7 P_{(mf)_{c,j,t}}^a\right)}$$

$$y_{c,5,t}^a = y_{c,3,t}^a + \frac{P_{(mf)_{c,4,t}}^a \left(ayss_{(f)_{c,t}}^a\right) - P_{(f)_{c,4,t}}^a \left(ayss_{(mf)_{c,t}}^a\right)}{P_{(mf)_{c,4,t}}^a \left(\sum_{j=5}^7 P_{(f)_{c,j,t}}^a\right) - P_{(f)_{c,4,t}}^a \left(\sum_{j=5}^7 P_{(mf)_{c,j,t}}^a\right)}$$

$$y_{c,6,t}^a = y_{c,5,t}^a + \frac{P_{(f)_{c,7,t}}^a \left(ayst_{(mf)_{c,t}}^a\right) - P_{(mf)_{c,7,t}}^a \left(ayst_{(f)_{c,t}}^a\right)}{P_{(mf)_{c,6,t}}^a P_{(f)_{c,7,t}}^a - P_{(f)_{c,6,t}}^a P_{(mf)_{c,7,t}}^a}$$

$$y_{c,7,t}^a = y_{c,5,t}^a + \frac{P_{(mf)_{c,6,t}}^a \left(ayst_{(f)_{c,t}}^a\right) - P_{(f)_{c,6,t}}^a \left(ayst_{(mf)_{c,t}}^a\right)}{P_{(mf)_{c,6,t}}^a P_{(f)_{c,7,t}}^a - P_{(f)_{c,6,t}}^a P_{(mf)_{c,7,t}}^a}$$

Il faut préciser que lorsque le dénominateur est nul, on peut retrouver la valeur de $y_{c,j,t}^a$ par l'équation suivante $y_{c,j,t-5}^{a-1} = y_{c,j,t}^a = y_{c,j,t+5}^{a+1}$, puisqu'il s'agit de la même cohorte.

2.3. Indices d'inégalité d'éducation pour des groupes agrégés

Répetons que l'erreur de pondération qu'on a pu déceler dans la base de Barro et Lee ne remet pas en question la qualité des données sur la répartition de l'éducation mais concerne plus particulièrement les moyennes d'années de scolarisation de la population âgée de 15 ans et plus et de 25 ans et plus. Pour calculer les indices d'inégalité nous avons donc utilisé les données de Barro et Lee sur la répartition de l'éducation et nous avons mené les corrections nécessaires pour les moyennes d'années de scolarisation.

Ceci étant, afin de calculer les indices d'inégalité d'éducation de la population âgée de 15 ans et plus, on pourrait opter pour une solution qui consiste à calculer la moyenne des « $y_{c,j,t}^a$ » avec les bonnes pondérations pour obtenir « $y_{c,j,t}^{15+}$ » puis l'utiliser par la suite pour l'évaluation de l'indice d'inégalité. Toutefois, cette solution consiste à agréger les données avant de calculer l'indice d'inégalité du groupe d'âge agrégé et ne permet pas de prendre en considération les inégalités inter-tranches d'âge ni d'exploiter la richesse des données désagrégées. Aussi, afin d'en tenir compte, nous utilisons une autre approche, en utilisant une formulation plus adaptée. Les indices d'inégalité considérés se déclinent alors comme suit.

2.3.1. Indice de Gini de l'éducation pour des groupes agrégés

$$Gini_{c,t}^G = \frac{\sum_{a \in G} \sum_{i=1}^n \sum_{j=1}^n n_{c,i,t}^a |y_{c,i,t}^a - y_{c,j,t}^a| n_{c,j,t}^a}{2 N_{c,t}^G \left[\sum_{a \in G} \sum_{i=1}^n n_{c,i,t}^a y_{c,i,t}^a \right]}$$

avec :

- $Gini_{c,t}^G$ l'indice de Gini de l'éducation du grand groupe d'âge "G"⁴ appartenant au pays "c" à la date "t".
- $a=1$ ($a'=1$) pour le groupe d'âge 15–19 ans, $a=2$ ($a'=2$) pour le groupe d'âge 20–24 ans, ..., et $a=13$ ($a'=13$) pour le groupe d'âge 75 et plus.
- i et j sont les niveaux d'éducation : $i=1$ pour les illettrés, $i=2$ pour le primaire inachevé, $i=3$ pour le primaire achevé, $i=4$ pour le secondaire inachevé, $i=5$ pour le secondaire achevé, $i=6$ pour le supérieur inachevé, and $i=7$ pour le supérieur achevé.
- $n_{c,i,t}^a$ représente la taille de la population de la tranche d'âge "a" du pays "c" ayant atteint le niveau d'éducation "i" à la date "t".
- $y_{c,i,t}^a$ traduit le nombre d'années de scolarisation cumulées de la tranche d'âge "a" du pays "c" pour atteindre le niveau d'éducation "j" à la date "t".

De même pour des groupes de pays, qu'ils soient considérés selon la zone géographique, le niveau de revenu ou de développement, des auteurs comme Földvári et Van Leeuwen (2014) calculent une moyenne des indices de Gini par pays au lieu de calculer un indice de Gini sur la distribution de l'ensemble des pays appartenant au groupe en question et ignorent de ce fait, les inégalités inter-pays. Pour rendre compte du niveau d'inégalités pour un groupe de pays, il nous paraît plus approprié de calculer un indice de Gini sur la distribution de l'ensemble des pays du groupe :

$$Gini_{R,t}^a = \frac{\sum_{c \in R} \sum_{c' \in R} \sum_{i=1}^n \sum_{j=1}^n n_{c,i,t}^a |y_{c,i,t}^a - y_{c',j,t}^a| n_{c',j,t}^a}{2 N_{R,t}^a \left[\sum_{c \in R} \sum_{i=1}^n n_{c,i,t}^a y_{c,i,t}^a \right]}$$

avec, outre les symboles déjà définis :

- $Gini_{R,t}^a$ l'indice de Gini de l'éducation du groupe d'âge "a" appartenant à la région "R" à la date "t".
- c (c') représente les pays appartenant à la région "R".
- n est le nombre des niveaux d'éducation retenus qui correspond ici à 7.
- $N_{R,t}^a$ représente la taille de la population du groupe d'âge "a" appartenant à la région "R" à la date "t". ($N_{R,t}^a = \sum_{c \in R} \sum_{i=1}^n n_{c,i,t}^a = \sum_{c \in R} N_{c,t}^a$).

On peut généraliser la formule du calcul de l'indice de Gini pour des groupes agrégés à la fois par pays et par grands groupes d'âge (par exemple, pour la population âgée de 15 ans et plus) tout en utilisant une formule qui, en plus de prendre en considération les inégalités inter-niveaux d'éducation des individus, tienne compte à la fois des inégalités inter-pays et inter-tranches d'âge :

⁴ Par exemple G peut être la population âgée entre 15 et 65 ans, de 25 à 65 ans, 15 à 24 ans, ou encore la population âgée de 15 ans et plus ou 25 ans et plus...

$$Gini_{R,t}^G = \frac{\sum_{c \in R} \sum_{c' \in R} \sum_{a \in G} \sum_{a' \in G} \sum_{i=1}^n \sum_{j=1}^n n_{c,i,t}^a |y_{c,i,t}^a - y_{c',j,t}^{a'}| n_{c',j,t}^{a'}}{2 N_{R,t}^G \left[\sum_{c \in R} \sum_{a \in G} \sum_{i=1}^n n_{c,i,t}^a y_{c,i,t}^a \right]}$$

où :

- $Gini_{R,t}^G$ est l'indice de Gini de l'éducation du grand groupe d'âge "G" appartenant à la région "R" à la date "t".
- $N_{R,t}^G$ représente la taille de la population du grand groupe d'âge "G" appartenant à la région "R" à la date "t". ($N_{R,t}^G = \sum_{c \in R} \sum_{a \in G} \sum_{i=1}^n n_{c,i,t}^a = \sum_{c \in R} N_{c,t}^G = \sum_{c \in R} N_{R,t}^a$).

2.3.2. Les indices d'entropie généralisée et d'Atkinson de l'éducation

En suivant la même approche adoptée pour le calcul de l'indice de Gini de l'éducation d'un groupe d'âge quinquennal d'un pays donné ou d'un groupe agrégé de pays et/ou de groupes d'âge, nous pouvons calculer d'autres indices d'inégalité comme ceux appartenant à la famille d'entropie généralisée ou d'Atkinson.

Les indices d'entropie généralisée et d'Atkinson de l'éducation sont :

$$GE(r)_{c,t}^a = \frac{N_{c,t}^{a(r-1)}}{r(r-1)} \left[\frac{\sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^r}{\left(\sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a \right)^r} - 1 \right]$$

$$\text{et } A(\varepsilon)_{c,t}^a = 1 - \frac{1}{\left[N_{c,t}^a \right]^{\frac{\varepsilon}{1-\varepsilon}}} \left[\frac{\sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^{(1-\varepsilon)}}{\sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a} \right]^{\frac{1}{1-\varepsilon}}$$

avec :

- $GE(r)_{c,t}^a$ l'indice d'entropie généralisée de l'éducation d'ordre r du groupe d'âge "a" appartenant au pays "c" à la date "t".
- $A(\varepsilon)_{c,t}^a$ l'indice Atkinson de l'éducation d'ordre ε du groupe d'âge "a" appartenant au pays "c" à la date "t".
- $a=1$ pour le groupe d'âge 15-19 ans, $a=2$ pour le groupe d'âge 20-24 ans, ..., et $a=13$ pour le groupe d'âge 75 et plus.
- n est le nombre des niveaux d'éducation retenus qui correspond ici à 7.
- i et j sont les niveaux d'éducation tel que: $i=1$ pour les illettrés, $i=2$ pour le primaire inachevé, $i=3$ pour le primaire achevé, $i=4$ pour le secondaire inachevé, $i=5$ pour le secondaire achevé, $i=6$ pour le supérieur inachevé, and $i=7$ pour le supérieur achevé.
- $n_{c,i,t}^a$ représente la taille de la population de la tranche d'âge "a" du pays "c" ayant atteint le niveau d'éducation "i" à la date "t".
- $N_{c,t}^a$ représente la taille de la population de la tranche d'âge "a" du pays "c" à la date "t" ($N_{c,t}^a = \sum_{i=1}^n n_{c,i,t}^a$).
- $y_{c,i,t}^a$ indique le nombre d'années de scolarisation cumulées de la tranche d'âge "a" du pays "c" pour atteindre le niveau d'éducation "j" à la date "t".

Les indices d'entropie généralisée et d'Atkinson de l'éducation peuvent être établis pour des groupes agrégés.

Grand groupe d'âge d'un pays

$$GE(r)_{c,t}^G = \frac{N_{c,t}^{G(r-1)}}{r(r-1)} \left[\frac{\sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^r}{\left(\sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a \right)^r} - 1 \right]$$

$$A(\varepsilon)_{c,t}^G = 1 - \frac{1}{\left[N_{c,t}^G \right]^{\left(\frac{\varepsilon}{1-\varepsilon} \right)}} \left[\frac{\left[\sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^{(1-\varepsilon)} \right]^{\left(\frac{1}{1-\varepsilon} \right)}}{\sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a} \right]$$

où :

- $GE(r)_{c,t}^G$ l'indice d'entropie généralisée de l'éducation d'ordre r du grand groupe d'âge "G" appartenant au pays "c" à la date "t".
- $A(\varepsilon)_{c,t}^G$ l'indice Atkinson de l'éducation d'ordre ε du grand groupe d'âge "G" appartenant au pays "c" à la date "t".
- $a=1$ ($a'=1$) pour le groupe d'âge 15-19 ans, $a=2$ ($a'=2$) pour le groupe d'âge 20-24 ans, ..., et $a=13$ ($a'=13$) pour le groupe d'âge 75 et plus.
- $N_{c,t}^G$ représente la taille de la population du grand groupe d'âge "G" du pays "c" à la date "t" ($N_{c,t}^G = \sum_{a \in G} \sum_{i=1}^n n_{c,i,t}^a = \sum_{a \in G} N_{c,t}^a$).

Groupe d'âge quinquennal d'une zone géographique

$$GE(r)_{R,t}^a = \frac{N_{R,t}^{a(r-1)}}{r(r-1)} \left[\frac{\sum_{c \in R} \sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^r}{\left(\sum_{c \in R} \sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a \right)^r} - 1 \right]$$

$$A(\varepsilon)_{R,t}^a = 1 - \frac{1}{\left[N_{R,t}^a \right]^{\left(\frac{\varepsilon}{1-\varepsilon} \right)}} \left[\frac{\left[\sum_{c \in R} \sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^{(1-\varepsilon)} \right]^{\left(\frac{1}{1-\varepsilon} \right)}}{\sum_{c \in R} \sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a} \right]$$

où :

- $GE(r)_{R,t}^a$ l'indice d'entropie généralisée de l'éducation d'ordre r du groupe d'âge "a" appartenant à la région "R" à la date "t".
- $A(\varepsilon)_{R,t}^a$ l'indice Atkinson de l'éducation d'ordre ε du groupe d'âge "a" appartenant à la région "R" à la date "t".
- c (c') les pays appartenant à la région "R".
- n est le nombre des niveaux d'éducation retenus qui correspond ici à 7.
- $N_{R,t}^a$ représente la taille de la population du groupe d'âge "a" appartenant à la région "R" à la date "t". ($N_{R,t}^a = \sum_{c \in R} \sum_{i=1}^n n_{c,i,t}^a = \sum_{c \in R} N_{c,t}^a$).

Grand groupe d'âge d'une zone géographique

$$GE(r)_{c,t}^G = \frac{N_{R,t}^{G(r-1)}}{r(r-1)} \left[\frac{\sum_{c \in R} \sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^r}{\left(\sum_{c \in R} \sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a \right)^r} - 1 \right]$$

$$A(\varepsilon)_{R,t}^G = 1 - \frac{1}{\left[N_{R,t}^G \right]^{\frac{\varepsilon}{1-\varepsilon}}} \left[\frac{\left[\sum_{c \in R} \sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^{(1-\varepsilon)} \right]^{\frac{1}{1-\varepsilon}}}{\sum_{c \in R} \sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a} \right]$$

où :

- $GE(r)_{R,t}^G$ l'indice d'entropie généralisée de l'éducation d'ordre r du grand groupe d'âge "G" appartenant à la région "R" à la date "t".
- $A(\varepsilon)_{R,t}^G$ l'indice Atkinson de l'éducation d'ordre ε du grand groupe d'âge "G" appartenant à la région "R" à la date "t".
- $a=1$ pour le groupe d'âge 15-19 ans, $a=2$ pour le groupe d'âge 20-24 ans, ..., et $a=13$ pour le groupe d'âge 75 et plus.
- n est le nombre des niveaux d'éducation retenus qui correspond ici à 7.
- $N_{R,t}^G$ représente la taille de la population du grand groupe d'âge "G" appartenant à la région "R" à la date "t". ($N_{R,t}^G = \sum_{c \in R} \sum_{a \in G} \sum_{i=1}^n n_{c,i,t}^a = \sum_{c \in R} N_{c,t}^G = \sum_{c \in R} N_{R,t}^a$).

L'annexe n°1 présente une synthèse de l'ensemble des formules proposées pour le calcul des indices d'inégalité éducative.

3. TENDANCES DES INÉGALITÉS ÉDUCATIVES AU NIVEAU MONDIAL

A partir de la base que nous avons construite, le tableau n° 3 présente l'évolution de l'inégalité mondiale de l'éducation mesurée à travers quatre indices d'entropie généralisée, l'indice de Gini et deux indices d'Atkinson.

Tableau 3. Evolution des inégalités mondiales d'éducation, 1950-2015

Année	Indices d'entropie généralisée				Indice de Gini	Indices d'Atkinson	
	GE(0)	GE(0.5)	GE(1)	GE(2)		A(0.5)	A(1)*
1950	13.61	1.23	0.81	0.78	0.64	0.52	0.37
1960	12.02	1.08	0.70	0.64	0.60	0.47	0.36
1975	9.34	0.83	0.53	0.44	0.52	0.37	0.35
1990	7.44	0.65	0.41	0.31	0.45	0.30	0.32
2005	4.85	0.43	0.27	0.19	0.35	0.20	0.25
2015	3.60	0.33	0.20	0.15	0.31	0.16	0.21

* L'indice d'Atkinson d'ordre 1 est calculé en ajoutant une année de scolarisation à l'ensemble de la distribution.

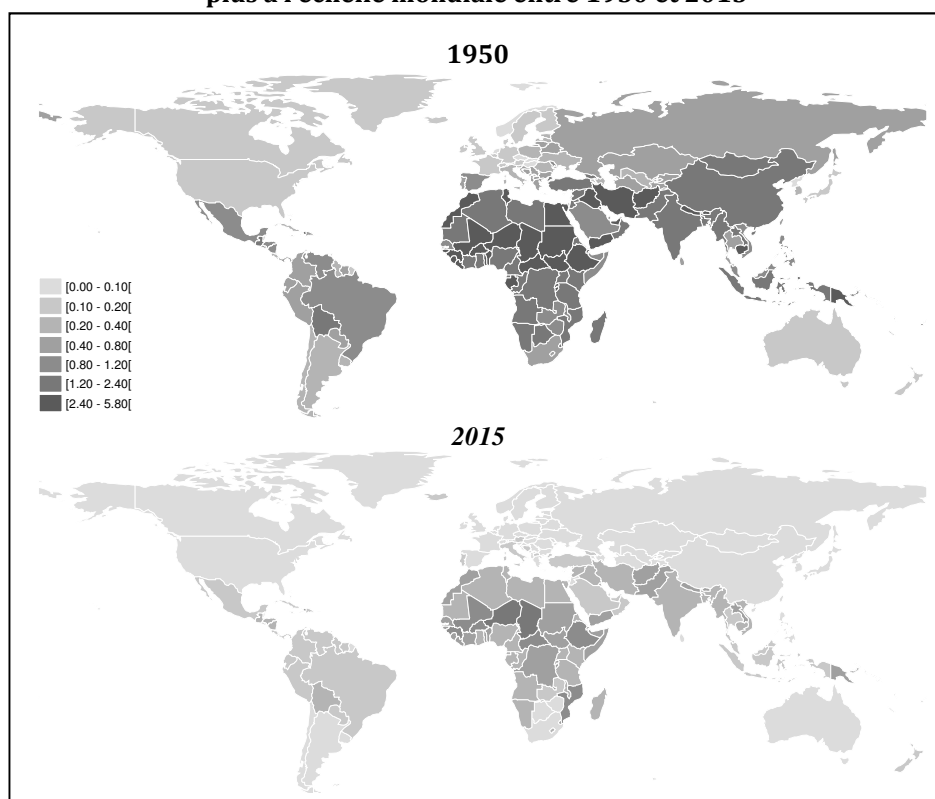
Source : Calcul des auteurs à partir de la base mondiale des inégalités d'éducation, 1950-2015.

Il apparaît que, durant ces soixante-cinq années, d'importants efforts ont été entrepris par les pays sur le plan de l'accès à l'éducation. Une baisse de l'inégalité mondiale de l'éducation s'observe en général quels que soient les indices présentés. Elle

s'est accompagnée d'une augmentation significative de la moyenne d'années de scolarisation : cette dernière a augmenté de 3.14 en 1950 à 8.49 années en 2015 pour la population âgée de 15 ans et plus. En parallèle, le taux d'analphabétisme a largement baissé : entre 1950 et 2015 ce taux est passé de 48% à 12% pour la population âgée de 15 ans et plus.

La figure n° 1 présente ainsi une carte de l'indice de Theil de l'éducation (GE(1)) à travers le monde en 1950 et en 2015. De manière générale, l'indice de Theil de l'éducation de la population mondiale âgée de 15 et plus a été divisé par quatre, il est passé de 0.81 en 1950 à 0.20 en 2015.

Figure 1 : Indice de Theil de l'éducation de la population âgée de 15 ans et plus à l'échelle mondiale entre 1950 et 2015



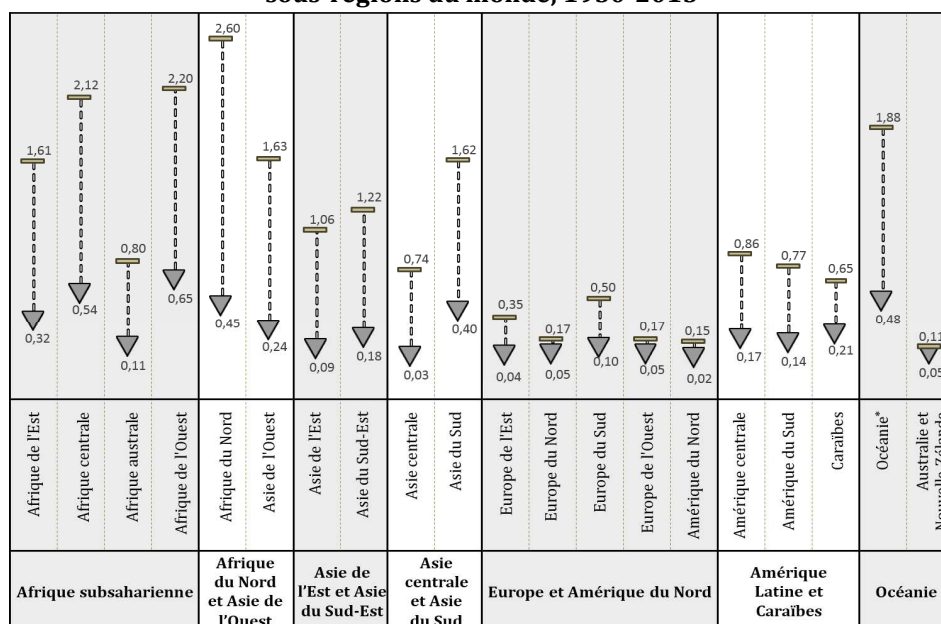
Source : Auteurs (Base mondiale des inégalités d'éducation, 1950-2015).

3.1. Les inégalités mondiales d'éducation par sous-régions

L'annexe n°2 présente les valeurs des différents indices d'inégalité entre 1950 et 2015 pour l'ensemble des 20 sous-régions du monde, en se référant au nouveau découpage de l'ONU du programme de développement durable 2015-2030. A l'exception de l'indice d'Atkinson d'ordre 1, $A(1)$, qui paraît revêtir une certaine spécificité, tous les autres indices montrent une tendance à la baisse des inégalités d'éducation. L'indice d'Atkinson d'ordre 1 dévoile plutôt une courbe en cloche au fil du temps dans le cas des sous-régions de l'Afrique « non australe », de l'Asie du sud et de l'Océanie. Ainsi, lorsque le taux d'analphabétisme est très élevé (supérieur à 90%), l'inégalité d'éducation mesurée par l'indice d'Atkinson $A(1)$ est atypiquement faible

alors que les autres indices considérés présentent des niveaux d'inégalité très élevés. Lorsque le niveau d'analphabétisme diminue au fil du temps, l'indice d'Atkinson A(1) augmente jusqu'à atteindre un certain pic⁵. Au-delà de ce pic, sa tendance globale rejoint celle des autres indices d'inégalité de l'éducation. Il convient de noter que cette tendance particulière est principalement expliquée par le mode de calcul de cet indice. En effet, l'indice d'Atkinson d'ordre 1 correspond au rapport entre la moyenne géométrique et la moyenne arithmétique de l'éducation. Étant donné que la moyenne géométrique n'accepte pas des valeurs nulles⁶ ($y_1 = 0$), la distribution de l'éducation subit au préalable une transformation avant le calcul de l'indice⁷, ce qui est de nature à affecter la sensibilité à l'extrémité inférieure de cette distribution.

Graphique n°1 : Changement de l'indice de Theil de l'éducation selon les sous-régions du monde, 1950-2015



* Océanie sauf Australie et Nouvelle-Zélande.

Source : Calcul des auteurs à partir de la base mondiale des inégalités d'éducation 1950-2015.

La réduction importante des inégalités dans l'éducation globalement observée montre notamment un certain mouvement de rattrapage des pays en retard d'éducation mais ne doit pas cacher l'existence encore de profondes disparités.

Comme le montrent la figure n° 1 et le graphique n°1, c'est le continent africain qui a enregistré la réduction la plus importante en matière d'inégalité d'éducation.

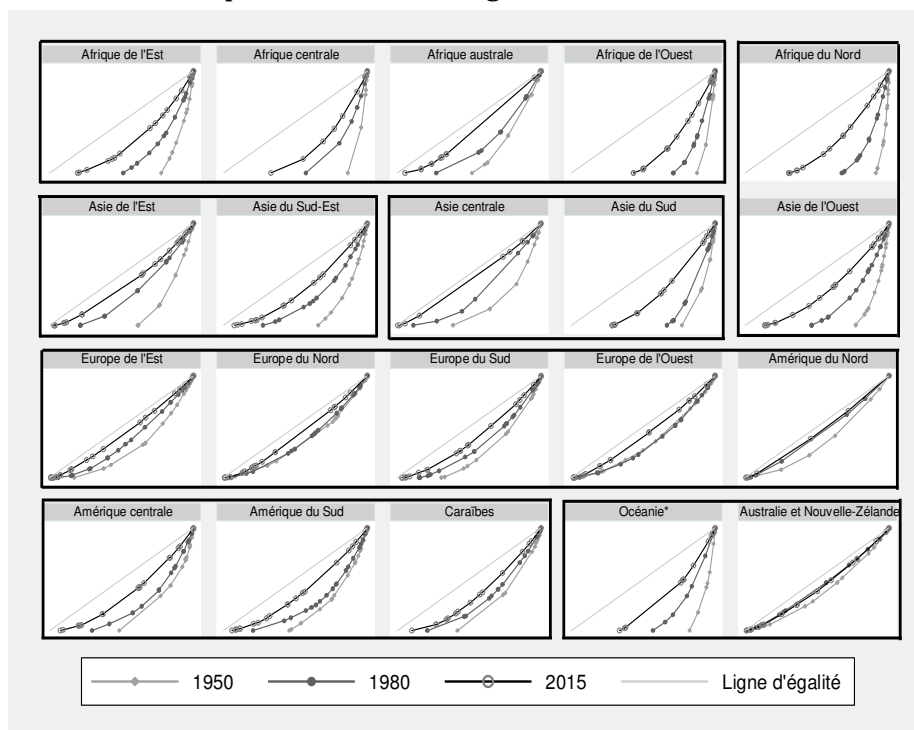
⁵ Ce point de retournement est de l'ordre de 0.36 pour l'Afrique de l'Est, de 0.39 pour l'Afrique centrale, de 0.46 pour l'Afrique de l'Ouest de 0.44 pour l'Afrique du Nord, de 0.43 pour l'Asie du Sud et de 0.42 pour l'Océanie (sauf Australie et Nouvelle-Zélande).

⁶ Contrairement à l'indice d'Atkinson d'ordre 1, l'indice de Theil même s'il n'est pas compatible avec les valeurs nulles, accepte néanmoins une limite finie lorsque les valeurs nulles sont considérées comme étant infiniment petites (y_1 tend vers 0).

⁷ Cette transformation consiste à ajouter une année de scolarisation à l'ensemble de la distribution de l'éducation pour éviter le problème des valeurs nulles ($y_i \rightarrow y_i + 1$). Cette même transformation a été adoptée par le PNUD pour évaluer la perte due à l'inégalité dans la « composité éducation » de l'indice de développement humain ajusté aux inégalités.

L'indice de Theil de l'éducation de l'Afrique du Nord en particulier est passé de 2.60 en 1950 à 0.45 en 2015, soit une réduction de 2.15 points. Les indices de Theil de l'éducation de l'Afrique de l'Ouest, de l'Afrique centrale et de l'Afrique de l'Est ont baissé significativement de 1.55, 1.58 et 1.29 points sur la même période. Le niveau d'inégalité d'éducation reste toutefois élevé en Afrique, sauf en Afrique australe qui présente en 2015 une distribution de l'éducation relativement équitable (indice de Theil de 0.11). Soulignons que cette région est la seule du continent africain à présenter une moyenne d'années de scolarisation supérieure à la moyenne mondiale (9.88 années). L'Asie et l'Océanie ont connu aussi une baisse sensible des inégalités éducatives depuis les années 1950-1960.

Graphique n°2 : La courbe de Lorenz de l'éducation de la population âgée de 15 ans et plus selon les sous-régions du monde, 1950-2015



* Océanie sauf Australie et Nouvelle-Zélande.

Source : Calcul des auteurs à partir de la base mondiale des inégalités d'éducation 1950-2015.

Par ailleurs, le graphique n°2 présente la courbe de Lorenz de l'éducation pour la population âgée de 15 ans et plus de chacune des vingt sous-régions du monde. Il apparaît nettement un déplacement de ces courbes entre 1950 et 2015 vers la bissectrice qui représente la ligne égalitaire.

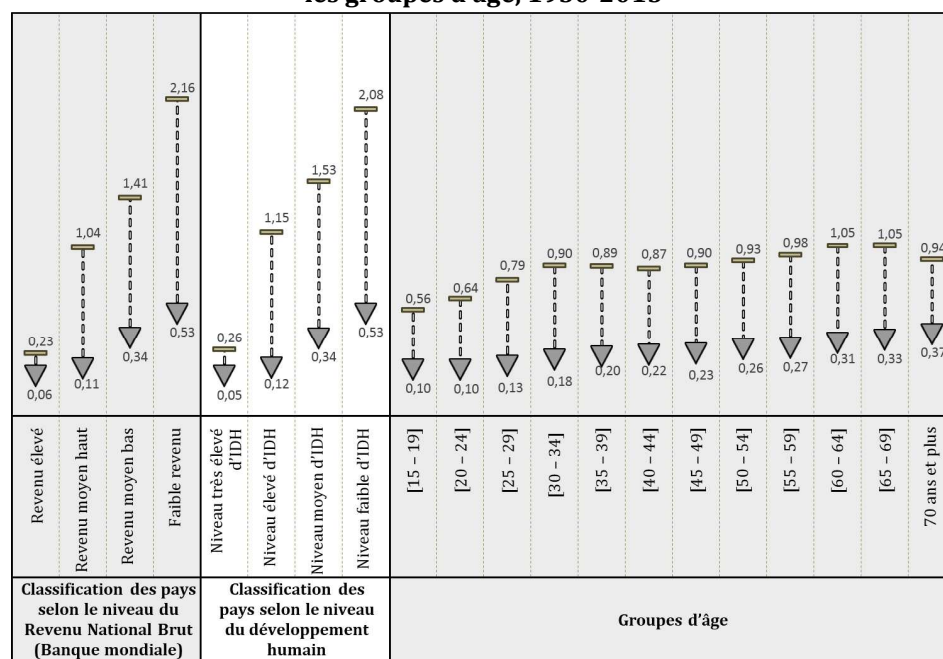
Toutefois, pour l'Amérique du Nord, l'Europe du Nord, l'Europe de l'Est, l'Europe de l'Ouest et l'Australie/Nouvelle-Zélande, les courbes de Lorenz se déplacent au cours du temps à travers la partie supérieure de la distribution. Ces courbes sont pratiquement reliées à l'origine du graphique, ce qui traduit pour ces pays un niveau d'alphabétisation relativement élevé dès le début des années 1950. Il semblerait que l'amélioration de l'accès à l'éducation dans les pays de ce groupe soit passée par une

proportion plus grande de la population ayant atteint le niveau d'enseignement supérieur. Pour tous les autres pays, les courbes se déplacent plutôt à travers la partie inférieure de la distribution de l'éducation. Ces pays ont beaucoup plus élargi l'éducation de base, c'est-à-dire la partie inférieure de la distribution de l'éducation. En conséquence, l'inégalité de l'éducation dans ces pays s'est réduite en raison notamment d'une baisse sensible de la proportion de la population n'ayant aucune scolarisation.

3.2. L'inégalité d'éducation selon le niveau de développement et les groupes d'âge

Le graphique n°3 présente une déclinaison de l'indice de Theil de l'éducation selon le niveau de revenu et le niveau de développement humain des pays entre 1950 et 2015. De manière générale, le groupe de pays qui dispose du revenu national brut (d'après la classification de la Banque mondiale), et du niveau de développement humain (selon la classification du PNUD) les plus élevés présente un niveau d'inégalité de l'éducation faible. Dès lors la réduction de cette même inégalité apparaît relativement limitée entre 1950 et 2015.

Graphique n°3 : Changement de l'indice de Theil de l'éducation selon le niveau de revenu, le niveau de développement humain et les groupes d'âge, 1950-2015



* Océanie sauf Australie et Nouvelle-Zélande.

Source : Calcul des auteurs à partir de la base mondiale des inégalités d'éducation 1950-2015.

En revanche l'indice de Theil de l'éducation des pays à faible niveau de développement humain est passé de 2.08 à 0.53, soit une baisse sensible de 1.55 points. Néanmoins, malgré cette réduction significative des inégalités éducatives, l'écart

entre pays à niveau de développement humain très élevé et ceux à niveau faible demeure assez important puisque la différence entre les deux indices de Theil est de l'ordre de 0.48.

Le graphique n°3 présente également l'indice de Theil de l'éducation au niveau mondial selon les différents groupes d'âge. Si la réduction des inégalités d'éducation est perceptible pour l'ensemble des groupes d'âge, le niveau d'inégalité le plus faible s'observe pour les tranches 15-19 ans et 20-24 ans. Ainsi c'est la population des jeunes qui a le plus bénéficié de la progression liée au développement de l'éducation de base tandis que la population âgée de 60 ans et plus constitue le groupe d'âge pour lequel le niveau d'inégalité est le plus élevé : l'indice de Theil mondial de l'éducation est passé de 0.56 en 1950 à 0.10 en 2015 pour le groupe d'âge 15-19 ans, et de 0.64 à 0.10 pour les 20-24 ans. Il est passé de 1.05 en 1950 à 0.31 en 2015 pour les 60-64 ans, à 0.33 pour les 65-69 ans, et de 0.94 à 0.37 pour les 70 ans et plus.

CONCLUSION

Cet article propose une nouvelle base de données, disponible en ligne, comportant plusieurs indicateurs caractéristiques des inégalités éducatives couvrant 146 pays sur la période 1950-2015. Par rapport aux autres bases existantes, nous avons cherché à améliorer sensiblement les estimations des indices d'inégalité d'éducation et de la moyenne d'années de la scolarisation. Nous avons montré que la pondération utilisée par Barro et Lee (2013), pour obtenir les durées cumulées de scolarisation d'une tranche d'âge agrégée, conduit à un biais dans l'estimation de la moyenne d'années de scolarisation. Ce biais peut même atteindre pour certains pays une différence de l'ordre d'une demi-année en termes de moyenne d'années de scolarisation. Notre base de données corrige cette erreur de pondération et relâche certaines hypothèses peu vraisemblables adoptées par plusieurs auteurs. Nous avons ainsi pris en compte les variations au cours du temps et entre les différentes classes d'âge des durées de scolarisation et l'hétérogénéité par rapport aux niveaux d'éducation au sein d'un même cycle d'enseignement. Nous avons ainsi amélioré la méthodologie du calcul des inégalités d'éducation.

En résultat, cette base offre de nouvelles estimations de l'inégalité mondiale de l'éducation par tranche d'âge et selon les sous-régions du monde. A l'exception de l'indice d'Atkinson d'ordre 1, dont l'interprétation liée à sa méthode de calcul est plus particulière, tous les indices d'inégalité présentés dans la base confirment une diminution conséquente de l'inégalité mondiale de l'éducation. Il apparaît que les réductions les plus importantes, repérées entre 1950 et 2015, sont principalement dues pour les pays en développement à une réduction de la population n'ayant aucune scolarisation et du taux d'analphabétisme, ce que montre le déplacement de la distribution de l'éducation dans la partie inférieure de la courbe de Lorenz. Les pays de l'Amérique du Nord, de l'Europe du Nord, de l'Europe de l'Est, de l'Europe de l'Ouest et de l'Australie/Nouvelle-Zélande présentent en revanche des courbes de Lorenz qui se déplacent au cours du temps à travers la partie supérieure de la distribution, correspondant essentiellement à une progression de la part de la population dans l'enseignement supérieur.

La base de données que nous proposons renseigne, de la manière la plus appropriée possible, sur le degré de scolarisation à travers différents indicateurs d'inégalité et peut aider à différentes études, mais certes ne renseigne pas sur la qualité de l'éducation et a fortiori sur les résultats, les performances scolaires et les « profits » de l'éducation.

REFERENCES

- Araujo C., Ferreira F., Schady N.**, 2004, Is the World Becoming More Unequal? Changes in the World Distribution of Schooling", *World Bank*. Washington, DC.
- Barro R.J., Lee J.W.**, 1993, International Comparisons of Educational Attainment, *Journal of Monetary Economics*, 32, 3, 363-394.
- Barro R.J., Lee J.W.**, 1996, International Measures of Schooling Years and Schooling Quality", *American Economic Review*, 86, 2, 218-223.
- Barro R.J., Lee J.W.**, 2001, International Data on Educational Attainment: Updates and Implications", *Oxford Economic Papers*, 53, 3, 541-563.
- Barro R.J., Lee J.W.**, 2013, A New Data Set of Educational Attainment in the World, 1950-2010", *Journal of Development Economics*, 104, 3, 184-198.
- Bartusevičius H.**, 2014, The Inequality-Conflict Nexus Re-Examined: Income, Education and Popular Rebellions, *Journal of Peace Research*, 51, 1, 35-50.
- Benaabdelaali W., Hanchane S., Kamal A.**, 2012, Educational Inequality in the World, 1950-2010: Estimates from a New Dataset, in John A. Bishop, Rafael Salas, ed., *Inequality, Mobility and Segregation: Essays in Honor of Jacques Silber*, *Research on Economic Inequality*, 20, Emerald Group Publishing Limited, Chapter 13, 337-366.
- Castelló A., Doménech R.**, 2002, Human Capital Inequality and Economic Growth: Some New Evidence, *The Economic Journal*, 112, 478, 187-200.
- Castelló A.**, 2008, On the distribution of education and democracy, *Journal of Development Economics*, 87, 179-190.
- Castelló A.**, 2010a, Channels Through Which Human Capital Inequality Influences Economic Growth, *Journal of Human Capital*, 4, 4, 394-450.
- Castelló A.**, 2010b, Inequality and Growth in Advanced Economies: An Empirical Investigation, *Journal of Economic Inequality*, 8, 3, 293-321.
- Cecchi D.**, 2004, Does educational achievement help to explain income inequality?, In A. Cornia, eds., *Inequality, Growth and Poverty in an Era of Liberalization and Globalization*, Oxford University Press, Chapter 4.
- Cecchi D., García-Peñalosa C.**, 2004, Risk and the distribution of human capital, *Economics Letters* 82, 53-61.
- Cohen D., Soto M.**, 2007, Growth and Human Capital: Good Data, Good Results, *Journal of Economic Growth*, 12, 1, 51-76.
- Cuaresma J.C., Samir K.C., Sauer P.**, 2013, Age-Specific Education Inequality, Education Mobility and Income Growth, *Welfare, Wealth and Work for Europe*, Working Paper no 6.
- Dessus S.**, 2000, Capital humain et croissance : le rôle retrouvé du système éducatif, *Economie publique* 6, 2, 95-114.
- Földvári P., Van Leeuwen B.**, 2014, Educational and income inequality in Europe, ca. 1870-2000, *Cliometrica*, 8, 3, 271-300.
- Galea S., Ahern J., Tracy M., Rudenstine S., Vlahov D.**, 2007, Education Inequality and Use of Cigarettes, Alcohol, and Marijuana, *Drug and Alcohol Dependence*, 90, 1, 4-15.
- Green A.**, 2011, Lifelong Learning, Equality and Social Cohesion, *European Journal of Education*, 46, 2, 228-243.
- Hori T.**, 2011, Educational Gender Inequality and Inverted U-Shaped Fertility Dynamics, *The Japanese Economic Review* 62, 1, 126-150.
- Johnes G.**, 2006, Education and Economic Growth, Working Paper 2006/019 Lancaster University Management School
- Lê F., Ahern J., Galea S.**, 2010, Neighborhood education inequality and drinking behavior, *Drug and Alcohol Dependence* 112, 1, 18-26.
- Lim A.S.K., Tang K.K.**, 2008, Human Capital Inequality And The Kuznets Curve, *The Developing Economies*, 46, 1, 26-51.
- López R., Thomas V., Wang Y.**, 1998, Addressing the Education Puzzle: the Distribution of Education and Economic Reforms, *World Bank Working Paper Series* No. 2031, World Bank, Washington, DC.

- Lutz W., Goujon A., Samir K. C., Sanderson W.**, 2007, Reconstruction of populations by age, sex and level of educational attainment for 120 countries for 1970–2000, *Vienna Yearbook of Population Research* 2007 : 193–235.
- McKenzie D., Rapoport H.**, 2007, Network Effects and the Dynamics of Migration and Inequality: Theory and Evidence from Mexico, *Journal of Development Economics*, 84, 1, 1–24.
- Meschi E. Scervini F.**, 2014, A New Dataset on Educational Inequality, *Empirical Economics*, 47, 2, 695–716.
- Morrisson C., Murtin F.**, 2010, The Kuznets Curve of Education: A Global Perspective on Education Inequalities, Centre for the Economics of Education, *London School of Economics*, CEE DP 116.
- Morrisson C., Murtin F.**, 2013, The Kuznets curve of human capital inequality: 1870–2010, *Journal of Economic Inequality*, 11, 3, 283–301.
- Patrawart K.**, 2010, Can Equality in Education Be a New Anti-Corruption Tool? Cross-Country Evidence, 1990-2005, *NACC Journal*, 3, 2, 127–157.
- Psacharopoulos G., Arriagada A. M.**, 1986, The Educational Composition of the Labour Force: An international Comparison, *International Labour Review*, 125, 5, 561–574.
- Sauer P.**, 2016, "The Role of Age and Gender in Education Expansion". *INEQ Working Paper Series*, 4. WU Vienna University of Economics and Business, Vienna.
- Thomas V., Wang Y., Fan X.**, 2001, Measuring Education Inequality: Gini Coefficients of Education., *Policy Research Working Paper*, N°. 2525, World Bank Institute
- Thomas V., Wang Y., Fan X.**, 2003, Measuring Education Inequality: Gini Coefficients of Education for 140 Countries (1960–2000), *Journal of Education Planning and Administration*, 17, 1, 5–33.

ANNEXE 1.

Synthèse des formules de l'entropie généralisée de l'éducation selon les différents cas de figure

Type	Entropie généralisée de l'éducation
Groupe d'âge quinquennal « a » d'un pays « c »	$GE(r)_{c,t}^a = \frac{N_{c,t}^{a(r-1)}}{r(r-1)} \left[\frac{\sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^r}{\left(\sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a \right)^r} - 1 \right]$
Grand groupe d'âge « G » d'un pays « c »	$GE(r)_{c,t}^G = \frac{N_{c,t}^{G(r-1)}}{r(r-1)} \left[\frac{\sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^r}{\left(\sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a \right)^r} - 1 \right]$
Groupe d'âge quinquennal « a » d'une zone géographique « R »	$GE(r)_{R,t}^a = \frac{N_{R,t}^{a(r-1)}}{r(r-1)} \left[\frac{\sum_{c \in R} \sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^r}{\left(\sum_{c \in R} \sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a \right)^r} - 1 \right]$
Grand groupe d'âge « G » d'une zone géographique « R »	$GE(r)_{c,t}^G = \frac{N_{R,t}^{G(r-1)}}{r(r-1)} \left[\frac{\sum_{c \in R} \sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^r}{\left(\sum_{c \in R} \sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a \right)^r} - 1 \right]$

Synthèse des formules de l'indice de Gini de l'éducation selon les différents cas de figure

Type	Indice de Gini de l'éducation
Groupe d'âge quinquennal « a » d'un pays « c »	$Gini_{c,t}^a = \frac{\sum_{i=1}^n \sum_{j=1}^n n_{c,i,t}^a y_{c,i,t}^a - y_{c,j,t}^a n_{c,j,t}^a}{2 N_{c,t}^a \sum_{i=1}^n n_{c,i,t}^a y_{c,i,t}^a}$
Grand groupe d'âge « G » d'un pays « c »	$Gini_{c,t}^G = \frac{\sum_{a \in G} \sum_{a' \in G} \sum_{i=1}^n \sum_{j=1}^n n_{c,i,t}^a y_{c,i,t}^a - y_{c,j,t}^{a'} n_{c,j,t}^{a'}}{2 N_{c,t}^G \left[\sum_{a \in G} \sum_{i=1}^n n_{c,i,t}^a y_{c,i,t}^a \right]}$
Groupe d'âge quinquennal « a » d'une zone géographique « R »	$Gini_{R,t}^a = \frac{\sum_{c \in R} \sum_{c' \in R} \sum_{i=1}^n \sum_{j=1}^n n_{c,i,t}^a y_{c,i,t}^a - y_{c',j,t}^a n_{c',j,t}^a}{2 N_{R,t}^a \left[\sum_{c \in R} \sum_{i=1}^n n_{c,i,t}^a y_{c,i,t}^a \right]}$
Grand groupe d'âge « G » d'une zone géographique « R »	$Gini_{R,t}^G = \frac{\sum_{c \in R} \sum_{c' \in R} \sum_{a \in G} \sum_{a' \in G} \sum_{i=1}^n \sum_{j=1}^n n_{c,i,t}^a y_{c,i,t}^a - y_{c',j,t}^{a'} n_{c',j,t}^{a'}}{2 N_{R,t}^G \left[\sum_{c \in R} \sum_{a \in G} \sum_{i=1}^n n_{c,i,t}^a y_{c,i,t}^a \right]}$

Synthèse des formules de la famille des indices d'Atkinson de l'éducation selon les différents cas de figure

Type	Indice d'Atkinson de l'éducation ($\epsilon \neq 1$)	
Groupe d'âge quinquennal « a » d'un pays « c »	$A(\epsilon)_{c,t}^a = 1 - \frac{1}{[N_{c,t}^a]^{\frac{\epsilon}{1-\epsilon}}}$	$\frac{\left[\sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^{(1-\epsilon)} \right]^{\frac{1}{1-\epsilon}}}{\sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a}$
Grand groupe d'âge « G » d'un pays « c »	$A(\epsilon)_{c,t}^G = 1 - \frac{1}{[N_{c,t}^G]^{\frac{\epsilon}{1-\epsilon}}}$	$\frac{\left[\sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^{(1-\epsilon)} \right]^{\frac{1}{1-\epsilon}}}{\sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a}$
Groupe d'âge quinquennal « a » d'une zone géographique « R »	$A(\epsilon)_{R,t}^a = 1 - \frac{1}{[N_{R,t}^a]^{\frac{\epsilon}{1-\epsilon}}}$	$\frac{\left[\sum_{c \in R} \sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^{(1-\epsilon)} \right]^{\frac{1}{1-\epsilon}}}{\sum_{c \in R} \sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a}$
Grand groupe d'âge « G » d'une zone géographique « R »	$A(\epsilon)_{R,t}^G = 1 - \frac{1}{[N_{R,t}^G]^{\frac{\epsilon}{1-\epsilon}}}$	$\frac{\left[\sum_{c \in R} \sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a (y_{c,j,t}^a)^{(1-\epsilon)} \right]^{\frac{1}{1-\epsilon}}}{\sum_{c \in R} \sum_{a \in G} \sum_{j=1}^n n_{c,j,t}^a y_{c,j,t}^a}$

Synthèse des formules de l'indice d'Atkinson de l'éducation d'ordre 1 selon les différents cas de figure

Type	Indice d'Atkinson de l'éducation ($\epsilon = 1$)	
Groupe d'âge quinquennal « a » d'un pays « c »	$A(\epsilon)_{c,t}^a = 1 - \frac{\prod_{j=1}^n (y_{c,j,t}^a + 1)^{\frac{n_{c,j,t}^a}{N_{c,t}^a}}}{\left[\sum_{j=1}^n \left(\frac{n_{c,j,t}^a}{N_{c,t}^a} \right) (y_{c,j,t}^a + 1) \right]}$	
Grand groupe d'âge « G » d'un pays « c »	$A(\epsilon)_{c,t}^G = 1 - \frac{\prod_{a \in G} \prod_{j=1}^n (y_{c,j,t}^a + 1)^{\frac{n_{c,j,t}^a}{N_{c,t}^G}}}{\left[\sum_{a \in G} \sum_{j=1}^n \left(\frac{n_{c,j,t}^a}{N_{c,t}^G} \right) (y_{c,j,t}^a + 1) \right]}$	
Groupe d'âge quinquennal « a » d'une zone géographique « R »	$A(\epsilon)_{R,t}^a = 1 - \frac{\prod_{c \in R} \prod_{j=1}^n (y_{c,j,t}^a + 1)^{\frac{n_{c,j,t}^a}{N_{R,t}^a}}}{\left[\sum_{c \in R} \sum_{j=1}^n \left(\frac{n_{c,j,t}^a}{N_{R,t}^a} \right) (y_{c,j,t}^a + 1) \right]}$	
Grand groupe d'âge « G » d'une zone géographique « R »	$A(\epsilon)_{R,t}^G = 1 - \frac{\prod_{c \in R} \prod_{a \in G} \prod_{j=1}^n (y_{c,j,t}^a + 1)^{\frac{n_{c,j,t}^a}{N_{R,t}^G}}}{\left[\sum_{c \in R} \sum_{a \in G} \sum_{j=1}^n \left(\frac{n_{c,j,t}^a}{N_{R,t}^G} \right) (y_{c,j,t}^a + 1) \right]}$	

Source : Auteurs.

ANNEXE 2.

Inégalité de l'éducation par sous-région, 1950-2015

Région	Année	Indices d'entropie généralisée				Indice de Gini	Indices d'Atkinson	
		GE(0)	GE(0.5)	GE(1)	GE(2)		A(0.5)	A(1)
Afrique de l'Est	1950	21.11	2.16	1.61	2.37	0.83	0.79	0.28
	1955	20.36	2.04	1.49	2.03	0.81	0.76	0.30
	1960	19.12	1.85	1.31	1.60	0.77	0.71	0.31
	1965	18.05	1.71	1.18	1.35	0.74	0.67	0.32
	1970	16.77	1.56	1.05	1.12	0.71	0.63	0.33
	1975	15.66	1.43	0.95	0.97	0.68	0.59	0.35
	1980	14.37	1.29	0.84	0.80	0.65	0.54	0.36
	1985	13.03	1.15	0.74	0.67	0.61	0.49	0.36
	1990	11.65	1.02	0.65	0.56	0.57	0.45	0.35
	1995	10.41	0.90	0.57	0.47	0.53	0.40	0.34
	2000	9.40	0.81	0.51	0.41	0.50	0.37	0.32
	2005	8.29	0.72	0.45	0.36	0.47	0.33	0.31
	2010	6.93	0.61	0.39	0.31	0.44	0.28	0.28
2015	5.61	0.51	0.32	0.26	0.40	0.24	0.26	
Afrique centrale	1950	23.20	2.57	2.12	4.30	0.90	0.87	0.22
	1955	22.47	2.41	1.92	3.46	0.87	0.84	0.24
	1960	21.65	2.26	1.73	2.79	0.85	0.81	0.27
	1965	20.63	2.08	1.55	2.25	0.82	0.77	0.30
	1970	19.39	1.90	1.37	1.80	0.78	0.72	0.32
	1975	17.76	1.69	1.18	1.41	0.74	0.67	0.35
	1980	16.03	1.49	1.01	1.10	0.70	0.61	0.37
	1985	14.46	1.32	0.87	0.86	0.66	0.55	0.39
	1990	12.95	1.17	0.77	0.72	0.62	0.50	0.38
	1995	11.83	1.06	0.69	0.63	0.60	0.46	0.37
	2000	11.54	1.03	0.67	0.60	0.59	0.45	0.37
	2005	11.18	1.00	0.65	0.57	0.58	0.44	0.37
	2010	10.99	0.98	0.63	0.55	0.57	0.43	0.37
2015	9.42	0.84	0.54	0.46	0.53	0.38	0.35	
Afrique australe	1950	14.69	1.28	0.80	0.68	0.62	0.54	0.42
	1955	13.83	1.19	0.74	0.61	0.59	0.51	0.41
	1960	13.34	1.14	0.71	0.58	0.58	0.49	0.41
	1965	12.19	1.06	0.66	0.55	0.57	0.46	0.39
	1970	10.94	0.95	0.60	0.50	0.54	0.42	0.37
	1975	9.39	0.82	0.51	0.41	0.50	0.37	0.33
	1980	7.79	0.69	0.43	0.34	0.46	0.31	0.30
	1985	5.84	0.54	0.35	0.29	0.42	0.25	0.25
	1990	3.77	0.36	0.24	0.18	0.34	0.17	0.21
	1995	2.00	0.22	0.15	0.12	0.27	0.11	0.15
	2000	4.42	0.39	0.24	0.17	0.32	0.18	0.23
	2005	3.34	0.30	0.18	0.13	0.27	0.14	0.19
	2010	2.11	0.21	0.13	0.09	0.22	0.10	0.14
2015	1.62	0.16	0.11	0.07	0.18	0.08	0.12	
Afrique de l'Ouest	1950	23.53	2.62	2.20	4.68	0.91	0.88	0.25
	1955	23.26	2.56	2.11	4.26	0.90	0.87	0.26
	1960	22.69	2.43	1.94	3.50	0.88	0.85	0.28
	1965	22.18	2.31	1.79	2.91	0.86	0.82	0.34
	1970	21.53	2.18	1.64	2.37	0.84	0.79	0.39
	1975	20.58	2.03	1.48	1.92	0.81	0.76	0.42
	1980	19.78	1.91	1.35	1.62	0.79	0.73	0.44
	1985	18.82	1.78	1.23	1.37	0.76	0.69	0.45
	1990	17.99	1.67	1.13	1.19	0.73	0.66	0.46
	1995	17.04	1.56	1.04	1.05	0.71	0.63	0.45
	2000	16.16	1.46	0.96	0.93	0.69	0.60	0.45
	2005	15.36	1.37	0.88	0.81	0.66	0.57	0.45
	2010	13.96	1.22	0.77	0.66	0.61	0.52	0.44
2015	12.35	1.05	0.65	0.53	0.56	0.46	0.42	

Région	Année	Indices d'entropie généralisée				Indice de Gini	Indices d'Atkinson	
		GE(0)	GE(0.5)	GE(1)	GE(2)		A(0.5)	A(1)
Afrique du Nord	1950	24.27	2.85	2.60	7.60	0.94	0.92	0.22
	1955	24.02	2.76	2.45	6.45	0.93	0.90	0.25
	1960	23.83	2.69	2.34	5.65	0.92	0.89	0.28
	1965	23.28	2.54	2.12	4.34	0.90	0.87	0.32
	1970	22.24	2.33	1.84	3.14	0.87	0.83	0.36
	1975	20.88	2.10	1.57	2.27	0.84	0.77	0.40
	1980	18.84	1.80	1.27	1.51	0.78	0.70	0.43
	1985	16.53	1.51	1.01	1.01	0.70	0.61	0.44
	1990	14.96	1.33	0.86	0.79	0.65	0.56	0.44
	1995	13.50	1.18	0.74	0.64	0.61	0.50	0.43
	2000	12.08	1.04	0.65	0.53	0.56	0.45	0.41
	2005	10.95	0.93	0.57	0.45	0.53	0.41	0.40
	2010	9.94	0.84	0.51	0.40	0.50	0.38	0.38
	2015	8.72	0.73	0.45	0.33	0.46	0.33	0.36
Asie de l'Ouest	1950	21.16	2.15	1.63	2.44	0.84	0.79	0.37
	1955	20.08	1.99	1.46	1.97	0.81	0.75	0.38
	1960	18.81	1.81	1.28	1.56	0.78	0.70	0.39
	1965	17.45	1.64	1.13	1.27	0.74	0.65	0.40
	1970	16.20	1.49	1.01	1.06	0.71	0.61	0.40
	1975	14.63	1.32	0.87	0.85	0.66	0.55	0.40
	1980	13.08	1.17	0.76	0.69	0.62	0.50	0.40
	1985	10.81	0.96	0.62	0.54	0.57	0.42	0.38
	1990	9.37	0.83	0.53	0.44	0.52	0.37	0.36
	1995	8.04	0.71	0.45	0.37	0.48	0.32	0.33
	2000	6.47	0.59	0.37	0.30	0.44	0.27	0.29
	2005	5.60	0.50	0.32	0.25	0.40	0.24	0.27
	2010	4.75	0.43	0.27	0.21	0.37	0.20	0.25
	2015	4.04	0.37	0.24	0.18	0.34	0.18	0.22
Asie de l'Est	1950	17.07	1.58	1.06	1.09	0.71	0.63	0.38
	1955	15.37	1.39	0.91	0.87	0.67	0.57	0.37
	1960	13.77	1.22	0.78	0.70	0.62	0.52	0.37
	1965	11.66	1.01	0.63	0.53	0.56	0.44	0.35
	1970	9.61	0.82	0.51	0.40	0.50	0.37	0.32
	1975	8.06	0.69	0.42	0.33	0.45	0.32	0.29
	1980	6.16	0.53	0.33	0.25	0.39	0.25	0.26
	1985	5.68	0.49	0.30	0.22	0.38	0.23	0.25
	1990	5.58	0.48	0.30	0.22	0.37	0.23	0.25
	1995	4.07	0.37	0.23	0.17	0.32	0.18	0.21
	2000	2.85	0.27	0.18	0.13	0.28	0.13	0.17
	2005	2.04	0.21	0.14	0.11	0.25	0.10	0.14
	2010	1.42	0.16	0.11	0.09	0.22	0.08	0.11
	2015	0.92	0.12	0.09	0.07	0.21	0.06	0.09
Asie du Sud-Est	1950	18.15	1.74	1.22	1.45	0.76	0.68	0.32
	1955	16.82	1.59	1.09	1.22	0.73	0.64	0.33
	1960	15.57	1.44	0.97	1.02	0.69	0.59	0.33
	1965	13.42	1.23	0.82	0.81	0.65	0.52	0.33
	1970	11.23	1.02	0.67	0.63	0.59	0.45	0.32
	1975	9.66	0.88	0.58	0.52	0.55	0.39	0.31
	1980	7.90	0.74	0.49	0.43	0.51	0.33	0.29
	1985	8.13	0.75	0.50	0.43	0.51	0.34	0.31
	1990	7.93	0.74	0.48	0.41	0.51	0.33	0.31
	1995	6.35	0.60	0.39	0.33	0.46	0.28	0.28
	2000	5.01	0.49	0.33	0.27	0.42	0.23	0.25
	2005	3.68	0.38	0.26	0.21	0.37	0.18	0.22
	2010	2.70	0.30	0.21	0.17	0.33	0.14	0.19
	2015	2.17	0.26	0.18	0.14	0.30	0.12	0.17

Région	Année	Indices d'entropie généralisée				Indice de Gini	Indices d'Atkinson	
		GE(0)	GE(0.5)	GE(1)	GE(2)		A(0.5)	A(1)
Asie centrale	1950	10.97	1.06	0.74	0.78	0.62	0.46	0.34
	1955	9.59	0.94	0.66	0.67	0.59	0.42	0.33
	1960	8.13	0.82	0.58	0.56	0.56	0.37	0.32
	1965	7.01	0.73	0.52	0.49	0.53	0.33	0.31
	1970	5.79	0.63	0.44	0.40	0.49	0.29	0.30
	1975	4.44	0.50	0.35	0.30	0.43	0.24	0.27
	1980	3.40	0.40	0.28	0.22	0.37	0.19	0.24
	1985	2.50	0.31	0.22	0.17	0.32	0.15	0.20
	1990	1.74	0.24	0.17	0.13	0.27	0.12	0.17
	1995	1.00	0.15	0.11	0.08	0.20	0.07	0.12
	2000	0.40	0.08	0.06	0.04	0.14	0.04	0.07
	2005	0.15	0.05	0.04	0.03	0.11	0.02	0.04
	2010	0.11	0.04	0.03	0.02	0.10	0.02	0.03
	2015	0.08	0.03	0.03	0.02	0.10	0.02	0.03
Asie du Sud	1950	20.88	2.14	1.62	2.53	0.84	0.78	0.28
	1955	20.58	2.10	1.57	2.38	0.83	0.77	0.28
	1960	20.21	2.03	1.50	2.16	0.82	0.76	0.29
	1965	19.92	1.98	1.45	1.99	0.81	0.75	0.32
	1970	18.96	1.84	1.30	1.61	0.78	0.71	0.35
	1975	18.82	1.79	1.25	1.43	0.76	0.70	0.39
	1980	18.57	1.73	1.18	1.25	0.74	0.68	0.42
	1985	16.58	1.49	0.98	0.94	0.68	0.61	0.42
	1990	14.84	1.31	0.83	0.74	0.63	0.55	0.41
	1995	13.81	1.20	0.76	0.65	0.61	0.51	0.42
	2000	12.72	1.08	0.67	0.54	0.57	0.47	0.43
	2005	10.83	0.91	0.55	0.42	0.51	0.40	0.39
	2010	9.63	0.80	0.48	0.36	0.47	0.36	0.37
	2015	8.20	0.68	0.40	0.29	0.43	0.31	0.34
Europe de l'Est	1950	4.96	0.51	0.35	0.31	0.43	0.24	0.23
	1955	3.43	0.39	0.28	0.25	0.40	0.19	0.20
	1960	2.91	0.36	0.26	0.23	0.38	0.17	0.19
	1965	2.43	0.31	0.23	0.20	0.36	0.15	0.18
	1970	1.93	0.27	0.20	0.17	0.33	0.13	0.16
	1975	1.86	0.25	0.18	0.15	0.31	0.12	0.16
	1980	1.78	0.22	0.16	0.12	0.28	0.11	0.15
	1985	1.70	0.20	0.14	0.11	0.26	0.10	0.14
	1990	1.56	0.18	0.12	0.09	0.23	0.09	0.13
	1995	1.14	0.14	0.10	0.07	0.20	0.07	0.11
	2000	0.55	0.09	0.06	0.05	0.17	0.04	0.07
	2005	0.39	0.07	0.05	0.04	0.15	0.03	0.06
	2010	0.30	0.06	0.05	0.04	0.14	0.03	0.05
	2015	0.26	0.05	0.04	0.03	0.13	0.02	0.04
Europe du Nord	1950	1.23	0.21	0.17	0.16	0.31	0.10	0.13
	1955	1.11	0.20	0.17	0.16	0.31	0.10	0.13
	1960	1.02	0.19	0.16	0.15	0.30	0.09	0.13
	1965	0.98	0.19	0.16	0.15	0.30	0.09	0.13
	1970	0.93	0.18	0.15	0.14	0.30	0.09	0.13
	1975	0.91	0.18	0.15	0.14	0.29	0.09	0.13
	1980	1.03	0.18	0.14	0.13	0.28	0.09	0.13
	1985	0.96	0.17	0.13	0.12	0.27	0.08	0.12
	1990	0.87	0.16	0.13	0.11	0.27	0.08	0.12
	1995	0.89	0.15	0.12	0.10	0.26	0.07	0.11
	2000	0.89	0.14	0.11	0.10	0.25	0.07	0.11
	2005	0.46	0.10	0.08	0.07	0.20	0.05	0.08
	2010	0.29	0.08	0.06	0.05	0.17	0.04	0.06
	2015	0.31	0.07	0.05	0.04	0.15	0.03	0.06
Europe du Sud	1950	8.69	0.78	0.50	0.43	0.50	0.35	0.29
	1955	8.12	0.73	0.47	0.40	0.49	0.33	0.29
	1960	7.37	0.67	0.43	0.36	0.47	0.30	0.28
	1965	6.58	0.60	0.38	0.32	0.44	0.28	0.27
	1970	5.78	0.53	0.34	0.29	0.42	0.25	0.25
	1975	5.10	0.48	0.31	0.26	0.40	0.22	0.24

	1980	4.38	0.42	0.28	0.22	0.38	0.20	0.23
	1985	4.32	0.40	0.26	0.21	0.36	0.19	0.23
	1990	4.00	0.38	0.24	0.19	0.35	0.18	0.22
	1995	3.41	0.33	0.21	0.17	0.33	0.16	0.20
	2000	2.79	0.28	0.18	0.14	0.30	0.13	0.18
	2005	1.49	0.18	0.13	0.10	0.26	0.09	0.13
	2010	1.34	0.17	0.12	0.10	0.25	0.08	0.12
	2015	0.95	0.14	0.10	0.08	0.23	0.07	0.10
Europe de l'Ouest	1950	1.73	0.23	0.17	0.16	0.31	0.11	0.14
	1955	0.66	0.15	0.13	0.13	0.28	0.07	0.10
	1960	0.64	0.16	0.14	0.14	0.29	0.08	0.11
	1965	0.73	0.17	0.14	0.14	0.29	0.08	0.11
	1970	1.04	0.19	0.16	0.15	0.30	0.09	0.13
	1975	1.09	0.20	0.17	0.16	0.31	0.10	0.14
	1980	1.09	0.21	0.18	0.16	0.32	0.10	0.14
	1985	1.49	0.23	0.18	0.16	0.32	0.11	0.16
	1990	1.79	0.24	0.17	0.14	0.30	0.11	0.16
	1995	1.57	0.20	0.14	0.11	0.27	0.10	0.15
	2000	1.29	0.17	0.12	0.09	0.24	0.08	0.13
	2005	1.10	0.12	0.08	0.06	0.19	0.06	0.10
	2010	0.55	0.08	0.06	0.05	0.16	0.04	0.07
2015	0.43	0.07	0.05	0.04	0.15	0.03	0.05	
Amérique du Nord	1950	0.78	0.18	0.15	0.13	0.29	0.09	0.13
	1955	0.75	0.17	0.14	0.12	0.28	0.08	0.12
	1960	0.68	0.15	0.13	0.11	0.26	0.08	0.12
	1965	0.56	0.12	0.10	0.08	0.22	0.06	0.10
	1970	0.48	0.09	0.07	0.06	0.19	0.05	0.08
	1975	0.42	0.07	0.06	0.05	0.16	0.04	0.06
	1980	0.31	0.06	0.04	0.03	0.14	0.03	0.05
	1985	0.35	0.06	0.05	0.04	0.15	0.03	0.05
	1990	0.38	0.07	0.05	0.04	0.15	0.03	0.06
	1995	0.21	0.04	0.03	0.03	0.12	0.02	0.04
	2000	0.17	0.03	0.03	0.02	0.11	0.02	0.03
	2005	0.16	0.03	0.03	0.02	0.11	0.02	0.03
	2010	0.13	0.03	0.02	0.02	0.10	0.01	0.03
2015	0.12	0.03	0.02	0.02	0.09	0.01	0.02	
Région	Année	Indices d'entropie généralisée				Indice de Gini	Indices d'Atkinson	
		GE(0)	GE(0.5)	GE(1)	GE(2)		A(0.5)	A(1)
Amérique centrale	1950	13.49	1.26	0.86	1.02	0.64	0.53	0.30
	1955	12.85	1.19	0.82	0.94	0.63	0.51	0.30
	1960	12.18	1.13	0.78	0.87	0.62	0.49	0.30
	1965	11.22	1.06	0.73	0.82	0.61	0.46	0.32
	1970	10.14	0.96	0.66	0.69	0.58	0.42	0.31
	1975	9.57	0.90	0.62	0.61	0.57	0.40	0.33
	1980	8.44	0.80	0.53	0.49	0.53	0.36	0.33
	1985	6.87	0.66	0.44	0.38	0.48	0.30	0.31
	1990	5.54	0.54	0.36	0.30	0.44	0.25	0.28
	1995	4.59	0.46	0.31	0.25	0.40	0.22	0.25
	2000	3.78	0.39	0.27	0.22	0.37	0.19	0.23
	2005	2.94	0.32	0.22	0.17	0.33	0.15	0.20
	2010	2.58	0.29	0.20	0.16	0.32	0.14	0.19
2015	2.24	0.25	0.17	0.14	0.30	0.12	0.17	
Amérique du Sud	1950	12.86	1.16	0.77	0.75	0.62	0.50	0.34
	1955	11.89	1.08	0.71	0.69	0.60	0.47	0.33
	1960	10.75	0.98	0.65	0.62	0.58	0.43	0.32
	1965	9.56	0.88	0.59	0.55	0.56	0.39	0.31
	1970	8.13	0.77	0.51	0.47	0.52	0.35	0.30
	1975	6.27	0.66	0.48	0.48	0.51	0.30	0.28
	1980	5.93	0.64	0.46	0.45	0.51	0.29	0.28
	1985	5.42	0.56	0.39	0.35	0.47	0.26	0.27
	1990	4.91	0.50	0.35	0.30	0.44	0.24	0.26
	1995	4.36	0.44	0.30	0.24	0.40	0.21	0.24
	2000	3.65	0.37	0.25	0.19	0.36	0.18	0.22
	2005	2.76	0.29	0.19	0.15	0.31	0.14	0.18
	2010	2.11	0.23	0.16	0.12	0.28	0.11	0.16
2015	1.74	0.20	0.14	0.11	0.26	0.10	0.14	

Caraïbes	1950	11.73	1.03	0.65	0.58	0.56	0.45	0.29
	1955	11.24	0.98	0.62	0.55	0.55	0.43	0.29
	1960	10.49	0.92	0.58	0.51	0.53	0.41	0.29
	1965	9.65	0.85	0.55	0.48	0.52	0.38	0.30
	1970	8.63	0.77	0.50	0.43	0.50	0.35	0.30
	1975	7.35	0.67	0.44	0.37	0.48	0.31	0.29
	1980	6.00	0.56	0.37	0.30	0.44	0.26	0.28
	1985	4.85	0.48	0.32	0.26	0.41	0.22	0.26
	1990	4.49	0.45	0.30	0.24	0.39	0.21	0.25
	1995	4.12	0.42	0.28	0.22	0.38	0.20	0.24
	2000	3.81	0.39	0.26	0.20	0.36	0.19	0.23
	2005	3.64	0.37	0.24	0.19	0.35	0.18	0.23
	2010	3.41	0.35	0.23	0.17	0.33	0.17	0.22
2015	2.96	0.31	0.21	0.16	0.32	0.15	0.21	
Océanie (sauf Australie et Nouvelle-Zélande)	1950	22.47	2.38	1.88	3.26	0.87	0.84	0.31
	1955	21.73	2.24	1.72	2.70	0.85	0.81	0.33
	1960	20.59	2.06	1.52	2.13	0.82	0.76	0.35
	1965	19.24	1.87	1.33	1.66	0.78	0.72	0.37
	1970	17.91	1.70	1.18	1.34	0.75	0.67	0.39
	1975	17.05	1.58	1.06	1.11	0.72	0.63	0.41
	1980	15.99	1.44	0.95	0.91	0.68	0.59	0.42
	1985	14.55	1.27	0.81	0.71	0.63	0.54	0.41
	1990	13.40	1.15	0.71	0.59	0.58	0.49	0.41
	1995	12.17	1.03	0.63	0.51	0.55	0.45	0.39
	2000	10.89	0.93	0.58	0.47	0.53	0.41	0.36
	2005	9.75	0.81	0.49	0.38	0.47	0.36	0.34
	2010	9.38	0.78	0.47	0.36	0.46	0.35	0.33
2015	9.59	0.80	0.48	0.37	0.47	0.36	0.34	
Australie et Nouvelle-Zélande	1950	0.50	0.13	0.11	0.10	0.25	0.06	0.09
	1955	0.45	0.12	0.10	0.09	0.25	0.06	0.09
	1960	0.42	0.12	0.10	0.09	0.24	0.06	0.09
	1965	0.39	0.11	0.09	0.08	0.22	0.05	0.08
	1970	0.35	0.09	0.08	0.07	0.20	0.05	0.08
	1975	0.35	0.07	0.06	0.05	0.16	0.04	0.06
	1980	0.30	0.06	0.05	0.04	0.14	0.03	0.05
	1985	0.34	0.06	0.05	0.04	0.15	0.03	0.05
	1990	0.32	0.07	0.05	0.04	0.16	0.03	0.06
	1995	0.39	0.07	0.06	0.05	0.16	0.04	0.06
	2000	0.32	0.07	0.05	0.04	0.16	0.03	0.06
	2005	0.29	0.06	0.05	0.04	0.15	0.03	0.05
	2010	0.29	0.06	0.05	0.04	0.16	0.03	0.05
2015	0.31	0.06	0.05	0.04	0.16	0.03	0.05	

Source : Auteurs, base mondiale des inégalités d'éducation 1950-2015.

Educational inequalities: A new world database (1950-2015)

Abstract - A new world database on educational inequalities is proposed, inspired by the dataset of Benaabdelali et al. (2012) and based on the 2016 version of the database of Barro and Lee (2013). Our database covers the period 1950-2015 for 146 countries and provides an estimation of educational inequality of the population by 5-year age intervals and by aggregated age groups, considering seven levels of education. In addition to the Gini index, our database presents several indices of the generalized entropy family and the Atkinson family. It provides more accurate data by correcting the weighting error made by Barro and Lee in calculating the average years of schooling which can lead to a significant bias for many countries. From this database a panorama of the reduction of educational inequalities at the global level is schematically presented. It is mainly linked in developing countries to the decline of the population with no schooling. Nevertheless, the gap remains high between countries with low levels of human development in some parts of the world and those with high human development.

Mots-clés

Educational inequalities
Theil index
Gini index
World database
