Working Paper

A RE-EXAMINATION OF THE EXPECTED YEARS OF SCHOOLING:

WHAT CAN IT TELL US?

Working Paper number 117 November, 2013

José Irineu Rangel Rigotti
Centre of Regional Development and Planning (CEDEPLAR) and
Department of Demography, Federal University of Minas Gerais (UFMG)

Diana Oya Sawyer
Laetícia Rodrigues de Souza
Clarissa Guimarães Rodrigues
International Policy Centre for Inclusive Growth, UNDP

International Policy Centre for Inclusive Growth

International Policy Centre for Inclusive Growth (IPC - IG)
United Nations Development Programme

SBS, Quadra 1, Bloco J, Ed. BNDES, 13º andar
70076-900 Brasilia, DF - Brazil
Telephone: +55 61 21055000

E-mail: ipc@ipc-undp.org ▪ URL: www.ipc-undp.org

The International Policy Centre for Inclusive Growth is jointly supported by the United Nations Development Programme and the Government of Brazil.

Working Papers are available online at www.ipc-undp.org and subscriptions can be requested by email to ipc@ipc-undp.org

Print ISSN: 1812-108X

A RE-EXAMINATION OF THE EXPECTED YEARS OF SCHOOLING:

WHAT CAN IT TELL US?

José Irineu Rangel Rigotti,[*] Diana Oya Sawyer,[**]

Laetícia Rodrigues de Souza[***] and Clarissa Guimarães Rodrigues[***]

1 INTRODUCTION

Expected Years of Schooling (EYS), one of the components of the education indicator in the Human Development Index (HDI) (UNDP, 2010), is a measure of the number of years of schooling a child at the start of his or her education is expected to receive, if current rates of enrolment are maintained throughout the child's life (UNESCO, 2009). The advantages of using this indicator are that it represents a measure that takes into account both stock and flow dimensions in the school system and is easy to calculate and does not require standardisation in comparisons involving countries with distinct age structures.

EYS is an important component of HDI rankings by country, since the higher its values, the higher a given country is ranked given the other components. However, in its current formulation it only takes into account enrolment rates, and overlooks the progression structure in the school system. Thus, if the indicator is intended to be a proxy for the knowledge accumulated in schools (Hanushek, 2013; Barakat, 2012; Rios-Neto and Guimarães, 2010), which the HDI's education dimension seeks to represent, results may be misleading, since the structure of school progression is not taken into consideration.

* Associate Professor , Centre of Regional Development and Planning (CEDEPLAR) and Department of Demography. Federal University of Minas Gerais (UFMG).

** Senior Researcher, International Policy Centre for Inclusive Growth, UNDP.

*** Associate Researcher, International Policy Centre for Inclusive Growth, UNDP.

The authors greatly appreciate the fine translation from Portuguese by Eric Sawyer as well as the relevant comments provided by Fabio Veras and Michael MacLennan of the UNDP International Policy Centre for Inclusive Growth.

This article proposes adjustments to the indicator, to bring it closer to the meaning of schooling, assuming that completion of successive school years can be considered an approximation of greater or lesser education. In addition to age and rate of enrolment, the adjustment also takes into consideration a third component, which is the grade of enrolment. This adjusted indicator for comparative purposes works better if calculated up to an age compatible with the completion of the last cycle in the school system. For ease of presentation, we shall label the HDI indicator EYS and the indicator adjusted by the contribution of delayed students, proposed here, Adjusted Expected Years of Schooling (AEYS).

2 EXPECTED YEARS OF SCHOOLING

EYS, at the age α at the start of the educational trajectory, are calculated by adding up specific rates of enrolment by age weighted by the respective amplitude of the age group, measured in year *t*, according to formula (1).

$$EYS_{\alpha} = \sum_{\alpha}^{\omega} n \times {}_{n}m_{x} \quad (1)$$

where

$${}_{n}m_{x} = \frac{{}_{n}f_{x}}{{}_{n}P_{x}}$$

and:

α = age at the start of school trajectory

ω = upper age limit

n = age interval

${}_{n}f_{x}$ = number of pupils between ages x and x+n enrolled in school, in year t

${}_{n}P_{x}$ = population between ages x and x+n, in year t

${}_{n}m_{x}$ = rate of enrolment of pupils between ages x and x+n, in year t

The relationship ($n \times {}_{n}m_{x}$), measured in year *t*, represents the rate of enrolment weighted by the corresponding age interval. It indicates the total amount of time the set of pupils between ages *x* and *x*+*n* were enrolled, assuming they remained in school throughout the year.

As an example, if the ${}_nm_x$ rate in 2012, of children between ages 10 and 14, were 0.60, by multiplying this by n=5, we reach 3.00. This means that 10-year-old children, before turning 11, were in school for a year at the rate of 0.60, and that 11-year-olds, before turning 12, were in school for a year at the same rate. The same rationale applies for 12-, 13- and 14-year-olds. The set of pupils in this age group contributed 5 x 0.60 — i.e. 3 years' worth of enrolment.

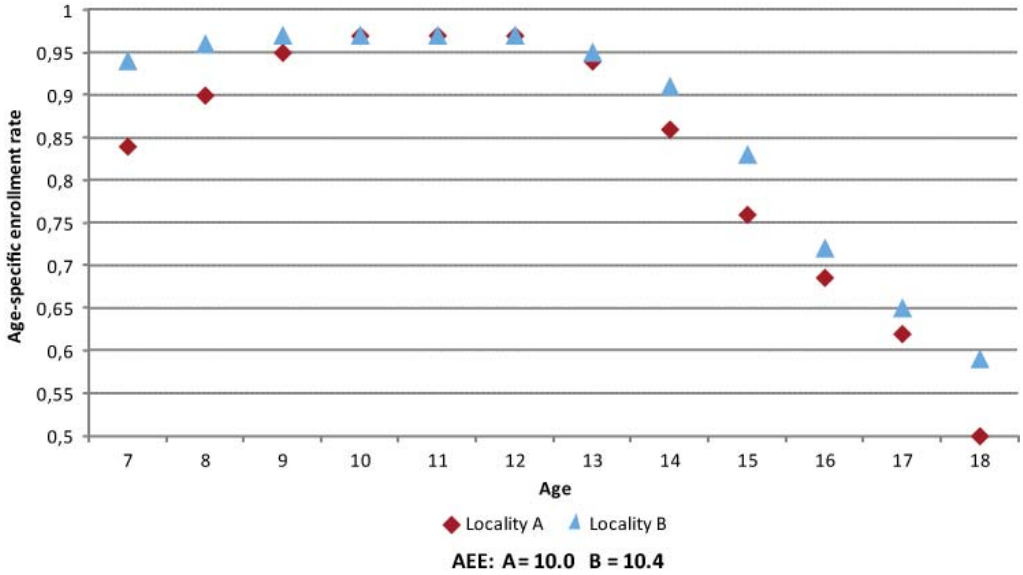
When the age interval is set to 1 (n=1), the EYS would be the sum of enrolment rates. This is the form employed by UNESCO and used in the composition of the international HDI.

One of the shortfalls of the EYS is that it does not capture enrolment structures of the age-specific rates, which in turn could represent different flows of promotions, repetitions, late entries and school drop-outs — therefore, creating comparative difficulties, especially if there are different policies for promotion to higher grades among regions under comparison.

Figure 1 illustrates enrolment structures in two locations with similar EYS values; 10 years of schooling in Locality A and 10.4 in Locality B. It should be noted that in B, nearly all children were in school at the official starting age, with high enrolment rates at early ages and significant decreases taking place only at the age of 15, corresponding to entry into high school. In Location A, children begin school late, with the enrolment peaks taking place between the ages of 10 and 12, and at the age of 15 there is a more significant decrease, possibly due to high rates of repetition and drop-out. These situations have distinct age-grade distortions that are not visible within EYS values. In a context of elevated age-grade gap indices, the indicator would overestimate expected years of schooling, since repeating students would have the same weight as regular students. A regular student contributes 1 year to average schooling each year he or she passes, thus having a weight of 1. The late student, in turn, contributes with a fraction of a year, which will be proportionally lower according to the age-grade gap.

FIGURE 1

Age-specific Enrolment Rates in Two Hypothetical Locations with Similar EYS



Source: Developed by the authors based on the 2010 Demographic Census.

3 ADJUSTED EXPECTED YEARS OF SCHOOLING

To overcome the problem of deviations not detected by EYS, weights are proposed for specific frequency rates according to students' contribution of schooling years, adding a new variable: grade. Thus, by comparing age and grade, it is possible to determine the weight of specific rates. Formula 2 calculates AEYS. It should be noted that henceforth n=1 will be considered.

$$AEEA_{\alpha} = \sum_{\mathbf{x}=\boldsymbol{\alpha}}^{\boldsymbol{\omega}} \sum_{i=1}^{z} \left(\frac{a_{i,x}}{r_{i,x}}\right) \times \frac{f_{i,x}}{p_x} \qquad (2)$$

i = grade of pupils aged x

z = highest grade finished by pupils at age x

$a_{i,x}$ *= years of schooling concluded by pupils, up to grade i at age x*

$r_{i,x}$ *= years of schooling a regular pupil would have concluded by grade i at age x*

$f_{i,x}$ *= number of pupils enrolled in grade i at age x*

p_x *= population at age x*

Each of these weighted rates is simply a substitution for *n* in equation (1) by the contributed years of schooling, adjusted by the age-grade distortion. The adjusted indicator is a better representation of the contribution, at each age, to EYS during the (ω-α) years in school.

The ratio $\left(\frac{a_{i,x}}{r_{i,x}}\right)$ is the adjustment factor, and its values adjusted to the Brazilian school system can be found in Table A1 in the Appendix. The indicators EYS and AEYS were calculated using these factors, between the ages of 6 and 18, for Brazilian states. Data refer to school attendance according to the 2010 Demographic Census (IBGE, 2012). Once indicators were found, their values were ranked to place the highest value in the first position and the lowest in 27th position. Results are presented in Table 1 and Figure 2.

4 FINDINGS

Table 1 shows that EYS in Brazil calculated from the age of 6 to 18 varied between 9.10 years in Acre and 10.27 in Piauí. After adjusting according to the age-grade gap, the lowest value remained in Acre, although with a far lower value (7.52), while the highest was São Paulo, with 9.02 years of schooling. What draws the most attention to these findings is the disparity of positions in the state rankings when comparing EYS and AEYS. Notably, Piauí went from first to 16th position, while São Paulo went from 16th to first. Possible causes for this inversion are high rates of repetition in Piauí, keeping students in school for more years, as well as the higher level of promotion in São Paulo, due to that state's policy for promotion in cycles, in which students are not retained in one grade even if they fail exams but, rather, are promoted between multi-year cycles. This means that grade promotion is automatic within each cycle.

TABLE 1

Expected Years of Schooling (EYS) and Adjusted Expected Years of Schooling (AEYS) and Respective Positions in Ranking, Brazilian States, 2010

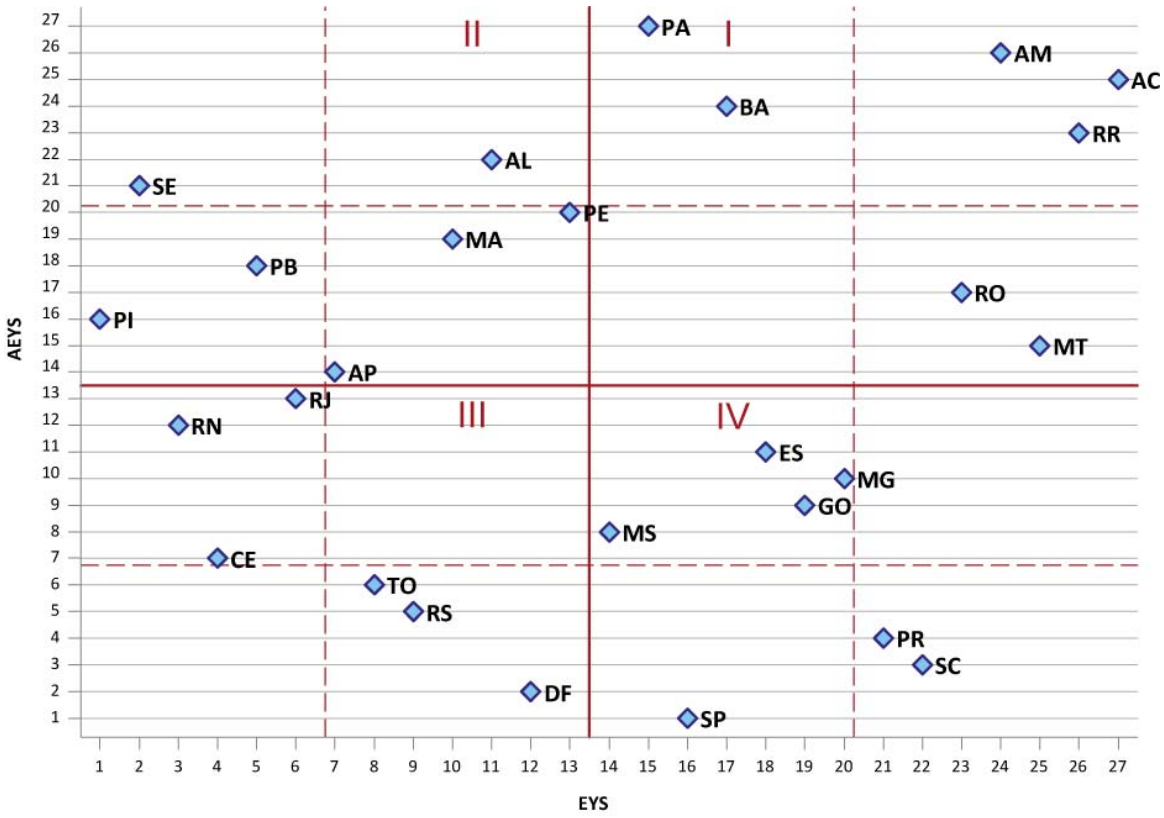
Brazilian states	EYS		AEYS		EYS - AEYS	
	Value	Ranking	Value	Ranking	Value	Ranking
Piauí	10.27	1	8.22	16	2.05	-15
Sergipe	10.16	2	8.03	21	2.13	-19
Rio Grande do Norte	10.12	3	8.43	12	1.69	-9
Ceará	10.12	4	8.68	7	1.44	-3
Paraíba	10.04	5	8.15	18	1.89	-13
Rio de Janeiro	10.03	6	8.36	13	1.67	-7
Amapá	9.96	7	8.32	14	1.64	-7
Tocantins	9.95	8	8.69	6	1.26	2
Rio Grande do Sul	9.94	9	8.77	5	1.17	4
Maranhão	9.92	10	8.14	19	1.78	-9
Alagoas	9.91	11	7.99	22	1.92	-11
Distrito Federal	9.87	12	8.82	2	1.05	10
Pernambuco	9.84	13	8.13	20	1.71	-7
Mato Grosso do Sul	9.81	14	8.65	8	1.16	6
Pará	9.79	15	7.52	27	2.27	-12
São Paulo	9.78	16	9.02	1	0.76	15
Bahia	9.75	17	7.73	24	2.02	-7
Espírito Santo	9.65	18	8.44	11	1.21	7
Goiás	9.64	19	8.58	9	1.06	10
Minas Gerais	9.60	20	8.49	10	1.11	10
Paraná	9.54	21	8.78	4	0.76	17
Santa Catarina	9.54	22	8.82	3	0.72	19
Rondônia	9.51	23	8.16	17	1.35	6
Amazonas	9.37	24	7.53	26	1.84	-2
Mato Grosso	9.24	25	8.22	15	1.02	10
Roraima	9.17	26	7.81	23	1.36	3
Acre	9.10	27	7.54	25	1.56	2

Source: Developed by the authors based on the 2010 Demographic Census.

Figure 2 depicts the overall changes in the ranks of the Brazilian states by plotting them according to their rankings by AEYS and EYS. In general, the four quadrants group the states according to their respective ranks before and after adjustment.

FIGURE 2

Position of States According to Expected Years of Schooling (EYS) and Adjusted Expected Years of Schooling (AEYS) — Brazil, 2010



Source: Developed by the authors based on the 2010 Demographic Census.

Quadrant I corresponds to states that were lower ranked than the halfway mark and remained there. Among those states, there is no clear pattern of improvement or worsening within the rankings. Thus, Mato Grosso (MT) and Rondônia (RO) improved in the ranking, whereas Bahia (BA) and Pará (PA) went down, and Roraima (RR), Acre (AC) and Amazonas (AM) were in the bottom places and remained there after the adjustment.

Quadrant II includes states with worsened ranks after the age-grade adjustment. These are predominantly in the North-East; exceptions are Amapá (AP) from the North, Rio Grande do Norte (RN) and Bahia (BA), both from the North-East but placed in Quadrant III. The following states stand out: Piauí (PI), Paraíba (PB) and Sergipe (SE), which went from 1st, 5th and 2nd places, respectively, to 16th, 18th and 21st.

Quadrant III contains states with ranks higher than the halfway mark (13.5) and which remained there. As in Quadrant I, there is no clear pattern of improving or worsening positions. They are: Distrito Federal (DF), Rio Grande do Sul (RS), Tocantins (TO), Ceará (CE), Rio Grande do

Norte (RN) and Rio de Janeiro (RJ). The following states improved: DF, from 12th to 2nd; RS, from 9th to 5th; TO, from 8th to 6th. Those that went down in rank were: CE, from 4th to 7th, RN, from 3rd to 12th, and RJ, from 6th to 13th.

Quadrant IV of the chart shows states that moved up in the ranking after the adjustment was made. These states were in positions higher than 13 and shifted to positions under 13. These states are in the South and South-East, with the exception of: Rio de Janeiro (RJ), which is in the South-East but is placed in Quadrant III, and Goiás (GO) and Mato Grosso do Sul (MS), which are in the Center-West region. The following states stand out: São Paulo (SP), Santa Catarina (SC) and Paraná (PR) due to the positions they previously occupied, 16th, 22nd and 21st, shifting into 1st, 3rd and 4th places, respectively.

This exercise clearly shows that Brazil, still characterised by a large age-grade gap, should not use EYS to rank states. The proposed adjustments are capable of reorganising states into a ranking that appears more plausible considering the development levels in those that changed for better or for worse. Regarding educational policies, states in Quadrants I and II, in addition to RJ and RN (from Quadrant III) are worthy of attention. Those in Quadrant I with very low positions before and after the adjustment might indicate inefficient attraction and maintenance of children in the school system. Those in Quadrant II and RJ and RN from Quadrant III deserve attention due to their high age-grade gaps assessed by the high differences between EYS and AEYS.

5 CLOSING COMMENTS

The ambiguity created by calculation of EYS may inappropriately inflate HDI results. The use of AEYS seeks to adapt the concept of expected years of schooling to the reality of developing countries and regions, where school systems evolve, to attenuate some of their main shortcomings, such as high repetition rates and late starts. In these contexts, while the proposed indicator does not give late years the same weight as those pupils who are in the appropriate grades for their ages, it is not detrimental to comparisons among different regions and periods.

In fact, when a school system manages to improve its flow of students from one grade to another by reducing repetition, the EYS may go down if promotion is not improved in parallel. In a situation such as this, a reading of the indicator could transmit the wrong idea of a deteriorating situation, when in truth the region would have improved the efficiency of its teaching system. This may currently be taking place in Brazil. Conceptually, the proposed indicator should be interpreted as an expectation — i.e. the average number of years of schooling a cohort entering school should reach, if the school system remains unaltered with regards to repetition, promotion and drop-out, from start to finish in the school trajectory.

The AEYS bear similarities with the indicator of 'average years of study', traditionally used in many demographic studies. With no changes in the school system regarding transition rates (promotion, repetition and drop-out), average years of study for a population precisely at the age of 18 would be exactly the same as the indicator proposed in this article. However, in situations with changing school flows, this is not what happens, as is the case in Brazil, particularly since the mid-1990s, when increases in promotion and decreases in repetition and drop-out took place. In this situation, the average years of study for the population

at the age of 18 is lower than the EYS, since it reflects a past experience, different from the one that new generations entering school will have. Therefore, calculation of EYS seeks to more adequately reflect more recent conditions of the current education system.

From an operational standpoint, calculation of the adjusted indicator has the advantage of not requiring information about years of study completed by those out from school, which is fundamental, considering that in Brazil the Demographic Census stopped collecting such data in 2010. The interpretation of synthetic cohorts assumes, for the generation entering the school system, the same current rates of promotion, repetition and drop-out. In other words, the prior experience of those out of school is portrayed by cohorts at younger ages, attending school during the period of time in question.

ANNEX

1 METHODOLOGICAL NOTES

MEASUREMENT OF PERIOD, REAL COHORT AND SYNTHETIC COHORT

The EYS concept involves transporting rates calculated in year *t* — i.e. period rates for a synthetic cohort — which would have the same period rates observed in year *t* during the school trajectory. In this cohort, the example mentioned in the text "if the rate ${}_nm_x$ in 2012, of children between the ages of 10 and 14, were 0.60, by multiplying this by n=5, we reach 3.00", the product ($n \,.\, {}_5m_{10} = 5 \,.\, 0.60 = 3$) can be interpreted as:

1. 60 per cent of the cohort, on average, was enrolled in school for 5 years, between the ages of 10 and 14.
2. The cohort was enrolled, on average, 60 per cent of the 5 years — i.e. 3 years, between the ages of 10 and 14.

The interpretation of interest to us is the second one, since the sum of average years, at successive ages, yields the average of the total enrolled years this cohort would have completed on finishing its school trajectory.

The difference between EYS in a real cohort and a synthetic cohort is that a real cohort is prospectively followed. The actual years of schooling are observed and known from age α to age ω. However, this procedure requires a long time of observation over the span of the cohort's time in school. Alternatively, it is possible to carry out a retrospective reconstitution of this cohort, but data used to do so would be outdated for older individuals.

The synthetic cohort borrows measurements of several real cohorts, observed during a period *t*, projected as those it would have for ages at the respective times. Since period measurements can be taken in a single year and refer to data that are updated for that year, the synthetic cohort makes it possible to overcome the difficulties of real cohorts of lengthy observation times and outdated data. EYS can also be calculated for periods as short as a single year. Its measurements can also be updated yearly, once for each synthetic cohort, without the need for concern about what happened to the previous cohort, except for comparative purposes to observe short- or long-term changes in EYS.

A real cohort provides the synthetic cohort with a measure of a single age or age group. A real cohort starting school in 2010, at the age of 7 years, will contribute to the synthetic cohort, in 2010, with the rate of enrolment of 7-year-old children; contributions from other ages are given by real cohorts starting in previous years. In 2011, the synthetic cohort will have the contribution of the real cohort at age 8 (age 7 in the previous year); the age 7 one will be given by the real cohort starting in 2011 and those of other ages by real cohorts starting before 2010. Figure A1 illustrates this fact. This example is necessary to show that adjustment factors are not cumulative in a synthetic cohort, since measurements for successive ages in the real cohorts are part of different EYS.

FIGURE A1

Diagram Representing the Composition of Multicohorts, in the Period Measurement and in the Synthetic Cohort

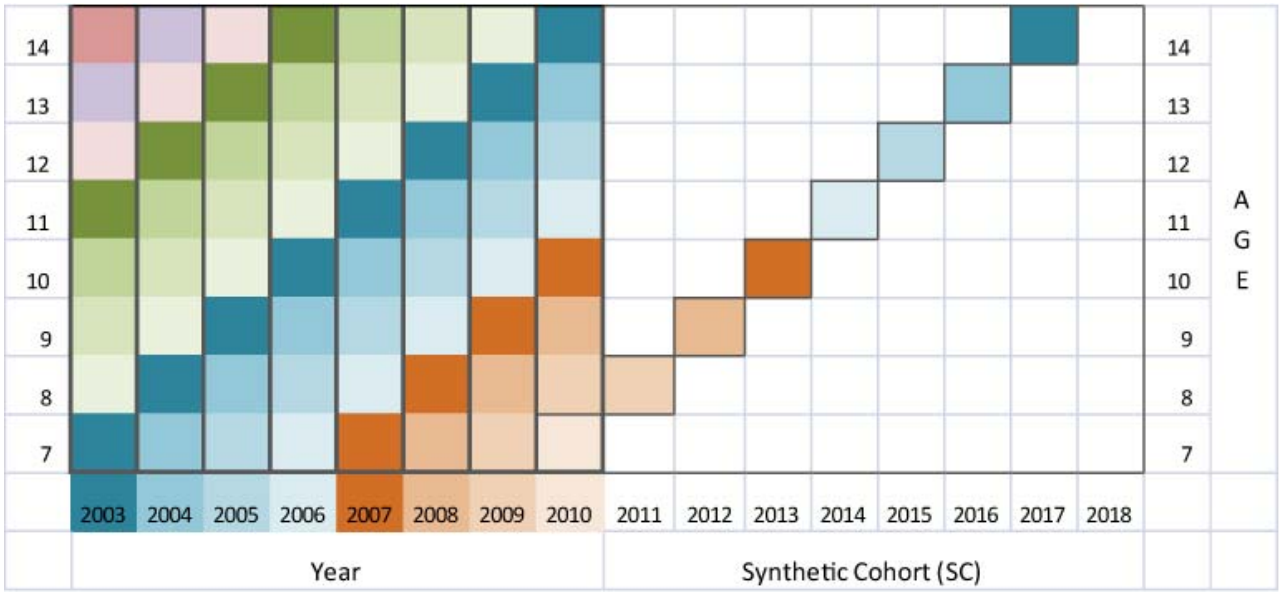


Figure A1 represents in each coloured diagonal cohorts joining school, between 2003 and 2010, at 7 years of age.* When making period measurements of enrolments in 2010, values for children in cohorts between 2003 and 2010 will be used. Each cohort contributes to measurement of a single, specific age. For example, 2003 contributes to the 14-year-old cohort. This same cohort, at the age of 13, contributed to the 13-year-old measurement for 2009 and will be part of another EYS. Period measurements for 2010 will be calculated for a synthetic cohort that supposedly has the same rates per age during the school trajectory. It should be noted that in a situation with constant cohort rates over an extended period, the values for the real and the synthetic cohorts are the same.

2 EYS AND AEYS

To calculate the AEYS, enrolment data for each age must be disaggregated by late school years, considering the grade of enrolment and the grade the pupil should be in. Data in this table are multiplied by the appropriate adjustment factors. The AEYS is obtained by dividing the adjusted enrolment date by the respective mid-year population. The EYS, on the other hand, utilises the value of one for the factor for all enrolment data, resulting in the division of the enrolled students at a certain age, regardless of grade, by the respective mid-year population.

* Figure A1 differs from the classic Lexis Diagram, and, for purposes of presentation, we may assume that measurements were taken at the end of each year, when all students had turned the respective ages.

In Brazil, tabulation of school attendance, in grades and according to age, resulted in a percentage of students considered young for their grades. The decision was to consider them regular students (the expected age for each respective grade), assuming that different local education systems still in existence throughout the country could be the cause.

TABLE A1

Contribution of School Years (adjustment factor) by Age and Years of Delay in the Appropriate Grades

CURRENT SCHOOL SYSTEM = Basic Education in 9 years													
Age	CUMULATIVE contribution in years of schooling by age	Years of delay in the appropriate grade for the age											
		Regular	1 year delayed	2 year delayed	3 year delayed	4 year delayed	5 year delayed	6 year delayed	7 year delayed	8 year delayed	9 year delayed	10 year delayed	11 year delayed	Average
6	0.00	0.00	-	-	-		-	-	-	-	-	-	-	
7	1.00	1.00	0.00	-	-	-	-	-	-	-	-	-	-	
8	2.00	1.00	0.50	0.00	-	-	-	-	-	-	-	-	-	0.50
9	3.00	1.00	0.67	0.33	0.00	-	-	-	-	-	-	-	-	0.50
Average														0.50

PREVIOUS SCHOOL SYSTEM = Basic Education in 8 years													
Age	CUMULATIVE contribution in years of schooling by age	Years of delay in the appropriate grade for the age											
		Regular	1 year delayed	2 year delayed	3 year delayed	4 year delayed	5 year delayed	6 year delayed	7 year delayed	8 year delayed	9 year delayed	10 year delayed	11 year delayed	Average
10	3.00	1.00	0.67	0.33	0.00	-	-	-	-	-	-	-	-	0.50
11	4.00	1.00	0.75	0.50	0.25	0.00	-	-	-	-	-	-	-	0.50
12	5.00	1.00	0.80	0.60	0.40	0.20	0.00	-	-	-	-	-	-	0.50
13	6.00	1.00	0.83	0.67	0.50	0.33	0.17	0.00	-	-	-	-	-	0.50
14	7.00	1.00	0.86	0.71	0.57	0.43	0.29	0.14	0.00	-	-	-	-	0.50
15	8.00	1.00	0.88	0.75	0.63	0.50	0.38	0.25	0.13	0.00	-	-	-	0.50
16	9.00	1.00	0.89	0.78	0.67	0.56	0.44	0.33	0.22	0.11	0.00	-	-	0.50
17	10.00	1.00	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.00	-	0.50
18	11.00	1.00	0.91	0.82	0.73	0.64	0.55	0.45	0.36	0.27	0.18	0.09	0.00	0.50
Avearge														0,50

Table A1 contains the adjustment factors for students between the ages of 6 and 18. A change in the Brazilian school system took place in 2006, shifting the starting age from 7 to 6 years. The table incorporates this change, separating adjustment factors by ages in the new and previous school system, for ease of application. There would be only one table for situations in which the system remains constant, with 6 as the enrolment age, and adjustment factors, starting at the age of 11, should be shifted to the lower age.

REFERENCES

Barakat, B. (2012). 'A Second Look at "School-Life Expectancy"', *International Journal of Educational Development*, Vol. 32, No. 4: 564–574.

Hanushek, E. (2013). *Economic Growth in Developing Countries: The Role of Human Capital*. Stanford, CA, Stanford University, <http://hanushek.stanford.edu/sites/default/files/publications/Education%20and%20Economic%20Growth.pdf> (accessed 23 October 2013).

IBGE (2012). *Censo Demográfico 2010: Educação e Deslocamento: resultados da amostra*. Brasília, Instituto Brasileiro de Geografia e Estatística.

Rios-Neto, E.L.G. and R.R.M. Guimarães (2010). 'The demography of education in Brazil: inequality of educational opportunities based on Grade Progression Probability (1986–2008)', *Vienna Yearbook of Population Research*, Vol. 8: 283–312.

UNDP (2010). *Human Development Report 2010 – 20th Anniversary Edition. The Real Wealth of Nations: Pathways to Human Development*. New York, United Nations Development Programme.

UNESCO Institute for Statistics (2009). *Education Indicators: Technical Guidelines*. Paris, United Nations Educational, Scientific and Cultural Organization, <http://uis.unesco.org/Library/Documents/eiguide09-en.pdf> (accessed 23 October 2013).

International Policy Centre for Inclusive Growth (IPC - IG)
United Nations Development Programme

SBS, Quadra 1, Bloco J, Ed. BNDES, 13º andar
70076-900 Brasilia, DF - Brazil
Telephone: +55 61 2105 5000

E-mail: ipc@ipc-undp.org ▪ URL: www.ipc-undp.org