Journal of Advanced Research in Social Sciences and Humanities
Volume 2, Issue 3 (221-229)
DOI: https://dx.doi.org/10.26500/JARSSH-02-2017-0307

Quantification of Educational Inequality through the Application of Gini Coefficient in Educational Indices

SATOSHI WATARAI*, **SAKI TOMITA**
Graduate School of Frontier Science, The University of Tokyo, Tokyo, Japan
Department of Public Affairs, Mitsubishi Research Institute, Tokyo, Japan

Abstract

Aim: This paper aims to show why it is important to keep an eye on educational disparities across countries and how to do so.
Methodology: The paper uses the Gini coefficient of education, a method suggested by Hojo (2009) and Thomas, Wang, and Fan (2001). The degree of inequality in educational attainment was measured using this technique, and its historical development in Japan, Korea, and the United States was followed. Then, the Gini index was contrasted with the mean years people spend in school to see how the two measures evolved over time.
Findings: The results showed that while educational inequality increased at varying rates across the three countries, the mean number of years people spent in school rose steadily. Second, it discovered that the Gini coefficient of Korea and the United States differed significantly, even though both countries have a sizable population that is highly educated.
Novelty/Implications: This is a rare paper because it not only makes a case for why it's important to pay attention to educational inequality but also explains what it is and how it relates to a more standard measure of educational achievement. In addition, the Gini coefficient was evaluated through a series of comparisons with other results.

Keywords: Educational Inequality, Gini Coefficient, Economics of Education, Educational Development, Educational Attainment

Received: 6 February 2017 / **Accepted:** 13 April 2017 / **Published:** 30 June 2017

INTRODUCTION

Many people have speculated that a country's educational system is directly related to its social and economic development (Kurosaki and Khan 2006; Venkataraman 2009; Coombs and Ahmed 1974). In this light, measuring a nation's educational attainment is crucial because it serves as a starting point for further research and discussion. Still, it's a freshly opened field of research with plenty of room for development. Specifically, much methodological refinement is still required before it can be used as a pragmatic tool. In other words, there is a pressing need to address the issue of how to standardize the measurement of national education levels so that they can be used in meaningful comparisons across nations and societies. In empirical analyses, mean years of schooling are currently the most popular indicator of education level (Wail, Hanchane, and Kamal 2011). However, a "mean" is only an approximation of the prevalence of education and does not consider the distribution within the specified population. Most significantly, it fails to capture the underlying inequality in the group. The Gini coefficient has been proposed as a practical method for accounting for this aspect of educational attainment in papers by Hojo (2009) and Thomas et al. (2001). The Atkinson index was also used to measure educational inequality as part of the Inequality-adjusted Human Development Index (IHDI) developed by the United Nations Development Programme (UNDP) in 2010 (UNDP, 2010).

The relationship between the average number of years spent in school and the Gini coefficient is still poorly understood, despite the growing interest in educational inequality. Inequality's progression is assumed to not always occur next to the average level when its importance is argued. Discussing the index's utility and applicability requires verifying this assumption. Specifically, we can look to the theoretical foundation of development economics, where it was argued that economic growth does not directly lead to eradicating poverty, much like the initial criticisms of

*Corresponding author: Satoshi Watarai
†Email: 0796301174@edu.k.u-tokyo.ac.jp

the Trickle Down Theory (Aghion and Bolton, 1997). We're interested in seeing if this claim holds true in education, specifically whether an increase in the average number of years a population has spent in school correlates with the lessening of educational disparities.

This paper's goal is to compare the educational inequality trend to the average years people spend in school to better understand how it has developed over time. It'll discuss how inequality evolves over time and diverges from the average person's education level. It will use a dataset from the pre-war era to compare the two indices over time and across different macro trends.

Section 2 will then detail the research methods used in this investigation. The trend of the two indices will be presented in Section 3, followed by a discussion of their implications. The final thoughts are presented in Section 4.

METHODOLOGY

Data Source

Among few datasets on education stock, there are useful data such as that of Barro and Lee (2013) offering the mean years of schooling of 146 countries over 50 years. Despite the fact that the paper is frequently referred by their co-researchers, there is an important limitation to this data whose estimate is calculated from completed level of education and thus does not reflect years of schooling attained by those who dropped out or repeated years. Consequently, mean years of schooling are underestimated in their dataset. Such limitation of this dataset has also been pointed out by previous studies (Godo 2011). In response to this limitation of the dataset by Barro and Lee (2013), Godo (2011) has provided an estimated dataset with an improved accuracy and detail. This study will use this estimated data provided by Godo (2011).

Although the dataset covers only 3 countries, the estimate is calculated from the actual enrolment and population data both of which are objectively measured, making it more accurate than an estimated value which is calculated from completed level of education. Furthermore, the time span covered by this dataset extends to 80 years, dating back to 1920 until 2000. Since the dataset with greater accuracy and longer time span is preferred when examining the behaviour of indices, this study uses the estimated data of Godo (2011) as its data source. The data in Godo (2011) is presented in Table 1. In order to make the educational level comparable, the designated full years of schooling was framed according to the US's educational system, in which 1st to 8th grades is primary education, 9th to 12th grades is secondary education, and everything over 12th grade is tertiary education.

Table 1: Mean years of schooling by levels of education: Japan, Korea, and the US

	Japan			Korea			US		
Year	Primary	Secondary	Tertiary	Primary	Secondary	Tertiary	Primary	Secondary	Tertiary
1920	4.1	0.17	0.022	0.63	0.0012	0.0003	7.7	0.42	0.18
1930	5.1	0.40	0.073	0.80	0.0075	0.0011	8.0	0.80	0.26
1940	5.7	0.69	0.12	1.1	0.022	0.0028	8.3	1.2	0.33
1950	6.3	1.1	0.17	2.1	0.16	0.028	8.4	1.7	0.45
1960	6.9	1.6	0.24	2.8	0.34	0.071	8.5	2.2	0.61
1970	7.3	2.1	0.37	3.9	0.64	0.14	8.4	2.7	0.89
1980	7.6	2.6	0.58	5.4	1.2	0.20	8.3	3.1	1.4
1990	7.8	3.0	0.75	6.5	1.9	0.47	8.2	3.4	1.8
2000	8.0	3.3	1.0	7.1	2.5	0.86	8.1	3.6	2.2

Source: Based on Godo (2011)

Gini Coefficient of Education

This study uses Gini coefficient of education as an index for educational inequality. If we follow a geometrical interpretation, Gini coefficient is determined by the Lorentz curve, which plots the percentage of the total schooling years of the population on vertical axis that is cumulatively earned by the bottom X% of the population. The graph demonstrates what percentage of the total schooling years of population is distributed to the bottom X% of population. Therefore, the equidistribution line represents the perfectly equal distribution of the total schooling years. Gini coefficient of education is the ratio of the area between the equidistribution line and the Lorentz curve to the area under the equidistribution line. As a consequence of definitions, Gini coefficient falls in the range between 0 and 1, and the smaller Gini coefficient is, more equal distribution of acquired years of schooling is in the population. Algebraically, the index can be defined as follows:

$Educ.Gini = 1 - 2\int_0^1 L(x)dx$

$Y = L(x)$ is Lorentz curve

Processing of the Dataset

Godo (2011) offers the estimated datasets of mean schooling years by levels of education, but since Lorentz curve is acquired from the percentage of cumulative population, Gini coefficient cannot be calculated without this data. In order to extract a cumulative population, it is assumed that at primary, secondary, and tertiary education respectively, the individual had either fully completed or had not enrolled at all. The population is then allocated into either case according to the value of mean years of schooling. For example, if the mean schooling years of elementary education is α years, that of secondary education is β years, and that of tertiary education is γ years, we assume that 0 year of schooling is distributed to the individuals in $(1 - \frac{\alpha}{8})$ of total population, 8 years of schooling are distributed to the individuals in $(\frac{\alpha}{8} - \frac{\beta}{4})$ of total population, 12 years of schooling are distributed to the individuals in $(\frac{\beta}{4} - \frac{\gamma}{4})$ of total population, and 16 years of schooling are distributed to the individuals in $(\frac{\gamma}{4})$ of total population. Through this manipulation, Lorentz curve is drawn without distortion in the estimated value of mean schooling years by level of education. See Appendix A for the Proof. In our study, designated full years of schooling for tertiary education is defined as 4 years. Although in reality, it is necessary to consider graduate school, this assumption does not affect the result significantly as the maximum value was 2.2 years for the US in 2000. Since estimate in Godo (2011) includes extra schooling years attained by students who repeated years, the mean years of schooling in primary level has exceeded 8 years in the US between 1940 and 2000. In such case, one must be mindful of the fact that the value of the population ratio allocated to 0 year of schooling becomes negative when it is processed in the same manner as the other years. Because the population is allocated either to 0 or 8 years of schooling according to the mean value, the population ratio of those who attained 0 year, $(1 - \frac{\alpha}{8})$, takes a negative value when mean years of schooling in primary level exceeds 8 years. As it is impossible for this value to be negative under the empirical assumption, the population exceeding 8 years was shifted to secondary education. This additional data process was adopted as it does not distort the overall mean years of schooling. The processed data are presented in Table 2. The population ratio by attained levels of education is presented in Table 3.

Table 2: Processed mean years of schooling by levels of education: Japan, Korea, and the US

	Japan			Korea			US		
Year	Primary	Secondary	Tertiary	Primary	Secondary	Tertiary	Primary	Secondary	Tertiary
1920	4.1	0.17	0.022	0.63	0.0012	0.0003	7.7	0.42	0.18
1930	5.1	0.40	0.073	0.80	0.0075	0.0011	8.0	0.80	0.26
1940	5.7	0.69	0.12	1.1	0.022	0.0028	8.0	1.5	0.33
1950	6.3	1.1	0.17	2.1	0.16	0.028	8.0	2.1	0.45
1960	6.9	1.6	0.24	2.8	0.34	0.071	8.0	2.7	0.61
1970	7.3	2.1	0.37	3.9	0.64	0.14	8.0	3.1	0.89
1980	7.6	2.6	0.58	5.4	1.2	0.20	8.0	3.4	1.4
1990	7.8	3.0	0.75	6.5	1.9	0.47	8.0	3.6	1.8
2000	8.0	3.3	1.0	7.1	2.5	0.86	8.0	3.7	2.2

Source: Based on Godo (2011)

Table 3: Population ratio by attained levels of education: Japan, Korea, and the US

	Japan				Korea				US			
Year	0 Years	8 Years	12 Years	16 Years	0 Years	8 Years	12 Years	16 Years	0 Years	8 Years	12 Years	16 Years
1920	0.49	0.47	0.037	0.0055	0.92	0.078	0.00023	0.000075	0.038	0.86	0.060	0.045
1930	0.36	0.54	0.082	0.018	0.90	0.098	0.0016	0.00028	0.0	0.80	0.14	0.065
1940	0.29	0.54	0.14	0.030	0.86	0.13	0.0048	0.00070	0.0	0.63	0.29	0.083
1950	0.21	0.51	0.23	0.043	0.74	0.22	0.033	0.0070	0.0	0.48	0.41	0.11
1960	0.14	0.46	0.34	0.060	0.65	0.27	0.067	0.018	0.0	0.33	0.52	0.15
1970	0.088	0.26	0.51	0.15	0.51	0.33	0.13	0.035	0.0	0.23	0.55	0.22
1980	0.050	0.30	0.51	0.15	0.33	0.38	0.25	0.050	0.0	0.15	0.50	0.35
1990	0.025	0.23	0.56	0.19	0.19	0.34	0.36	0.12	0.0	0.10	0.45	0.45
2000	0.0	0.18	0.58	0.25	0.11	0.26	0.41	0.22	0.0	0.075	0.38	0.55

Source: Calculated based on Table 2

RESULTS AND DISCUSSION

Mean Years of Schooling in Japan, Korea, and the US

The first subsection will examine the results of mean years of schooling in Japan, Korea, and the US. Table 4 shows the progression of the index in the three countries between 1920 and 2000. In all three countries, the mean years have increased monotonically. As for Korea, whilst the result is less than a year until 1930, it surpasses 10 years in 2000. On the other hand, in the US where the result was already 8.3 years at 1920, the change is not as radical as it marks 14.0 years at 2000. This trend agrees with those presented in previous studies such as Barro and Lee (2013) and Ministry of Education, Culture, Sports, and Technology, Japan (2005). According to the datasets of Barro and Lee (2013), the value of mean years of schooling of Korea which was 5.02 in 1955 declined to 4.20 in 1960. However, since this is the only exception to the monotonic trend in the datasets, its overall trend agrees with our result. Godo (2011) also pointed out the possibility that this figure might be a statistical error as this is the only exception to the trend.

Table 4: Mean years of schooling: Japan, Korea, and the US

Year	Japan	Korea	US
1920	4.3	0.6	8.3
1930	5.6	0.8	9.1
1940	6.5	1.1	9.8
1950	7.6	2.3	10.5
1960	8.7	3.3	11.3
1970	9.8	4.8	12.0
1980	10.7	6.9	12.8
1990	11.5	9.0	13.5
2000	12.3	10.5	14.0

Source: Based on Godo (2011)

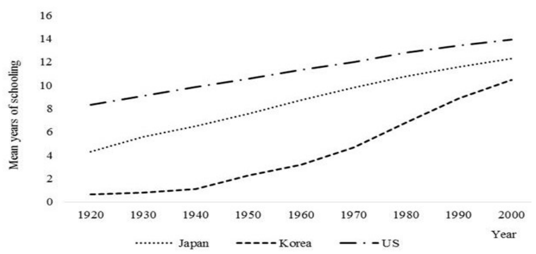

Figure 1. Mean years of schooling: Japan, Korea, and the US

Source: Based on Table 4

Educational Gini Coefficient in Japan, Korea, and the US

Table 5 and Figure 2 show the historical progression of the Gini coefficient of the three countries over 80 years dating back to pre-war (1920-2000). This demonstrates that in both Japan and Korea, there has been an extremely rapid decline in the inequality level. This compares to the US whose level has stayed largely the same. Moreover, there are even points at which there is a slight increase in its value. These trends agree with the results of Thomas et al. (2001) and Ziesemer (2016). Since both of Thomas et al. (2001) and Ziesemer (2016) are based on the datasets of Barro and Lee (2013), the figure of Korea in 1965 is the only exception to the trend.

Table 5: Gini coefficient of education: Japan, Korea, and the US

Year	Japan	Korea	US
1920	0.51	0.92	0.10
1930	0.41	0.90	0.097
1940	0.36	0.87	0.085
1950	0.30	0.76	0.091
1960	0.24	0.69	0.086
1970	0.20	0.57	0.096
1980	0.17	0.41	0.10
1990	0.13	0.30	0.097
2000	0.11	0.23	0.090

Source: Calculated based on Table 3

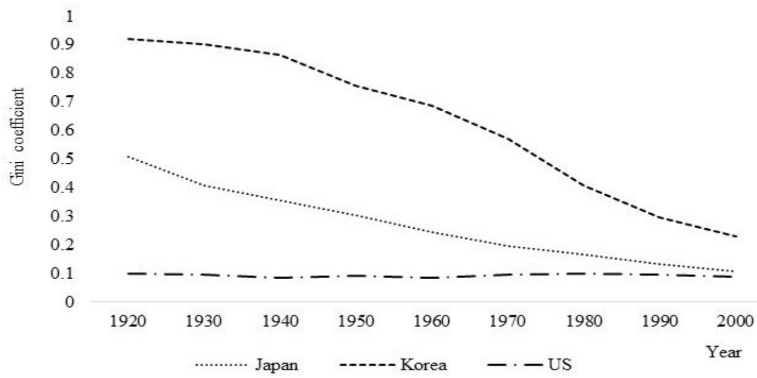

Figure 2. Gini coefficient of education: Japan, Korea, and the US

Source: Based on Table 5

Comparison

The result presented in Table 4 and 5 shows that the relation between mean years of schooling and educational Gini coefficient is not uniform in the three case study countries. In fact, it has proved that the latter does not progress adjacent to the former. In the case of Japan and Korea, both mean years of schooling and Gini coefficient of education have demonstrated gradual improvement, signifying that the decrease in the inequality of educational attainment has occurred simultaneously with the increase in the population's acquired years of schooling. On the other hand, in the US, the results between 1960 and 1980 demonstrate that whilst the population had acquired increased years of schooling, there was, in fact, an increase in the inequality level.

Firstly, this implies that the increase in the population's schooling years does not necessarily occur from the lower level. In other words, the increase in mean years of schooling does not always signify that there are more children enrolled for primary education. It could just as well reflect a situation where the population which had completed secondary education proceeded to tertiary education. As mentioned by Holsinger (2005), a country may invest disproportionately in tertiary education, consequently skewing the distribution of educational attainment. If the population attaining primary school education rests unchanged, such progress would, in fact, lead to an enlarged inequality in educational attainment in that society. In case of the US, after primary education became prevalent, the growth rate of tertiary education was higher than that of secondary education which had led to this widening inequality.

The second implication the results have is that the level of the mean value itself could affect how it relates to Gini coefficient of education. Looking at the progression of mean years of schooling in Japan and Korea, one is able to see that their attainment level increases substantially. Whereas the mean value is 4.3 years in Japan, and 0.6 years in Korea in 1920, the level soars up to 12.3 years and 10.5 years in 2000. In contrast, in the US, the mean level already reaches 8.3 years in 1920 which increases to 14.0 years by 2000. In both Japan and Korea, access to education was still very limited in 1920. The mean value is both below 6 years, signifying that an average person had not finished primary education. The high level of inequality in these countries reflects a society where the majority of the population has barely had basic level of education. This observed trend agrees with that which has been pointed out for other underdeveloped societies. For example, Mesa (2007) indicated that educational Gini index and mean years of schooling of regions and provinces in the Philippines are negatively associated with each other. Holsinger (2005) also showed that Vietnam successfully increased its educational attainment whilst also mitigating educational inequality. On the other hand, in the US, where the basic level of education was already prevalent, the increase of the schooling years began to occur at different levels in which its effect on the two variables was not as tangible as in the case of Japan and Korea.

There are also several limitations to the study which should be pointed out. Firstly, this study has found that there is a limitation in the Gini coefficient itself, which is mentioned in previous studies (Mellor 1989; Bellu and Liberati 2006; Kwok 2010). Although Gini coefficient has been the most popular method (De Maio 2007) to quantify inequality due to its mathematical nature, its estimate of inequality is liable to the level at which this inequality occurs (Yildiz and Kayili 2015). The comparison between the Gini index of Korea and the US is a good example to demonstrate this. As mentioned in the Result and Discussion, when we calculate the mean years of schooling at each educational level, it is 0.63 years in primary education and 0.0012 years in secondary education for Korea in 1920, signifying that most people had not completed primary education. On the other hand in the US, the mean value was 7.7 years in primary education and 0.42 years in secondary education in 1920, meaning that an average person's educational attainment was slightly below completion of primary education. The Gini index is 0.92 for Korea, and 0.10 for the US, representing a large difference in the inequality level between the two countries.

However, one must be cautious that Gini coefficient is an index that serves to demonstrate the degree of disparity irrespective of the average level. In other words, it should not question at which level the society is equal, it should only observe the distribution. Korea in 1920 may be limited at the average level of educational attainment. However, if the majority of the population had not finished primary education, it could also be described as largely equal, just as much as the case of the US in which the majority of the population had completed primary school but had not been to secondary school. It has, therefore, been revealed through the results that Gini coefficient is affected by the absolute level of educational attainment, and thus does not always reflect inequality. Consequently, inequality is overestimated in societies where there is a high representation of population in the lowest percentile.

Secondly, the number of countries and the timespan are also limitations of our study. This paper only analysed Japan, Korea, and the US, and its data only covered up to 2000. In future studies, we wish to enlarge our scope both geographically and chronologically in order to apply our findings to more diverse samples to validate the claim.

CONCLUSION, RECOMMENDATIONS AND IMPLICATIONS

This paper discussed how the progress of inequality in educational attainment differs from that of the prevalence of education. It differs from previous papers in that it did not only argue the importance of focusing on educational inequality but also discussed its nature and relation to other more prevalent educational indices. Firstly, it found that inequality is not always mitigated as a result of the prevalence of education, but in fact, an increased prevalence could lead to a rise in inequality. Secondly, it was also implied that such case is more viable in societies with higher level of mean years of schooling. Thirdly, it has also made a methodological finding that there is a limitation of Gini coefficient in depicting inequality for its tendency to overestimate the inequality level in a society where there is a high representation of the lowest percentile. This paper has also made a methodological contribution by proposing a data processing method to subtract the size of population of the respective educational level from limited information. The method is also useful in that it does not distort the mean years of schooling when doing so. This has opened a path for quantitative analysis in many more countries where raw data on education are often scarce.

Limitations and Future Research Directions

The empirical analysis of our paper demonstrates that there is importance of observing the educational inequality when discussing the educational attainment as part of social development. The variable has the ability to present an aspect that is not necessarily portrayed by the other prevalent indices of the domain, especially when comparing societies at different development levels. Though it should also be kept in mind, how and when this index is applied for its aforementioned limitation. In future work, first of all, we will enlarge our sample both geographically and chronologically in order to verify the index with further accuracy. We will also seek to find more specific conditions in which the application of educational Gini coefficient is appropriate to contribute to further refinement of methodological development.

REFERENCES

Aghiton, P., and Bolton, P. 1997. "A Theory of Trickle-Down Growth and Development." ***The Review of Economic Studies*** 64(2): 151-172.

Barro, R. J., and Lee, J. W. 2013. "A New Data Set of Educational Attainment in the World, 1950-2010." ***Journal of Development Economics*** 104: 184-198.

Bellu, L. C., and Liberati, P. 2006. "Inequality Analysis: The Gini Index." Retrieved August 04, 2017 (https://goo.gl/qvZRa7).

Coombs, P. H., and Ahmed, M. 1974. ***Attacking Rural Poverty: How Nonformal Education Can Help.*** Washington, DC, WA: International Bank for Reconstruction and Development.

De Maio, F. G. 2007. "Income Inequality Measures." ***Journal of Epidemiology & Community Health*** 61(10): 849-852.

Godo, Y. 2011. "Estimation of Average Years of Schooling for Japan, Korea and the United States." Retrieved August 04, 2017 (https://goo.gl/VPQXhy).

Hojo, M. 2009. "Inequality in Japanese Education: Estimation Using the Gini Education Coefficient." ***Japanese Economy*** 36(3): 3-27.

Holsinger, D. B. 2005. "Inequality in the Public Provision of Education: Why It Matters." ***Comparative Education Review*** 49(3): 297-310.

Kurosaki, T., and Khan, H. 2006. "Human Capital, Productivity, and Stratification in Rural Pakistan." ***Review of Development Economics*** 10(1): 116-134.

Kwok, K. C. 2010. ***Income Distribution of Hong Kong and the Gini Coefficient***. Hong Kong, China: The Government of Hong Kong.

Mellor, J. 1989. Dramatic Poverty Reduction in the Third World; Prospects and Needed Action. Retrieved April 16, 2014 (https://goo.gl/knqB6y).

Mesa, E. P. 2007. "Measuring Education Inequality in the Philippines." ***Philippine Review of Economics*** 44(2): 33-70.

Ministry of Education, Culture, Sports and Technology Japan. 2005. ***Japanese Education at a Glance***. Tokyo, Japan: The Ministry of Education, Culture, Sports and Technology Japan.

Thomas, V., Wang, Y., and Fan, X. 2001. ***Measuring Education Inequality: Gini Coefficients of Education*** (Vol. 2525). Washington, DC, WA: World Bank Publications.

United Nations Development Programme. 2010. ***The Real Wealth of Nations: Pathways to Human Development***. New York, NY: The United Nations Development Programme.

Venkataraman, B. 2009. "Education for Sustainable Development." ***Environment: Science and Policy for Sustainable Development*** 51(2): 8-10.

Wail, B., Hanchane, S., and Kamal, A. 2011. "A New Data Set of Educational Inequality in the World, 1950-2010: Gini Index of Education by Age Group." Retrieved October, 2011 (https://goo.gl/JZMhDw)

Yildiz, F. U., and Kayili, G. 2015. "Examining the Effects of Computer-Assisted Preschool Educational Activities on Children's Intuitive Mathematical Ability." ***Journal of Advances in Humanities and Social Sciences*** 1(1): 1-8.

Ziesemer, T. 2016. "Gini Coefficients of Education for 146 Countries, 1950-2010." ***Bulletin of Applied Economics*** 3(2): 1-8.

Appendix A

Assertion

Allocating population through the method mentioned in paper doesn't distort mean years of schooling by levels of education.

Proof

Let

X = Population;
A years = Desinated full years of schooling for primary education;
B years = Desinated full years of schooling for secondary education;
C years = Desinated full years of schooling for tertiary education;
α_a years = Mean years of schooling in primary education;
β_a years = Mean years of schooling in secondary education; and
γ_a years = Mean years of schooling in tertiary education.

Now allocated population by attained years of schooling is calculated as follows:

0 year X $\left(1 - \frac{\alpha_a}{A}\right)$
A years X $\left(\frac{\alpha_a}{A} - \frac{\beta_a}{B}\right)$
B years X $\left(\frac{\beta_a}{B} - \frac{\gamma_a}{C}\right)$
C years X $\left(\frac{\gamma_a}{C}\right)$

Let

α_e years = Culculated mean years of schooling in primary education based on allocated population
β_e years = Culculated mean years of schooling in secondary education based on allocated population
γ_e years = Culculated mean years of schooling in tertiary education based on allocated population

Now α_e is calculated as follows:

$$\alpha_e = \frac{0.X.\left(1-\frac{\alpha_a}{A}\right)+A.\left\{X.\left(\frac{\alpha_a}{A} - \frac{\beta_a}{B}\right) + X.\left(\frac{\beta_a}{B} - \frac{\gamma_a}{C}\right) + X.\frac{\gamma_a}{C}\right\}}{X} = \alpha_a$$

In the same way,

$\beta_e = \beta_a$ and
$\gamma_e = \gamma_a$

From the above, the assertion is proven.