

Revisiting Educational Kuznets Curve: An Analysis of Educational Inequality based on Absolute and Relative Inequality Measures

Takahiro Akita
International University of Japan

March 2023

IUJ Research Institute
International University of Japan

These working papers are preliminary research documents published by the IUJ research institute. To facilitate prompt distribution, they have not been formally reviewed and edited. They are circulated in order to stimulate discussion and critical comment and may be revised. The views and interpretations expressed in these papers are those of the author(s). It is expected that the working papers will be published in some other form.

Revisiting Educational Kuznets Curve: An Analysis of Educational Inequality based on Absolute and Relative Inequality Measures

Takahiro Akita*
Professor Emeritus
International University of Japan

Abstract

This paper shows that even if educational inequality is measured by the Gini coefficient, we could produce an educational Kuznets curve under certain conditions. Using the Barro and Lee data set, it examines the evolution of the education Gini coefficient under different sets of conditions. The paper presents an educational Kuznets curve when the proportion of those with no formal education is very small or when they receive informal education equivalent to a small amount of formal education.

Key words: educational Kuznets curve, educational expansion, education Gini coefficient

JEL codes: I24; I25; O15

* The author is grateful to the Japan Society for the Promotion of Science (Grant-in-Aid for Scientific Research No. 18K01589).

1. Introduction

When educational inequality is measured by variance (or standard deviation) of years of education, it exhibits an inverted U-shaped pattern with respect to educational expansion; that is, at the early stages of educational expansion, educational inequality first rises, but after reaching a peak, it starts to decline (Ram 1990; Lam and Levison 1992; De Gregorio and Lee 2002; Meschi and Scervini 2014).¹ This inverted U-shaped pattern is usually termed the educational Kuznets curve.

As an alternative measure of educational inequality, Thomas, Wang and Fan (2000) introduced the education Gini coefficient, which is an extension of the Gini coefficient for income distribution. Since then, many researchers have used the Gini coefficient to measure inequality in educational attainment. Unlike the studies based on variance (or standard deviation), they mostly found a downward-sloping pattern rather than an inverted U-shaped pattern with respect to educational expansion, that is, educational inequality declines monotonically with educational expansion (Castello and Domenech 2002; Lin 2007; Lim and Tang 2008; Hojo 2009; Fordvari and van Leeuwen 2011; Morrisson and Murtin 2013; Agrawal 2014; Meschi and Scervini 2014; Banzragch, Mizunoya and Bayarjargal 2019; Shukla and Mishra 2019; Castello-Climent and Domenech 2021).

This paper will explore conditions under which the Gini coefficient could produce an educational Kuznets curve. Using the Barro and Lee data set, the paper will also examine the evolution of the education Gini coefficient under different sets of conditions (Barro and Lee 2013).

2. Measuring educational inequality

Consider a country with four levels of education: 0 no formal education; 1 primary education; 2 secondary education; and 3 tertiary education. Let e_i and p_i be the cumulative years of education of i th level of education and the proportion of people who have completed i th level of education as their highest level, respectively ($i = 0, 1, 2,$ and 3). Then, mean years of education for the country is given by $\mu = \sum_{i=0}^3 p_i e_i$, where $\sum_{i=0}^3 p_i = 1$. To measure educational inequality, variance and the Gini coefficient can be used. These inequality measures satisfy anonymity principle, the principle of population independence and the Pigou-Dalton transfer principle (Fields 2001). Additionally, the Gini coefficient satisfies the principle of mean independence. Thus, it is a *relative* inequality measure. On the other hand, variance does not satisfy the principle of mean independence. Variance (or standard deviation) is an *absolute* inequality measure.

Using e_i , p_i and μ , variance and the Gini coefficient can be defined, respectively, by

$$V = \sum_{i=0}^3 p_i (e_i - \mu)^2, \quad (1)$$

$$G = \frac{1}{2\mu} \sum_{i=0}^3 \sum_{j=0}^3 p_i p_j |e_j - e_i| = \frac{1}{\mu} \sum_{i=0}^3 \sum_{j>i}^3 p_i p_j |e_j - e_i|. \quad (2)$$

Since we have $e_0 < e_1 < e_2 < e_3$, the Gini coefficient (equation 2) is rewritten as follows.

$$G = \frac{1}{\mu} \sum_{i=0}^3 \sum_{j>i}^3 p_i p_j (e_j - e_i). \quad (3)$$

The Gini coefficient ranges between 0 (perfect equality) and 1 (perfect inequality).

¹ Educational expansion refers to increasing mean years of education.

When educational inequality is measured by variance, we can obtain the following proposition.

Proposition 1.

$$\begin{aligned} V_0 &= p_0(e_0 - \mu)^2 + (1 - p_0)((\mu^* - \mu)^2 + V_1) \quad 0 \leq V_0 \\ &= \frac{p_0}{1-p_0}(e_0 - \mu)^2 + (1 - p_0)V_1. \end{aligned} \quad (4)$$

V_0 is the variance when all the education levels are included ($V_0 = V$ in equation 1), while V_1 is the variance when the no formal education group is excluded. V_1 can be given by

$$\begin{aligned} V_1 &= \sum_{i=1}^3 p_i^* (e_i - \mu^*)^2 \\ &= \frac{p_1^*}{1-p_1^*} (e_1 - \mu^*)^2 + (1 - p_1^*)V_2 \quad 0 \leq V_1, \end{aligned}$$

where $p_i^* = \frac{p_i}{\sum_{j=1}^3 p_j} = \frac{p_i}{1-p_0}$ ($i = 1, 2, \text{ and } 3$), $\mu^* = \sum_{i=1}^3 p_i^* e_i$, and $1 = \sum_{i=1}^3 p_i^*$. V_2 is the variance when the no formal and primary education groups are excluded.

When $p_0 = 1$, we have $p_1 = p_2 = p_3 = 0$ and $\mu = e_0$; thus, we have $V_0 = 0$. On the other hand, when $p_0 = 0$, we have $\mu = \mu^*$ and $V_0 = V_1 \geq 0$. We now have the following proposition.

Proposition 2. Under the assumption that p_1^* , p_2^* and p_3^* are constant regardless of the value of p_0 , that is, μ^* and V_1 are constant, V_0 rises as μ increases, but after reaching a peak at $p_0 = \frac{1}{2} \left(1 - \frac{V_1}{(e_0 - \mu^*)^2}\right)$, V_0 declines as μ increases ($e_0 \leq \mu \leq \mu^*$). When $\mu = e_0$, $V_0 = 0$, while when $\mu = \mu^*$, $V_0 = V_1$. That is, educational inequality, as measured by variance, exhibits an inverted U-shaped pattern with respect to educational expansion (that is, educational Kuznets curve).

Proof. Because we have $e_0 - \mu = (1 - p_0)(e_0 - \mu^*)$, equation 4 can be rewritten as $V_0 = (1 - p_0)p_0(e_0 - \mu^*)^2 + (1 - p_0)V_1$. Differentiating this equation with respect to p_0 , we obtain $\frac{\partial V_0}{\partial p_0} = (1 - 2p_0)(e_0 - \mu^*)^2 - V_1$ under the assumption that μ^* and V_1 are constant regardless of the value of p_0 . Because $\frac{\partial^2 V_0}{\partial p_0^2} = -2(e_0 - \mu^*)^2 < 0$, V_0 takes the maximum at $p_0 = \frac{1}{2} \left(1 - \frac{V_1}{(e_0 - \mu^*)^2}\right)$.

When educational inequality is measured by the Gini coefficient, we can obtain the following proposition by expanding equation 3.

Proposition 3.

$$G_0 = p_0 \left(1 - \frac{e_0}{\mu}\right) + (1 - p_0) \left(1 - \frac{p_0 e_0}{\mu}\right) G_1 \quad 0 \leq G_0 \leq 1. \quad (5)$$

G_0 is the Gini coefficient when all the education levels are included ($G_0 = G$ in equation 3), while G_1 is the Gini coefficient when the no formal education group is excluded. G_1 can be given by

$$\begin{aligned} G_1 &= \frac{1}{\mu^*} \sum_{i=1}^3 \sum_{j>i}^3 p_i^* p_j^* (e_j - e_i) \\ &= p_1^* \left(1 - \frac{e_1}{\mu^*}\right) + (1 - p_1^*) \left(1 - \frac{p_1^* e_1}{\mu^*}\right) G_2 \quad 0 \leq G_1 \leq 1, \end{aligned}$$

where $p_i^* = \frac{p_i}{\sum_{j=1}^3 p_j} = \frac{p_i}{1-p_0}$ ($i = 1, 2, \text{ and } 3$), $\mu^* = \sum_{i=1}^3 p_i^* e_i$, and $1 = \sum_{i=1}^3 p_i^*$. G_2 is the Gini coefficient when the no formal and primary education groups are excluded.

When the no formal education group is given 0 year of education ($e_0 = 0$), equation 5 is given by

$$G_0 = p_0 + (1 - p_0)G_1. \quad (6)$$

In other words, G_0 is a weighted average of 1 and G_1 . The following proposition presents the evolution of G_0 with respect to μ when $e_0 = 0$.

Proposition 4. Under the assumption that $e_0 = 0$ and that p_1^* , p_2^* and p_3^* are constant regardless of the value of p_0 , that is, μ^* and G_1 are constant, G_0 declines monotonically as μ increases and reaches G_1 when $\mu = \mu^*$. When $\mu = 0$ ($p_0 = 1$), G_0 is undefined; thus, we have $1 > G_0 \geq G_1 \geq 0$ for $0 < \mu \leq \mu^*$.

Now, suppose that those without any formal education receive some informal education equivalent to α years of formal education where $e_1 > e_0 = \alpha > 0$. Then, when $p_0 = 1$, we have $\mu = \alpha$ and $G_0 = 0$. On the other hand, when $p_0 = 0$, we have $\mu = \mu^*$ and $G_0 = G_1 \geq 0$. The following proposition presents the evolution of G_0 with respect to μ when $e_0 = \alpha > 0$.

Proposition 5. Under the assumptions that p_1^* , p_2^* and p_3^* are constant regardless of the value of p_0 , that is, μ^* and G_1 are constant and that $e_0 = \alpha > 0$ and $\frac{\mu^* - e_0}{\mu^* + e_0} > G_1$, G_0 rises as μ increases, but after reaching a peak, G_0 declines as μ increases ($0 < e_0 \leq \mu \leq \mu^*$). When $\mu = e_0 = \alpha > 0$, $G_0 = 0$, while when $\mu = \mu^*$, $G_0 = G_1 \geq 0$. That is, educational inequality, as measured by the Gini coefficient, exhibits an inverted U-shaped pattern with respect to educational expansion (that is, educational Kuznets curve).

Proof. Because we have $\mu - e_0 = (1 - p_0)(\mu^* - e_0)$ and $\mu - p_0 e_0 = (1 - p_0)\mu^*$, equation 5 is rewritten as

$$G_0 = \frac{1}{\mu} [\mu^* G_1 + ((\mu^* - e_0) - 2\mu^* G_1)p_0 + (\mu^* G_1 - (\mu^* - e_0))p_0^2]$$

where $\mu = \mu^* - (\mu^* - e_0)p_0$. Differentiating this equation with respect to p_0 , we obtain

$$\frac{\partial G_0}{\partial p_0} = \frac{1}{\mu^2} \left[(\mu^* - (\mu^* - e_0)p_0) \left((\mu^* - e_0) - 2\mu^* G_1 + 2(\mu^* G_1 - (\mu^* - e_0))p_0 \right) + (\mu^* - e_0)(\mu^* G_1 + ((\mu^* - e_0) - 2\mu^* G_1)p_0 + (\mu^* G_1 - (\mu^* - e_0))p_0^2) \right].$$

When $p_0 = 1$ (or $\mu = e_0 = \alpha$), we have $\left. \frac{\partial G_0}{\partial p_0} \right|_{p_0=1} = -\frac{\mu^* - e_0}{e_0} < 0$ because $\mu^* - e_0 > 0$. On the

other hand, when $p_0 = 0$ (or $\mu = \mu^*$), we have $\left. \frac{\partial G_0}{\partial p_0} \right|_{p_0=0} = \frac{(\mu^* - e_0) - (\mu^* + e_0)G_1}{\mu^*}$. Therefore, if

$\frac{\mu^* - e_0}{\mu^* + e_0} > G_1$, then G_0 rises as μ increases, but after reaching a peak, G_0 declines as μ increases ($0 < e_0 \leq \mu \leq \mu^*$). When $\mu = e_0 = \alpha > 0$, $G_0 = 0$, while when $\mu = \mu^*$, $G_0 = G_1 \geq 0$.

This proposition implies that even if educational inequality is measured by the Gini coefficient, educational inequality would exhibit an inverted U-shaped pattern with respect to educational expansion when those with no formal education receive informal education equivalent to a small amount of formal education.² Coady and Dizioli (2018) assigned 1 year of education for the no formal education category ($e_0 = 1$) to estimate the education Gini. Based on the Barro and Lee data set on educational attainment, they found evidence of an inverted U-shaped pattern with respect to educational expansion.

² We can obtain a similar proposition for the coefficient of variation (CV), which is defined as the ratio between standard deviation and mean. CV is another relative inequality measure.

The question arises whether G_1 changes as p_0 changes from 0 to 1. Based on the Barro and Lee data set for population aged 25 years and over between 1950 and 2015, we estimate G_1 by assigning 6, 12 and 16 years for the primary, secondary and tertiary education groups, respectively. Figure 1 presents a scatter plot between the education Gini coefficient excluding the no education group (G_1) and the proportion of the no education group (p_0). G_1 ranges between 0.01 and 0.24 for $0 \leq p_0 \leq 1$. Correlation coefficient between these two variables is almost zero (-0.07), that is, G_1 is independent of p_0 . These observations could support the validity of Propositions 2, 4 and 5.

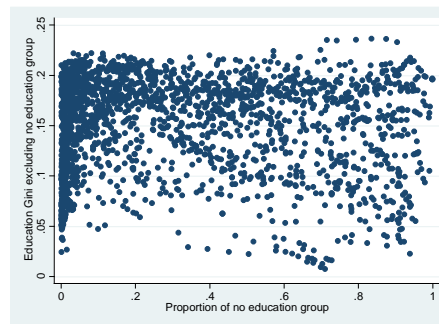


Figure 1: Relationship between the Education Gini (G_1) and the Proportion of No Education Group (p_0). (Note) Sample size = 2,043. (Source) Barro and Lee (2013).

Some studies introduced the concept of human capital and measured human capital inequality by the Gini coefficient. In their studies, the number of years of education is converted to human capital, usually, using the following equation (Lim and Tang 2008; Morrisson and Murtin 2013).³

$$h_i = \exp(re_i), \quad (7)$$

where h_i and r are human capital for education level i and returns to formal education, respectively. Here, for simplicity, returns to formal education is assumed to be constant across education levels. If this equation is used to obtain human capital for each education level, then $h_i > 0$ for $e_i \geq 0$ ($i = 0, 1, 2,$ and 3). In particular, we have $h_0 = 1$ when $e_0 = 0$. Therefore, based on Proposition 5, human capital inequality, as measured by the Gini coefficient, could exhibit an inverted U-shaped pattern with respect to the expansion of human capital. The studies using equation 7, in fact, observed an inverted U-shaped pattern.

We could conclude that when educational inequality or human capita inequality is measured by the Gini coefficient (or other relative inequality measures), the existence of an inverted U-shaped pattern depends on whether the no formal education group is given a positive value as years of education or human capital. As long as the value is greater than zero, an inequality value should be zero if all people do not have any formal education (that is, perfect equality), and it is likely to have an inverted U-shaped pattern.

3. Educational Kuznets curve: An empirical evidence

Using the Barro and Lee data set on educational attainment for population aged 25 years and over between 1950 and 2015, we estimate standard deviation and the Gini coefficient by

³ Though Castello and Domenech (2002) and Castello-Climent and Domenech (2021) employed the concept of human capital, they used years of education as a proxy for human capital; thus, they observed a downward sloping pattern for the education Gini.

assigning 0, 3, 6, 9, 12, 14 and 16 years for the no, incomplete primary, primary, incomplete secondary, secondary, incomplete tertiary and tertiary education groups, respectively.⁴

If educational inequality is measured by standard deviation, there is an educational Kuznets curve (Figure 2). On the other hand, if it is measured by the Gini coefficient, there is no educational Kuznets curve, but a downward sloping curve (Figure 3).

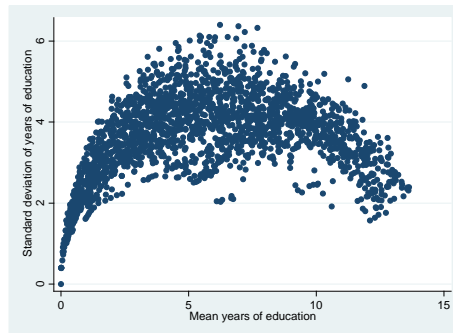


Figure 2: Education Standard Deviation vs Mean Years of Education including All Education Groups: Years of Education for No Education is Zero ($e_0 = 0$). (Note) Sample size = 2,043. (Source) Same as Figure 1.

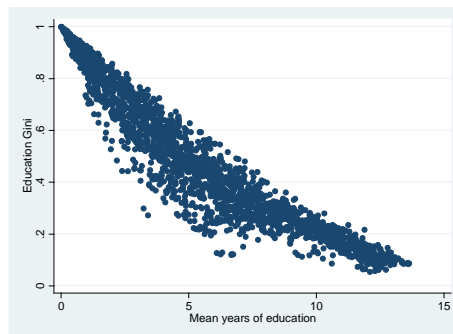


Figure 3: Education Gini vs Mean Years of Education including All Education Groups: Years of Education for No Education is 0 ($e_0 = 0$). (Note) Sample size = 2,043. (Source) Same as Figure 1.

On the other hand, even if educational inequality is measured by the Gini coefficient, there is an educational Kuznets curve when the no education group is given 1 year of education ($e_0 = 1$) (Figure 4).

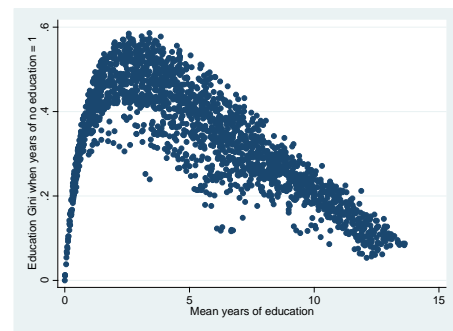


Figure 4: Education Gini vs Mean Years of Education including All Education Groups: Years of Education for No Education is 1 ($e_0 = 1$). (Note) Sample size = 2,043. (Source) Same as Figure 1.

⁴ Each country has a different length of schooling for each education level. But, in this paper, we assume for simplicity that all countries share the same length for each education level.

When we focus on a sample of countries where the proportion of those with no formal education is smaller than 10 % (sample size = 743), then we find an educational Kuznets curve even if the no education group is given 0 year of education ($e_0 = 0$) (Figure 5). A panel data regression analysis using the fixed effects model confirms the existence of an educational Kuznets curve.⁵ On the other hand, if we exclude the no education group and focus on educated groups, then we also observe an educational Kuznets curve (Figure 6).

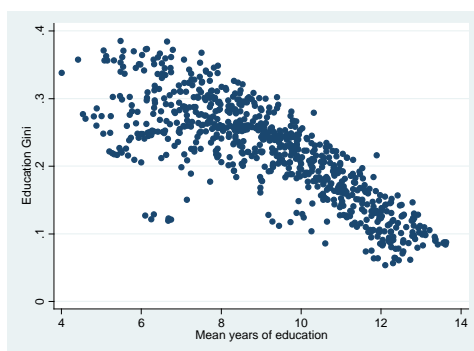


Figure 5: Education Gini vs Mean Years of Education including All Education Groups where Years of Education for No Education is 0 ($e_0 = 0$) but Proportion of No Education Group Smaller than 10% ($p_0 < 0.1$). (Note) Sample size = 743. (Source) Same as Figure 1.

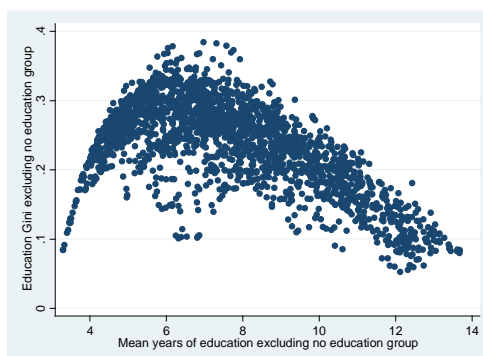


Figure 6: Education Gini vs Mean Years of Education excluding No Education Group. (Note) Sample size = 2,043. (Source) Same as Figure 1.

4. Conclusion

This paper showed that even if educational inequality is measured by the Gini coefficient, we could produce an educational Kuznets curve under certain conditions. Using the Barro and Lee data set of educational attainment, we also presented an educational Kuznets curve when the proportion of those with no formal education is very small or when they receive informal education equivalent to a small amount of formal education.

⁵ The following is the fixed effects model used in the panel data regression analysis. $G_{it} = \beta_0 + \beta_1\mu_{it} + \beta_2\mu_{it}^2 + \alpha_i + u_{it}$, where G_{it} is the education Gini coefficient, μ_{it} is mean years of education, and α_i is the country specific effects. With a sample of countries with $p_0 < 0.1$, we obtain $\hat{\beta}_1 = 0.030 > 0$ and $\hat{\beta}_2 = -0.003 < 0$. Both $\hat{\beta}_1$ and $\hat{\beta}_2$ are statistically significant at the 1% level with the t-values of 6.79 and -13.10, respectively. This implies that after controlling for country fixed effects, there is an educational Kuznets curve. As a country experiences educational expansion, the proportion of individuals with no formal education decreases. Thus, the country would go through a rising phase of the educational Kuznets curve.

References

- Agrawal, T. (2014). "Educational Inequality in Rural and Urban India." *International Journal of Educational Development*, 34, 11-19.
- Banzragch, O., Mizunoya, S., and Bayarjargal, M. (2019). "Education Inequality in Mongolia: Measurement and Causes." *International Journal of Educational Development*, 68, 68-79.
- 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.
- Castello, A., and Domenech, R. (2002). "Human Capital Inequality and Economic Growth: Some New Evidence." *The Economic Journal*, 112, C187-C200.
- Castello-Climent, A., and Domenech, R. (2021). "Human Capital and Income Inequality Revisited." *Education Economics*, 29(2), 194-212.
- De Gregorio, J. D., and Lee, J. W. (2002). "Education and Income Inequality: New Evidence from Cross-country Data." *Review of Income and Wealth*, 48(3), 395-416.
- Fields, G. S. (2001). *Distribution and Development: A New Look at the Developing World*, The MIT Press, Cambridge, MA.
- Fordvari, P., and van Leeuwen, B. (2011). "Should Less Inequality in Education Lead to a More Equal Income Distribution?" *Education Economics*, 19(5), 537–554.
- Hojo, M. (2009). "Inequality in Japanese Education: Estimation using the Gini Education Coefficient." *Japanese Economy*, 36(3), 3-27.
- Lam, D., and Levison, D. (1992). "Declining Inequality in Schooling in Brazil and Its Effects on Inequality in Earnings." *Journal of Development Economics*, 37, 199-225.
- Lim, A. S. K., and Tang, K. K. (2008). "Human Capital Inequality and the Kuznets Curve." *The Developing Economies*, 46(1), 26-51.
- Lin, C. (2007). "Education Expansion, Educational Inequality, and Income Inequality: Evidence from Taiwan, 1976–2003." *Social Indicators Research*, 80, 601–615.
- Meschi, E., and Scervini, F. (2014). "Expansion of Schooling and Educational Inequality in Europe: the Educational Kuznets Curve Revisited." *Oxford Economic Papers*, 66(3), 660–680.
- Morrisson, C., and Murtin, F. (2013). "The Kuznets Curve of Human Capital Inequality: 1870–2010." *The Journal of Economic Inequality*, 11, 283–301.
- Ram, R. (1990). "Educational Expansion and Schooling Inequality: International Evidence and Some Implications." *The Review of Economics and Statistics*, 72(2), 266-274.
- Shukla, V. & Mishra, U. S. (2019). "Educational Expansion and Schooling Inequality: Testing Educational Kuznets Curve for India." *Social Indicators Research*, 141:1265–1283.