

## Income Inequality and Education in South Asia: Empirical Investigation Under Persistence, Endogeneity and Heterogeneity

Faisal Munir<sup>1</sup> and Sami Ullah<sup>2</sup>

### **Abstract**

*This study empirically investigates the relationship between income inequality and education expansion that is, increasing average years of schooling and reducing inequality of schooling by taking a panel of South Asian countries over the period from five years interval from 1980 to 2015. South Asian economies are consistently studied in the inequality literature and world development report 2018 also focused on the issue of educational concerns and elaborated a high inequality of learning exist in South Asia. Arellano-Bond dynamic panel GMM estimation techniques are used by following existing literature to address issues of persistence and endogeneity, we find a large, positive, statistically significant and stable relationship between average years of schooling and income inequality in South Asian economies. The relationship between income inequality and average years of schooling is positive, consistent with constant or increasing returns to additional years of schooling. Another positive and statistically significant relationship exists between educational cost and income inequality is small and not always statistically significant, we find a statistically significant negative relationship between income inequality and younger cohorts. Statistical tests indicate that our dynamic estimators are consistent and that our identifying instruments are valid. Policy concerns suggest that education expansion will continue to be inequality reducing. This role will diminish as countries develop, but it could be enhanced through a stronger focus on reducing inequality in the quality of education.*

**Key words:** Income Inequality, Education Expansion, Arellano-Bond, Dynamic Panel GMM.

### **Introduction**

The persistence of high and, in many countries, rising income inequality over recent decades is a growing concern for policy makers worldwide, and has received increasing attention both from economists and in public debate (Dabla-Norris et al., 2015). Literature suggested that rising inequality has been attributed to a range of factors, including the globalization and liberalization of factor and product markets; skill-biased technological change; increases in labor force participation by low-skilled workers; declining top marginal income tax rates; increasing bargaining power of high earners; and the growing share of high-income couples and single-parent households (Alvaredo et al., 2013). However, many of these factors have also had beneficial effects on growth and poverty reduction in many economies (Chen and Ravallion, 2010).

The main focus of current study is on the relationship between education expansion and income inequality by taking sample of South Asia. Expansion of education is often seen as an important policy instrument for restricting rising income inequality over the medium term. Not only is education expansion viewed as being important for promoting economic growth (Barro, 2013; Hanushek, 2013), but it can also help to break the intergenerational transmission of poverty and reduce inequality of opportunity (Corak, 2013), which in turn

---

<sup>1</sup> PhD. Scholar of Economics at University of Gujrat and Lecturer in Gujrat institute of Management sciences, PMAS Arid Agriculture University.

<sup>2</sup> Assistant Professor in Economics at University of Gujrat.

reduces future income inequality. Reducing income inequality through education expansion would also reduce the need for fiscal redistribution through distortionary fiscal policies such as progressive income taxes or means-tested transfers. So, from this perspective, education expansion has a “win-win” potential to simultaneously achieve both efficiency and equity objectives.

The paper extends the existing empirical literature in a number of dimensions. First, it expands the econometric analysis to address key estimation challenges not addressed in the existing literature, more specifically the issues of the endogeneity of the education and income inequality relationship and the persistence of income inequality over time with special reference to South Asia. Second, it is unique study in the south Asian case, no previous evidence found in literatures which determine the subject matter of current study. Finally, it also allows for heterogeneity in the relationship between education expansion and income inequality across working-age groups since there is evidence that education and experience are complementary inputs in human capital formation so that returns to education, and thus income inequality, can be expected to increase with working age.

Further sections of the paper provided comprehensive discussion on relevant literature, complete methodology of empirical and theoretical validation of the subject and empirical findings using appropriate econometric techniques of analysis. Conclusion is given after the discussion on findings which includes some policy implications.

## Literature Review

The relationship between education expansion and income inequality is frequently studied phenomena in historical as well as in recent literature. There are number of studies which are in-line with the current study, such as Gregorio (2002) estimated the relationship between education expansion and income inequality in penal of countries on different stages of development and found that education expansion has positive relationship with income inequality and there is negative relationship between cohort years of schooling and income inequality, the study does not taken in account the problem of persistence and endogeneity issues with empirical estimation and concludes that as country develop the relationship between education and income inequality become stronger and positive. Recently the study conducted by Coady and Dizioli (2017) encountered the issues ignored in previous literature, they used penal of developed emerging and developing country on five years interval data to estimate the relationship between education expansion and income inequality under persistence, endogeneity and heterogeneity problem. They estimated Arellano Bond difference GMM, System GMM and long difference estimates to have consistent and efficient parameters. The concludes that education expansion has positive significant relationship with income inequality and average years schooling has also positive impact on income inequality and cohort schooling years are negatively associated with income inequality.

There are a number of studies which elaborated the importance of income inequality concerns to development some studies like Clements et al. (2015) and Dabla-Norris et al. (2015) have investigated the importance of rising income inequality to different economic and social issues. Many other studies have determine the important implications of education expansion on income inequality by following human development models. Studies of Colclough et al. (2010) and Climent and Doménech (2014) determined that an increase in the average level of schooling will unambiguously result in higher income inequality in economies in the process of development. Similarly, Psacharopoulos and Patrinos (2004) concluded that if the return to an extra year of schooling is higher at higher levels of schooling so that, then an increase in the average level of schooling will also unambiguously result in higher income inequality. However, if returns are lower at higher levels of education, as suggested by much of the

empirical literature. Further the literature on methodological issues in estimating the relationship of education expansion and income inequality is followed the Arellano Bond (1991, 1998), Arellao Bover (1995) and Blundell and Bond (1998) which provides discussion of GMM estimations which are important in current study.

In case of south Asian only limited literature is found almost zero literature found on the same subject as previously data and estimations issues were prominent in empirical evidence validations. So, current study is vital contribution in existing literature for South Asia.

## Methodology

The study follows Gregorio (2002) and Coady and Dizioli (2017) by taking standard theoretical framework for analyzing the relationship between education expansion and income inequality is the traditional human capital model. This model implies that the distribution of income (or earnings) is determined by both the level and distribution of education (or schooling) across the population. Using this model, earnings ( $Y$ ) of an individual with  $S$  years of schooling can be approximated as

$$\log Y_s = \log Y_0 + rS + u$$

Where  $Y_0$  is the earnings of individuals with zero formal education,  $r$  is the rate of return to an additional year of schooling, and  $u$  captures other factors that influence earnings independent of education. The dispersion of earnings across individuals in a population can then be written as follows, with bar superscript denoting mean values:

$$\begin{aligned} \text{Var}(\log Y_s) = & \bar{r}^2 \text{Var}(S) + \text{Var}(r) \text{Var}(S) + \bar{S}^2 \text{Var}(r) + 2\bar{r}\bar{S} \text{Cov}(r, S) + \text{Cov}(r, S)^2 \\ & + \text{Var}(u) + 2\text{Cov}(rS, u) \end{aligned}$$

Therefore, an increase in education inequality,  $\text{Var}(S)$ , keeping the average level of schooling and other factors constant, unambiguously results in higher income inequality—i.e., the first two terms are unambiguously positive. However, the impact on income inequality of increasing the average level of schooling,  $\bar{S}$  keeping other factors constant, will depend on the relationship between  $r$  and  $S$ , i.e.  $\text{Cov}(r, S)$  i.e., on the combined effect of the third and fourth terms. If the return to an extra year of schooling is constant across levels of schooling, so that  $\text{Cov}(r, S)=0$ , then an increase in the average level of schooling will unambiguously result in higher income inequality. Similarly, if the return to an extra year of schooling is higher at higher levels of schooling (Colclough and others, 2010; Climent and Doménech, 2014), so that  $\text{Cov}(r, S)>0$ , then an increase in the average level of schooling will also unambiguously result in higher income inequality. However, if returns are lower at higher levels of education, as suggested by much of the empirical literature (Psacharopoulos and Patrinos, 2004), so that  $\text{Cov}(r, S)<0$ , then this will attenuate the increase in income inequality and, if sufficiently negative, may actually result in an increase in average schooling leading to a net decrease in income inequality.

To test the empirical relationship between income inequality and the average level of education and education inequality, we use the following country-panel specification:

$$I_{it} = \beta_0 + \beta_1 S_{it} + \beta_2 E_{it} + \beta_3 X_{it} + \alpha_i + \varepsilon_{it}$$

Where subscripts refer to country  $i$  and year  $t$  respectively,  $I$  is a measure of income inequality,  $S$  is average years of education,  $E$  is a measure of education cost,  $X$  denotes other variables that impact income inequality independently of education outcomes,  $\alpha$  captures unobserved time-invariant country-fixed effects, and  $\varepsilon$  captures other unobserved determinants that can vary across countries and time periods. The study further developed the model by incorporating more variables as follows

$$\begin{aligned} INEQ_{it} = & \beta_0 + \beta_1 S_{it} + \beta_2 E_{it} + \beta_3 GDP_{it} + \beta_4 GDP_{it}^2 + \beta_5 INF_{it} + \beta_6 HDI_{it} + \beta_7 HEX_{it} \\ & + \beta_8 DC_{it} + \beta_9 TO_{it} + \beta_{10} YP_{it} + \beta_{11} OP_{it} + \beta_{12} UP_{it} + \alpha_i + \varepsilon_{it} \end{aligned}$$

The data sources for the key income inequality and education variables used in the analysis are as follows,

**Table 1 Description of variables and data sources**

<b>Variable</b>	<b>Description</b>	<b>Data source</b>
<b>INEQ</b>	Income inequality (GINI coefficient)	WDI (2018), Bastagli et al. (2012)
<b>Average Schooling years</b>	Mean years of Schooling	UNDP (2017)
<b>GDP</b>	Gross domestic product (real growth)	WDI (2018)
<b>Domestic credit</b>	Domestic credit to private sector (% of GDP)	WDI (2018)
<b>Inflation</b>	Yearly % change in CPI	WDI (2018)
<b>HDI</b>	Human Development index	UNDP (2017)
<b>Health expenditures</b>	Net foreign direct investment inflow (% of GDP)	WDI (2018)
<b>Education expenditures</b>	Government expenditures on education (% of GDP)	WDI (2018)
<b>Population (below 15)</b>	Proportion of population aged below 15	WDI (2018)
<b>Population (above 65)</b>	Proportion of population aged 65 and above	WDI (2018)
<b>Urbanization</b>	Urban Population (% to total population)	WDI (2018)
<b>Trade openness</b>	Total trade % of GDP	WDI (2018)

By following previous literature, the paper extends the estimation strategy to address two econometric issues, namely, persistence and endogeneity:

**Persistence of Income Inequality:** Income inequality tends to change only slowly over time with very little within-country variation over the sample period, suggesting that there may be some, possibly unobserved, slowly-changing factors that explain this persistence. For example, this state dependence could reflect factors that prevent intergenerational mobility so that it is harder for a person born poor to achieve social mobility than for a person born in the middle class (Corak, 2013). If these unobserved factors are correlated with education outcomes, then the estimated OLS and fixed-effects coefficients can be biased.

**Endogeneity of Education Outcomes:** Any observed relationship between education outcomes and income inequality may reflect reverse causation, i.e., current income inequality also affects current educational attainment and its dispersion. Therefore, any unobserved factors that affect income inequality and also education outcomes can bias the estimated relationship between education outcomes and income inequality.

To address these two issues, we use dynamic panel estimation techniques. To control for persistence, it is common to include past income inequality levels as an additional independent variable. However, by construction, this implies that the exogeneity assumption in the fixed-effects estimator is violated so that fixed-effects estimates are then biased (Nickell, 1981). To address this problem, Arellano and Bond (1991) suggest using a first-differenced GMM (Diff-GMM) estimator that also deals with the endogeneity problem by first differencing the data and then deploying suitably lagged values of the independent and

dependent variables as instruments. <sup>3</sup>However, Blundell and Bond (1998) show that the Diff-GMM estimator suffers from the weak instrument problem when the number of time periods is small and that this bias is exacerbated when the time series are persistent. Building on Arellano and Bover (1995), the system GMM estimator (Sys-GMM) developed by Blundell and Bond (1998) addresses this weak instrument problem by exploiting level restrictions which remain informative even in the presence of persistence. Thus, where the number of time periods is small and in the presence of persistence, Sys-GMM estimator can produce dramatic efficiency gains over the basic Diff-GMM estimator.<sup>4</sup> For this reason, our preferred model is the Sys-GMM estimator.

## Results and Discussion

The study follows the prescribed econometric analysis to reach the conclusion and to give valid and justifiable policy suggestions. In the first step of findings we applied simple data descriptive analysis to understand the trends and locative measures of data used in analysis.

**Table 2 Descriptive Statistics**

	Mean	Median	Maximum	Minimum	Std. Dev.	Obs.
<b>GINI</b>	35.684	34.300	49.400	26.400	5.053	56
<b>Schooling years</b>	4.089	3.400	10.900	0.800	2.556	56
<b>Education expenditures</b>	3.159	2.966	7.362	0.937	1.299	56
<b>Inflation</b>	7.742	7.042	26.145	-1.175	5.315	56
<b>Domestic credit</b>	25.248	23.119	64.749	2.425	14.105	56
<b>HDI</b>	0.492	0.489	0.766	0.245	0.124	56
<b>Health expenditures</b>	4.316	3.826	9.525	2.135	1.657	56
<b>GDP</b>	4.824	5.397	11.731	-13.129	4.105	56
<b>Population (below 15)</b>	37.070	38.501	47.267	21.243	6.541	56
<b>Population (above 65)</b>	4.212	4.022	9.299	2.461	1.245	56
<b>Urbanization</b>	24.170	24.883	45.536	6.091	8.862	56
<b>Trade openness</b>	65.748	46.174	358.660	13.040	56.875	56

Descriptive statistics includes the Mean, Median, Standard deviation, minimum and maximum values. These statistics are given by combining the data of 7 South Asian countries named; Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan and Sri Lanka. To elaborate more specifically the data of income inequality is given in figure 2 to make comparison of all countries in panel.

<sup>3</sup> This approach is typically seen as superior to that suggested by Anderson and Hsiao (1982) that includes the dependent variable lagged two periods as an independent variable in the differenced equation, which results in biased coefficients when the number of time periods is small.

<sup>4</sup> Note also that the implicit assumption in Sys-GMM is that independent variables are predetermined (or weakly exogenous), depending only on past values of income inequality. For this reason, in addition to the Hausman test, we also do the AR(2) test suggested by Arellano and Bond (1991).

**Figure 1 Income Inequality in South Asia**

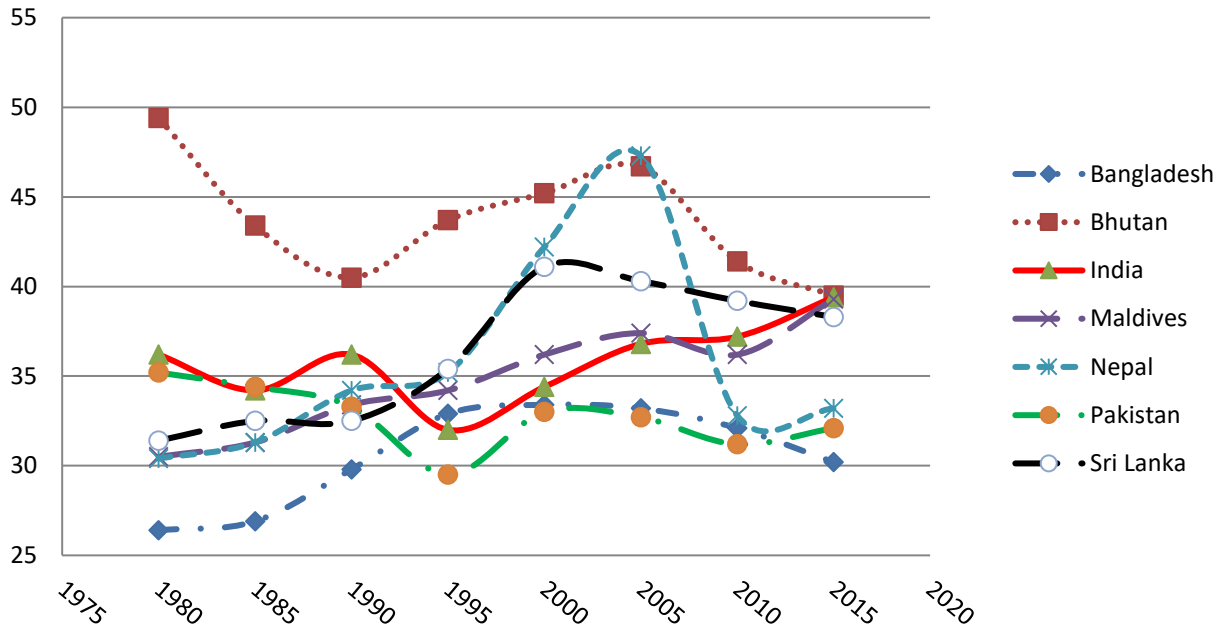


Figure 1 plots the data of income inequality of South Asian economies used in empirical analysis. There is overall persistent and increasing trend in income inequality in all economies except Bhutan.

**Figure 2 Average years of Schooling in South Asia**

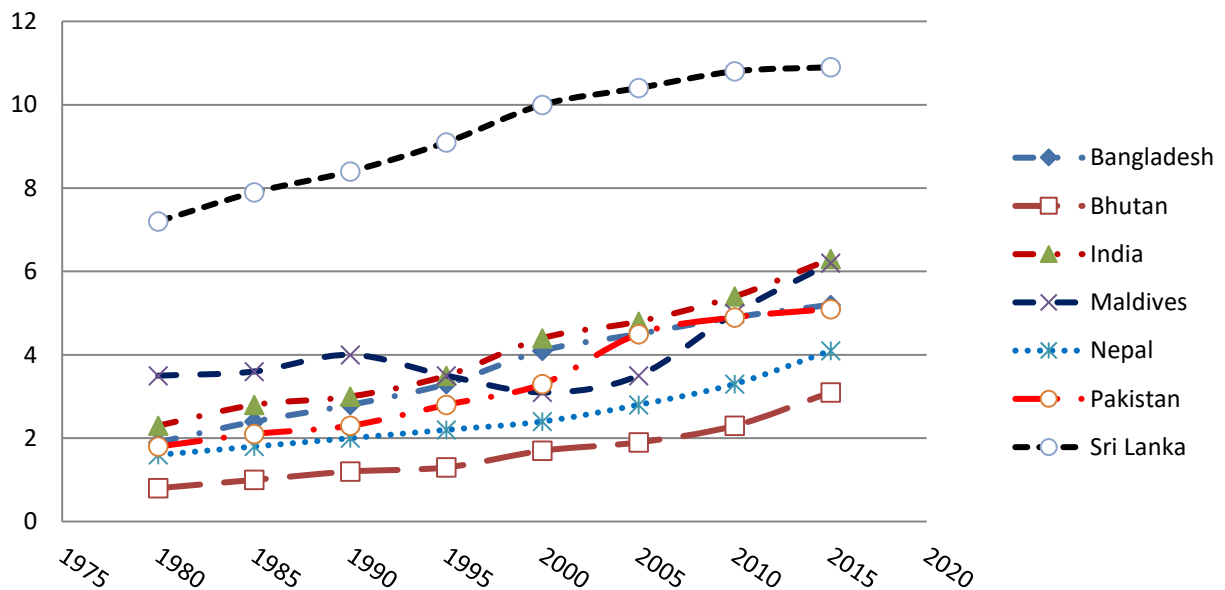


Figure 2 shows the mean years of schooling in south Asian economies that shows increasing trend in all the countries, the data is sourced from United Nations estimates.

In the next step to estimate parameters to address empirical evidence the study applied Arellano Bond dynamic model GMM. The results of different models are given in table 3.

**Table 3 Results using Arellano-Bond Dynamic Panel GMM for Income Inequality**

VARIABLES	Diff-GMM Income Inequality	Sys-GMM(1) Income Inequality	Sys-GMM(2) Income Inequality
Lag(1) Income inequality	0.855** (0.372)	0.636*** (0.134)	0.529*** (0.166)
Schooling years	4.960 (3.732)	1.869*** (0.669)	1.482*** (0.331)
Education	0.673 (1.404)	1.479** (0.697)	1.426** (0.661)
Inflation	-0.520** (0.243)	-0.284* (0.144)	-0.198 (0.153)
Population(below15)	0.766 (0.551)	0.256*** (0.0638)	0.218*** (0.0751)
Domestic Credit			-0.122** (0.0532)
Health			0.737 (0.658)
Trade openness			-0.0262 (0.0200)
GDP			0.0728 (0.207)
GDP <sup>2</sup>			0.0199 (0.0199)
Population(65 and above)	-5.545* (3.106)	-0.789 (0.992)	
HDI	28.73 (38.32)	-6.969 (10.69)	
Instruments for first differences equation Standard	D.(EDU HE UP)	D.(EDU HE UP)	D.(EDU HEX UP TO)
F-Statistics	1.55	948.92	641.10
Prob > F	0.182	0.000	0.000
Arellano-Bond test for AR(1)	Pr > z = 0.031	Pr > z = 0.008	Pr > z = 0.005
AR(2)	Pr > z = 0.732	Pr > z = 0.612	Pr > z = 0.879
Sargan test of overid. restrictions (Prob > chi2)	0.534	0.668	0.806

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 4 provides the results of difference GMM and system GMM in three different models, the first model is estimated with Arellano Bond AR (1) and AR (2) in difference GMM. The results of model one are not significant as the f-test is not significant, however model estimated under sys-GMM is significant and we can report these results as AR(1) test is significant and AR(2) is not significant so the persistence of income inequality is only at AR(1). Average schooling years have significant and positive statistical relationship with income inequality in sys-GMM results. Educational is also significant and positive in contributing income inequality. Other variables used in the model inflation and domestic credit are significantly reducing income inequality. Proportion of young population ages less

than 15 is also significant and contributing in increase in income inequality. HDI and GDP are not statistical significant in our analysis. However, the signs of the parameters are according to existing literature.

## Conclusion

This paper presents empirical evidence on the relationship between education expansion and income inequality in South Asia by using data from 1980 to 2015 with five year interval. It extends the existing literature in a number of dimensions with special reference to South Asia panel. First, it addresses key econometric issues ignored in the existing literature related to the need to allow for the persistence of income inequality and the endogeneity of education and inequality outcomes, both of which require the use of dynamic panel analysis. Second, the analysis tests for heterogeneity in these relationships across country income groups as well as across different age cohorts. Finally, the paper uses reliable income inequality for five year interval that addresses concerns about the quality of the income inequality data currently widely used in the literature, and also extends the period of the analysis.

The analysis demonstrates clearly the importance of controlling for persistence, endogeneity and heterogeneity. When dynamic panel estimation techniques are applied, the positive relationship between education inequality and income inequality becomes substantially larger, statistically significant and stable across the various estimators. This is consistent with our theoretical insights based on the human capital model and confirms that education expansion has significant impact on income inequality through decreasing with multiple stages of development. However, the relationship between income inequality and schooling levels is found to be positive but small and not always statistically significant. Statistical tests indicate that our dynamic estimators are consistent and that our identifying instruments are valid. The study concludes that education expansion along with persistent income inequality and other determinants is significant in income inequality. The policy makers are suggested to design appropriate policy to capture this issue into median to long run for controlling and reducing income inequality through education expansion in countries of South Asia.

## References

- Acemoglu, D. 1997, “Matching, Heterogeneity, and the Evolution of Income Distribution,” *Journal of Economic Growth*, Vol. 2 (March), pp. 61–92.
- Ahluwalia, M. S., 1976, “Income Distribution and Development: Some Stylized Facts,” *American Economic Review*, Papers and Proceedings, Vol. 66 (May), pp. 128–35.
- Alvaredo, F., Atkinson, A., Piketty, T. and Saez, E., 2013, “The Top 1 Percent in International and Historical Perspective,” *Journal of Economic Perspectives*, Vol. 27, No. 3, pp. 3–20.
- Anand, S. and Ravi, K., 1993, “The Kuznets Process and the Inequality-Development Relationship,” *Journal of Development Economics*, Vol. 40 (February), pp. 25–52.
- Arellano, M. and S. Bond. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58. pp. 277 – 297.
- Bastagli, F., Coady, D. and Gupta, S. (2012). “Income Inequality and Fiscal Policy” *IMF Staff Discussion Note*. SDN/12/08.
- Barro, R. 1996, “International Measures of Schooling Years and Schooling Quality,” *American Economic Review*, Papers and Proceedings, Vol. 86 (May), pp. 218–23.
- Barro, R., 2013, “Education and Economic Growth”, *Annals of Economics and Finance*, Vol.14(2), pp. 301–28.



- Becker, G. S. and Barry R. C. 1966, “Education and the Distribution of Earnings,” *American Economic Review, Papers and Proceedings*, Vol. 56 (May), pp. 358–69.
- Benabou and Roland. 1994, “Education, Income Distribution, and Growth: The Local Connection,” NBER Working Paper 4798, National Bureau of Economic Research, Cambridge, Massachusetts.
- Coady, D. and Dizioli A., (2017). *Income Inequality and Education Revisited: Persistence, Endogeneity, and Heterogeneity. IMF Working Paper, WP/17/126.*
- Corak, M. 2013, “Income Inequality, Equality of Opportunity, and Intergenerational Mobility,” *Journal of Economic Perspectives*, Vol. 27 (No. 3, Summer), pp. 79–102.
- Dabla-Norris, E., Kalpana., Nujin, S., Frantisek, R., and Evridiki, T. 2015, “Causes and Consequences of Income Inequality: A Global Perspective”, *IMF Staff Discussion Paper SDN/15/13* (Washington: International Monetary Fund).
- De Gregorio, Jose and Se-Jik Kim, 2000, “Credit Markets with Differences in Abilities: Education, Distribution and Growth,” *International Economic Review*, Vol. 41 (August), pp. 579–607.
- Galor, Oded and Joseph Zeira, 1993, “Income Distribution and Macroeconomics,” *Review of Economic Studies*, Vol. 60 (January), pp. 35–53.
- Hanushek, E., 2013, “Economic Growth in Developing Countries: The Role of Human Capital”, *Economics of Education Review*, 37, 204-212.
- Jenkins, S. 2015. “World Income Inequality Databases: An Assessment of WIID and SWIID,” *Journal of Economic Inequality*, Vol. 13 (December), pp. 629–71.
- Kuznets, Simon, 1955, “Economic Growth and Income Inequality,” *American Economic Review*, Vol. 45 (March), pp. 1–28.
- LaGarda, G. and Kevin G., and Jennifer L. 2016, “Capital Openness and Income Inequality: A Cross-Country Study and The Role of CCT in Latin America,” *Munich Personal RePEc Archive papers* (June).
- Montenegro, C. and Harry. 2014, “Comparable estimates of returns to schooling around the world”, *Policy Research Working Paper WPS 7020*. Washington, DC: World Bank Group.
- Psacharopoulos, George, 1977, “Unequal Access to Education and Income Distribution: An International Comparison,” *De Economist*, Vol. 125, pp. 383–92.
- Psacharopoulos, G. and Maureen W., 1985, *Education for Development: An Analysis of Investment Choices*, Oxford University Press, New York.
- UNDP. (2016). *United Nation Development Program, Country Country Reports.*
- Winegarden, C. R., 1979, “Schooling and Income Distribution: Evidence from International Data,” *Economica*, Vol. 46, pp. 83–7.
- WDI, (2017). *World Development Indicators. World Bank Data base.*