WP/17/126

IMF Working Paper

Income Inequality and Education Revisited: Persistence, Endogeneity, and Heterogeneity

by David Coady and Allan Dizioli

INTERNATIONAL MONETARY FUND

 WP/17/126

IMF Working Paper

Fiscal Affairs Department

Income Inequality and Education Revisited: Persistence, Endogeneity, and Heterogeneity

Prepared by David Coady and Allan Dizioli

May 2017

Abstract

This paper presents new results on the relationship between income inequality and education expansion—that is, increasing average years of schooling and reducing inequality of schooling. When dynamic panel estimation techniques are used to address issues of persistence and endogeneity, we find a large, positive, statistically significant and stable relationship between inequality of schooling and income inequality, especially in emerging and developing economies and among older age cohorts. The relationship between income inequality and average years of schooling is positive, consistent with constant or increasing returns to additional years of schooling. While this positive relationship is small and not always statistically significant, we find a statistically significant negative relationship with years of schooling of younger cohorts. Statistical tests indicate that our dynamic estimators are consistent and that our identifying instruments are valid. Policy simulations 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.

JEL Classification Numbers: D30, H52, I24

Keywords: Education inequality, education levels, income inequality

Author's E-Mail Address: DCoady@imf.org; Allan.Gloe_Dizioli@ecb.europa.eu

Contents **Page**

Abstract 2

I. Introduction 3

II. Economic Theory and Empirical Estimation 4
A. Economic Theory 4
B. Empirical Estimation 5

III. Data Description 8

IV. Empirical Results 9

V. Policy Simulations 14
A. Education and Past Income Inequality 14
B. Education and Future Income Inequality 14

VI. Conclusions 15

References 20

Tables
1. Education and Income Inequality: OLS, SURE and Fixed Effects Estimates 10
2. Education and Income Inequality: Difference, GMM, Long Distance Estimates 12
3. Education and Income Inequality: Cohort Effects 13
4. Implications of Changing Education Outcomes for Income Inequality, 1990–2005 14
5. Implications of Changing Education Outcomes for Income Inequality, 2005–2025 15

Figures
1. Disposable Income Inequality by Region, 1980–2010 8
2. Education Levels and Education Inequality 9

Appendixes
1. Explanatory Variables and their Sources 17
2. Constructing a Measure of Education Inequality 19

Appendix Table
1. Summary Statistics 18

I. 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 (OECD, 2008; Clements and others, 2015; Dabla-Norris and others, 2015). 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 (OECD, 2008; Alvaredo and others, 2013; Hoeller, Joumard, and Koske, 2014). However, many of these factors have also had beneficial effects on growth and poverty reduction both nationally and globally (Chen and Ravallion, 2010; Milanovic, 2012).

The focus of this paper is on the relationship between education expansion and income inequality. Expansion of education is often seen as an important policy instrument for combating 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 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. Second, it uses a new database on income inequality that expands the period of analysis while recognizing the need to use comparable measures of income inequality. Thirdly, it allows for heterogeneity in the relationship between education expansion and income inequality across advanced and developing economies to capture possible differences in returns to education. 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.

The structure of the paper is as follows. In Section 2 we briefly discuss the conceptual framework underpinning the analysis of the impact of education expansion on income inequality and outline our empirical strategy for estimating this relationship. Section 3 discusses the data used in the analysis. Section 4 presents results based on these data and estimation methods, and compares them with the existing literature. Based on these results, Section 5 uses simulation analysis to discuss the implications of past and future changes in education outcomes for income inequality. Section 6 concludes.

II. Economic Theory and Empirical Estimation

A. Economic Theory

The 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[1]:

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

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:

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

Therefore, an increase in education inequality, 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. 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 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; Castelló-Climent and Doménech, 2014), so that 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 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.[2]

---

[1] Schooling is taken here as a proxy for human capital. More generally, it can also be viewed as a proxy for other forms of human capital accumulation such as on-the-job training. For a discussion of the importance of schooling quality, see Barro (2013) and Hanushek and Wobmann (2010). For a review of country-level estimates based on household survey data, see Montenegro and Patrinos (2014).

[2] Note that the total impact of education expansion on income inequality will depend on the relationship between average schooling and its dispersion. For instance, at the early stages of development, both the average level and dispersion of schooling are typically low. Expansion of education will initially tend to increase income inequality as a few more individuals gain higher education and earnings, but eventually lower the inequality of education and earnings as education becomes more widespread. The increase in the supply of high-education individuals will also tend to decrease the skilled wage premium and thus also income inequality. The net effect of education expansion on income inequality will therefore depend on the relative sign and magnitudes of these

(continued)

B. Empirical Estimation

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

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

where subscripts refer to country *i* and year *t* respectively, I is a measure of income inequality, E is average years of education, σ is a measure of education inequality, X denotes other variables that impact income inequality independently of education outcomes, α captures unobserved time-invariant country-fixed effects, and ε captures other unobserved determinants that can vary across countries and time periods.

The data sources for the key income inequality and education variables used in the analysis are as follows (see Appendix 1 for details on the other explanatory variables included in the regression)**:**

- *Income Inequality* (I): For our analysis we use the Gini coefficient for disposable income inequality as our dependent variable since this is the inequality index that is most widely available and used in the related literature. We use an updated Gini coefficient database based on that assembled by Bastagli, Coady and Gupta (2012), which emphasized the need for comparability of country inequality estimates across time. Gini estimates are taken at five-year intervals.

- *Average Education* (E): The average level of education of a country is taken as the average years of school attainment for the population aged 25 and over from Barro and Lee (2013). These data are collected from census and survey information in five-year intervals, as compiled by UNESCO, Eurostat, and other sources. Average education is constructed based on the distribution of education attainment in the population over age 25, by five-year age groups and, for most cases, in six attainment categories: no formal education, incomplete primary, complete primary, incomplete secondary, complete secondary, and complete tertiary.

- *Education Inequality* (σ): The inequality of education in a country is taken as the Gini coefficient of years of education for a given five-year interval based on educational attainment data from Barro and Lee (2013). Appendix 2 discusses the construction of this Gini in more detail.

Our point of departure is the papers by De Gregorio and Lee (2002), Castelló-Climent and Doménech (2014), and Dabla-Norris and others (2015)—henceforth referred to as DGL, CCD, and DNO, respectively. DGL used country-panel data for around 70 advanced and developing economies covering the five-year periods from 1965 to 1990. They estimated the relationship between education outcomes and income inequality (as captured by the Gini

"composition" and "compression" effects, and is more likely to be positive at low levels of development and education attainment (Knight and Sabot, 1983).

coefficient for disposable income) using the technique of seemingly unrelated regressions (SURE). Exploiting only the cross-country variation, they assumed that coefficient estimates were common across panels and thus ignored country-fixed effects. If these assumptions are valid then SURE provides more efficient estimates than OLS. As with previous attempts to estimate the relationship between education and income inequality[3], they found that income inequality increases with education inequality and decreases with the average level of education. However, the positive coefficient on inequality of education was statistically insignificant when controls for a country's GDP per capita and the level of public social spending were added.

Similar results were found by CCD, who extended the database to 2010 and used a fixed-effects estimation model to control for unobserved factors (such as historical factors, institutions, or culture) specific to a country that affect the level of income inequality but do not change over time. They also used the Gini measure of inequality from the SWIID database (unlike DGL who used the WIID database). Since there are valid concerns about the use of these Gini coefficients, in the current paper we use the Gini database constructed by Bastagli, Coady and Gupta (2012), which emphasizes the importance of comparability of inequality measures over time.[4] CCD also controlled for other factors such as technological change (i.e, a time trend and the share of high-technology exports in total exports) and globalization (i.e., proxies for trade and financial openness). Reflecting their primary focus on "the race between education and technological change" they also include variables for the ratio of the average years of tertiary to primary education in the population 25 years and older as well as the Gini coefficient for education outcomes. When controlling for fixed effects, the authors find a positive relationship between inequality of education outcomes and income inequality, although this is not always significant. They also find a negative relationship between the relative supply of skills, which we interpret as their proxy for the level of education, and income inequality, although again this is not significant in all specifications.

DNO also used a fixed-effects estimation model with average years of education and education inequality as dependent variables along with other dependent variables similar to those used in CCD. Their analysis does not find any statistically significant relationship between the income Gini and education inequality and education levels. They find a negative but insignificant relationship between education inequality and income inequality and a positive (negative) but insignificant relationship for education level in advanced (emerging) countries. They do however find a significantly positive (negative) relationship between education level and the share of income accruing to the top (middle) income decile.

This paper extends the estimation strategy to address two remaining 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

[3] For reviews of past studies, see Psacharopoulos and Woodhall (1985, pp264-70) and Ram (1989).

[4] The use of country fixed-effects estimators will control for time-invariant differences in income inequality measures across countries, e.g. due to the use of different household surveys based on incomes, expenditures or consumption. See Jenkins (2015) for a detailed and critical discussion of the issues that arise with the use of the WIID and SWIID income inequality databases.

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.[5] 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.[6] For this reason, our preferred model is the Sys-GMM estimator.

Recent research has shown that Sys-GMM may equally suffer from the weak instrument problem, particularly when the time series is large and when substantial unobserved heterogeneity exists (Hayakawa, 2006; Bun and Windmeijer, 2010). We therefore complement our Sys-GMM estimates by using the approach adopted by Barro and Lee (2010), which instruments a cohort's schooling with the educational attainment of its parents (Ins-GMM). Finally, to further test the robustness of our estimates, we employ the long-difference instrumental variables estimator proposed by Hahn, Hausman and Kuersteiner (2006), which has

[5] 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.

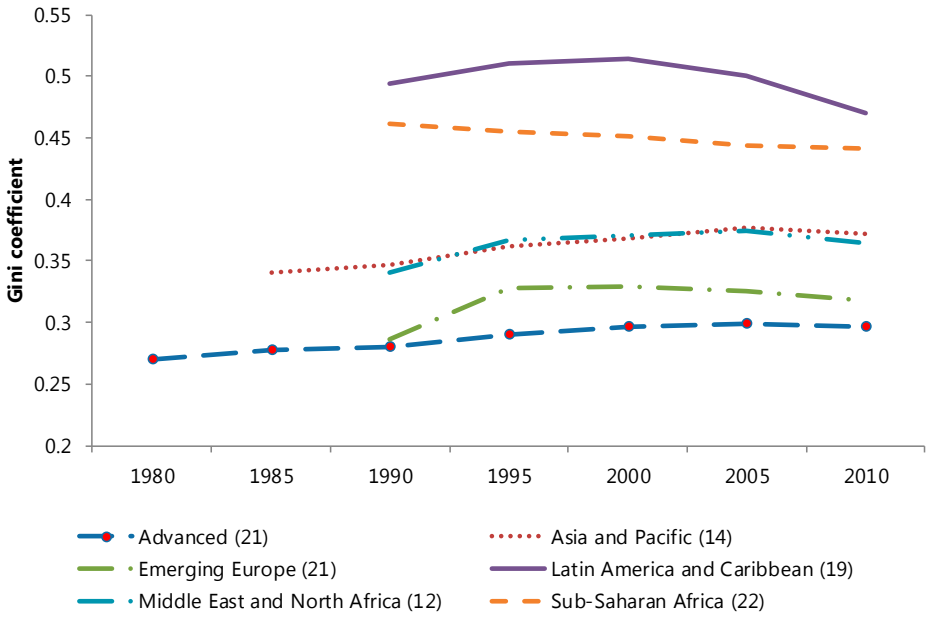
[6] Note also that the implicit assumption in Sys-GMM is that independent variables are predetermined (or weakly exogenous), depending only on past values of income inequality. For example, when a family decides on the choice of education in year t, it takes into account income developments up to this year and does not anticipate future income developments. This assumption can be tested with a Hausman test. The assumption of weak exogeneity also implies a lack of autocorrelation in the error terms. Testing for lack of second-order correlation in the difference equation is therefore equivalent to testing the validity of the weak exogeneity assumption. For this reason, in addition to the Hausman test, we also do the AR(2) test suggested by Arellano and Bond (1991).

been shown to be much less biased and more efficient than conventional implementations of the Sys-GMM when the number of time periods is small and in the presence of persistence.

III. Data Description

As indicated, the measure of income inequality used in the analysis is the Gini coefficient for disposable income, taken from the database constructed by Bastagli, Coady and Gupta (2012). More specifically, we use Gini coefficients for five-year intervals from 1980 to 2010. Based on this, Figure 1 shows the profile of regional average income inequality over the last four decades. Whereas Gini coefficients are available from 1980 for advanced economies, for many other regions they are only available from 1990 onwards.

Figure 1. Disposable Income Inequality by Region, 1980–2010



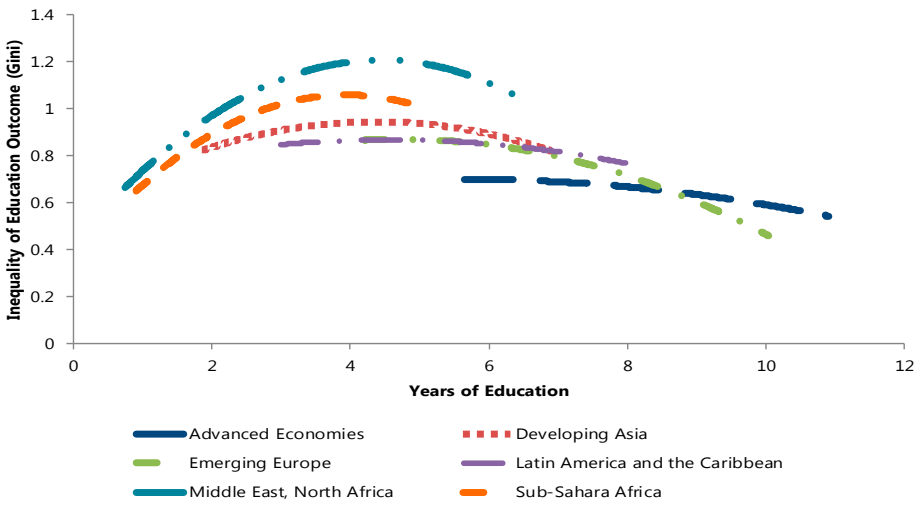
Note: Gini coefficients are regional averages of country-level estimates at 5-year intervals.
Source: Updated by authors based on Bastagli, Coady and Gupta (2012).

Average income inequality is lowest in advanced countries (ADV) but has increased steadily over the last four decades. In Emerging European countries (EE), income inequality exhibits a step increase between 1990 and 1995, a period of substantial structural change associated with transition to market economies, but has stayed relatively stable (with a slight decrease) since then. Average income inequality in the Asian and Pacific (AP) and in the Middle East and North Africa (MENA) regions are higher than in ADV and EE and have also increased steadily over recent decades. More striking is the relatively large level of income inequality in

countries in sub-Saharan Africa (SSA) and especially in Latin America and the Caribbean (LAC).[7] In SSA, while income inequality has been declining slowly it still remains among the highest in the world. Similarly, although average income inequality in LAC has declined sharply since its peak in 2000 it still has the most unequal income distribution in the world.

Figure 2 plots the inequality of education against average years of schooling for different country groupings. As expected, the data suggest an inverted-U relationship with education inequality increasing over lower levels of schooling before declining. However, the precise relationship seems to differ across country groups. Noticeable is the relatively sharp increase in education inequality in MENA and SSA countries as education expands from initial low levels. As a result, at around an average of 4 years of schooling, education inequality was substantially higher in MENA and SSA than in AP or LAC. Therefore, MENA and SSA appear to have experienced a much more unequal early expansion of education than in other regions, although there is some evidence that education inequality is beginning to decrease sharply as access to education expands further.[8]

Figure 2. Education Levels and Education Inequality



Note: Lines are based on non-parametric regressions.
Source: Based on data from Barro and Lee (2013).

IV. Empirical Results

To facilitate comparison with the existing literature, we start by estimating the above equation using SURE and fixed-effects estimators. Table 1 (column 1) shows the correlations between both education levels and education inequality and income inequality based on an OLS estimator. Whereas income inequality is lower at higher education levels (consistent with

[7] Note that income inequality measures in both LAC and SSA are typically based on consumption or expenditure inequality, which tends to be lower than the income inequality measures typically used in ADV and EE. Therefore, the gap between ADV and other country groups may actually be higher.

[8] Whereas in ADV and EE much of the expansion in education levels between 1950 and 2010 came from a decreasing share of the working population with primary education, in other regions, but especially in MENA and SSA, it has been driven by a decreasing share of the working population with no formal education.

decreasing returns to education), income inequality decreases with education inequality, although this relationship is insignificant. Together, variation in these variables explains 22 percent of the total variation in income inequality. Most of the variation in income inequality is explained by cross-country variation as opposed to changes over time. For instance, when time dummies are added (column 2) the explained variation increases to just 30 percent. However, when country-dummies are also added the explained variation increases to 87 percent. While both coefficients are negative and significant when time dummies are included, both are insignificant when country dummies are added highlighting the importance of controlling for country-fixed effects.

Table 1. Education and Income Inequality: OLS, SURE and Fixed Effects Estimates

	OLS Simple	OLS Time	OLS Time and Country	OLS All	SURE	FEALL
	(1)	(2)	(3)	(4)	(5)	(6)
Gini Education	-2.909	-18.094*	0.000	-5.847	-1.498	3.300
	(0.77)	(0.09)	(1.00)	(0.55)	(0.88)	(0.68)
Schooling	-1.716***	-3.101***	-0.395	-0.797	-0.980*	-0.795*
	(0.00)	(0.00)	(0.58)	(0.14)	(0.09)	(0.09)
GDP				9.322**	11.858***	7.696**
				(0.03)	(0.01)	(0.03)
GDP^2				-0.571**	-0.647***	-0.398**
				(0.02)	(0.01)	(0.02)
Openness				-7.479**	-2.945	3.411
				(0.03)	(0.15)	(0.25)
Social Protection				-0.499***	-0.491***	-0.040
				(0.00)	(0.00)	(0.65)
Population over 65				0.256	0.409*	0.419**
				(0.31)	(0.07)	(0.03)
Population less than 15				0.349**	0.523***	0.185
				(0.02)	(0.00)	(0.15)
Inflation				0.001***	-0.017	0.000***
				(0.00)	(0.52)	(0.00)
Urban				0.069**	0.003	-0.145**
				(0.04)	(0.95)	(0.02)
Capital Account Openness				0.045**	0.016	0.005
				(0.03)	(0.35)	(0.73)
Credit Growth				0.031**	0.021**	0.018***
				(0.02)	(0.03)	(0.00)
N	873	873	873	418	32	418
Adjusted R^2	0.22	0.29	0.85	0.78	0.82	0.17

Note: p-values in parentheses: * p<0.1 ** p<0.05 *** p<0.01

When other determinants of income inequality are added to the specification (column 4), the coefficient signs on both education level and education inequality remain negative but the coefficients become smaller in absolute terms (compared to column 2) and insignificantly different from zero. The explained variation decreases to 79 percent reflecting the absence of country dummies. Estimating this relationship using SURE results in a smaller (in absolute terms) and still insignificantly negative coefficient on education inequality, but the negative coefficient on education level becomes larger in absolute terms and significant. Applying fixed effects results in the negative coefficient on education levels decreasing in absolute terms and remaining significant. Although the coefficient on education inequality now becomes positive it remains insignificant.

Table 2 presents the results when we attempt to control for endogeneity and persistence. With the first-difference estimator, the coefficient on education levels remains negative, decreases in absolute terms compared to the fixed effects estimate but now becomes insignificant. Consistent with theory, the coefficient on education inequality remains positive although still insignificant. Under both difference GMM (Diff-GMM) and system GMM (Sys-GMM) the positive coefficient on education inequality increases around six fold and becomes significant at the 10 and 5 percent levels respectively, while the coefficient on education levels becomes positive but remains insignificant.[9] When we instrument education levels and education inequality using parents' schooling levels (Ins-GMM) the coefficient on education inequality remains positive, large and significant, while the positive coefficient on education level now becomes significant. Therefore, the coefficients on education inequality and levels remain quite stable across all the GMM estimators. The Arellano Bond tests for the serial correlation of the disturbances indicate that our GMM estimators are consistent. Moreover, both the Sargan and Hansen tests for over-identifying restrictions do not reject the hypothesis that our instruments are valid in our GMM estimations. Under the long-distance estimator, the positive coefficient on education inequality becomes smaller and insignificant while the coefficient on education levels remains positive but is smaller and insignificant. However, the insignificance found when running the long distance estimator could be explained by the reduction in the number of observations that we have to incur to implement this estimation.

When we allow these coefficients to differ between advanced and other countries, we find that the positive coefficient on education inequality is halved in advanced while the coefficient on education level remains positive and insignificant in both sets of countries. This result suggests that the impact of education inequality on income inequality is not homogeneous across different levels of development and shows that reducing education inequality is an even more important policy for developing countries while becoming less important as countries develop.

[9] We conducted a Wald test to assess the importance of keeping both education inequality and level in the regressions. We rejected the null of the test at 1 percent significance so decided to keep education level in the regressions.

Table 2. Education and Income Inequality: Difference, GMM, Long Distance Estimates

	First Difference	Diff-GMM	Sys-GMM	Ins-GMM	Long Difference	Sys-GMM
	(1)	(2)	(3)	(4)	(5)	(6)
Gini Education	6.213	36.816*	34.445**	30.466**	12.101	31.038**
	(0.42)	(0.05)	(0.02)	(0.03)	(0.18)	(0.03)
Schooling	-0.390	1.087	1.310	1.453**	0.194	0.965
	(0.36)	(0.32)	(0.13)	(0.05)	(0.73)	(0.24)
Gini Education Advanced						-15.740**
						(0.02)
Schooling Advanced						0.123
						(0.52)
GDP	11.566**	13.420*	11.582***	11.464***	6.986*	14.077***
	(0.03)	(0.06)	(0.00)	(0.00)	(0.06)	(0.00)
GDP^2	-0.599**	-0.714*	-0.694***	-0.703***	-0.463**	-0.830***
	(0.03)	(0.06)	(0.00)	(0.00)	(0.03)	(0.00)
Openness	4.481**	4.755*	0.820	0.325	2.953	-0.860
	(0.04)	(0.07)	(0.78)	(0.91)	(0.19)	(0.67)
Social Protection	-0.068	-0.120	-0.236	-0.162	-0.115	-0.239
	(0.50)	(0.53)	(0.16)	(0.27)	(0.26)	(0.12)
Population over 65	0.299	0.382	0.157	0.014	0.373*	0.004
	(0.23)	(0.12)	(0.56)	(0.95)	(0.07)	(0.98)
Population less than 15	0.132	0.010	0.028	0.103	0.120	0.033
	(0.44)	(0.97)	(0.81)	(0.32)	(0.31)	(0.82)
Inflation	0.001***	0.000***	0.000*	0.000*	0.000	0.001**
	(0.00)	(0.00)	(0.06)	(0.05)	(0.64)	(0.01)
Urban	-0.124	-0.195*	-0.019	-0.025	-0.094	0.001
	(0.14)	(0.09)	(0.51)	(0.41)	(0.18)	(0.97)
Capital Account Openness	0.014	0.014	0.025*	0.026*	0.010	0.033**
	(0.28)	(0.43)	(0.08)	(0.08)	(0.48)	(0.01)
Credit Growth	0.017**	0.014	-0.001	0.004	0.005	0.008
	(0.01)	(0.20)	(0.90)	(0.56)	(0.56)	(0.33)
L.gini_net		0.416**	0.644***	0.593***	0.588***	0.503***
		(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
N	322	309	402	402	164	402
R^2	0.10				0.37	
# of Instruments		47	51	53		73
AR(1) Test p-val.		0.046	0.002	0.002		0.006
AR(2) Test p-val.		0.133	0.163	0.150		0.145
Hansen J Test p-val.		0.136	0.188	0.160		0.850
Sargan Test p-val.		0.182	0.517	0.454		0.060

Note: p-values in parentheses: * p<0.1 ** p<0.05 *** p<0.01

The literature on the returns to education levels points to a "fanning out" of these returns with age (a proxy for years of experience) consistent with the returns to education accruing later in an individual's working career. Therefore, inequality in years of education is more likely to translate into higher income inequality among older workers than among younger workers (Table 3). To test this, we add education inequality for older workers as an additional explanatory variable. Consistent with the literature on returns to education, we find that higher inequality in education levels among older workers (keeping inequality of education for all workers constant) is associated with a substantially larger level of income inequality. In addition, the acceleration of the skill-biased technological change means that higher education levels for the young could reduce the income gap between older and younger generations and thus

income inequality. We test this by including average education for the young as an additional explanatory variable and find that higher education levels for the young are indeed associated with lower income inequality.

Table 3. Education and Income Inequality: Cohort Effects

	OLS	FE	First Difference	Diff-GMM	Sys-GMM	Ins-GMM
	(1)	(2)	(3)	(4)	(5)	(6)
Gini Education	2.102	-0.990	1.256	7.043	5.225	1.719
	(0.55)	(0.83)	(0.80)	(0.47)	(0.37)	(0.67)
Schooling	-0.317	-0.268	-0.792	1.076	1.986***	1.958***
	(0.50)	(0.58)	(0.16)	(0.27)	(0.00)	(0.01)
Gini Education Old	-6.720	21.803	10.180	72.386	63.903*	69.535*
	(0.79)	(0.37)	(0.68)	(0.23)	(0.08)	(0.09)
Schooling Young	-0.188	-0.901	-0.689	-1.342	-2.830**	-2.250**
	(0.86)	(0.11)	(0.23)	(0.27)	(0.01)	(0.04)
GDP	10.337**	8.340**	12.837**	12.232	12.327***	11.999***
	(0.01)	(0.03)	(0.03)	(0.14)	(0.00)	(0.00)
GDP^2	-0.631***	-0.406**	-0.665**	-0.627	-0.731***	-0.734***
	(0.01)	(0.03)	(0.03)	(0.15)	(0.00)	(0.00)
Openness	-7.426**	4.342	4.782**	4.917*	1.318	1.390
	(0.04)	(0.11)	(0.03)	(0.10)	(0.59)	(0.59)
Social Protection	-0.488***	-0.003	-0.073	-0.413*	-0.167	-0.155
	(0.00)	(0.97)	(0.53)	(0.09)	(0.21)	(0.28)
Population over 65	0.298	0.353**	0.355	0.746***	0.012	-0.020
	(0.33)	(0.05)	(0.16)	(0.01)	(0.96)	(0.94)
Population less than 15	0.358**	0.054	0.119	0.012	0.171	0.252**
	(0.03)	(0.65)	(0.49)	(0.96)	(0.22)	(0.02)
Inflation	0.001***	0.000*	0.001***	0.001***	0.001*	0.001
	(0.00)	(0.08)	(0.00)	(0.00)	(0.07)	(0.12)
Urban	0.059*	-0.101	-0.099	-0.222**	-0.017	-0.025
	(0.08)	(0.14)	(0.21)	(0.03)	(0.58)	(0.49)
Capital Account Openness	0.047**	0.014	0.014	0.010	0.030**	0.032**
	(0.02)	(0.31)	(0.29)	(0.52)	(0.01)	(0.01)
Credit Growth	0.034**	0.020***	0.017**	0.013	0.008	0.011
	(0.01)	(0.00)	(0.02)	(0.11)	(0.36)	(0.25)
N	395	395	301	283	374	374
# of Instruments				45	50	52
AR(1) Test p-val.				0.374	0.007	0.008
AR(2) Test p-val.				0.149	0.189	0.162
Hansen J Test p-val.				0.375	0.627	0.533
Sargan Test p-val.				0.357	0.389	0.473

Note: p-values in parentheses: * p<0.1 ** p<0.05 *** p<0.01

V. POLICY SIMULATIONS

In this section, we use the results from our preferred Sys-GMM estimation to simulate the impacts of education inequality and level on income inequality to get a sense of their quantitative importance in determining income inequality. We start by analyzing how much of the change in income inequality over the last fifteen years from 1990 to 2005 can be attributed to changes in education outcomes. We then analyze how changing education outcomes over the subsequent two decades might impact income inequality.

A. Education and Past Income Inequality

Trends in income inequality between 1990 and 2005 varied substantially across regions. On average in the sample, inequality decreased in both SSA and MENA countries while other regions experienced increases with the Gini increasing by over 5 points in both AP and EE (Table 4). Across all regions, decreases in the inequality of education reduced income inequality, ranging from a 1.5 point decrease in ADV and EE to a 4.8 point decrease in MENA. However, increases in the level of education increased income inequality across all regions by 1.5 to 2.2 points. While the net effect of changing education outcomes was to slightly increase income inequality in ADV, the effect was negative in all other regions. These results highlight the important role education investments can play in mitigating increases in income inequality, in particular the key role played by education expansion strategies that emphasize more equal access to education

Table 4. Implications of Changing Education Outcomes for Income Inequality, 1990–2005

Regions	Inequality in 1990	Inequality in 2005	Change in Income inequality	From education inequality (absolute change)	From years of schooling (absolute change)	Total change in Income inequality from education	Number of countries
Advanced Economies	28.1	30.4	2.3	-0.9	1.5	0.5	32
Emerging Europe	25.3	33.0	7.7	-2.9	1.5	-1.4	11
Latin America & the Caribbean	46.4	48.0	1.5	-2.8	1.9	-0.9	21
Asia & Pacific	33.5	38.6	5.1	-3.8	1.9	-1.8	15
Middle East & North Africa	38.7	37.2	-1.6	-4.8	2.2	-2.7	7
Sub-Saharan Africa	49.4	43.6	-5.8	-3.6	1.7	-1.9	17

Source: Simulations based on Table 2, Column 6.

B. Education and Future Income Inequality

Table 5 presents projections of the impact of future changes in education levels and education inequality on income inequality from 2005 to 2025. The average level of education is projected forward based on known education levels for current working cohorts while future working cohorts are assumed to have the same education outcomes of the youngest current working cohort (i.e., those aged 20–25 years in 2005). These changes in education levels are then used to project the impact on education inequality. The results indicate that changes in education outcomes will continue to have a dampening impact on income inequality over this period in most regions, with the exception of ADV, where it will increase income inequality by 0.1 points. Similar to the period 1990–2005, continued decreases in education inequality will have

a dampening impact on income inequality while increasing education levels will increase income inequality, with both impacts being smaller than for the earlier shorter period. However, the results for ADV suggest that the income inequality reducing impact of education can be expected to decrease as the level of education increases, and the inequality of education decreases, in emerging and developing economies and converges to levels observed in advanced economies. This highlights the importance of also focusing on reducing the inequality of education quality (e.g., as captured by cognitive skills) to enhance the income inequality reducing impact of education expansion.

Table 5. Implications of Changing Education Outcomes for Income Inequality, 2005–2025

Regions	Inequality in 2005	From education inequality (% improvement)	From years of schooling (% improvement)	Total	Number of countries
Advanced Economies	30.4	-0.7	1.2	0.1	35
Emerging Europe	33.0	-1.4	1.1	-0.1	12
Latin America & the Caribbean	48.0	-2.9	1.4	-1.3	25
Asia & Pacific	38.6	-3.0	1.3	-1.5	23
Middle East & North Africa	37.2	-4.1	1.6	-2.3	18
Sub-Saharan Africa	43.6	-3.2	1.1	-2.0	29

Source: Simulations based on Table 2, Column 6.

VI. CONCLUSIONS

This paper presents new results on the relationship between education expansion and income inequality. It extends the existing literature in a number of dimensions. 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 a new database on income inequality 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 reduces income inequality through decreasing the inequality of education. 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 that our identifying instruments are valid.

Our policy simulations confirm that the net impact of education expansion over the last fifteen years has been to reduce income inequality, especially in emerging and developing economies. Although the magnitude of the net impact on income inequality varies across

emerging and developing economies, it is always inequality reducing. This reduction reflects decreasing education inequality, which is only partly offset by the inequality-increasing effects of rising education levels. In advanced economies, education expansion is associated with a net increase in income inequality. This reflects the relatively smaller impact of decreasing education inequality at the lower levels of education inequality observed in advanced economies being offset by the income inequality-increasing impact of rising levels of education (consistent with constant or increasing returns to additional years of education).

Projections forward over two decades suggest that education expansion will continue to have an inequality-reducing impact in emerging and developing economies. Even though the inequality-reducing impact of falling education inequality is offset by the inequality-increasing impact of rising education levels, the net impact is still inequality reducing. The inequality-increasing impact of education expansion in advanced economies suggests that the inequality-reducing role of education expansion in emerging and developing economies will diminish as these countries develop. Therefore, other policies will also be needed to address rising income inequality. Among these, it is likely that focusing on reducing the inequality of education quality (e.g., improving cognitive skills) can help to enhance the role of education expansion as a force for reducing income inequality.

Appendix 1. Explanatory Variables and their Sources

The explanatory variables used in this paper were selected based on factors identified in the literature as important determinants of income inequality. In this appendix, we briefly discuss the variables and their sources.

Level of income. We test the Kuznets inverted-U hypothesis. According to this hypothesis, there is a positive correlation between income inequality and per capita income at low levels of income, which eventually becomes negative after a country reaches a specific level of development. In order to test this hypothesis, we use the log of ppp per capita income and its square. We use the variable "GDP per capita, ppp" in the World Bank's World Development Indicators (WDI).

*Social public spending***.** Spending on social programs is expected to be progressively targeted and thus to reduce inequality. Government social expenditure is obtained from IMF, Government Finance Statistics Yearbook. It includes pension and other welfare benefits. For most countries data is available from 1990, but advanced economies have available data from 1970.

Trade openness. According to the Heckscher-Ohlin and the Stolper-Samuelson theoretical framework, greater trade openness should increase the relative demand and prices for unskilled labor and lead to a more equal distribution of wages in low-skilled-labor abundant countries and a more unequal distribution of wages in high-skilled labor abundant countries. We use the commonly used variable (exports + imports) / GDP to capture a country's degree of trade openness.

Capital account openness. By relaxing credit supply constraints, capital openness is expected to reduce income inequality if institutions are strong (LaGarda et al, 2016). In this paper, we use the variable "cap100" of the IMF's Annual Report on Exchange Restrictions to reflect the degree of openness of the capital account. Higher values of the indices represent greater openness; lower values of the indices represent greater restrictiveness.

Credit to the private sector. Similar to capital account openness, greater availability of credit should minimize the amount of credit-constrained people in a country and reduce income inequality. In our paper, we use the variable "credit to the private sector/GDP" from the IMF's Government Finance Statistics Yearbook.

Dependency ratio. Theoretical economic models posit that, all other things being equal, an ageing population and a rising dependency ratio tend to increase income inequality (Von Weizsäcker, 1995). We use population data from the World Bank's WDI. Specifically, we use the proportion of the elderly (above 65 years old) and the children (below 16 years old) in the population.

Inflation. Several empirical and theoretical studies have analyzed the relationship between inflation and income inequality. The data on these studies have consistently shown a positive correlation between these two variables. In this paper, we use average annual CPI inflation as measured in the IMF's Government Finance Statistics Yearbook.

Urban population. Reflecting the gap in income between rural and urban populations, different urbanization levels can lead to differences in income inequality. We use the percentage of the population living in urban areas in the World Bank's WDI.

Appendix Table 1. Summary Statistics

	Mean	SD	Obs
GDP per capita PPP (constant 2000 US$)	5617.89	8872.32	1588
Social spending (%GDP)	6.30	6.10	813
Trade openness ((Exports+Imports)/GDP)	0.22	0.16	775
Capital account openness (Cap100)	58.53	29.38	1134
Credit to the private sector (%GDP)	37.75	38.02	1400
Elderly population (% total population)	6.05	4.06	1985
Children (% total population)	35.47	10.05	1985
Inflation (% average)	37.35	412.93	1267
Urban (% population)	48.57	25.40	2150

Appendix 2. Constructing a Measure of Education Inequality

There is no measure of education inequality in Barro and Lee database, which is the source for our education data. Thus, we estimate each country's education inequality using the standard method to calculate income inequality from a Lorenz curve.

In order to implement this method, we need a disaggregated measure of schooling attainment. The education attainment data provided in the Barro and Lee database is presented as the fraction of the population with no education, and with incomplete and complete primary, secondary and tertiary schooling. Thus, the first step in the estimation of the education Gini for education inequality is to assign an average education level to each of these categories of education achievement.

While assigning a specific average to these categories might be considered somehow arbitrary, the estimation results did not change much when we assigned slightly different numbers to these categories. Specifically, we assigned a value of 1 for the category no education, a value of 4 for incomplete and a 7 for complete primary education. We assigned a value of 10 for incomplete and a 13 for complete secondary education. Finally, we assigned a value of 16 for incomplete and 19 for complete tertiary education.

Finally, with these average numbers for each education category, we can calculate total number of years of education for different segments of each country's population. Using these total years of education, we simply apply the concept of the Lorenz curve using the population of each segment as weights.

REFERENCES

Acemoglu, Daron, 1997, ''Matching, Heterogeneity, and the Evolution of Income Distribution,'' *Journal of Economic Growth*, Vol. 2 (March), pp. 61–92.

Adelman, Irma and Cynthia T. Morris, 1973, *Economic Growth and Social Equity in Developing Countries*, Stanford University Press, Stanford, California.

Ahluwalia, Montek S., 1976, ''Income Distribution and Development: Some Stylized Facts,'' *American Economic Review*, Papers and Proceedings, Vo. 66 (May), pp. 128–35.

Alvaredo, F., A. Atkinson, T. Piketty, and E. Saez, 2013, "The Top 1 Percent in International and Historical Perspective," *Journal of Economic Perspectives*, Vol. 27, No. 3, pp. 3–20.

Anand, Sudhir and Ravi Kanbur, 1993, ''The Kuznets Process and the Inequality-Development Relationship,'' *Journal of Development Economics*, Vol. 40 (February), pp. 25–52.

Atkinson, Anthony B., 1997, ''Bringing Income Distribution in from the Cold,'' *Economic Journal*, Vol. 107 (March), pp. 297–321.

Barro, Robert and Jong-Wha Lee, 1994, ''Sources of Economic Growth,'' *Carnegie-Rochester Series on Public Policy, Vol.* 40 (June), pp. 1–46.

———, 1996, ''International Measures of Schooling Years and Schooling Quality,'' *American Economic Review*, Papers and Proceedings, Vol. 86 (May), pp. 218–23.

Barro, Robert, 1999, ''Inequality, Growth, and Investment,'' mimeo, Harvard University, Cambridge, Massachusetts.

———, 2013, "Education and Economic Growth", Annals of Economics and Finance, Vol. 14(2), pp. 301–28.

——— and Jong-Wha Lee, 2015, *Education Matters: Global Schooling Gains from the 19th to the 21st Century,* Chapter 2 (London: Oxford University Press).

Becker, Gary S. and Barry R. Chiswick, 1966, ''Education and the Distribution of Earnings,'' *American Economic Review*, Papers and Proceedings, Vol. 56 (May), pp. 358–69.

Becker, Gary S. and Nigel Tomes, 1986, ''Human Capital and the Rise and Fall of Families,'' *Journal of Labor Economics*, Vol. 4 (July), S1–S39.

Benabou, Roland, 1994, ''Education, Income Distribution, and Growth: The Local Connection,'' NBER Working Paper 4798, National Bureau of Economic Research, Cambridge, Massachusetts.

———, 1993, ''Inequality and Growth,'' *NBER Macroeconomics Annual*, pp. 11–92.

Bulir, Ales, 1998, ''Income Inequality: Does Inflation Matter?'' IMF Working Paper WP/98/7 (Washington: International Monetary Fund).

Castelló-Climent, Amparo and Rafael Doménech, 2014, Working Paper 12/28, BBVA Research.

Chenery, Hollis B. and Moshe Syrquin, 1975, *Patterns of Development*, *1950–1970,* Oxford University Press for the World Bank, London.

Chiswick, Barry R., 1971, ''Earnings Inequality and Economic Development,'' *Quarterly Journal of Economics*, Vol. 85, pp. 21–39.

Corak, Miles, 2013, "Income Inequality, Equality of Opportunity, and Intergenerational Mobility," *Journal of Economic Perspectives*, Vol. 27 (No. 3, Summer), pp. 79–102.

Dabla-Norris, Eras, Kalpana Kochhar, Nujin Suphaphiphat, Frantisek Ricka, and Evridiki Tsounta, 2015, "Causes and Consequences of Income Inequality: A Global Perspective", IMF Staff Discussion Paper SDN/15/13 (Washington: International Monetary Fund).

De Gregorio, Jose, 1995, ''Comment,'' in Rudi Dornbusch and Sebastian Edwards (eds), *Stabilization*, *Economic Reform and Growth* (Chicago: University of Chicago Press).

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.

De Gregorio, Jose and Jong-Wha Lee, 1999, ''Education and Income Distribution: New Evidence from Cross-Country Data,'' Development Discussion Papers No. 714 (Cambridge, Massachusetts: Harvard Institute for International Development), www.hiid.harvard.edu_pub_ddps.html).

Deininger, Klaus W. and Lyn Squire, 1996, ''A New Data Set Measuring Income Inequality,'' *World Bank Economic Review*, Vol. 10 (September), pp. 565–91.

———, 1998, ''New Ways of Looking at Old Issues: Inequality and Growth,'' *Journal of Development Economics*, Vol. 57 (December), pp. 259–87.

Durlauf, Steven N., 1996, ''A Theory of Persistent Income Inequality,'' *Journal of Economic Growth*, Vol. 1 (March), pp. 75–93.

Gallop, J., Steven Radelet, and Andrew Warner, 1998, ''Economic Growth and Income of the Poor,'' mimeo (Cambridge, Massachusetts: Harvard Institute for International Development).

Galor, Oded and Daniel Tsiddon, 1997, ''The Distribution of Human Capital and Economic Growth,'' *Journal of Economic Growth*, Vol. 2 (March), pp. 93–124.

Galor, Oded and Joseph Zeira, 1993, ''Income Distribution and Macroeconomics,'' *Review of Economic Studies*, Vol. 60 (January), pp. 35–53.

Gottschalk, Peter and Timothy M. Smeeding, 1997, ''Cross-National Comparisons of Earnings and Income Inequality,'' Journal of Economic Literature, Vol. 35 (June), pp. 633–87.

Hanushek, E., and L. Wobmann, 2010, "Education and Economic Growth", International Encyclopedia of Education, Vol. 2, pp. 245–52.

Hanushek, E., 2013, "Economic Growth in Developing Countries: The Role of Human Capital", *Economics of Education Review*, 37, 204-212.

Jenkins, Stephen, 2015. "World Income Inequality Databases: An Assessment of WIID and SWIID," *Journal of Economic Inequality*, Vol. 13 (December), pp. 629–71.

Knight, John B. and Richard H. Sabot, 1983, ''Educational Expansion and the Kuznets Effect,'' *American Economic Review*, Vol. 73 (December), pp. 1132–6.

Kuznets, Simon, 1955, ''Economic Growth and Income Inequality,'' *American Economic Review*, Vol. 45 (March), pp. 1–28.

LaGarda, Guillermo and Kevin Gallagher, and Jennifer Linares, 2016, "Capital Openness and Income Inequality: A Cross-Country Study and The Role of CCT in Latin America," *Munich Personal RePEc Archive papers (June)*.

Li, Hongyi, Lyn Squire, and Heng-fu Zou, 1998, ''Explaining International and Intertemporal Variations in Income Inequality,'' *Economic Journal*, Vol. 108 (January), pp. 26–43.

Londoño, Juan Luis, 1996, ''Poverty, Inequality, and Human Capital Development in Latin America, 1950–2025,'' World Bank Latin American and Caribbean Studies 2 (Washington: World Bank).

Marin, A. and George Psacharopoulos, 1976, ''Schooling and Income Distribution,'' *Review of Economics and Statistics*, Vol. 58, pp. 332–8.

Montenegro, Claudio and Harry Patrinos, 2014, "Comparable estimates of returns to schooling around the world", Policy Research Working Paper WPS 7020. Washington, DC: World Bank Group.

Organisation for Economic Co-operation and Development (OECD), 2008, *Growing Unequal? Income Distribution and Poverty in OECD Countries*. Paris: OECD Publishing.

Psacharopoulos, George, 1977, ''Unequal Access to Education and Income Distribution: An International Comparison,'' *De Economist*, Vol. 125, pp. 383–92.

Psacharopoulos, George and Maureen Woodhall, 1985, *Education for Development*: *An Analysis of Investment Choices*, Oxford University Press, New York.

Ram, Rati, 1984, ''Population Increase, Economic Growth, Educational Inequality, and Income Distribution: Some Recent Evidence,'' *Journal of Development Economics*, Vol. 14 (April), pp. 419–28.

———, 1989, ''Can Educational Expansion Reduce Income Inequality in Less-Developed Countries?'' *Economics of Education Review*, Vol. 8(2), pp. 185–9.

Sarel, Michael, 1997, ''How Macroeconomic Factors Affect Income Distribution: The Cross-Country Evidence,'' IMF Working Paper WP_97_152 (Washington: International Monetary Fund).

Winegarden, C. R., 1979, ''Schooling and Income Distribution: Evidence from International Data,'' *Economica*, Vol. 46, pp. 83–7.