Schooling Inequality, Returns to Schooling, and Earnings Inequality: Evidence from Brazil and South Africa

David Lam
University of Michigan
davidl@umich.edu

Murray Leibbrandt
University of Cape Town
Murray.Leibbrandt@uct.ac.za

Arden Finn
University of Cape Town
Aj.Finn@uct.ac.za

Prepared for submission to the 10th IZA/World Bank Conference on Employment and Development

Bonn, Germany

June 5-6, 2015

David Lam is Professor of Economics and Research Professor in the Population Studies Center at the University of Michigan. Murray Leibbrandt is Professor of Economics and Director of the Southern African Labour Development Unit (SALDRU) at the University of Cape Town. Arden Finn is Researcher in SALDRU and a PhD student at the University of Cape Town.

Schooling Inequality, Returns to Schooling, and Earnings Inequality: Evidence from Brazil and South Africa

Abstract

Human capital models imply that both the distribution of education and returns to education affect earnings inequality. Decomposition of these "quantity" and "price" components have been important in understanding changes in earnings inequality in developed and developing countries. This paper provides theoretical and empirical analysis of the interactions of schooling inequality, returns to schooling and earnings inequality. We focus on two main questions. What is the relationship between inequality in schooling and inequality in earnings? How do changes in returns to schooling affect earnings inequality when returns differ by schooling level? We derive new analytical results that are used to guide empirical analysis of changes in earnings inequality in Brazil and South Africa. While both countries have had declines in schooling inequality, only Brazil has translated those into declines in earnings inequality. In South Africa, rising returns to schooling at the top have offset equalizing changes in the schooling distribution.

Introduction

The link between education and the distribution of income has long been fundamental to research on inequality. Theoretical models and a vast body of empirical evidence point to a large explanatory role for education in the distribution of income, especially the distribution of labor earnings. Standard human capital models imply that both the distribution of education and the returns to education will affect earnings inequality. Decomposition of these two components, often referred to as the "quantity" and "price" components, have played an important role in understanding changes in earnings inequality in both high-income countries and developing countries (for example Juhn, Murphy, Pierce 1993, for the United States, and World Bank 2011 for Latin America).

The goal of this paper is to advance our understanding of both the theory and the empirical evidence regarding the interactions of schooling inequality, returns to schooling and earnings inequality. We focus on two main questions. First, what is the relationship between inequality in schooling and inequality in earnings? As shown by Lam and Levison (1992), it is theoretically possible to generate increases in earnings inequality by expansions of schooling that decrease schooling inequality. This phenomenon of declining inequality in schooling associated with rising earnings inequality in earnings seems to have been the case for Brazil and may actually be quite common during the early stages of economic development. Improvements in the schooling distribution appear to eventually become equalizing, however. We elaborate on these issues from a theoretical perspective below, and discuss how they apply empirically to the cases of Brazil and South Africa.

The second issue we consider is how changes in returns to schooling affect earnings inequality when returns differ by the level of schooling. What happens, for example, if the returns to completing grade 8 increase while returns to all other grades remain constant? A common feature of labor markets in developing countries has been for returns to schooling to change at different rates (and even in different directions) at different levels of schooling. Returns to university may have increased, for example, at the same time that returns to secondary schooling declined. In this context it can be misleading to generalize about whether the change in average returns to schooling has been equalizing or disequalizing. As we will show, and as makes sense intuitively, increases in returns to schooling at low grades may actually be inequality reducing, while increases in returns to schooling at high grades are inequality increasing. We develop a general framework for analyzing these issues, and derive some simple analytical results about the impact of returns to schooling at different levels of schooling on earnings inequality.

These results provide a useful framework for empirical analysis. They call attention to an interesting summary statistic that has not previously been studied – the year of schooling which separates equalizing from disequalizing increases in returns to schooling. In the case of the variance of log earnings, for example, we show that this is the level of schooling at which mean log earnings is earned. Our analytical results demonstrate that increases in returns to schooling above this level will be disequalizing, while increases in returns below this level will be equalizing. This level of schooling also provides a benchmark for understanding how changes in the distribution of schooling affect earnings inequality. Changes in the schooling distribution that shift the distribution toward the schooling level of mean log earnings will be equalizing, while shifts away from that schooling level (in either direction) will be disequalizing.

We use this framework to guide empirical analysis of schooling inequality, returns to schooling, and earnings inequality in Brazil and South Africa in recent decades. These two countries competed for many years for the dubious distinction of being the most unequal country in the world. Brazil has experienced declining inequality in recent years, however, while South Africa has experienced persistently high inequality. Both countries have excellent microdata that make it possible to look closely at the distribution of schooling and the distribution of earnings. In Brazil we are able to track both distributions from 1976 to the present using the annual labor market survey. In South Africa we have a consistent labor market series from 1994 to the present.

This paper begins by presenting some of our key theoretical results. We then lead into some examples of empirical analysis that are guided by the theoretical results. Finally, we use counterfactual simulations to analyze how changes in schooling distributions and returns to schooling can explain why Brazil and South Africa took different paths in the evolution of their earnings inequality in the past few decades.

Theoretical Links Between Schooling Inequality and Earnings Inequality

We begin our theoretical analysis with a simplified version of the standard human capital earnings equation. Leaving experience and other determinants of earnings aside for now, the logarithm of the *i*th worker's earnings can be expressed as

$$\log Y_i = \alpha + \beta S_i + \mu_i \tag{1}$$

where Y_i is earnings, S_i is years of schooling, and μ_i is a residual uncorrelated with schooling. Given Equation (1), the variance of log earnings, a standard mean-invariant measure of wage inequality, is

$$V(\log Y) = \beta^2 V(S) + V(u) \tag{2}$$

where *V* denotes variance. This simple result demonstrates an important point about the link between schooling inequality and earnings inequality. If the relationship between schooling and earnings is log-linear as in (1), then earnings inequality (as measured by the log variance) is a linear function of the variance in schooling. This has a number of important implications that are often neglected in discussions of the link between the distribution of schooling and the distribution of earnings. Suppose, for example, that we could double the schooling of every worker, holding returns to schooling constant. This would quadruple the variance in years of schooling and thus quadruple the "explained" component of earnings inequality. If we measure inequality in schooling by some standard mean-invariant measure of inequality, this doubling of schooling would imply no change in schooling inequality. Alternatively, giving each worker one additional year of schooling would unambiguously reduce schooling inequality, but would have no effect on earnings inequality.

Lorenz dominance in schooling distributions and earnings distributions

In order to look at the relationship between schooling inequality and earnings inequality in a fairly general way, consider a linear transformation of the schooling distribution, mapping some initial distribution S' into a new distribution S''

$$S_i'' = \gamma + \delta S_i' \tag{3}$$

Even with simple transformations such as Equation (3), we can generate changes in the schooling distribution that imply unambiguous *reductions* in schooling inequality and unambiguous *increases* in earnings inequality, using the criterion of Lorenz dominance.

> ***Proposition 1:*** Given the earnings generation process in Equation (1) and the linear transformation of the schooling distribution in Equation (3), any transformation in which $\gamma > 0$ and $\delta > 1$ will cause the schooling distribution to become unambiguously more equal and the earning distribution to become unambiguously less equal, in the sense that S'' Lorenz dominates S' and Y' Lorenz dominates Y''.

To prove Proposition 1, it is useful to observe that we will have Lorenz dominance whenever the proportional difference in the schooling (or earnings) of any two randomly drawn individuals in the distribution is smaller in the Lorenz dominating distribution. That is

$$S'' \text{ Lorenz dominates } S' \text{ if } \frac{S_j''}{S_i''} < \frac{S_j'}{S_i'}, \forall(i,j) \text{ s.t. } S_j' > S_i' \tag{4}$$

Given the transformation in Equation (3), the change in the ratios of any two schooling levels is

$$\frac{S_j''}{S_i''} - \frac{S_j'}{S_i'} = \frac{\gamma+\delta S_j'}{\gamma+\delta S_i'} - \frac{S_j'}{S_i'} \tag{5}$$

Inspection of Equation (5) indicates that the difference will be negative for any $\gamma > 0$, with the value of δ affecting the magnitude but not the sign of the difference for any $\delta > 0$. This implies that

$$S'' \text{ Lorenz dominates } S' \text{ for any } \gamma > 0 \text{ and } \delta > 0 \tag{6}$$

Turning to the earnings distribution, it is useful to begin by pointing out the simple special case is of an additive shift in schooling such that $\gamma > 0$ and $\delta = 1$ (for example, giving every person one additional year of schooling). This implies an unambiguous reduction in the inequality of schooling by the criterion of Lorenz dominance. Since this leaves the variance of schooling unchanged, however, it is clear from Equation (2) that the variance in log earnings will be unchanged. This lack of change in earnings inequality is not limited to the log variance measure. Since an additive increase in schooling will cause a multiplicative increase in each person's income, any measure of inequality will be unaffected. Put another way, the additive shift in schooling implies an additive shift in the logarithm of earnings, which is equivalent to simply multiplying the earnings distribution by a constant, a shift that would leave all measures of earnings inequality unchanged. Another simple illustrative case is a multiplicative transformation in schooling, with $\gamma = 0$ and $\delta > 1$ (for example, increasing every individual's schooling by 10 percent). This will have no effect on inequality in schooling, with the Lorenz curves identical for S' and S''. It will increase the variance of schooling by δ^2, however, so the log variance of earnings will increase. Once again, the result is much more general than the log variance. In order to see this, it is useful to move to the general case in which $\gamma \neq 0$ and $\delta \neq 1$, comparing inequality in earnings before and after the change in schooling.

Following the approach above, consider the ratio of earnings for two individuals in each of the two schooling distributions. Consider two individuals *i* and *j* with initial schooling levels $S_j' > S_i'$ and income levels $Y_j' > Y_i'$. If the earnings ratio Y_j/Y_i increases when schooling is changed from S' to S'', for all possible pairings *i* and *j*, then the new earnings distribution will be unambiguously less equal by the criterion of Lorenz dominance. That is,

$$Y' \text{ Lorenz dominates } Y'' \text{ if } \frac{Y_j''}{Y_i''} > \frac{Y_j'}{Y_i'}, \forall (i, j) \text{ s.t. } S_j' > S_i' \tag{7}$$

Since the logarithm is a monotonic transformation, we can also express the Lorenz dominance condition as $log\left[Y_j''/Y_i''\right] > log\left[Y_j'/Y_i'\right]$. If earnings are generated as in Equation (1), and the schooling transformation is given by Equation (3), the difference in log earnings between *i* and *j* after the transformation (assuming that the return to schooling β and the residuals μ_i and μ_j remain constant) is

$$\begin{aligned} logY_j'' - logY_i'' &= \alpha + \beta\left(\gamma + \delta S_j'\right) + \mu_j - [\alpha + \beta(\gamma + \delta S_i') + \mu_i] \\ &= \beta\delta\left(S_j' - S_i'\right) + \mu_j - \mu_i \end{aligned} \tag{8}$$

The difference in log earnings before the transformation will be $\beta\left(S_j' - S_i'\right) + \left(\mu_j - \mu_i\right)$, so the change in the difference in log earnings will be $\beta\left(S_j' - S_i'\right)(\delta - 1)$. Using the condition in Equation (7), this implies that

$$Y' \text{Lorenz dominates } Y'' \text{ if } \delta > 1 \tag{9}$$

This holds for all values of γ. Combining the results in Equation (6) and Equation (9) gives the result in Proposition 1.

Proposition 1 was derived assuming the log-linear relationship between schooling and earnings of Equation (1). While this is a very standard assumption, with strong empirical support, it is important to note that similar results will exist whenever there is a convex relationship between schooling and earnings. It is the convexity in general, not the specific exponential relationship, that leads to the result that an unambiguous reduction in schooling inequality can lead to an unambiguous increase in earnings inequality. The linear transformation in schooling is used simply for analytical simplicity. Obviously if we can generate distributions in schooling that have opposite effects on schooling inequality and earnings inequality using these simple linear transforms, we can do the same with much more general transformations of the schooling distribution.

Opposing trends in schooling inequality and income inequality are far from being just a theoretical possibility. They may be fairly common in the process of economic development. Brazil's experience, for example, is consistent with a pattern in which improvements in schooling inequality coincided with increases in income inequality. As shown by Lam and Levison (1992), the trend across cohorts in Brazil for cohorts born between 1925 and 1950 was for mean schooling to rise at a slightly faster rate than the standard deviation. Schooling inequality was thus declining over this period, as measured by the coefficient of variation and as indicated by constantly improving Lorenz curves for schooling. Since the variance of schooling was rising, however, these improvements in schooling inequality did not translate into improvements in earnings inequality. The "explained variance" in the log variance of earnings, $\beta^2 V(S)$, rose steadily across cohorts, helping contribute to continued high inequality in Brazil. As shown below, the variance of schooling has peaked among more recent cohorts in Brazil, suggesting that this component could contribute to reductions in earnings inequality in the future.[1]

While there is intuitive appeal to the notion that a more equal distribution of schooling should lead to a more equal distribution of earnings, there is clearly no theoretical reason to expect such a relationship to hold. What might be considered unambiguous improvements in the distribution of schooling (as indicated, for example, by stochastic dominance), could plausibly

[1] Ram (1990) shows with cross-national data that the standard deviation of schooling tends to follow an inverted-U pattern in relation to mean schooling, with a peak when the mean is around seven years.

lead to increased inequality in earnings. The fundamental reason for this is that earnings are very likely to be a convex function of schooling, the simple log-linear wage equation being just one simple example of such convexity. Any convex relationship between schooling and earnings will tend to produce the result that proportional increases in schooling will increase earnings inequality. This point will be important in analyzing the link between schooling inequality and earnings inequality in Brazil and South Africa.

Generalizing the relationship between schooling and earnings

The results above assume that there is a single rate of return to schooling that applies to all levels of schooling. One of the important recent patterns in returns to schooling in developing countries, however, is the emergence of convex returns to schooling, with returns increasing at higher levels of schooling (especially post-secondary) at the same time that they have fallen at intermediate levels of schooling.

This pattern complicates what we mean when we consider the relationship between returns to schooling and earnings inequality. What happens to inequality if, for example, we increase the return to grade 8, holding returns at other grades constant. What if we increase returns to grade 4 or grade 12? This section provides an analytical way to answer to these questions.

Consider a very general model of the relationship between schooling and earnings

$$y_i \equiv logY_i = \alpha + \sum_j \beta_j S_{ji} + \mu_i \tag{10}$$

where Y_i is earnings, y_i is the log of earnings S_{ji} is a 0,1 indicator for whether person *i* in the *j*th schooling category (which could be single years of schooling in the most general case, but could also be larger categories), and μ_i is a residual uncorrelated with schooling.[2] Denote mean log earnings as $\bar{y}$ and mean log earnings for schooling level *j* as $\bar{y}_j$. Note that β_j is the multiplicative shift in earnings for group *j* relative to the omitted category. We will call an increase in β_j an increase in returns to schooling for schooling group j. For example, an increase in β_9 by .01 will increase the earnings of individuals with 9 years of schooling by 1 percent relative to all other groups, holding earnings of all other schooling groups constant. Note that we may want to think of an increase in returns to schooling at grade 9 as increasing the earnings of everyone with schooling greater than or equal to 9 years. For simplicity (and in order to make the results

[2] Note that nothing about this analysis requires that these be schooling categories. The same results would apply to any other categories, such as age, region, or gender.

relevant to any kind of groupings), we begin with the specification in (10) in which changes in β_j only affect earnings for group j.

The following proposition describes the relationship between returns to schooling and earnings inequality.

Proposition 2: If $\hat{s}$ is a level of schooling for which $\bar{y}_j > \bar{y}, \forall_j > \hat{s}$ and $\bar{y}_j < \bar{y}, \forall_j < \hat{s}$, then increases in β_j for $j > \hat{s}$ will increase the variance of log earnings, and increases in in β_j for $j < \hat{s}$will increase the variance of log earnings.

To prove Proposition 2, note that the variance of log earnings for this more general model can be written as:

$$V(\log Y) = \sum_j \beta_j^2 V(S_j) - 2\sum_j \sum_{k \neq j} \beta_j \beta_k p_j p_k + V(\mu), \tag{11}$$

where p_j is the proportion in schooling category j. Since the S_j. terms only take on values of 0 or 1, $V(S_j) = p_j(1 - p_j)$. What happens to inequality if we increase one of the β terms? This is still an increase in returns to schooling, but is only an increase for one category of schooling (relative to some arbitrary omitted category) and no longer translates necessarily into an increase in inequality. We take the derivative of Equation (11) with respect to β_1, which could arbitrarily be assigned to any schooling category and thus is completely general:

$$\frac{\partial V(\log Y)}{\partial \beta_1} = 2\beta_1 p_1 - 2\beta_1 p_1^2 - 2p_1 \sum_{j \neq 1} \beta_j p_j = 2p_1 \left[\beta_1 - \beta_1 p_1 - \sum_{j \neq 1} \beta_j p_j\right] \tag{12}$$

Note that:

$$\beta_1 p_1 + \sum_{j \neq 1} \beta_j p_j = \bar{y} - \alpha, \text{where } \bar{y} = E(\log y), \text{and } \alpha + \beta_1 = \bar{y}_1, \text{where } \bar{y}_1 = E(\log y \,|S_1 = 1).$$

Substituting into (12), the result simplifies to:

$$\frac{\partial V(\log Y)}{\partial \beta_1} = 2p_1[\bar{y}_1 - \bar{y}], \tag{13}$$

or

$$dV(\log Y) = d\beta_1 * 2p_1[\bar{y}_1 - \bar{y}]. \tag{14}$$

The result is very intuitive. Increasing β_1 will increase the earnings of the first schooling category (arbitrarily defined) relative to the omitted category, and therefore relative to every other category as well. This will be equalizing if the first category had a mean below the overall mean and will be disequalizing if its mean was above the overall mean. The magnitude of the change will depend on how far the group's mean is above or below the overall mean, and on the relative size of the group. For example, if the group's mean of log earnings was 0.1 below the overall log mean (in other words, a difference of approximately 10%), and if the group was 10% of the income earning population, an increase in β_1 of 0.01 would reduce the variance of log earnings by 2*0.1*0.1*0.01=0.0002.

If we calculate this derivative for every single year of schooling, Equation (14) calls attention to a statistic that we do not ordinarily calculate – the year of schooling for which mean log earnings is equal (or closest to equal) to overall mean log earnings. Suppose there is a level of schooling $\hat{s}$ such that $\bar{y}_i > \bar{y}, \forall i > \hat{s}$ and $\bar{y}_i < \bar{y}, \forall i < \hat{s}$. Then increasing returns to schooling for all years below $\hat{s}$ is equalizing, and increasing returns for years above $\hat{s}$ is disequalizing. This is the result in Proposition 2.

It is also interesting to consider whether the year of schooling at which mean log earnings is reached is less than or greater than mean years of schooling. That is, is $\hat{s} - \bar{s}$ positive, negative, or zero? It is easy to see that in the simple linear Mincer earnings equation, $\hat{s} - \bar{s}$, since mean log earnings will be earned by someone with mean schooling, abstracting from other variables such as age and experience. More generally, however, $\hat{s}$ could be greater or less than $\bar{s}$, depending on whether returns to schooling are concave or convex in schooling. If returns are convex, as they have been in South Africa and in many other developing countries in recent years, then $\hat{s} > \bar{s}$ - the year of schooling associated with mean log earnings is above mean schooling. This means that an increase in returns to schooling will be equalizing even for some years above mean schooling. We look at this empirically below for Brazil and South Africa.

Another interesting question is what happens when we change the distribution of schooling. One simple way to model this is to imagine shifting people from some arbitrary group 2 to some arbitrary group 1, so that $dp_2 = -dp_1$, or, equivalently, $\partial p_2 / \partial p_1 = -1$.

$$\frac{\partial V(\log Y)}{\partial p_1} = \beta_1^2 - 2\beta_1^2 p_1 + (\beta_2^2 - 2\beta_2^2 p_2)\frac{\partial p_2}{\partial p_1} - 2\beta_1 \sum_{j \neq 1} \beta_j p_j - 2\beta_2 \sum_{j \neq 2} \beta_j p_j \frac{\partial p_2}{\partial p_1}$$

$$= (\beta_1 + \alpha)^2 - 2\beta_1 \bar{y} - (\beta_2 + \alpha)^2 + 2\beta_2 \bar{y}$$

$$= \bar{y}_1^2 - 2\beta_1\bar{y} - \bar{y}_2^2 - 2\beta_2\bar{y}$$
$$= (\bar{y}_1 - \bar{y})^2 - (\bar{y}_2 - \bar{y})^2 \quad (15)$$

The result is once again very intuitive. Shifting the population from one group to another will be disequalizing if the second group has mean log earnings that are further from the mean (in absolute value) than the first group. For example, if mean log earnings for group 2 is 0.2 above overall mean log earnings, while mean log earnings for group 1 is 0.1 below the overall mean, then shifting 10% of the population from group 2 to group 1 will change the variance of log earnings by $(0.1^2 - 0.2^2)$*0.1=(0.01-0.04)*0.1= -0.003. As above, an interesting point of reference is the level of schooling corresponding to mean log earnings. The generalization of Equation (15) is that changes in the schooling distribution that push the distribution toward the level of schooling with mean log earnings will tend to be equalizing, while changes in the distribution that push the distribution away from the level of schooling with mean log earnings will tend to be disequalizing. Note that if returns to schooling are convex then the critical level of schooling will be higher than mean schooling.

Note that the result in (15) can be applied to any variance. We have applied it to the variance of log earnings, which is a mean-adjusted measure of inequality. We could use it to talk about inequality in schooling by noting that we will reduce inequality if we reduce the variance while raising the mean. Using the result in (15), we will do this for the distribution of schooling if we shift people upward in the distribution so that we raise the mean, while on average moving people closer to the mean.

Other measures of inequality

The results above apply to the variance of log earnings, one measure of inequality. We can also consider what happens to other measures of inequality when the returns to schooling change, continuing to assume that the fundamental relationship between schooling and earnings is given by the flexible log earnings function in Equation (10). One measure that has a simple analytical result is the Generalized Entropy Measure GE(0), which can be written as

$$GE(0) = \frac{1}{n}\sum_{i=1}^{n} log\left(\frac{\bar{Y}}{Y_i}\right) \quad (16)$$

where $\bar{Y}$ is the mean of earnings (not the mean of log earnings). Taking the derivative with respect to some β_1:

$$\frac{\partial GE(0)}{\partial \beta_1} = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial log\bar{Y}}{\partial \beta_1} - \frac{1}{n}\frac{\partial logY_i}{\partial \beta_1}$$

$$= p_1 \frac{\bar{Y}_1}{\bar{Y}} - p_1$$

$$= p_1 \left(\frac{\bar{Y}_1}{\bar{Y}} - 1 \right) \quad (17)$$

where p_1 is, as above, the proportion of the population in schooling group 1. Note that (17) will be positive when $\bar{Y}_1 > \bar{Y}$ and will be negative when $\bar{Y}_1 < \bar{Y}$. That is, if group 1 has mean earnings above (below) the overall mean, then an increase in β_1 will increase (decrease) inequality as measured by GE(0). The interesting difference from the result for the log variance is that the sign now depends on the difference between the group's mean earnings relative to overall mean earnings, whereas the result for the log variance depends on the difference between the group's mean log earnings and the overall mean log earnings. This means that we will also be interested in the level of schooling at which mean earnings is reached. As we will see below, the level of schooling corresponding to mean earnings will typically be higher than the level of schooling corresponding to mean log earnings, given the convex relationship between schooling and earnings.

Empirical Evidence from Brazil and South Africa

In this section of the paper we take the theory to the data and compare the experiences faced by Brazil and South Africa over the last few decades. We use earnings and education data from household and labor market surveys in both countries – Brazil's PNAD (1976-2012) and South Africa's Post Apartheid Labour Market Series (PALMS – 1994-2011). All earnings data are reported in real terms, and all data are weighted so as to be nationally representative. We begin by discussing how the distribution of schooling evolved in both countries, before plotting the path of earnings inequality. We then substantiate some of the propositions developed in the theory section of the paper by plotting how the schooling level associated with mean log earnings changed over time, and how this links to changing returns to schooling in both countries. We conclude by simulating the effect on aggregate earnings inequality of a 0.01 increase in the return to schooling at each year of schooling, while keeping the distribution of schooling constant.

Figure 1 presents the cumulative distributions of schooling attainment for the labor force aged 25 to 60 in Brazil and South Africa over the 1995 to 2011 period. Vertical lines have been superimposed to represent completed secondary education – 11 years in Brazil and 12 years in South Africa. It is striking to see how quickly average educational attainment increased in both countries. For Brazil, approximately 50% of the labor force had up to 4 years of education in

1995. By 2011 this had reduced to about 28%. Larger jumps are evident between 3 and 4, 7 and 8, and 10 to 11 years of schooling. Improvements in the average level of schooling are in evidence for South Africa too, though the country started off with a higher mean average education than Brazil in 1995.

Figure 2 allows us to unpack these CDFs a little more, as we can see exactly where the gains and losses took place over the period. In Brazil, the proportion of the labor force with education up to completed primary schooling decreased rapidly, while the share with completed secondary education expanded, particularly from 2000 onwards. The trend for South Africa shows a slow increase in the matric and tertiary share of the labor force, matched by a decline in the share of those with no education, or primary education.

Figure 1 Distributions of schooling over time, Brazil and South Africa

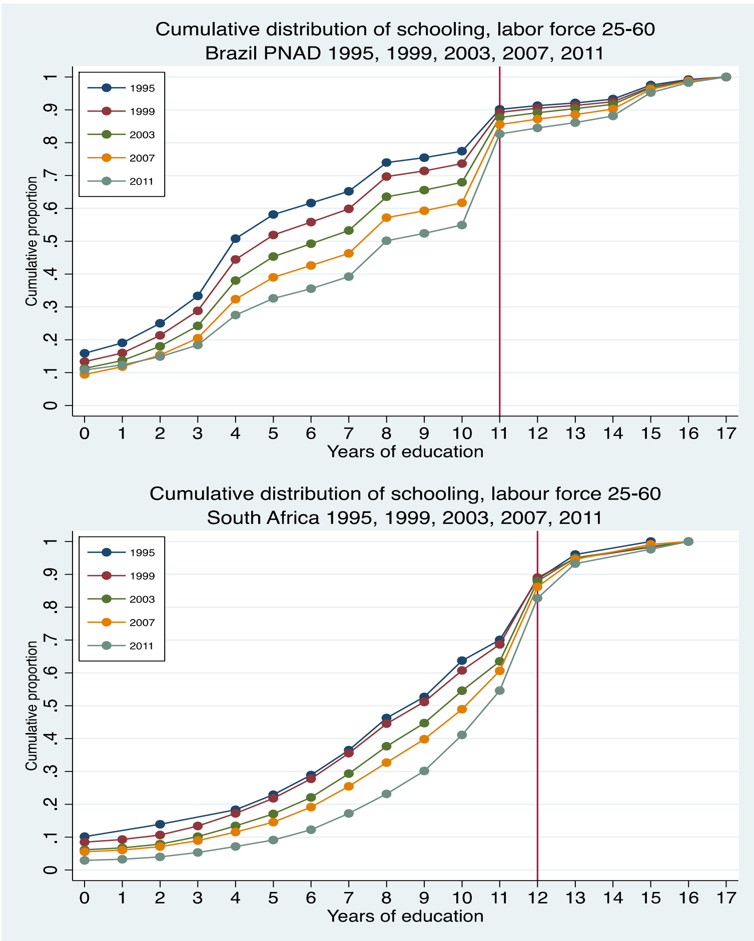


Figure 2 Educational composition of labor force over time, Brazil and South Africa

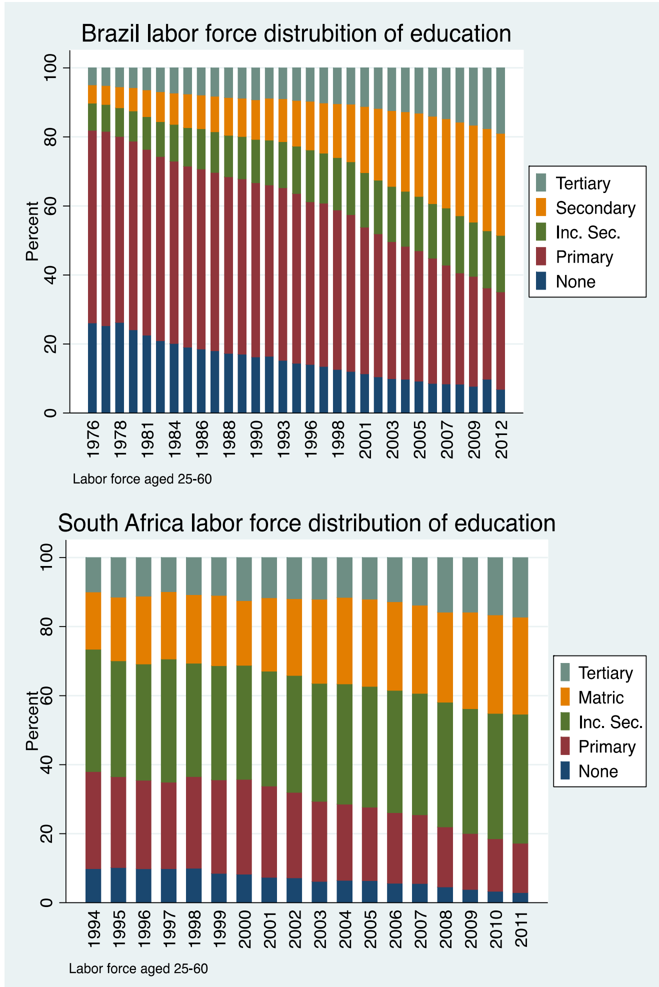


Figure 3 shows some key measures of the distribution of education for the working-age population in the two countries. Both countries have had rapid increases in mean education, but South Africa's mean in 1994 was already higher than Brazil had reached by 2008. The coefficient of variation, a simple mean-adjusted measure of education inequality, is shown on the same scale for both countries, revealing that Brazil had much higher level of education inequality than South Africa in the 1990s. The standard deviation, a key determinant of earnings inequality in the standard human capital earnings equation, has been fairly similar and relatively constant in the two countries, although it has declined more in South Africa than in Brazil.

Figure 3 Mean, standard deviation and coefficient of variation of years of education, Brazil and South Africa

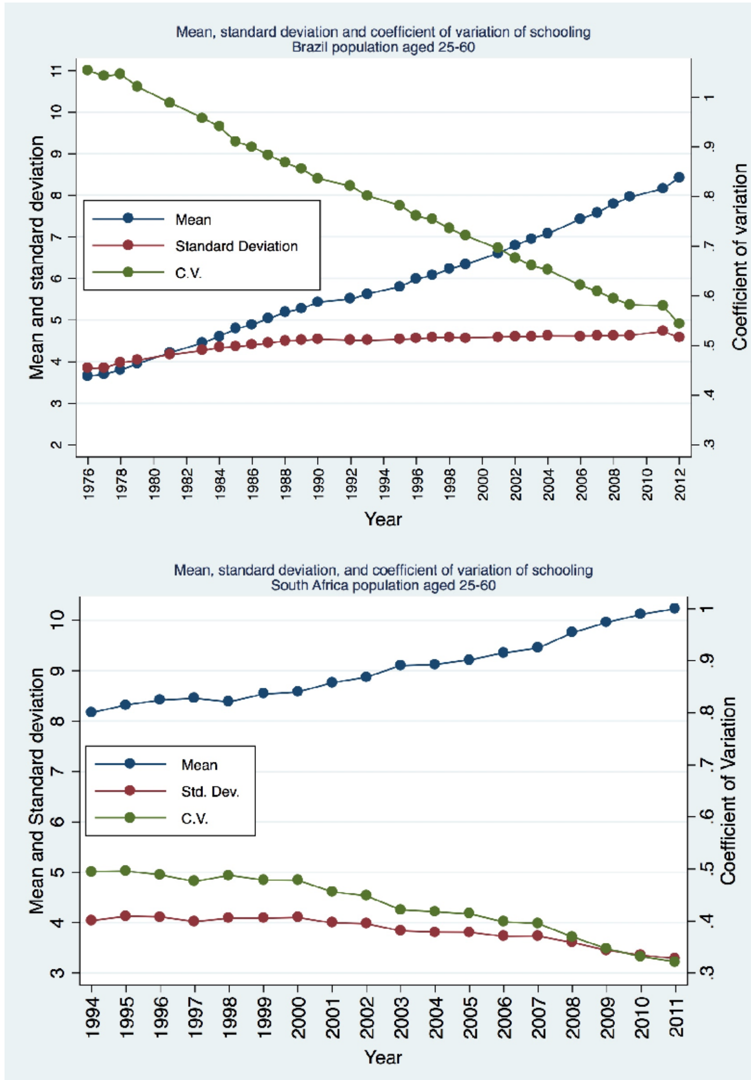
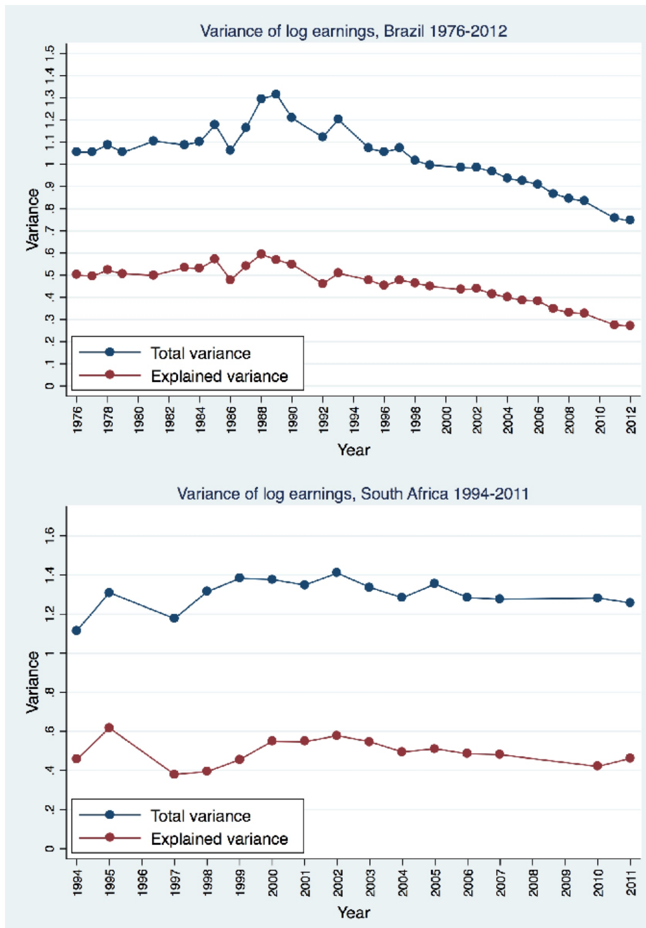


Figure 4 shows the variance of log earnings for the sample of all men and women with positive earnings in Brazil and South Africa for the total period of our samples. Note that the overall level of earnings inequality in South Africa is similar to the level in Brazil before it began to experience a decline in inequality beginning around 1990. The Figure also shows explained variance based on a regression that includes schooling dummies for each year of schooling along with age and age squared. Once again we see that the level of explained variance is fairly similar in the two countries. The R^2 in these earnings regressions is over 0.4, much higher than is typically found in similar earnings regressions in the U.S. (Lam and Levison 1992). The fact that the explained variance closely tracks the decline in earnings inequality in Brazil is important. It means that some combination of the change in the distribution of schooling and the change in returns to schooling account for most of the decline in earnings inequality in Brazil. As we will see, one of the puzzles in South Africa is why declines in inequality in schooling have not translated into similar declines in earnings inequality, and this is something that the earlier theory section seeks to address.

Figure 4 Total and explained variance of log earnings, Brazil and South Africa



In Figure 5 we see the evolution of two different earnings inequality measures – the generalized entropy (0) and Gini coefficient – over time for both countries. As with the total variance in the previous figure, Brazil experienced a sharp increase in inequality in the late 1980s and early 1990s, with the mid 1990s bring a sustained and significant decline. The pattern for post-apartheid South Africa is very different, with earnings inequality slightly higher at the end of the period than it was at the start – the Gini coefficient on South African earnings in 2011 is approximately the same as Brazil's in 1976. The fact that earnings inequality increased in South Africa despite a significant decrease in the inequality of schooling attainment means that the structure of the returns to education must have changed, and it is to this that we now turn.

Figure 6 shows the statistic that we argue is key to understanding the relationship between returns to schooling and earnings inequality – the level of schooling at which mean log earnings is reached. The figure also shows the mean level of schooling for those with positive earnings. Mean schooling over the period rose from 4.5 to 9 years, while the education level associated with mean log earnings rose from 3.5 years to 10.8 years. Note that in Brazil the level of schooling corresponding to mean log earnings was below the mean level of schooling until around 1985, then increases well above mean schooling in later years. This crossover indicates that returns to schooling went from being concave to convex in schooling. Given our analytical results above, we can see that increases in returns to schooling in the intermediate levels, say around seven years of schooling, would have been disequalizing until the late 1990s, after which they would have been equalizing. The South African data do not go as far back as those of Brazil, and differ in that the education level of mean log earnings and the mean years of education do not cross at any point. In fact, the gap between the two widened between 1994 and 2011. There was a two year increase in the mean level of education of positive earners over the period (from just over 8 years to 10.5 years), while the education level of mean log earnings jumped from 9 to 11.7. An increase in the return to schooling in the 9 to 11 range would have been disequalizing in the 1990s, but would have been equalizing from 2007 onwards.This pattern for South Africa suggests that it must have been the case that returns to education increased more in higher schooling categories (those above the mean) relative to lower categories, and this is confirmed in the next figure.

Figure 5 Mean years of schooling and schooling of mean log earnings, Brazil and South Africa

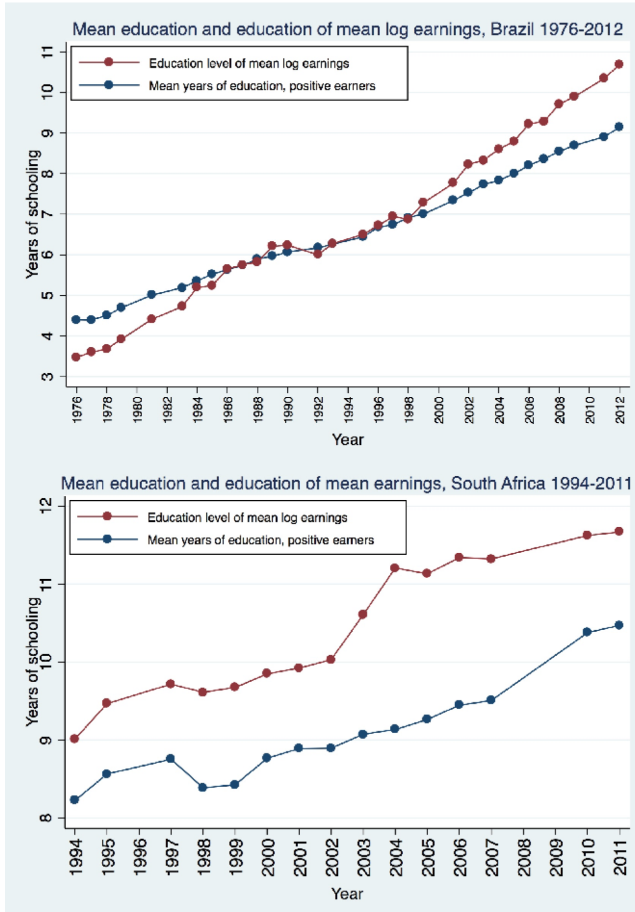
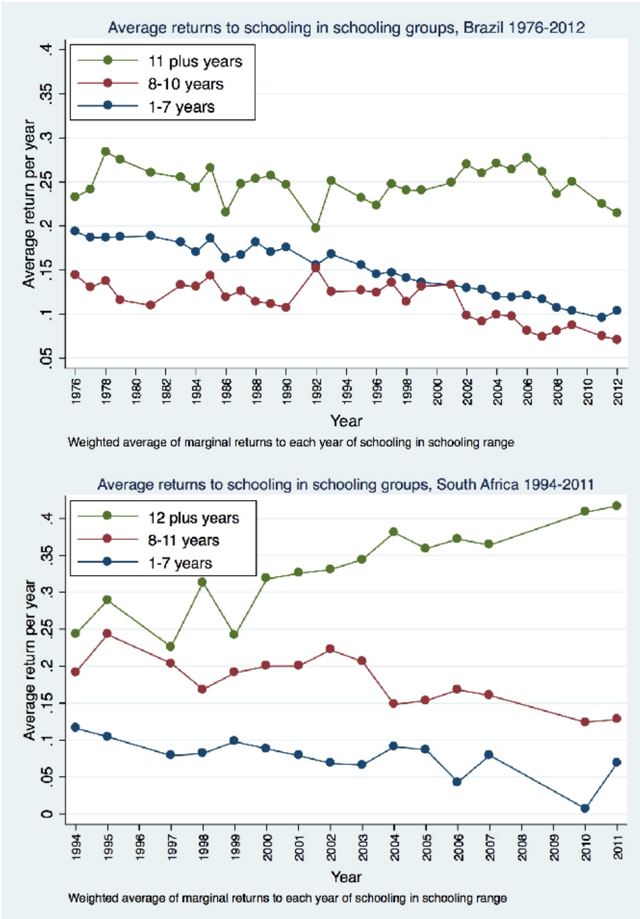


Figure 7 shows what has happened to returns to schooling in the top, middle, and bottom of the schooling distribution, using cutoffs for each country that reflect key schooling breaks (Grade 11 is the end of secondary in Brazil, while Grade 12 is the end of secondary in South Africa). We see several key differences in the patterns for the two countries. South Africa has seen a dramatic increase in returns to grade 12 and above since 1994. Our simulations indicate that this is the main factor explaining why improvements in schooling inequality have not led to decreases in earnings inequality. At the same time, the declines in returns to grade 9-11 have had a mixed impact. Based on our analytical results and what was presented in Figure 6, we see that declines in returns to grades 9-11 would have been equalizing in the 1990s, but became disequalizing by the mid 2000s.

In contrast, Brazil has had relatively constant returns to the highest levels of education. This has meant that improvements in the distribution of education have been translated into declines in inequality. The declining returns to intermediate levels of schooling are more complicated, as in South Africa. They were equalizing for much of the period, but may have been somewhat disequalizing in more recent years. The increase in returns at the top end of the schooling distribution in Brazil was not as marked as in South Africa, and the country witnessed a slight decrease in the returns to this category in recent years. Our theoretical findings suggest that the declines in returns to the 1 to 7 category should be disequalizing. This is indeed the case, but the effect is tempered by the fact that the relative weighting of this category is small and decreasing over time as the average education level of the population rises.

Figure 6 Average returns to schooling in schooling groups, Brazil and South Africa



Looking at the returns to education for each year of schooling provides a useful but somewhat limited view of the total interaction between the distribution of schooling and the distribution of earnings. What we are really interested in – as motivated in the theory section – is how total earnings inequality would change if we increased the returns to a particular year of schooling while holding the return at all other levels and the distribution of schooling itself constant. Recall from Equation (14) that, for log variance, this is a combination of a) how far that year of schooling's mean is from the overall mean, and b) the relative size of the group. We can already use Figure 6 to tell us where overall earnings inequality would go up and down if we increased a particular year's returns to education, the question now is about the magnitude of that change.

The following three figures plot what happens to overall inequality for a 0.01 increase in the returns to a given level of schooling, *ceteris paribus*. We do this for 1995, 2002 and 2011 for Brazil and South Africa. The horizontal line at 0 on the y-axis is the crossover point at which increasing returns begin to increase, rather than decrease, earnings inequality. For log variance this crossing point always corresponds with the education level at which mean log earnings is realized. Three other inequality measures are also presented in the Figures. These are the Gini coefficient, and the two generalized entropy measures of inequality. Recall that for the generalized entropy (0) measure, the crossing point corresponds to the education level at which the mean level of earnings is reached, rather than the mean level of log earnings as we saw with the log variance measure. The convex relationship between education and earnings means that the crossover point for log variance must be before that of the GE(0) measure, and this is confirmed in all of the figures below.

The mean education levels of those in the labor force in 1995 for Brazil and South Africa were 6 years and 8.3 years, respectively. For both countries, the largest impact on decreasing inequality is to raise the returns to zero years of education, which made up 14.4% and 10.1% in Brazil and South Africa's labor forces in 1995, respectively.

In the 1995 to 2002 to 2011 period the effect of increasing the returns to completed secondary education in Brazil flipped from having an inequality-increasing to an inequality-decreasing effect. Increasing the returns to any year of primary schooling lowered inequality in all years, and this is to be expected, given how the mean years of educational attainment increased over the period. By 2011 we see that raising the returns to completed secondary education leads to lower earnings inequality for all measures but log variance. This can go some way to explaining why overall inequality fell in Brazil over the period – the proportion of the labor force that had completed secondary education more than doubled between 1995 and 2011 (from 23% to 47%), and the education level of mean log earnings rose to about 11 years.

The most notable aspect of the South African profile is the fact that the inequality-increasing factor for the log variance measure almost doubled for tertiary education over the 1995 to 2011 period. This reflects the increasing returns to tertiary education that were highlighted in Figure 7, as well as a fact that this category made up an increasing proportion of the labor force.

In general, the profile from 0 to six years of education became flatter over the period, as a smaller and smaller percentage of the labor force found itself in the no schooling or primary schooling categories. Interestingly, increasing the returns to completed secondary education would have increased aggregate earnings inequality in 1995 and 2002, but reduced it in 2001 for all measures except log variance. The inequality impact of increasing returns to grade 11 showed a similar pattern by increasing inequality in 1995 but decreasing inequality by 2011 for the log variance measure.

The major shift in the educational composition of the South African labor force over the period was a relative shift out of primary and into completed secondary and tertiary education, the latter of which was far more disequalizing in 2001 than it was in 1995. The proportion in the incomplete secondary education category remained relatively constant. This stands in contrast to the Brazilian experience, which witnessed a large drop in the proportion of the labor force in the no schooling and primary schooling categories (rather than just primary) and an increase in the secondary and tertiary categories. This serves to highlight the fact that the starting positions of both countries are important in understanding why the inequality dynamics differed from the mid 1990s to 2011. In 1994 two thirds of the Brazilian labor force had no schooling or primary schooling compared to under 40% for South Africa. By 2011 the proportion in these two categories had dropped 20 37% and 18% in Brazil and South Africa, respectively. Although average educational attainment in the labor forced increased rapidly for both countries, Brazil started at a much lower base.

Conclusion

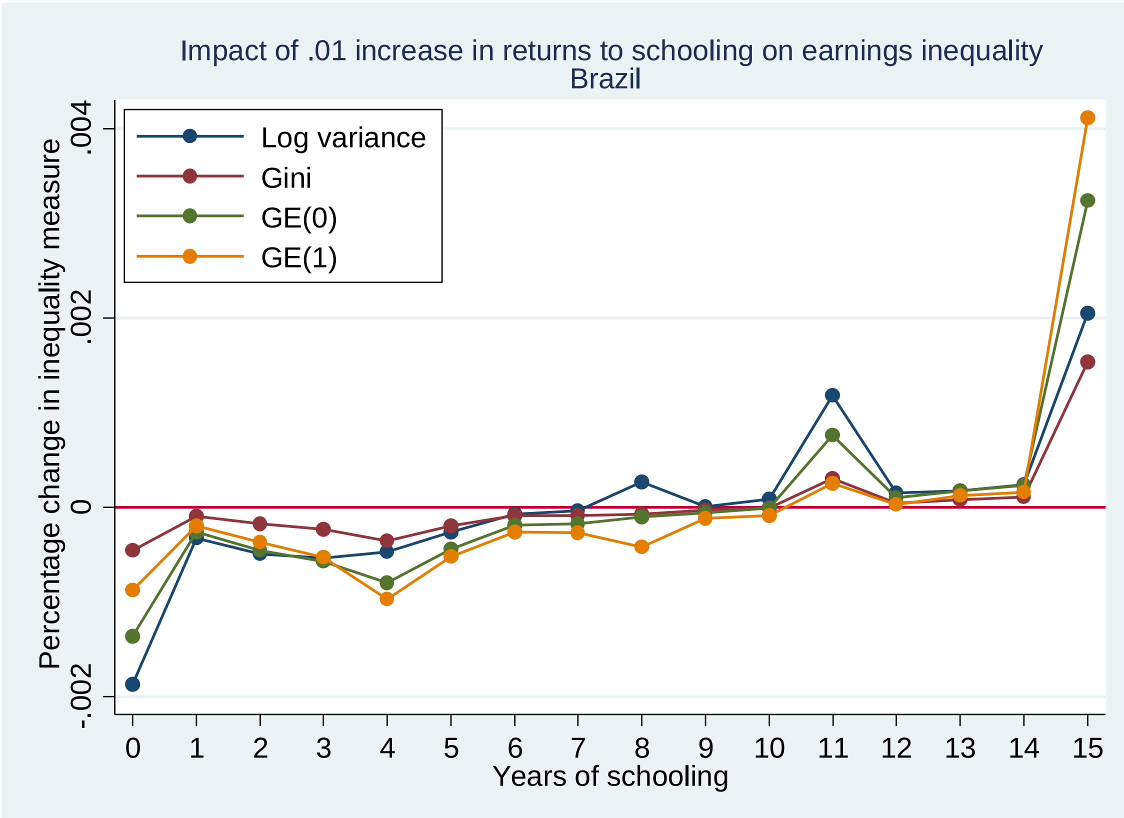
In this paper we sought to shed new light on the theoretical and empirical relationship between schooling inequality, returns to schooling and earnings inequality. We presented new theoretical results that call into interest a previously under-emphasized measure – the level of education associated with mean log earnings. Our empirical section applied the theory to two countries that experienced different earnings inequality dynamics over the last few decades. While Brazil's aggregate earnings inequality fell in recent years, South Africa's increased. In order to better understand why the evolution of inequality differed in these countries we focused our attention on the different educational composition of labor markets and the different trajectories of the returns to education across the schooling distribution over time.

Schooling inequality declined substantially in both countries over time. However, this did not lead to a decline in earnings inequality in South Africa, though it did eventually translate into lower earnings inequality in Brazil. That said, the educational composition of both labor forces were very different in the mid 1990s, with Brazil’s average level of education (and education level associated with mean log earnings) far lower than South Africa’s.

Returns to schooling changed across the distribution of education for both countries. In South Africa, returns increased at the top, and declined in the middle and lower parts of the distribution. In Brazil, the premium to the top of the educational distribution in terms of increased returns was comparatively smaller, and, in fact, began to decline in the last decade. The impact of changes in the returns to schooling depends on the level of schooling associated with mean log earnings. Increasing returns in the middle of the distribution (or for the median worker) would have been disequalizing in past years, but are now equalizing. Decreasing returns in the middle of the schooling distribution have contributed to rising inequality in South Africa, thereby compounding the impact of rising returns at higher levels of education.

The composition of the Brazilian labor force changed quite extensively over the years, and the bottom part of the distribution in 2012 looked similar to the South African distribution in 1994. It will be interesting to track the trajectory of earnings inequality as the level of mean education attainment converges between the two countries.

Figure 7 Impact of a 0.01 increase in returns to schooling on earnings inequality, Brazil and South Africa 1995



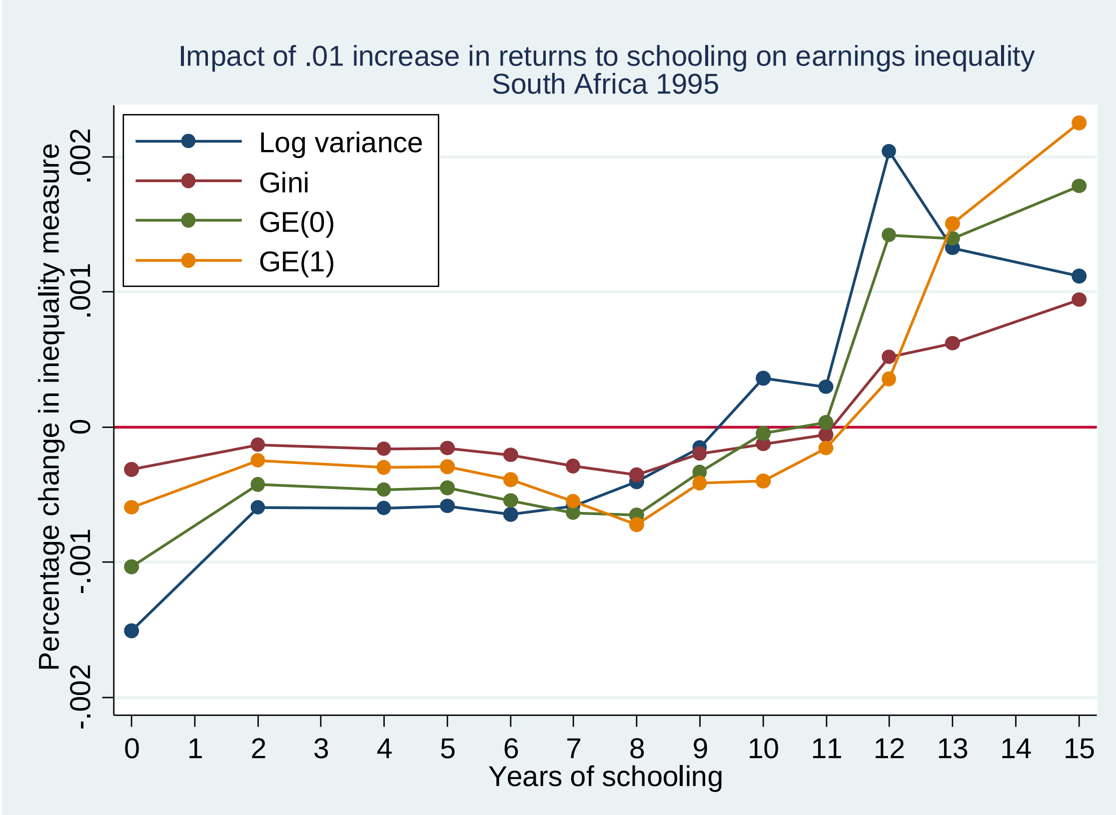


Figure 8 Impact of a 0.01 increase in returns to schooling on earnings inequality, Brazil and South Africa 2002

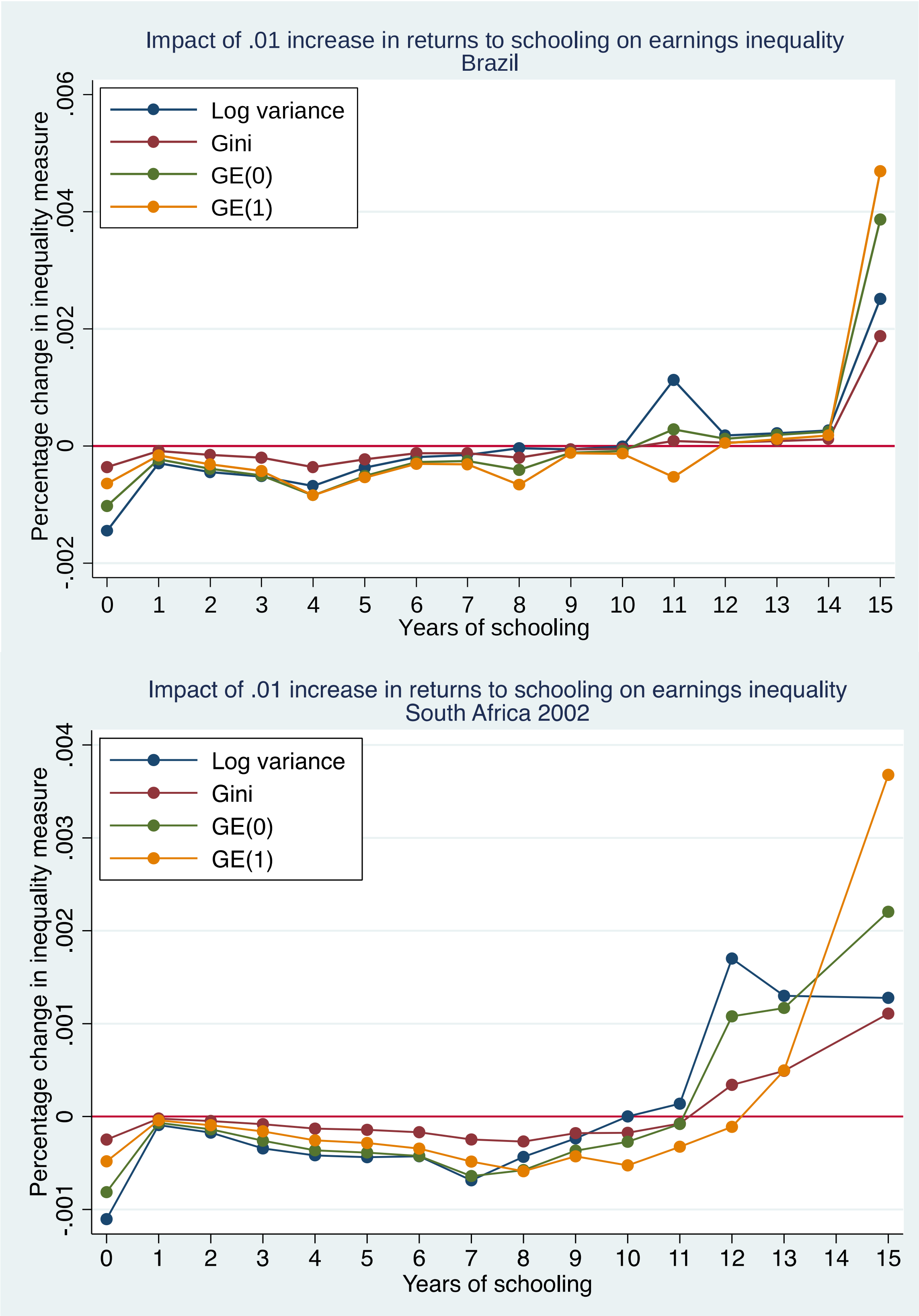


Figure 9 Impact of a 0.01 increase in returns to schooling on earnings inequality, Brazil and South Africa 2011

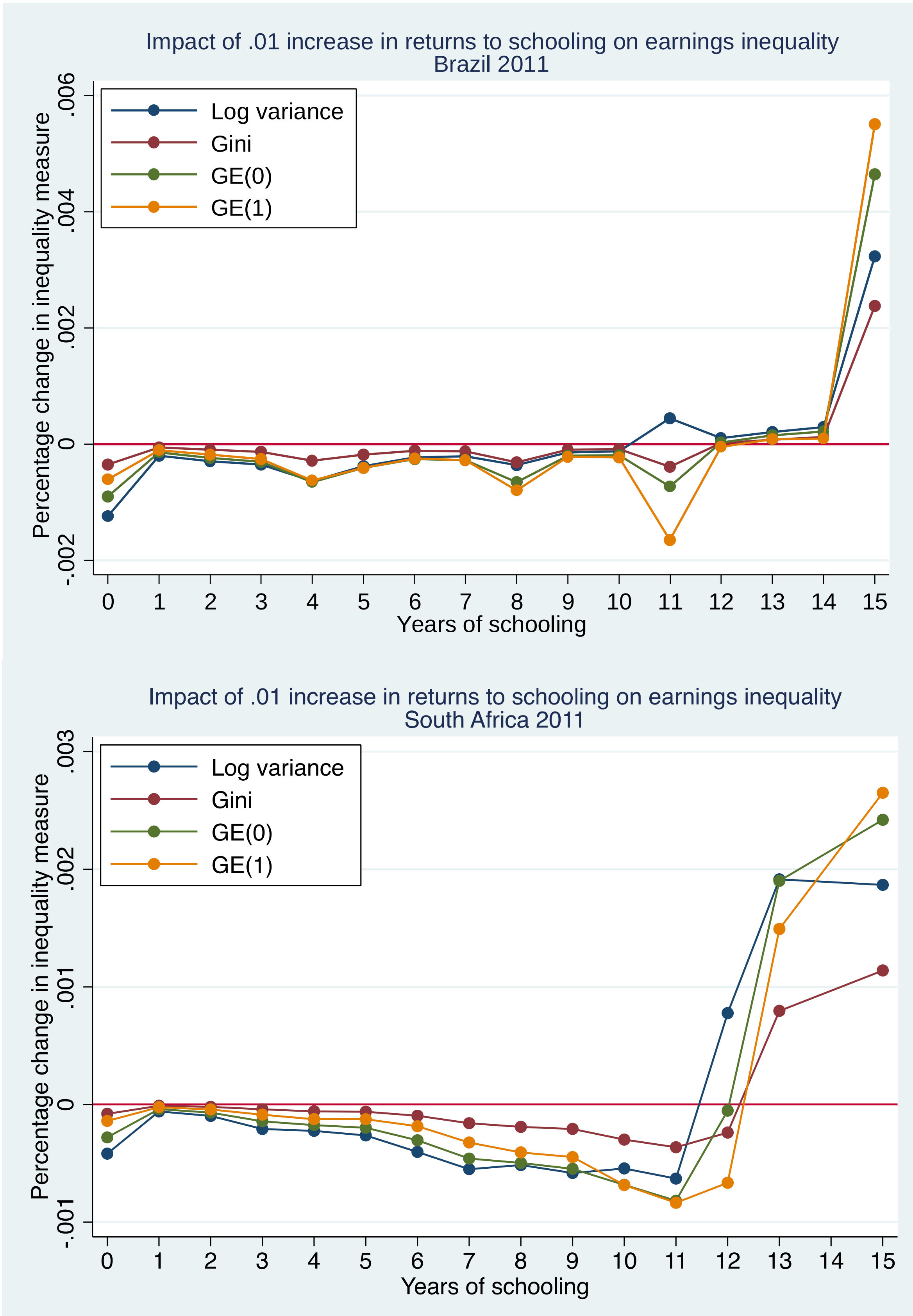
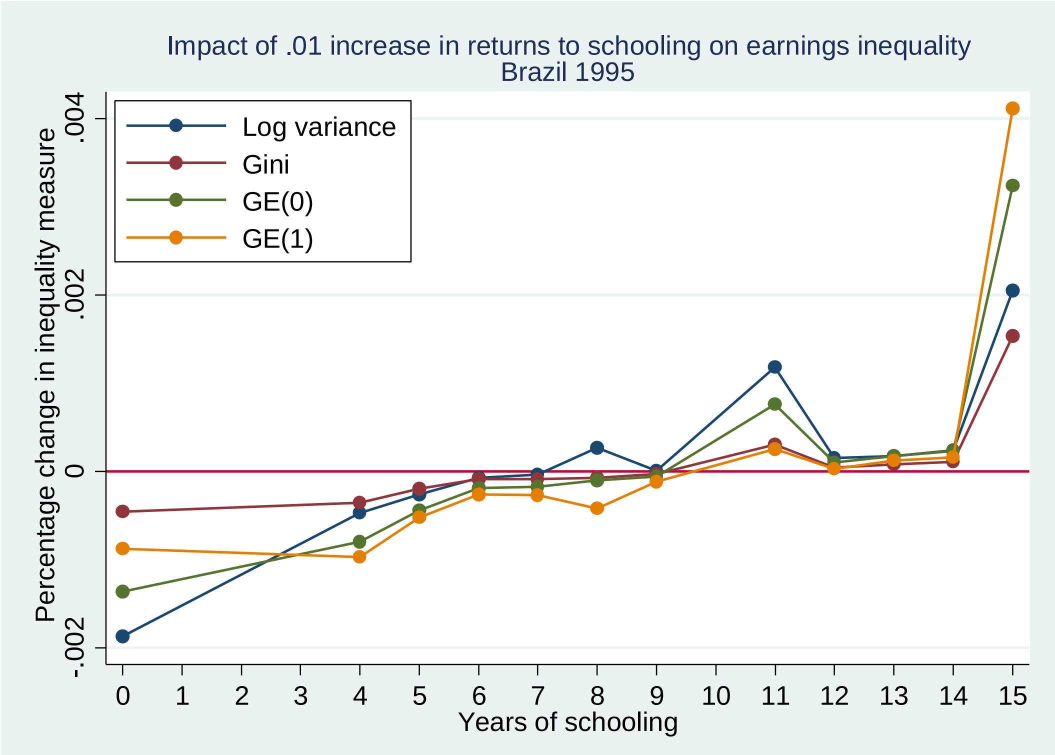


Figure 10 Alternative presentation, 1994/1995



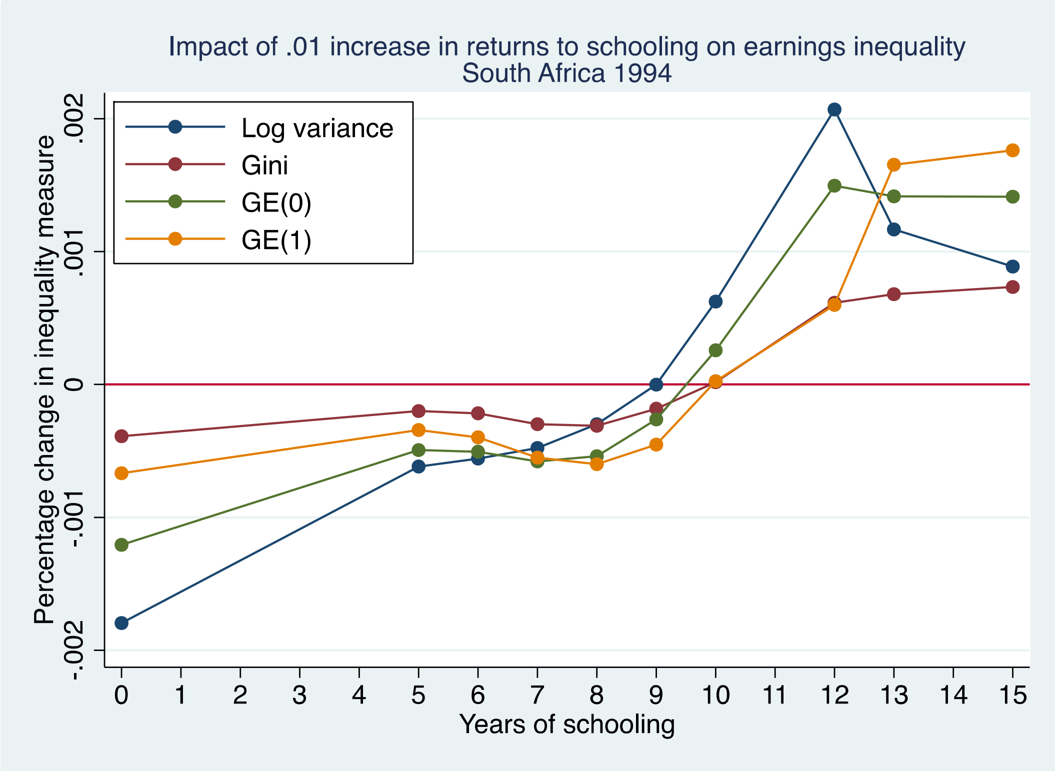
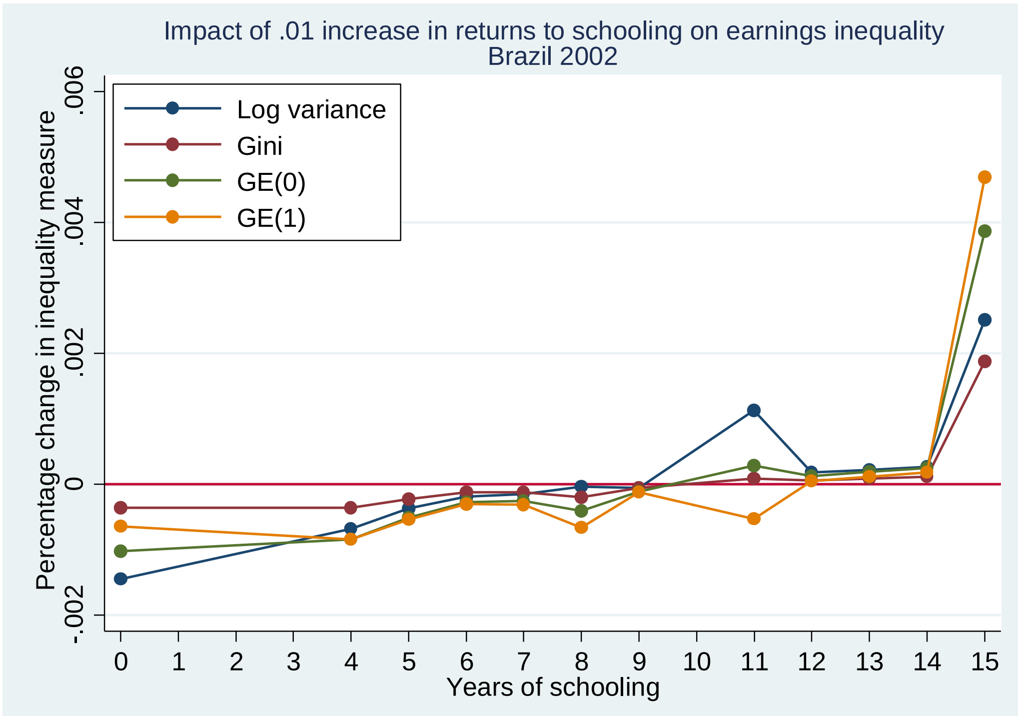


Figure 11 Alternative presentation, 2002



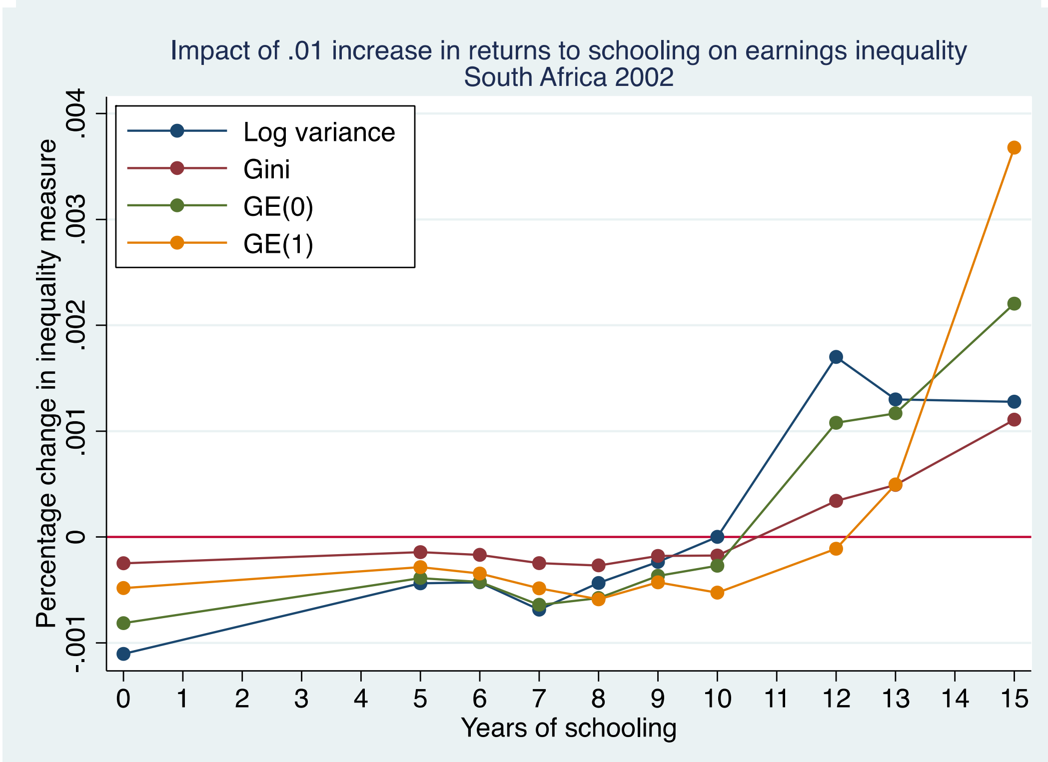
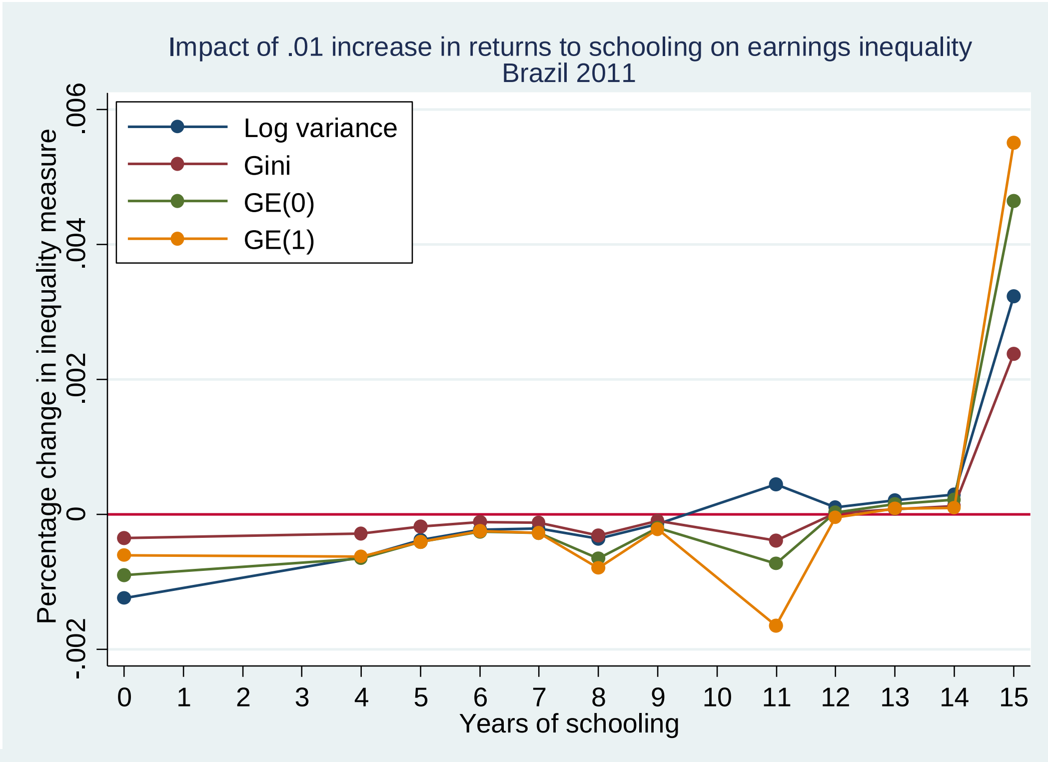
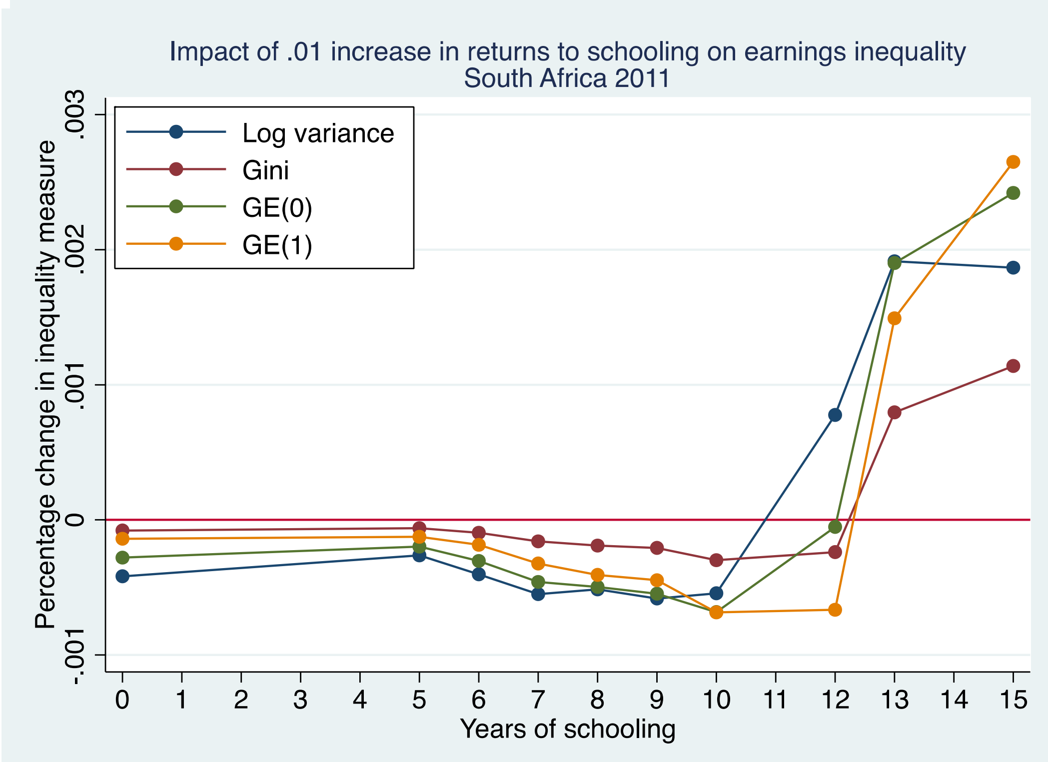


Figure 12 Alternative presentation, 2011





References

Azevedo, João Pedro, María Eugenia Dávalos, Carolina Diaz-Bonilla, Bernardo Atuesta, Raul Andres Castañeda. (2013) "Fifteen Years of Inequality in Latin America: How Have Labor Markets Helped." World Bank Policy Research Working Paper #6384.

Anderson, Kermyt, Anne Case and David Lam. 2001. "Causes and Consequences of Schooling Outcomes in South Africa: Evidence from Survey Data." *Social Dynamics* 27(1): 1-23.

Bourguignon, F., Ferreira, F. and Lustig, N. (2005). "The microeconomics of income distribution dynamics in East Asia and Latin America", Washington DC: Oxford University Press and The World Bank.

Goldin, Claudia, and Lawrence Katz (2008) *The Race between Education and Technology*. Harvard University Press.

Juhn, Chinhui, Kevin M. Murphy and Brooks Pierce (1993) "Wage Inequality and the Rise in Returns to Skill." *Journal of Political Economy*, 101(3), pp. 410-42.

Knight, J. B. and Sabot, R. H. (1983). "Educational Expansion and the Kuznets Effect". *American Economic Review* 73(5): 1132–36.

Lam, David, Cally Ardington, and Murray Leibbrandt. (2011). "Schooling as a Lottery: Racial Differences in Progress through School in Urban South Africa," *Journal of Development Economics*, 95(2):121-136.

Lam, David, and Deborah Levison (1992) "Declining Inequality in Schooling in Brazil and Its Effect on Inequality in Earnings," *Journal of Development Economics*, 37: 199-225

Moll, Peter (1996) "The Collapse of Primary Schooling Returns in South Africa, 1960-90," *Oxford Bulletin of Economics and Statistics*, 58: 185-209.

Mwabu, Germano, and T. Paul Schultz, 2000. "Wage Premia for Education and Location of South African Workers, by Gender and Race," *Economic Development and Cultural Change* 48(2), 307-334.

Paes de Barros, Ricardo and David Lam (1996) "Income and Education Inequality and Children's Schooling Attainment in Brazil," in Nancy Birdsall and Richard Sabot (eds.), *Opportunity Foregone: Education in Brazil*, Washington: Inter-American Development Bank.

Ram, Rati (1990) "Educational Expansion and Schooling Inequality: International Evidence and Some Implications," *Review of Economics and Statistics*, 72(2): 266-273.

Reis, J. Almeida dos and Barros, R. Paes de (1991). "Wage Inequality and the Distribution of Education: A Study of the Evolution of Regional Differences in Inequality in Metropolitan Brazil." *Journal of Development Economics* 36: 117–43.

World Bank (2011) *A Break with History: Fifteen Years of Inequality Reduction in Latin America*. Washington, DC: April, USA.