

“Education and Inequality: The South African Case”

David Lam*
University of Michigan
davidl@umich.edu

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

Julia Garlick
Yale University
Julia.Garlick@gmail.com

Nicola Branson
University of Cape Town
Nicola.Branson@gmail.com

* David Lam is Professor of Economics and Research Professor in the Population Studies Center at the University of Michigan. Murray Leibbrandt is a Professor of Economics and Director of the Southern Africa Labour and Development Research Unit (SALDRU) at the University of Cape Town. Julia Garlick is a graduate student in Economics at Yale University. Nicola Branson is a senior research at SALDRU. Support for this research was provided by the U.S. National Institute of Child Health and Human Development (Grants R01HD39788 and R01HD045581), the Fogarty International Center of the U.S. National Institutes of Health (D43TW000657), the Andrew W. Mellon Foundation and the Centre for Development Enterprise. Murray Leibbrandt acknowledges the Research Chairs Initiative of the Department of Science and Technology and National Research Foundation for funding his work as the Research Chair in Poverty and Inequality.

Abstract

Following the international literature, income inequality decompositions on data from contemporary South Africa show that the labour market is the key driver of overall household inequality. In order to understand one of the channels driving this labour market inequality, we use national household survey data to review changing returns to education in the South African labour market over the last 15 years; with a focus on both the returns to getting employment as well as the earnings returns for those that have employment. We show that South Africa has experienced a skills twist with the employment and earnings returns to matric and post-secondary education rising and the returns to levels of education below this remaining constant. Then, based on a regression based decomposition of earnings inequality, we show that this has had a major impact on earnings inequality. Indeed, the increase in returns to post-secondary education has directly counteracted the equalising gains that have been made by increased educational attainment, resulting in consistent levels of inequality over time.

South Africa has very high levels of unemployment thus making it imperative to incorporate the job allocation process into the analysis of earnings inequality. This is not a trivial analytic challenge and while the analysis in this paper is a promising start, the challenge to a more integrated analysis of employment and earnings remains. For the conference paper we will extend the analysis by simulating a full labour market process from employment rationing through to earnings for those with employment. This will allow us to derive a series of “full labour market earnings distributions” at difference points over the post-apartheid period and also to assess the role of education through a series of counterfactual simulations.

1. Introduction

The need to invest in human capital has been recognised in development economics for a long time. While other fashions have come and gone the case for such investment has grown stronger over time. Increased globalization and the consequent changing international demand for labour patterns have strengthened this case in general but have changed it too. A dominant outcome of these patterns is that they have increased inequality within developing economies and a focus on inequality is an interesting and useful prism through which to view the contemporary case for investment in human capital in general and in South Africa specifically. Human capital typically includes both skills and health. These two aspects often have different causes and consequences, and both have extensive literature devoted to them. This paper will consider only education, to provide thorough coverage rather than a broad overview.

The paper starts with a review of international thinking on education and inequality. This highlights the importance of forces coming from the labour market and looks at debates that have emerged about education policies in response to this situation. The next section turns to a review of the South African empirical evidence. The analysis is presented in section 4. First we show how the distribution of schooling has improved, with increases in schooling for non-whites resulting in a smaller racial gap. Next, we show that earnings inequality has remained consistently high since the end of apartheid. Taking advantage of the large national surveys collected for every year between 1997 and 2007, we estimate employment and earnings regressions that can be used to analyze the determinants on earnings inequality. We find offsetting effects related to education. On the one hand, the improvements in the distribution of schooling have tended to reduce overall inequality. On the other hand, increased returns to schooling for those with some post-secondary education, have tended to increase inequality.

The key insight from this discussion is that both more and better education is being demanded of the South African education system in order for it to become a more egalitarian force in the labour market and in general. Unlike much of the developing world, South Africa's school enrolment rates are very high into the middle secondary school years. The challenge is to build on the post-apartheid successes in increasing average years of schooling for the population so that most youth complete secondary schooling. This is doable. The

harder issue, which arises in the international milieu and in South Africa too, is the centrality of improving the quality of education.

2. Review of international literature

While there is a general consensus that investment in human capital is good both for growth as well as for the reduction in poverty and inequality, there is much disagreement over the precise quantitative effects of human capital on growth, poverty and inequality and over the most efficient forms of investment in education. At minimum though, there is agreement that it is important for governments to intervene to increase the quantity and quality of human capital being produced (World Development report, 2006 2007). While this is true of all forms of human capital we focus on education specifically.

Galor and Moav (2004) present a theoretical model that clarifies the problem of investment in education. During traditional industrialisation periods, the highest returns are to physical capital. The returns to physical capital are the same regardless of its distribution or who owns it. As the wealthy have higher savings and investment rates, economies can achieve high efficiency with highly unequal income distributions. However, there are definite limits on the returns to education per individual – three workers with twelve years of schooling produce a higher level of output than one worker with thirty-six years – so the distribution of human capital within the population affects the productivity of, and thus the returns to, human capital. As the economy develops, the stock and complexity of physical capital increases and the returns to human capital increase, due to complementarity between human and physical capital. Despite this, the majority of the population, facing liquidity constraints, cannot invest in human capital as they lack finances and access to credit, so the economy remains at a sub-optimal point.

Galor and Moav (2004) describe three ways out of this trap: wages can rise, reflecting the increasing returns to human capital and increasing the ability of the workers to save and invest in education; governments can intervene in recognition of efficiency goals to provide subsidized education; or the market can recognise the returns to education and choose to provide finance to those who are constrained by its absence. As this happens, human capital

becomes more widely spread and inequality decreases. The authors argue that traditionally economies used a mixture of these three options to move from capital intensity to skill intensity. The Galor and Moav (2004) model predicts with success the long run patterns of inequality observed within countries in the past. The implications of this model are clear: if action is not taken by some sector to address the liquidity constraints facing the lower income deciles, human capital and income will remain unequally distributed and education investments will occur at sub-optimal levels for growth.

Klasen (2002) focuses on gender inequality in education but arrives at similar implications. Discrimination, whether on income or gender or racial grounds, results in a sub-optimal allocation of education. If education is distributed on any grounds other than merit, some able students will not have access to higher levels of education and their places will be taken by less able students from the favoured group. The less able students will become less able workers, lowering the average productivity of educational investments. As human and physical capital is complementary, this results in lower returns to investment in physical capital. At a minimum, this will result in lower growth for a given level of investment, and, other things being equal, will decrease total investment in the economy. Further, as these less able students require more resources to achieve the same education levels, there will be a smaller equilibrium number of educated workers in the economy, again decreasing efficiency. Thus, even if we are interested only in the efficiency of overall production, the quantity and distribution of education available to the population is of concern.

However, this is not our only or even main interest. Development economics is deeply concerned with questions of poverty and income distribution, and the role of human capital in these areas is extremely large. Birdsall and Londono (1997) find that controlling for education and asset inequality removes the significance that is traditionally assigned to income inequality in hindering growth. This implies that the role attributed to income inequality is really a proxy for the effects of asset and education inequality. This supports the theoretical arguments referenced above. While inequality has a general negative effect on growth, its effect on the poor is particularly severe. The lower income quintiles typically experience lower growth in the presence of inequality than they would otherwise, and lower growth than the mean for their economies (Bourguignon, 2004). Thus, while growth might benefit the poor in the long run, in the short-run it contributes to increasing inequality. As inequality dampens growth, this creates a feedback loop which harms all members of the

economy. This would not be such a concern from an ethical point of view if there was substantial change in the composition of the lower income quintiles. However, most developing countries have quite rigid economic stratifications. In the majority of cases, the poor remain poor and the rich remain rich. This holds true between generations, too (World Development Report, 2006). Economic status appears to be inherited, meaning that the children of the poor will generally remain poor, in line with Galor and Moav's (2004) model.

The poor are also especially vulnerable to the composition of change in the globalised world. Wage inequality in general is rising in the world, with increases between but also within the groups of educated and uneducated workers. The gap between skilled and unskilled workers is being increased by technological change, as the returns to education increase. This force drives rising inequality among educated workers too, as ability and certain types of education are more productive (Heckman et al. 1998; Goldin & Katz, 2008). Inequality among uneducated workers is driven more strongly by randomness, as unskilled workers are less able to shift between industries and technologies. This makes them more vulnerable to shocks to a particular technology or industry. Workers are aware of this, and thus the demand for education among the poor and unskilled workers is driven by risk aversion as well as by perceived high returns to education (Gould et al., 2001).

Thus, the literature on international development makes a strong case for the fact that inequality in access to education plays a central role in perpetuating and generating inequality in labour market earnings and in income inequality more generally. This case is strengthened in the increasingly globalised world.

3. Review of the South African literature within these international dynamics

This section of the paper moves from the international literature and the international context to the South African situation. It provides a brief review of the empirical South African evidence on changes in inequality and returns to education over the past two decades. In doing so, it becomes clear that the missing link in the literature is the change in the

relationship between education and earnings inequality. This forms the motivation for our analysis in section 4.

The preceding international review emphasised the labour market as the central link between education and changes in national inequality. There is strong evidence to affirm the fact that, in South Africa too, it is labour market developments that dominate changes in broader inequality. Leibbrandt et al. (2010) review the post-apartheid empirical work on the relative impact on household income inequality of different income sources including wage income, state transfers and remittances. They then bring this work up to date using data from the 2008 National Income Dynamics Study. Two key points emerge. First, from the start of the post-apartheid period onwards, the relative success (or lack thereof) of household members in the labour market is the dominant driver of their position in the household income distribution. Second, there are two prongs to the labour market's role; namely, whether the members of the household have employment at all and then, for those with such earnings, their position within the distribution of labour market earnings.

Alongside South Africa's high income inequality, the country has long had one of the highest levels of earnings inequality in the world. Current research has focused on earnings inequality differences between subgroups defined by race, gender and location (Bhorat, 2004; Leite et al., 2006; Lam and Leibbrandt, 2004; Leibbrandt et al., 2007, 2010 and others). Earnings inequality is shown to have increased in the late 1990s and remained fairly stable thereafter. Unemployment is shown to be a key driver of inequality, with the increase in labour force participation and subsequent unemployment in the late 1990s contributing to increased inequality during this time. While between racial group inequality persists, within racial group inequality contributes a larger share to overall inequality and has increased substantially over the post apartheid years. Earnings inequality within the African population group is highest and has increased the most over time. Inequality within urban areas is also shown to have increased as workers moving to the cities to find jobs increase the variance in earnings among urban dwellers. Rural earnings inequality has decreased.

The literature on returns to education focusing on racial and gender discrimination is also well established (Mwabu and Schultz, 1996; Moll, 1998; Ntuli, 2007; Burger and Jafta, 2006; Burger, 2011; Burger and van der Berg, 2011; Lam et al., 2011). Returns to matric and post secondary education are high (Keswell and Poswell, 2004; Branson et al., 2009; Lam et al.,

2011). Over time, however, matric returns in terms of earnings have remained fairly static, with returns to post schooling (especially degree qualifications) increasing. This has widened the relative reward to post secondary education (Branson et al., 2009; Lam et al., 2011).

In addition to high earnings inequality, South Africa has a large unemployment problem (Kingdon and Knight, 2001). A crucial connection between education and inequality in South Africa is the role of education in determining who is employed. With a large pool of potential workers, education may influence employers' decisions about job allocations. Branson et al. (2009) show that matric and post secondary education improve access to the job market, with the probability of employment improving substantially on completion of matric with a further improvement for post secondary schooling. Given the high levels of unemployment in South Africa, measures of inequality need to account for differential access to employment. Tragenna (2009) investigates this between 2001 and 2007 and finds that the vast majority of earnings inequality can be attributed to the rate of unemployment. In addition, she shows that changes in unemployment, account for the majority of the change in inequality over the period (Tragenna, 2009).

It is particularly important to emphasise both of these prongs. International literature tends to focus on the link between education and the inequality of earnings for those that have earnings. However, given South Africa's acute and rising unemployment problem, a South African discussion has to balance this prong with one linking levels of education to whether or not a person is employed and has any earnings at all. We give attention to both of these issues by analysing changes in relationships between years of schooling and the probability of employment as well as the distribution of labour earnings in South Africa from 1997 to 2007.

4. Empirical evidence on the relationship between education, employment, earnings and inequality

The empirical analysis is presented in this section in three parts. First, we look at changes in the distribution of education. Second, we assess changes in the returns to education in terms of both earnings and employment probability. The final section ties the analysis together by assessing the extent to which the changes in the schooling distribution and the returns, explain the changes, or lack thereof, in inequality over the period.

4.1 Data

We make use of data from a series of large household surveys collected by Statistics South Africa. In the 1990s the October Household Survey (OHS) was collected annually from 1994 to 1999. In 2000 the OHS was replaced, in part, by the Labour Force Survey (LFS), a rotating panel collected twice each year. Given concerns about data quality in the 1995 and 1996 OHS, we begin our analysis with the 1997 OHS, using the 1998 and 1999 OHSs as well¹. We use the September rounds of the Labour Force Survey from 2000 to 2007, giving us data for a full decade of dramatic social, political, and economic change. We use a set of weights (cross entropy weights) calibrated to a consistent demographic model over time (Branson, 2010) in order to avoid errors resulting from inconsistent representation across the surveys. Our analysis focuses on men and women aged 25-59, the prime years for employment in the labour force.

Table 1 shows the sample size of men and women aged 25-59 for each year for each of the four major population groups in South Africa – African (black), coloured, Indian (Asian), and white. Statistics South Africa continues to collect self-identified population group data using the same classifications used under apartheid. We will use these classifications in our analysis. As seen in Table 1, we have roughly 40,000 observations in each year. Using the

¹ Note that Kerr and Wittenberg (2012) show that there is an important break in the series between the 1998 OHS and subsequent surveys that resulted from changes in survey practices. These changes resulted in a substantial increase in the proportion of people living in single person households or in backyard shacks over this period. Caution therefore needs to be taken in interpreting changes over this period.

cross entropy sample weights, the South African population is roughly 75% African, 10% coloured, 3% Indian, and 12% white, with a slight increase in the proportion African and a slight decrease in the proportion white over the decade. The table also presents the size of the labour force with the weighted percentage working with positive earnings. It shows that employment rates declined between 1997 and 2002, before starting to increase once more.

4.2 Analysis

4.2.1 Changes in the distribution of schooling

Figure 1 presents cumulative distributions of schooling for men and women aged 25-59 for 1997, 2002, and 2007, with separate estimates for Africans and whites. The cumulative distributions provide a detailed summary of the changes across the full distribution of years of schooling. The most noticeable feature of Figure 1 is the large racial differences in schooling. For all the years, white men and women have more years of education than their African counterparts². Another important feature of Figure 1 is that the distribution of schooling is very similar for men and women in both racial groups. As pointed out by Anderson et al. (2001) and Lam et al. (2011), women go through school somewhat faster than men in all racial groups in South Africa, and women end up with slightly higher schooling attainment.

Figure 1 shows improvements in schooling over time. Table 2 confirms that the education distribution has changed significantly. The table presents Kolmogorov-Smirnov tests of the equality of the educational attainment distributions in 1997 versus 2002 and 2002 versus 2007 for Africans and whites separately. This null hypothesis is rejected for both comparison within the African group, reflecting that the educational distribution changed significantly between 1997 and 2002 and 2002 and 2007. There was also a statistically significant shift in the education distributions between 1997 and 2002 among whites. Note that much of this improvement reflects schooling that was completed before the end of apartheid³.

² Fewer than 10% of white men and women had less than nine years of schooling in 1997 (with little change over time), while almost 60% of African men and women had less than nine years of schooling in 1997.

³ Respondents aged 25-59 at the survey date would have completed their schooling 5-40 years prior to the survey date. Thus the majority of respondents would have completed their schooling before to the end of apartheid.

Examining Figure 1, we see that the percentage of African men with less than nine years of schooling, for example, fell from 61% in 1997 to 47% in 2007. The percentage of African men with less than 12 years of schooling fell from 83% in 1997 to 73% in 2007. Similar changes are evident for women. Improvements in completed secondary school have been less dramatic. The percentage of men completing 12 years of education has remained much the same, while African and white women have both seen a slight improvement. The percentages of white men and women with more than 12 years of schooling have both shifted slightly during the decade, but the major changes for African men and women lie below grade 12. As we will see below, there are very high returns to post-secondary schooling in terms of both earnings and employment. The lack of improvement in post-secondary school completion among Africans is thus important in understanding the lack of progress in closing the racial gap in earnings.

4.2.2 Changes in the racial, gender and schooling premium

Key to this paper is the link between the changes in the distribution of schooling and changes in the distribution of earnings. Figure 2 shows the relationship between years of schooling and log monthly earnings in the top panel and between years of schooling and the probability of employment in the bottom panel for African and white men in 1997 and 2007⁴.

The figure shows that returns to schooling are relatively modest until the completion of secondary school. The earnings returns to completing grade 12 (versus grade 11) for Africans are very high in both years. Returns to post-secondary education are even higher and have increased over time. African men with two years of post-secondary schooling have earnings that are 40% higher than African men with grade 12 in 1997 and 70% higher in 2007. This means that in 2007 a man with 14 years of education earned almost twice as much as one

⁴ Whites are only shown above grade 8 due to low observation numbers. The figure is based on separate regressions for each population group and year using education indicators and age and age squared controls. The education indicators take on a value of one if the respondent completed that level of education or more, and zero otherwise. Fourteen years of education reflect diploma-type qualifications including technical university training and other post-secondary education short of a university degree. Fifteen years of education are equivalent to university degrees or anything beyond. The figure shows real log earnings for each education group relative to Africans with seven years of schooling (end of primary school) in the given year.

with 12 years of education⁵. The difference between Africans and whites is also stark, even after controlling for schooling. In fact, Africans with 15 years of education (a completed university degree) have similar earnings returns to whites with grade 12 (completed secondary). This difference persists over time.

The lower panel of Figure 2 presents a similar picture for returns to education in terms of attaining employment. Post-secondary education plays a strong role in predicting employment for Africans. African males with a diploma or degree are close to 20 percentage points more likely to be employed than African males with grade 7. Interestingly, the premium for employment from matric is relatively small (4 percentage points in 2007) and only significant in 2007.

Figure 2 suggests that while there is a large white premium in both earnings and employment⁶, there have been no substantial changes in the racial premium between 1997 and 2007. Figure A2 in the appendix presents an equivalent figure to Figure 2 for males versus females. The sample is restricted to Africans. The gender gap in earnings returns shows little change over time, with the gap larger at lower levels of education. There have, however, been decreases in the male premium for employment between 1997 and 2007, particularly for schooling levels below the post-secondary level.

4.2.3 Decomposing Earnings Inequality

Figure 3 shows three standard measures of earnings inequality for each year from 1997 to 2007 – the standard deviation of log earning, Theil's L⁷ and the Gini coefficient. In addition, the Gini coefficient is calculated both in the standard way for respondents with positive earnings and for a broader sample that includes the unemployed as zero earners. In each year the measures in figure 3a are calculated using the full sample of men and women from all population groups. Although there is some variations across years, the overall picture from all measures is that there has been very little change in inequality over the period. Figure 3b

⁵ $e^{0.70} \approx 2$

⁶ Figure A1 in the appendix presents a similar figure including coloured respondents and shows that coloureds also have a labour market premium over Africans.

⁷ The Generalized Entropy measure with $\alpha = 0$, or mean log deviation.

presents the inequality measures calculated within African and white groups. A similar pattern of persistent levels of inequality is evident.

While overall inequality may be fairly stable, inequality could be increasing or decreasing for specific groups. It is thus important to assess the role of schooling, race and gender in explaining these aggregate earnings trends. Earnings inequality can be examined by looking at the variance of log income – the greater the variance, the higher the inequality. Variance can be split into explained and unexplained (residual) components. The top panel of Figure 4 shows the decomposition of the variance into explained and residual components. When explained variance falls, the explanatory power of our analysis is lower, as the variables we are examining can explain less of the observed inequality. Table A1 in the appendix shows the effect of various determinants on the log of earnings in 1997, 2002 and 2007⁸.

As was evident in Figure 3a, Figure 4 shows that the total variance of log earnings remained remarkably stable between 1997 and 2007. Interestingly, while the residual variance mimics this stability over the period, the explained variance increases steadily between 1997 and 2002, from 0.54 to 0.81, and decreases thereafter to reach 0.65 in 2007.

The changes observed in the explained variance could result from changes in the distribution of characteristics of the population (Xs) or changes in the coefficients on key characteristics in the regressions. A useful technique for assessing these changing roles of various factors to inequality is the creation of counterfactuals. We use the information obtained from regressions in Table A1 to assess how much of the change in explained inequality observed is due to changes in the distribution of variables in the population, and how much is due to changes in the returns associated with education, race and gender. In Figure 1 we saw that the schooling distribution became more equal over time, with a smaller racial gap and with declining inequality in schooling among Africans. This would have a potentially equalizing effect on the variance of earnings. At the same time, the returns to schooling increased for those with post-secondary education. This could potentially have a

⁸ Table A1 shows Ordinary Least Squares regressions using the log of earnings as the dependent variable, with the sample including all men and women aged 25-59 with positive earnings. The independent variables include education indicators, racial and a male dummies and a quadratic function of age. The education indicators take on a value of one if the respondent completed that level of education or more, and zero otherwise. Diploma includes technical university training and other post-secondary education short of a university degree, degree includes university degrees and anything beyond university. We have estimated identical regressions for every year from 1997 to 2007, but display only three years for brevity.

disequalizing effect. In addition, the racial gaps appear to have increased, creating another potential disequalizing effect. We use counterfactual simulations to try to identify the contribution of these effects to overall inequality.

The bottom panel of Figure 4 summarizes one set of simulations that use 1997 as a baseline⁹. The top solid black line shows the simulated variance of log earnings if we hold the distribution of all characteristics at their 1997 level and then use the regression coefficients for each year to predict earnings. As can be seen in the figure, if the 1997 characteristics stayed constant but returns to characteristics altered, the shape of the inequality trend would be very similar to that evident in the explained variance curve in the top figure; an increase from 1997 to 2002 followed by a thereafter. The second line in the bottom panel of Figure 4 changes only the coefficients for race and schooling, leaving the age and male coefficients fixed as estimated in 1997. The line tracks very closely the top line where all coefficients are allowed to change, demonstrating that it is changes in the race and schooling coefficients that explain most of the pattern shown in the top line. Indeed, observing the third line, where only the schooling coefficients are allowed to change and all other coefficients at their 1997 level, it is evident that the pattern of increase between 1997 and 2002 followed by a decline thereafter is a function of changing coefficients on the schooling variables. The fourth line, where only the race coefficients vary over time, is remarkably flat. The results for the third and fourth simulations imply that it was the change in schooling coefficients that generated the steep increase in inequality shown in the top line of Figure 4. Importantly, while the race coefficients increase the overall variability, they contribute equally in each year and do not have an impact on the shape of the inequality changes.

The bottom line in Figure 4 presents the counterfactual in a different way; i.e., the coefficients are held at their 1997 level and the distribution of characteristics are allowed to change in every year. In other words, we take the actual distribution of characteristics and predict earnings using the 1997 regression coefficients. The results of this simulation are quite interesting. They show a steady decline in inequality from 2000 to 2007, with no evidence of the sharp increase shown in the other simulations. This implies that the changes in characteristics had an equalizing effect on earnings. Since the racial composition remained

⁹ Note that this is essentially an exercise in simulating the explained variance component, since residual variance is ignored. The variance for each year is constructed by using the regression coefficients for each year to predict earnings for the 1997 sample, then taking the variance of the predicted earnings.

almost constant, as shown in Table 1, it was the improvements in the schooling distribution that generated this decrease in inequality.

The coefficients from these regression are weighted estimates across all population groups. One may be concerned that unobserved heterogeneity between groups, for example different liquidity constraints or intertemporal rates of substitution, would reduce the usefulness of these estimates. Figure 4b and 4c present estimates from a similar exercise for Africans and whites separately.

- Smaller share of overall variance in earnings can be explained by schooling, gender and age within population group.

The shape of the African curve closely resembles that in the full population group. The estimate of earnings variation that is explained by schooling, gender and age within the African group increases between 1997 and 2002 and then decreases. Estimates of the explained variance? are far lower (0.1 compared to 0.3-0.5 for Africans) and very stable within the white group. Similar to the overall population, the African simulations show that changes in the coefficients on schooling are the main drivers of the increase in explained inequality between 1997 and 2002 and the decline thereafter. One difference is that when coefficients are fixed at 1997 levels and the distribution of the schooling, gender and age variables are allowed to change, the trend, unlike the decline evident in the overall population, is stable. Similar stability is evident within the white group. Thus the increase in educational attainment over the period evident in Figure 1 is just offset by the increase in inequality caused by shifts in the coefficients.

Taken together then, these simulations imply that there were important but offsetting effects on earnings inequality in South Africa between 1997 and 2007. Changes in the distribution of characteristics, primarily improvements in African schooling, tended in and of themselves to reduce earnings inequality. If coefficients (i.e., the returns to the difference years of schooling) had remained at their 1997 level, the improvements in the schooling distribution would have caused a 6% decline in the log variance of earnings¹⁰. This was offset by the fact

¹⁰ The improvements generate a 13 % decline in the explained component of the log variance of earnings according to the decompositions (dropped from 0.54 in 1997 to 0.48 in 2007). Since the explained component is about 45% of total variance (see the R^2 values in Table A1), this would have implied about a 6% decline in the log variance of earnings.

that the coefficients in the earnings regressions increased earnings inequality. Most importantly, increased returns to schooling at the post-secondary level, a factor that tended to pull the highest earners even farther away from those in the middle of the distribution.

5. Summary and discussion

South Africa's large national household surveys allow us to get a good picture of what has happened to both the distribution of education and the distribution of earnings since the end of apartheid. Taking advantage of what we consider the most reliable data on the labour market, we have analyzed data from the 1997, 1998, and 1999 October Household Surveys in combination with the Labour Force Surveys (September rounds) from 2000 to 2007. We show that there continue to be large racial gaps in education, although African schooling for the 25-59 population improves in important ways between 1997 and 2007. There is very little change in the proportion with post-secondary schooling, however, and this is an important factor in the large racial gap in earnings. We show that returns to post-secondary schooling in terms of both earnings and employment are very large and have risen substantially in the last decade.

Overall earnings inequality has remained remarkably stable between 1997 and 2007. If anything, inequality rose marginally in the 1990s and then declined slightly after 2000. We estimate earnings regressions for every year in order to clarify the determinants of these changes in inequality. We show that the changes in the distribution of characteristics, primarily the improvements in the schooling distribution, would have caused a decline in inequality (as measured by the variance of log earnings) of roughly 6% had wage returns to various characteristics remained constant. The reason that we see no change in inequality is that the improvements in schooling, which had an equalizing effect, were offset by changes in the returns associated with education characteristics, particularly.

While there have been significant improvements in average levels of education of the South African population over the post-apartheid period, these improvements leave the bulk of the South African population with incomplete secondary education. If the labour market had remained as it was in the early post-apartheid years, there would have been a high return to these improvements as seen through declining earnings inequality. However, the labour

market has not stayed the same. In line with trends elsewhere in the world, South Africa has experienced a skills twist with the returns to matric and post-secondary education rising and the returns to levels of education below matric falling sharply. This is particularly cruel in the South African case as the fall in returns has directly counteracted the gains that have been made in years of secondary schooling.

Our analysis of the changing effect of education on the probability of employment shows that the same forces are at work here too. The relative returns to any post-secondary qualification in terms of a higher probability of employment have risen between 1997 and 2007. That said, very few black South Africans reach the post-secondary level, so these returns serve to increase within-race inequality without necessarily decreasing overall inequality. The relative returns to completed matric in terms of accessing formal employment did not rise, though they remained high.

A major education policy focus in the international literature is the need to increase primary school enrolments. This is not a particular concern in South Africa. Indeed, South Africa enjoys close to universal primary school enrolment and we have shown that the post-apartheid era has been characterised by even further advances in the average years of schooling being acquired by South African youth. Rising returns to complete secondary and post-secondary education make a case for the continuation and extension of such policies to encourage the completion of secondary school and access to post-secondary school education. It seems that the average young South African must acquire at least complete secondary schooling to be rewarded in the labour market.

Yet our empirical work also raised issues that caution against an exclusive focus on years of education. First, there is the fact that South Africa combines one of the highest levels of inequality in the world with these high levels of primary and secondary school enrolment. This suggests that somewhere along the schooling path, something goes wrong. High education levels ought to result in lower inequality and high growth. Instead, South Africa has exceptionally high income inequality, high unemployment and relatively low growth rates. Second, it is a concern that our data suggest that, even controlling for years of education, white, coloured, and Indian labour market participants continue to receive strong premia relative to Africans both in terms of employment and earnings. On initial examination, this suggests persistence of racial discrimination in the South African labour

market. However, in post-apartheid South Africa this seems unlikely and Chamberlain and van der Berg (2002) show that controlling for school quality removes a great deal of this discrimination. This implies that a significant component of the discrimination in the South African labour market can be attributed to differences in characteristics of workers rather than explicit discrimination; i.e., racial disadvantage based on high inequality of functional years of education.

Indeed, school quality has been identified as a major contributor to labour market outcomes. The World Bank identifies South Africa as one of the countries which faces achievement rather than attainment challenges in education (World Development Report, 2007), a criticism which is widely repeated throughout the literature and acknowledged by the South African government. School quality aside, the pool of matriculants has grown with little change in the proportion of the population completing post-secondary education. Breier and Mabizela (2008) show that 80 percent of students who passed the senior certificate in 2002-2004 did not continue on to a post-secondary institution in the year after completing school. While the majority passed without university endorsement, these learners are eligible to attend alternative higher education institutions. Our research (Lam et al., 2010) shows that all forms of post secondary education are highly valued in the labour market. Policies that address blockages in access to post-secondary education could therefore have implications for inequality reduction.

In concluding it is important to locate the discussion of curricula and the school environment within a broader focus on human capital as a whole. Cognisance needs to be given to Carneiro and Heckman's (2002) work showing that educational disadvantages created by early childhood educational, familial and social environments are near permanent and very hard to reverse later on. Recent work on the progress through secondary schools in Cape Town (Lam et al., 2011) has strongly endorsed these findings in the South African context and eligibility and success in higher education is strongly determined by these foundational factors.

There are no obvious quick fixes to the education system that will make it an egalitarian driver in South African society. Increasing access to education is the one domain in which we have made progress but, apparently, without much of a return. We need to push on with this until most South Africans have at least completed secondary schooling. However,

accumulating years of education is not an end in itself. It is the increased potential and productivity that each extra year of schooling adds that is the real social return to investment in education. Without these increases in South Africa's human capital, increases in years of schooling are without value. This is why there is such intense focus on ensuring that young learners walk into school in a state that is receptive to quality education and that the schooling and higher education milieu then provides this quality education.

References

ANDERSON, K., CASE, A. and LAM, D. (2001). Causes and Consequences of Schooling Outcomes in South Africa: Evidence from Survey Data, *Social Dynamics*, 27(1):1-23.

BHORAT, H. (2004) Labour Market Challenges in the Post-Apartheid South Africa, *South African Journal of Economics*, 72 (5): 940-977.

BIRDSALL, N. and LONDONO, J. (1997). Asset Inequality Matters: An Assessment of the World Bank's Approach to Poverty Reduction, *The American Economic Review*, 87(2): 32-37.

BOURGUIGNON, F. (2004). The poverty-growth-inequality triangle, *Indian Council for Research on International Economic Relations, New Delhi working papers: 125*.

BRANSON, N. (2010). Cross entropy weights OHS 1994-LFS 2007 September.

<http://www.datafirst.uct.ac.za/catalogue3/index.php/catalog>

BRANSON, N, LEIBBRANDT, M & ZUZE, T. (2009). The demand for tertiary education in South Africa, *Final report to the Centre for Higher Education Transformation*.

BREIER, M. and MABIZELA, M. (2008). Higher education. In Human Resources development review 2008: Education, employment and skills in South Africa. Eds Andre Kraak and Karen Press. HSRC press.

BURGER, C. and VAN DER BERG, S. (2011). Modelling cognitive skills, ability and school quality to explain labour market earnings differentials, *Stellenbosch Economic Working Papers: 08/11*.

BURGER, R. (2011). Estimating the shape of the South African schooling-earnings profile, (Doctoral dissertation). Oriel College, Oxford University.

BURGER, R and JAFTA, R. (2006). Returns to Race: Labour Market Discrimination in Post-Apartheid South Africa, *University of Stellenbosch Working Paper*.

CARNEIRO, P. and HECKMAN, J. (2002). “The Evidence on Credit Constraints in Post-Secondary Schooling” , *The Economic Journal*, 112(482) : 705-734.

CASE, A. and YOGO, M. (1999). Does School Quality Matter? Returns to education and the characteristics of schools in South Africa, *NBER Working Paper 7399*.

CHAMBERLAIN, D. and VAN DER BERG, S. (2002). Earnings Functions, Labour Market Discrimination and Quality of Education in South Africa, *Stellenbosch Economic Working Papers 02/02*.

GALOR, O. and MOAV, O. (2004). From Physical to Human Capital Accumulation: Inequality and the Process of Development, *The Review of Economic Studies*, 71(4): 1001-1026.

GOLDIN, C. and KATZ, L. (2008). The Race between Education and Technology, *Harvard University Press*.

GOULD, E., MOAV, O. and WEINBERG, B. (2001). Precautionary Demand for Education, Inequality, and Technological Progress, *Journal of Economic Growth*, 6(4): 285 – 315.

HECKMAN, J., LOCHNER, L. and TABER, C. (1998). Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents, *Review of Economic Dynamics*, 1 : 1 – 58.

KESWELL, M. and POSWELL, L. (2004). Returns to education in South Africa: A retrospective sensitivity analysis of the available evidence, *South African Journal of Economics*, 72(4): 834-860.

KINGDON, G. and KNIGHT, J. (2001). What have we learnt about unemployment from microdatasets in South Africa?, *Social Dynamics*, 21(1): 79 – 95.

KLASEN, S. (2002). Low Schooling for Girls, Slower Growth for All? Cross-Country Evidence on the Effect of Gender Inequality in Education on Economic Development, *The World Bank Economic Review*, 16(3): 345 – 373.

LAM, D., ARDINGTON, C. and LEIBBRANDT, M. (2011). Schooling as a Lottery: Racial Differences in School Advancement in Urban South Africa, *Journal of Development Economics*, 95(2): 121-136.

LAM, D., ARDINGTON, C., BRANSON, N., GOOSTREY, K. AND LEIBBRANDT, M. (2010). Credit Constraints and the Racial Gap in Post-Secondary Education in South Africa, *Prepared for the 2010 meeting of the Population Association of America Dallas Texas.*

LAM, D. and LEIBBRANDT, M. (2004). What's Happened to Inequality in South Africa since the End of Apartheid, *mimeo, University of Cape Town.*

LEIBBRANDT, M., WOOLARD, I., FINN, A. and ARGENT, J. (2010). Trends In South African Income Distribution and Poverty Since the Fall of Apartheid, *OECD Social, Employment And Migration Working Papers: 101.*

LEIBBRANDT, M., WOOLARD, I. and WOOLARD, C. (2007). Poverty and Inequality Dynamics in South Africa: Post-Apartheid Developments in the Light of the Long-Run Legacy. *Prepared for the IPC- DRCLAS workshop, Brasilia 11-13 January 2007.*

LEITE, P.G., MCKINLEY, T. and OSORIO, R.G. (2006). The Post Apartheid Evolution of Earnings Inequality in South Africa, 1995-2004, *International Poverty Centre UNDP Working Paper: 32.*

MOLL, P. (1998). Primary Schooling, Cognitive Skills and Wages in South Africa, *Economica*, 65(258): 263 – 284.

MWABU, G. and SCHULTZ, P. (1996). Education Returns Across Quantiles of the Wage Function: Alternative Explanations for Returns to Education by Race in South Africa, *The American Economic Review*, 86(2):335 – 339.

NTULI, M. (2007). Exploring Gender Wage Discrimination in South Africa, 1995-2004: A Quantile Regression Approach. *IPC Working Paper Series: 56*.

STATISTICS SOUTH AFRICA (2000-2007). Labour Force Survey Data.

STATISTICS SOUTH AFRICA (1997-1999). October Household Survey Data.

TREGENNA, F. (2009). The Relationship Between Unemployment and Earnings Inequality in South Africa, *CWPE: 09/07*.

WORLD DEVELOPMENT REPORT (2007). Development and the Next Generation, *World Bank Publications*, 2 – 21.

WORLD DEVELOPMENT REPORT (2006). Equity and Development, *World Bank Publications*, 76 – 88.

Table 1: Size of sample and labour force with percentage by population group and for positive earners

Survey	Total	African		Coloured		Indian		White		# in Labour Force	% Positive Earning
	N	N	Weighted percent	N	Weighted percent	N	Weighted percent	N	Weighted percent		
1997 OHS	49101	36363	73.1	6894	9.8	1442	3.1	4402	14.0	28279	78.1
1998 OHS	30043	22196	73.4	3779	9.7	860	3.1	3157	13.7	19009	72.9
1999 OHS	40798	30390	73.6	5114	9.7	1099	3.1	4141	13.4	27442	71.0
2000 LFS	40338	31063	74.0	4941	9.7	1026	3.1	3236	13.1	30046	71.6
2001 LFS	41456	31293	74.3	5122	9.7	1073	3.1	3910	12.8	29713	69.6
2002 LFS	39807	29431	74.6	5033	9.7	1182	3.1	4122	12.5	28632	66.9
2003 LFS	38875	28570	74.9	4878	9.7	1158	3.1	4228	12.3	27054	69.0
2004 LFS	41717	30707	75.0	6239	9.7	949	3.1	3784	12.0	27396	69.5
2005 LFS	41421	30765	75.1	6172	9.8	1001	3.0	3406	11.8	28429	71.3
2006 LFS	40659	30390	75.2	6239	9.8	984	3.0	2998	11.7	28498	71.8
2007 LFS	40167	30200	75.4	6055	9.8	996	3.0	2852	11.5	27279	74.5

Notes to Table 1: Sample sizes and percentages of men and women aged 25-59 for each year for each of the four major population groups in South Africa. The table also presents the size of the labour force with the weighted percentage working with positive earnings. September rounds of the Labour Force Surveys used. Weighted using cross entropy weights.

Table 2: Kolmogorov-Smirnov tests of the equality of the educational attainment distributions in 1997 and 2002 and 2002 and 2007

	African		White	
	K-S statistic	p-value	K-S statistic	p-value
1997-2002	0.056	0.00	0.090	0.00
2002-2007	0.051	0.00	0.041	0.07

Notes to Table 2: Kolmogorov-Smirnov test statistics and significance levels of the test of whether the educational distributions in 1997 vs 2002 and 2002 vs 2007 are equivalent, presented separately for Africans and whites.

Table 3: Total variance in log earnings and the proportion explained by observable characteristics

	All South Africans		Africans only		Whites only	
	Total Variation	% Explained	Total Variation	% Explained	Total Variation	% Explained
1997	1.29	41.7	0.98	29.6	1.08	11.6
1998	1.41	45.2	1.11	31.0	0.97	17.6
1999	1.67	42.0	1.34	29.0	1.27	14.7
2000	1.43	50.4	1.14	34.5	0.73	23.7
2001	1.41	52.0	1.12	38.5	0.79	21.0
2002	1.53	52.8	1.26	41.6	0.79	19.9
2003	1.44	52.7	1.18	40.5	0.74	21.3
2004	1.38	49.9	1.10	36.9	0.86	21.1
2005	1.50	47.0	1.23	38.1	1.08	19.2
2006	1.40	49.5	1.15	37.7	0.78	21.5
2007	1.38	49.4	1.14	37.9	0.81	30.3

Notes to Table 3: The table presents the total variation in log monthly earnings in each year and the proportion that can be explained by observable characteristics for the full population, Africans only and whites only. Observable characteristics are schooling, gender, age and race in the full population estimates and schooling, gender and age in the African only and white only estimates. Sample include respondents 25-29 in employment with positive earnings.

Table 3: Total variance in log earnings and the proportion explained by observable characteristics

	All South Africans		Africans only		Whites only	
	Total Variation	% Explained	Total Variation	% Explained	Total Variation	% Explained
1997	9.03	20.7	9.42	12.8	3.62	11.0
1998	10.25	21.7	10.44	12.5	3.47	9.6
1999	9.99	19.8	10.01	11.9	4.75	10.0
2000	10.36	21.1	10.34	12.1	4.23	12.2
2001	11.23	21.8	11.04	13.9	4.91	14.5
2002	11.76	21.9	11.50	14.5	4.97	9.5
2003	11.35	20.8	11.24	13.5	4.74	10.0
2004	11.11	19.9	11.02	13.4	4.89	8.7
2005	11.23	21.1	11.13	15.1	5.08	13.6
2006	11.21	21.0	11.31	13.6	3.57	9.4
2007	10.54	19.8	10.68	12.9	3.49	13.0

Notes to Table 3: The table presents the total variation in log monthly earnings in each year and the proportion that can be explained by observable characteristics for the full population, Africans only and whites only. Observable characteristics are schooling, gender, age and race in the full population estimates and schooling, gender and age in the African only and white only estimates. Sample include respondents 25-29 in the labour force, the unemployment are included as zero earners.

Table A1

Variable	Log earnings		
	1997	2002	2007
Years of schooling at least:			
1 year	-0.0467 [0.115]	0.0437 [0.0902]	-0.301** [0.134]
2 years	0.105 [0.148]	0.0963 [0.109]	0.194 [0.122]
3 years	0.0315 [0.106]	-0.0946 [0.0784]	0.145 [0.106]
4 years	0.133** [0.0568]	0.151*** [0.0547]	0.00333 [0.0771]
5 years	0.100* [0.0519]	0.0586 [0.0513]	-0.0950 [0.0824]
6 years	0.0516 [0.0468]	0.113** [0.0490]	0.192*** [0.0714]
7 years	0.0975*** [0.0357]	0.0127 [0.0408]	0.0508 [0.0510]
8 years	0.157*** [0.0296]	0.249*** [0.0363]	0.182*** [0.0480]
9 years	0.112*** [0.0328]	0.0970** [0.0383]	0.0876** [0.0432]
10 years	0.166*** [0.0334]	0.223*** [0.0379]	0.0559 [0.0415]
11 years	0.157*** [0.0292]	0.151*** [0.0372]	0.260*** [0.0447]
12 years	0.221*** [0.0263]	0.321*** [0.0327]	0.260*** [0.0425]
Diploma	0.379*** [0.0283]	0.635*** [0.0276]	0.686*** [0.0428]
Degree	0.126*** [0.0466]	0.309*** [0.0449]	0.534*** [0.0705]
Population group (African omitted):			
White	0.779*** [0.0234]	0.885*** [0.0252]	0.817*** [0.0450]
Indian	0.458*** [0.0363]	0.597*** [0.0345]	0.638*** [0.0646]
Coloured	0.156*** [0.0169]	0.374*** [0.0212]	0.419*** [0.0334]
Male	0.494*** [0.0136]	0.530*** [0.0152]	0.510*** [0.0239]
Age	0.0786*** [0.00671]	0.106*** [0.00738]	0.0595*** [0.0108]
Age squared	-0.000789*** [0.000083]	-0.00108*** [0.000091]	-0.000541*** [0.000135]
Constant	4.137*** [0.133]	3.425*** [0.148]	5.066*** [0.206]
Observations	21,808	19,214	20,042
R-squared	0.417	0.529	0.494

Notes to Table A1: The table shows Ordinary Least Squares regressions using the log of earnings as the dependent variable. The independent variables include education indicators, racial and a male dummies and a quadratic function of age. The education indicators take on a value of one if the respondent completed x or more years of education, and zero otherwise. Diploma includes technical university training and other post-secondary

education short of a university degree, degree includes university degrees and anything beyond university. The sample is men and women aged 25-59 with positive earnings.