

INTERGENERATIONAL PERSISTENCE OF EDUCATIONAL STATUS IN SOUTH AFRICA

University of Cape Town

Benedict Rhodes

RHDBEN003

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Table of contents

INTRODUCTION	2
LITERATURE REVIEW	3
METHODOLOGICAL APPROACH	9
DATA	11
EMPIRICAL ANALYSIS	16
CONCLUSION	23
BIBLIOGRAPHY	24
APPENDIX	26

INTRODUCTION

The study of the correlation between the incomes of parents and adult children has been a topic of interest for social scientists and economists for years. This topic has garnered significantly more interest in recent years, leading to the use of new data and improved econometric techniques to analyse patterns of persistence in earnings. Unsurprisingly, developing countries tend to lack such information, restricting this line of research. Only recently has there been an emergence of quality longitudinal data for researchers to begin understanding how these long run trends were formed.

South Africa is a country steeped in inequality due to a discriminatory racial history, which makes investigations into social mobility especially difficult. The disparities in access to income earning opportunities is one of the biggest contributors to the inequality in the country, and those who are educated find upward occupational mobility easier than others. Few South African studies have examined the intergenerational transmission of status of any kind. While the limitations in doing so are understandable, a deeper understanding of the extent of these inherited outcomes would be necessary for any policy implementation with the intention of correcting the systematic discrimination of the past. Policies which aim to address inequality should understand its origin. Market orientated policies rely too heavily on the notion that equal access to education and occupation is sufficient, they fail to consider the effect of inherited status that explain the outcomes of generations. Policies to address these inherited inequalities are crucial. (Burns and Keswell, 2012)

With panel data being relatively new, the most comprehensive research of intergenerational earnings mobility was conducted by Piraino (2015). The paper estimated the degree of earnings persistence in South Africa and further explored the link between it and an index of inequality of opportunity. It found that the inequality of opportunity index was high, stating that this may partially be explained by the high levels of intergenerational transmission of earnings. Further, the paper suggested that the low levels of intergenerational mobility are comparable to those of other high inequality emerging economies like Brazil and China.

The role of education in shaping intergenerational earnings mobility should not be overlooked. In Leibbrandt et al (2016), the educational attainment of children was added to the estimation of intergenerational elasticity of earnings, which resulted in a reduction in earnings elasticity

of 41% and 39% for fathers and mothers respectively at the mean. This goes to show the importance of parental income on the educational attainment of children and the significant effect education has on mobility in general.

This paper will investigate the intergenerational persistence of educational status in South Africa, and contribute to the existing literature on educational mobility using the most recent panel data available. The research considers both father-son and mother-daughter pairs over the last ten years and yields interesting results, displaying a clear increase in educational mobility in terms of the estimated regression and correlation coefficients for both father-son and mother-daughter pairs. However, decomposing this result into educational cohorts, the distribution of the increase in educational mobility is not experienced uniformly, with a more mobile education system predominantly falling on the children of parents with a high school level of education. Children whose parents had no education and those whose parents were educated at a tertiary level experienced increases in the persistence of educational status. These results have serious policy implications as the average level of education has increased, yet these increases have not been experienced equally and are dependent on family background.

The rest of the paper proceeds as follows. Section 2 will review the relevant literature. Section 3 will explain the methodological approach. Section 4 will explore the data and present a brief analysis of the descriptive statistics. Section 5 will present the empirical findings and conclusions will follow in Section 6.

LITERATURE REVIEW

The very nature of social mobility makes it an intriguing field of study, with defining metaphors like the American Dream being the foundation of our understanding of the concept. However, the theoretical analysis and measurement of intergenerational persistence comes with a significant degree of difficulty, owing to the lack of quality data.

In most papers, the measures of persistence arise from the standard procedure in which the estimated elasticity comes from least squares estimation of a log-linear regression of father's earnings on son's earnings, using age controls for both generations. Accommodations need to be made when recognising the measurement issues that a specific study is facing. If one does not make these accommodations, it will be unclear whether the differences between estimates

exist because of actual differences in intergenerational mobility, or because of differences in the earnings measures, age ranges or other sample selection criteria used.

The most common empirical specification of the relationship would be given by the following model:

$$Y_i^c = \alpha + \beta Y_i^p + \varepsilon_i$$

In this equation, the dependent variable Y_i^c is a logged long run earnings measure for the child (offspring), and the Y_i^p variable is the same for the parent. The estimated value of β is the measure of intergenerational elasticity (IGE), and represents the level of earnings persistence, with $(1 - \beta)$ representing intergenerational mobility, a value between 0 and 1. (Piraino, 2015) In a study by Solon (2002), cross-country differences in intergenerational earnings mobility were compared for a large set of countries. Empirical methodologies and data may have differed, so the direct comparison of IGE estimates for different countries may be difficult because of measurement issues. However, there is knowledge to be derived from trends in the literature. As stated in the paper, comparisons across countries may yield valuable clues about how income status is transmitted generation to generation, and why the level of persistence varies across countries. The paper listed a number of results from different research endeavours, and found more equal societies like Sweden and Finland displayed higher levels of mobility than other developed nations like the UK and the US. This fed into the interest of the broader connection between inequality and mobility. This association was not an exact mapping as Canada was a country that displayed relatively high levels of inequality, yet its measures of intergenerational mobility were considerable. There were few studies on developing nations, however the two that were present corroborated the conjecture that developing countries have high levels of intergenerational transmission.

This connection between inequality and mobility is an interesting branch in the mobility research. The emerging body of research before Corak (2013) seemed to have suggested that existing income inequality would make family background play a more important role in determining the adult outcomes of young people, relative to their own hard work. Using this notion as a springboard, Corak (2013) went on to define the inverse relationship between greater inequality and lower levels of income mobility as “The Great Gatsby Curve”. The different countries and their respective IGEs are placed on the curve and the relationship

observed. Corak notes this value does not consider the ‘inequality of opportunity,’ which is a measurement that aims to remove the influence of factors over which individuals have no control. Indices of inequality of opportunity are highly correlated with indicators of intergenerational mobility in terms of income and education. This makes the Great Gatsby curve a signal, rather than a direct measurement of the relationship between inequality and equality of opportunity. Inequality of opportunity is the missing link between the concepts of income inequality and social mobility; if high levels of inequality decrease mobility, it is probable that opportunities for economic advancement are more unequally distributed amongst children.

There is limited published research investigating inequality and social mobility in South Africa. In an early study, Hertz (2001) utilised data from just one region of South Africa, and as such the study is not representative of the country and was deemed unsuitable for generalized statements. The most comprehensive study for South Africa, calculating estimates for intergenerational mobility as well as an inequality of opportunity index for the country, was done by Piraino (2015). IGE ranged between 0.62 and 0.68 depending on the variables used to predict father’s earnings, and placed South Africa on the ‘Great Gatsby Curve’. Its placing was in line with the finding that countries with high levels of inequality tend to have lower levels of mobility. These estimates were comparable to those of other developing countries, with South Africa’s inequality of opportunity index found to be high relative to the other studied countries.

However, as mentioned earlier, education consistently presents as one of the most important channels in the intergenerational transmission of economic status. The 1953 Bantu Education Act centralized control of black education, and by 1975 the average expenditure on white education was 15 times higher than the average expenditure on their black counterparts (Thomas, 1996). There has been significant growth in educational attainment in South Africa in recent years. Comparing two cohorts separated by 30 years, Van der Berg (2007) shows not only that the mean educational attainment levels increased and the coefficient of variation declined, but also that the variance declined amongst blacks. Convergence both across and within race groups has improved, but the distribution of this benefit has not been uniform and the least educated were found to be those excluded in upward social mobility.

The effects of education on mobility are experienced by a lot of other countries as well, investigating these effects would better inform our understanding of South African educational mobility going forward, especially when looking at the experiences of other developing countries. A study by Azam and Bhatt (2015) in India serves as a good starting point, as India, in some respects, is similar to South Africa. It also has a history of group discrimination by the caste system which stratified society, allowing unequal access to opportunities in economic and social spheres. Studies on mobility were also limited, as quality longitudinal data was lacking, leaving studies to use cross-sectional data and co-residence to identify father-son pairs. This inevitably leads to large losses in observations and issues of sample selection bias. Azam and Bhatt (2015) used a representative survey which relaxed the need for co-residence of father-son pairs, and use educational attainment as a measure of economic status and examine its transmission over generations. The study used two popular measures of persistence, a correlation and regression coefficient.

The paper found a declining trend in intergenerational educational persistence when using the regression coefficient, implying greater mobility for more recent cohorts. This signifies that increases in attainment are driven primarily by children of less educated parents. This declining trend however is not corroborated when using the correlation coefficient as the mobility measurement. When broken down further, because the variance of father's education increased in relation to that of son's, the steady trend displayed by the correlation coefficient was a result of persistence decreasing at the low end of the distribution and increasing at the top of fathers education. Public policies have thus been effective in increasing educational mobility for the children of parents with the lowest levels of education. However, these policies struggle to move sons of the least educated to the highest levels of education with caste gaps still posing a challenge, even with affirmative action policies. The paper acknowledged that investigating the transmission amongst women would be an interesting avenue of research, however this was hindered by a lack of data.

A study in a developing country in which data of the educational attainment of up to four generations of a given family exists for Malaysia. (Lillard and Willis, 1994). This is another multiracial country in which the majority group had limited access in comparison to the two minority groups, with these gaps converging after the implementation of race-based policies, seeing an increase in educational attainment for all groups. Lillard and Willis (1994) found that parents' education significantly influenced that of their children, as is common for most

countries. Further, they found that the introduction of measured time varying economic, demographic and environmental factors decreased this intergenerational effect. The effects of family environments, measured by housing quality, school availability and urban residence, showed significant positive effects on schooling. In line with other literature, a mother's education had a bigger effect on daughters, with the same being true for fathers and sons. Interestingly, the research explored the competition for resources among same-sex siblings, with a girl's education declining as the number of sisters she had increased, and the same result for boys with additional brothers.

This last result was in line with the findings of Blake (1985), investigating the effect of the number of male siblings and their educational mobility in American society. The paper found a large interaction between a father's education and sibsize (the number of siblings one has) on son's schooling, implying the upward mobility experienced by men in general was almost entirely owing to the effect of men from small and medium sibsizes. This led to the overstating and underestimation of life chances of men from large and small sibsizes respectively. The paper emphasized the importance of considering the number of siblings and family size when analysing social stratification and mobility.

Chevalier, Denny & McMahon (2003) examined the educational mobility between fathers and their children for a panel of 20 countries. Using a variety of methodologies, the authors show a negative relationship between mobility and educational inequality, with mobility tending to be higher for men when compared to women. With regards to marginal effects, it found increased access to tertiary education does not necessarily imply increased mobility. A more recent panel study of mobility in Europe and North America by Blanden, Gregg and Machin (2005), focused on measuring the role of education in the process of mobility. The results show how the relationship between family income and children's higher education attainment has grown, and the increased university participation disproportionately benefitting those from affluent backgrounds. The authors suggest a causal relationship between income and educational attainment, highlighting the effects of family income and background on educational mobility.

The objective of this paper is to analyse the educational mobility in South Africa over time, but as the literature suggests, parents' educational attainment affects that of their children, and other factors such as household size and geography also play a role when furthering our

understanding of intergenerational persistence. Before the advent of the NIDS data set, South Africa did not have a representative panel data set, so when analysing the differences in educational mobility between black and white South Africans, Nimubona and Vencatachellum (2007) employed a pseudo-panel made up of five cross sectional October household surveys in an attempt to avoid suffering attrition, as was observed in previous non-representative data sets. The findings concluded that poor black children had the lowest levels of intergenerational educational mobility, contrasting with the result for poor whites. Perhaps even more surprisingly, the paper found mobility for females was greater than that of males, a result not shared by many developing countries.

Louw, Van Der Berg and Yu (2007) also investigated the issue of mobility in South Africa at about the same time. The study took a more qualitative analytical approach, showing a strong improvement in educational attainment for blacks over the period considered, decreasing the black-white attainment gap. Two-thirds of the progress in attainment however was ascribed as being because of improved educational mobility, but while these results were positive, race still posed a barrier when it came to higher levels of education; as blacks significantly underachieved upon matriculation.

The relationship between race and education garners little attention outside of the US and other developed countries, and South Africa, as a society with inequality as its inheritance, deserved such a study. Beutel and Anderson (2008) aimed to investigate the relationship between race and educational expectations. The study used data from the first (2002) and third (2005) waves of the Cape Area Panel Study (CAPS), a longitudinal study of youth and their families that is representative of metropolitan Cape Town to estimate this relationship, and the model used descriptive and multivariate OLS regression techniques. The results were interesting as educational expectations were high regardless of race, but highest for blacks when socioeconomic and other factors are controlled. Coloureds and whites are more similar than blacks regarding expectations of parents and children.

A more recent South African study on the intergenerational persistence of educational status is that of Burns and Keswell (2012). They used two waves of the KwaZulu Income Dynamics Study (KIDS) panel data set and examined changes in the educational attainment of three successive generations. Only the educational attainment of co-resident grandparents, parents and children were observed, so there were obvious selection issues as non-co-resident members

of any three-generation family were not observed. This was still an informative study, using a Galton estimation approach it showed successive generations have experienced increases in average educational attainment. These gains were however more significant for the second generation than the third, implying increased persistence for subsequent generations. The mothers' education began to affect the outcomes of children more than that of fathers over time, mainly in the second and third generation, suggesting educational mobility is lower for daughters than sons, a result contrasting with that of Nimubona and Vencatachellum (2007).

Leibbrandt et al (2016), probably the most recent study of mobility in South Africa, followed the methodology outlined in Piraino (2015) and went on to investigate why earnings inequality remained high over generations to add depth to the understanding of intergenerational earnings in the South African context. It focused on the role of education, as it is known for its effectiveness in reducing inequality, and selection into employment. While the paper did not delve into the issue of educational quality in South Africa, the results made use of two decompositions of the intergenerational earnings elasticity, and as mentioned above, it found that education accounted for 40% of the elasticity, playing a more important role than occupational skill in the understanding of earnings persistence in the country.

METHODOLOGICAL APPROACH

While the IGE estimation procedure, first proposed in Björklund & Jäntti (1997) and implemented in Piraino (2015), has become the standard technique when estimating intergenerational mobility of income, this paper aims to investigate intergenerational educational mobility in South Africa. As such, an OLS approach is sufficient as defined below:

$$education^c = \beta_0 + \beta_1 education^p + \beta_{2i} Controls_i + \varepsilon$$

This is a standard OLS regression in which the highest levels of educational attainment of parents is regressed on that of their respective children. β_1 Is therefore the parameter of interest, from which we are able to gauge the level of persistence of educational status. Controls, represents a vector of variables with which parents' education is expected to be correlated, namely: age, age squared, race, a rural/urban dummy and a household size variable.

In the regression excluding the controls, β_1 is given by:

$$\beta_1 = \rho_{cp} \frac{\sigma_c}{\sigma_p}$$

Where σ_c and σ_p are the standard deviations of the child and parents schooling, and ρ_{cp} is the correlation between them. As such, it would be easy to interpret a declining value of β_1 as reduced educational persistence over time, but this change could be a by-product of a decreased $\frac{\sigma_c}{\sigma_p}$ value. As a way to gain more information about this potential change, and understand what the cause of the change is, the years of schooling of children and parents will be normalized with these standard deviations, and the respective ρ_{cp} values will be estimated.

The higher the value of ρ_{cp} , the more children's educational attainment is influenced by that of their parents, and a measure closer to 0 would indicate that children's schooling is less affected by parents' schooling. The difference between the β and ρ measurements, is that the β value will be affected by the variances. It will reflect possible changes in the inequality of educational outcomes between the two generations. The ρ coefficient factors out the cross-sectional dispersion of educational attainment in the two generations. It is arguably a better measurement when considering intergenerational educational mobility conditional on the overall dispersion of educational attainment for each generation.

These measures may evolve differently, and papers such as (Hertz et al. 2007) have depicted this, hence our incorporation of both measures. This has become standard practice, shedding more light on the reasons for the changes in educational mobility measures. As these measures evolve, owing to changes in the dispersion in schooling between the different generations, Checchi et al. (2013) argued that the changes in ρ not only account for the evolution of educational transmission, but also secular changes in schooling and compulsory education. Denoting the schooling of children and parents with c and p , they rewrote the correlation coefficient as follows:

$$\hat{\rho} = \sum_c \sum_p \underbrace{(c - E(c))(p - E(p))}_A \underbrace{P(c | p)}_B \underbrace{P(p)}_C = \sum_c \sum_p r_{c,p}$$

Where $c, p = 0,1,2,3$. E implies expected value, and P refers to probability. Thus $\hat{\rho}$ for each cohort is the sum of 16 elements.

This shows how ρ can change over time because of the change in the variance of child-parent education around their respective means (term A), because of changes in children's educational attainment conditional on parent education (term B), or because of changes in the unconditional distribution of parents' education (term C). Term A could vary because of uniform convergence towards higher levels of education. Term C could vary because of institutional changes that happen as a country develops and increase the compulsory level of education for parents over time. Checchi et al. (2013) therefore suggests term B is the policy relevant indicator of intergenerational persistence, the distribution of children's education conditional on that of their parents.

DATA

The analysis will be based on the National Income Dynamics Study (NIDS), which represents a milestone in South African panel data records as the first national longitudinal study in South Africa. The first wave consisted of about 7300 households which were nationally representative with subsequent waves attempting to re-interview the original group. If movements within country borders occurred, efforts to track these households were made. Using both individual and household questionnaires, information on a vast selection of variables was obtained. This paper will focus on the adult questionnaire, which includes the necessary variables as mentioned in the vector of controls, education levels, employment information etc. The most salient question however was one in which adults were asked retrospective questions on parental background. It is upon this key socio-economic information that a variable for parents' education will be built, which is especially important if parents are deceased or non-residents. The education variable for parents will be built for the parents of those aged between 18 and 45 for waves one and five, which spans 10 years. During the apartheid regime, large portions of the population were excluded from official statistics, so in constructing this variable of parents' education from these retrospective questions, the most reliable answers would come from individuals within this age range, while preserving a sufficient sample size.

The information of the education for both children and parents is then converted into a numerical value for educational attainment i.e. years of education. The South African schooling system can be broken down into primary school (Grades 1-7), high school (Grades 8-12) and tertiary education, which in this study implies 13-18 years of education. In tertiary education, typical undergraduate degrees take 3 years, honours degrees take one year, while Masters and

PhD programs have been grouped as two additional years of study. Diplomas and higher certificates are assumed to add one additional year of study. Masters and PhD programs can often be acquired over longer periods of time, however this two year assumption better fits the available data when converting educational attainment into years of education, and makes for easier comparison.

The race variable considered will partition the population into the standard four racial categories, African, Coloured, White and a collective Indian/Asian group. The urban dummy was constructed with both traditional and farm lands being classified as rural. A variable for average household size is considered, as literature such as that of Black, Devereux & Salvanes (2005) indicate negative effects of household size on the educational attainment of children. The empirical analysis will consider the intergenerational mobility of education for father-son and mother-daughter pairs. This approach is in line with what has been done in previous studies, especially that of Piraino (2015), conducted on South African data, as it has been the procedural norm to avoid investigating the potential effects of gender on mobility. After restricting the age of respondents, creating a variable for parental education and restricting the analytical sample to those with information on the control variables for men and women, it can be observed that the number of observations over the two waves varies, and the resulting panel is unbalanced.

Table 1 below reports some descriptive statistics for the analytical sample. Firstly, considering father-son pairs this sample consists of 2034 and 2808 observations in wave 1 and wave 5 respectively. Due to the age restriction imposed on the sample, the mean age is between 30 and 31 years which is relatively young. This means a son's education is generally observed after most formal schooling is completed. In income mobility studies, this is within the range that would reduce the effects of lifecycle bias. Considering the four race categories generally used in South African studies, the analytical sample in both wave 1 and 5 over-represents blacks and under-represents whites. This is especially true in wave 5, where in the overall sample whites aged 18-64 make up 8.2% of the weighted proportion, yet in the analytical sample they only make up 4.5% of the population. Much the same with blacks, initially representing 74.7% of the sample, and this transforming into 82.8% after selection.

The dummy variable for urban/rural also changes when the sample is restricted to respondents between the ages of 18-45, and the analytical sample is obtained. In wave 5 the sample initially

has 62.2% of the population being classified as living in an urban area, this decreases to 57.6. While this still leaves urban as the majority, this is not true for wave 1, in which urban initially refers to 55.4% of the sample and decreases to 48.6% after selection, leaving the dummy variable unrepresentative of the sample after selection. Mean household size also decreases for both waves of father-son pairs after restriction, however this difference is very small when considering the size of one standard deviation of the original household size.

The education variable considered is broken down into the following categories: (0) No schooling, (1) Primary school, (2) High school, and (3) Tertiary education. The proportion of sons in our sample that have no schooling is low, at 7.6% and decreases rather significantly to 1.5% over the ten-year period considered. Most sons report at least a high school level education, a value within the range of 8-12 years, and the proportion of those reporting a tertiary level of education more than doubles over the period.

This is reasonably representative of the original sample for wave 1, but in wave 5, the analytical sample slightly under-represents those with no schooling and increases the proportion of those reporting a high school level education. In contrast, most fathers have no schooling, or report some primary school level of attainment over the 10-year period, and in both waves the value of no schooling is under-represented. The vast increase in the levels of educational attainment of sons compared to their fathers is in line with the expectations of the literature, and the documented increase in education levels after the end of apartheid. Child mean education refers to the educational attainment levels of the respondents in the analytical sample. Those men and women between the ages of 18-45 who gave a response to the retrospective questions on parental educational attainment. The mean education for sons increased by more than a year over just a 10-year period. It must however be noted that of those sons or fathers who have no schooling in either wave, a disproportionate amount of them are black, while those reporting tertiary education are represented by a far higher percentage of whites than the percentage they represent in the analytical sample.

When mother-daughter pairs are considered, much the same patterns are observed. The analytical sample has 2799 and 3833 observations for wave 1 and 5 respectively when restricting age, with average age over the period considered around the 31-32-year area, also considering a relatively young female sample.

Table 1: *Descriptive statistics and sample proportions for adults aged 18-45 with parental educational attainment information*

SAMPLE RESTRICTION	MEN		WOMEN	
	WAVE 1	WAVE 5	WAVE 1	WAVE 5
MEAN AGE	30.9	30	32.2	31.6
CHILD MEAN EDUCATION	9.1	10.7	9.1	10.7
CHILD EDUCATION STANDARD DEVIATION	3.75	2.59	3.75	2.73
EDUCATIONAL CATEGORY:				
NO SCHOOLING	7.6	1.5	6.6	1.7
PRIMARY SCHOOL	18.8	7.8	21.2	8.1
HIGH SCHOOL	65.8	74.1	63.6	70.1
TERTIARY	7.8	16.6	8.6	19.9
RACE PROPORTIONS:				
AFRICAN	82.6	82.8	82.2	84.6
COLOURED	10.1	11	10.7	10.3
INDIAN/ASIAN	0.9	1.7	1.4	1.6
WHITE	6.4	4.5	5.7	3.5
PROPORTION URBAN	48.6	57.6	50.3	57
PROPORTION RURAL	51.4	42.4	49.7	43
MEAN HOUSEHOLD SIZE	4.1	4.4	4.7	4.7
PARENTS' MEAN EDUCATION	4.52	5.83	4.25	5.61
PARENTS' EDUCATION STANDARD DEVIATION	4.77	5.12	4.5	4.78
PARENTS' EDUCATIONAL CATEGORY:				
NO SCHOOLING	44.9	35.1	44	32.4
PRIMARY SCHOOL	22.9	21.1	26.1	26.6
HIGH SCHOOL	28.9	38.1	27	35.6
TERTIARY	3.3	5.7	2.9	5.4
NUMBER OF OBSERVATIONS	2034	2808	2799	3833

The race categories slightly over-represent black women and under-represent whites, and this is again more prevalent in wave 5 in which the percentage of whites decreases from 6.8% to 3.5% in the analytical sample.

The urban/rural dummy variable stays rather constant throughout the restriction process with very little change in the proportions, and urban remains the category with the larger proportion of people. Mean household size decreases for both waves, but again this decrease seems negligible when considering its magnitude in relation to the variance of household sizes.

The education variable is classified in the same way for mother-daughter pairs as for sons and their fathers. Those reporting no education is low, and this significantly decreases over the 10 year period from 6.6-1.7%. The majority of daughters report a high school level of educational attainment and those reporting tertiary levels also increased over the 10 year period with the same distribution of results, blacks disproportionately represented in the category of no schooling and whites in the tertiary cohort. When contrasting this distribution of educational attainment with that of their mothers, much like with father-son pairs, significant increases are observed. Most mothers have no schooling in both waves, and those mothers with either no schooling or only a primary school level represent an almost overwhelming proportion of the sample. This increase in attainment level was expected, and much like that of father-son pairs is in line with the expectations of literature and documented changes in attainment levels in South Africa.

We also note the way variance/standard deviation in education has changed over time for that of children and their parents. Decreasing by 1.16 and 1.02 years for sons and daughters respectively over the two waves, and the variances slightly increasing for both fathers and mothers over the period. This is to be expected for children as policy and regulation changed significantly over this period, increasing the general level of educational attainment and decreasing the variance. For adults, this result seems to be owing to disproportionately large shifts on the far ends of the distribution, with seemingly large increases in those attaining high school and tertiary levels, but not a very large shift in the mean. Large proportions of adults have no education, so variance is expected to increase when a minority gain access. The change seems to align with a tighter clumping around the mean with the respective category shifts, for children.

EMPIRICAL ANALYSIS

Table 2 below shows the estimation results of intergenerational educational mobility in South Africa. These estimates were derived using the estimation procedure outlined above, and repeated over the two waves considered, spanning a ten year period. The first method of estimation uses only the parents' education as a predictor of children's education. This is done for both father-son pairs and mother-daughter pairs over both waves. The second method, includes controls that, according to the literature, significantly affect a child's likelihood of increasing their level of educational attainment, like an urban/rural dummy and race.

What we observe from the output when only considering a father's education as the predictor of son's education, as seen in column 1, is an estimated intergenerational educational persistence of 0.339. This is relatively high, and suggests a father's education has both an economic and statically significant association with the educational attainment of sons. Over only a ten-year period, when these values are estimated for wave 5 of NIDS, this effect, while still being significant has reduced to a 0.17-year difference for every additional year of a father's education as seen in column 3. Considering the confidence intervals on the two coefficients, there is no overlap between the two coefficients implying a likely significant change.

When controls which have been known to influence children's education and cause bias in the estimated effect of fathers' education are included, the effect of fathers' education on that of their sons is reduced. The effect in wave 1 is now 0.24 and is 0.143 in wave 5 as shown in columns 2 and 4 respectively. Household size surprisingly has a small positive effect; however, this effect is insignificant. This effect changes in wave 5 to a significantly negative one, which is more in line with expectations. Living in an urban area has a significantly positive effect, increasing the average attainment for children 1.3 years in wave one, this decreases in wave 5 but its effect remains significant. The effect of age on educational attainment increases at a decreasing rate, with turning points at 24.5 and 30.9 years for wave 1 and 5 respectively, in line with the increases in educational attainment over those 10 years. When considering race, coloured is the only race group that signals lower levels of educational attainment when compared to blacks for both waves, and this becomes significant in wave 5. The effect for whites and Indian/Asians seems to be in line with expectations, with both groups having significantly higher levels for every year when compared to blacks. However, while these

values remain positive and significant in wave 5, they are lower. This speaks to the closing educational gap between blacks and whites post-apartheid.

Considering mother-daughter pairs, we note much the same pattern as for those of fathers and sons. When using only a mother's education as a predictor of daughters' potential attainment, the result is an estimated intergenerational educational persistence of 0.4 seen in column 5. This is quite a large value and is highly significant which implies for every additional year mothers are educated, daughters are expected to increase their schooling by 0.4 years. When this same estimation procedure is done ten years later, it reduces to 0.22 as seen in column 7. This suggests a decrease in educational persistence in South Africa for mothers and daughters over the period considered, and as the confidence intervals of the two coefficients do not overlap, a likely significant change as well.

After controlling for the selected variables of importance, in wave 1, there is a significant decrease, as these variables were causing an upwards bias of the effect of parents' education. The new estimate in column 6 falls to 0.3. Living in an urban area as opposed to rural, has a more positive effect on education as expected by over a year on average, and an extra person in the household significantly decreases daughter's years of education. Age's effect on education increases at a decreasing rate with a turning point at 22.73 years. With regards to race, being coloured is again associated with lower attainment compared to blacks, almost an entire year less for every year of black education, while Indian/Asians and Whites have significantly greater levels compared to blacks. With Indian/Asians having close to 2 years extra for every additional year for black child. Running the same regression on the data of wave 5, the persistence estimate drops to 0.18 from 0.22 in wave 5. This is significantly lower than the estimates from wave 1 as seen in column 8, when considering confidence intervals on coefficients.

This implies increased educational mobility in South Africa over the ten-year period which is in line with the signals of the descriptive statistics. As before, living in an urban area has a more positive effect on average attainment, and an extra member in the household leads to decreases holding everything constant. The effect of age increases at a decreasing rate having a turning point at 29.54. This result is also in line with expectations, as there have been documented increases in the years schooling over this period.

Table 2: Estimation results from regression of parents' educational attainment (in years education) on that of children.

SAMPLE RESTRICTION	MEN				WOMEN			
	WAVE 1 Fathers' education (1)	WAVE 1 Fathers' education & controls (2)	WAVE 5 Fathers' education (3)	WAVE 5 Fathers' education & controls (4)	WAVE 1 Mothers' education (5)	WAVE 1 Mothers' education & controls (6)	WAVE 5 Mothers' education (7)	WAVE 5 Mothers' education & controls (8)
FATHERS' EDUCATION	0.339*** (0.0157)	0.24*** (0.0181)	0.17*** (0.0089)	0.143*** (0.0099)				
RACE: COLOURED		-0.286 (0.253)		-0.925*** (0.149)		-0.971*** (0.199)		-0.779*** (0.134)
ASIAN/INDIAN		1.96*** (0.76)		0.899** (0.352)		1.748*** (0.504)		0.587* (0.32)
WHITE		1.482*** (0.335)		1.111*** (0.232)		0.789*** (0.29)		0.847*** (0.228)
URBAN		1.32*** (0.162)		0.784*** (0.0996)		1.373*** (0.131)		0.666*** (0.0867)
HOUSEHOLD SIZE		0.0287 (0.0248)		-0.027* (0.0147)		-0.0556** (0.0223)		-0.077*** (0.0136)
AGE		0.276*** (0.0761)		0.262*** (0.0488)		0.215*** (0.0662)		0.351*** (0.0439)
AGE SQUARED		-0.0056*** (0.00121)		-0.0042*** (0.0008)		-0.0047*** (0.001)		-0.0059*** (0.00069)
MOTHERS' EDUCATION					0.399*** (0.0138)	0.297*** (0.0157)	0.221*** (0.00853)	0.18*** (0.00935)
CONSTANT	7.559*** (0.103)	4.401*** (1.165)	9.681*** (0.0697)	5.758*** (0.748)	7.394*** (0.0857)	5.722*** (1.028)	9.433*** (0.0628)	4.883*** (0.688)
OBSERVATIONS	2034	2034	2808	2808	2799	2799	3833	3833

Standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Fathers' education is regressed on that of their sons, and mothers' education on their daughters for both wave 1 and 5, with and without the vector of controls. As some of the controls were categorical variables, the reference category is left out. For the race category, it is the black group and for the urban/rural variable, the urban coefficient is calculated in comparison to living in a rural area.

Regarding race however, all race groups seem to have an effect smaller than in wave one, including coloureds, while still having a negative sign.

These results are in line with the expectations of the descriptive statistics. The mean levels of education increased for both men and women, and the turning points for age increased for both as well, implying longer periods of attainment. When trying to understand the root cause for this decline in educational persistence, as mentioned in the methodology, the evolution of educational variance for both parents and children must be investigated. The measure to be used is the correlation coefficient defined above, as this measure factors out the potential trend in the variance of schooling. Because of changes in the variances of educational attainment for children and parents, the regression coefficients and correlation coefficients are bound to evolve slightly differently. When there are large discrepancies in their behaviour however, it may be indicative of an overriding change in relative variance, rather than an actual decrease in educational persistence.

Table 3: Regression and correlation coefficients in Wave 1 and 5.

SAMPLE RESTRICTION	MEN		WOMEN	
	WAVE 1	WAVE 5	WAVE 1	WAVE 5
PARENTS' SCHOOLING $\hat{\beta}$	0.339	0.17	0.4	0.22
PARENTS' SCHOOLING $\hat{\rho}$	0.431	0.336	0.48	0.385
SD IN CHILDS YEARS OF SCHOOLING	3.75	2.59	3.75	2.73
SD IN PARENTS' YEARS OF SCHOOLING	4.77	5.12	4.5	4.77

The decline in persistence is quite pronounced when analysing the regression coefficient, in both the father-son and mother-daughter case. This decline appears to be less clear when considering the correlation coefficients in Table 3 above, which appears to slightly decline for men, with a somewhat larger decrease for women. The ratio of standard deviations of parents and children is greater than 1 for all waves, leading to correlation coefficients greater than all estimated regression coefficients, with increasing levels of variance over time for both fathers and mothers. We would like to further investigate this decrease in the correlation coefficient

and uncover whether this change is uniformly due to all groups of children, regardless of their parents' education.

Table 4 shows the elements $r_{c,p}$ of the decomposition formula for waves 1 and 5 for father-son and mother-daughter pairs. The formula takes the estimated correlation coefficient, and further breaks it down into its contributing effects, to analyse which elements contribute the largest relative effect with regards to the effect of parents' educational attainment on that of their children. This will be done using the decomposition formula above in which the categories 0, 1, 2 and 3, represent no schooling, primary schooling, high school and tertiary levels of educational attainment respectively, (NE, PS, HS, TE). Line 21 reports the correlation coefficients, which is the sum of the absolute value contributions of each combination of child and parents' education and their relative contributions. This value represents the full 100% of the correlation coefficient. Considering males, table 3 shows a decrease in this value from 0.431 to 0.336 implying decreased persistence in educational status over time. A decomposition of these values will help illuminate whether this change is uniform for all cohorts; i.e. if educational status is now more mobile for all. Line 5 shows the total contribution of the group of sons whose fathers have no education. Over time this group has accounted for a large portion of the correlation coefficient, contributing an increase from 45% to 49% of the total correlation coefficient over the ten-year period considered.

The contribution from those whose fathers have primary school is negligible, and the contribution of those that have fathers with a high school education decreases from 39-20%. This decrease however is made up by a significant increase in the contribution from the group of sons with tertiary educated fathers, increasing its effect from 15-30% of the total value of the correlation coefficient over the ten-year period. Therefore, the decrease in the correlation coefficient over the ten-year period considered is primarily driven and experienced by the group of sons whose fathers have a high school level (Grade 8-12) of education. This means that sons whose fathers have a high school level of attainment, are more likely to achieve levels of educational attainment higher than their fathers, than the other cohorts. Mobility is being driven by this group.

Table 4: Decomposition of ρ as described in the methodology above.

LINE	COMBINATIONS	MEN		WOMEN	
		WAVE 1	WAVE 5	WAVE 1	WAVE 5
1	NE, PNE	0.167	0.068	0.187	0.083
2	PS, PNE	0.141	0.131	0.157	0.139
3	HS, PNE	-0.092	0.046	-0.102	0.04
4	TE, PNE	-0.024	-0,082	-0.028	-0.08
5	Total Contribution to the Correlation Coefficient of the Group of children with PNE	0.192 (45%)	0.163 (49%)	0.214 (45%)	0.182 (47%)
6	NE, PPS	-0.001	0.001	-0.001	0.000
7	PS, PPS	-0.004	0.007	-0.005	0.006
8	HS, PPS	0.007	0.004	0.008	0.005
9	TE, PPS	0.002	-0.009	0.002	-0.011
10	Total Contribution to the Correlation Coefficient of the Group of children with PPS	0.004 (1%)	0.003 (1%)	0.004 (1%)	0.000 (0%)
11	NE, PHS	-0.009	-0.003	-0.01	0.000
12	PS, PHS	-0.026	-0.022	-0.029	-0.015
13	HS, PHS	0.102	-0.039	0.113	-0.037
14	TE, PHS	0.102	0.133	0.114	0.145
15	Total Contribution to the Correlation Coefficient of the Group of children with PHS	0.169 (39%)	0.069 (20%)	0.188 (39%)	0.093 (24%)
16	NE, PTE	0.000	0.000	0.000	0.000
17	PS, PTE	0.000	0.000	0.000	0.000
18	HS, PTE	0.019	-0.008	0.022	-0.006
19	TE, PTE	0.047	0.11	0.052	0.116
20	Total Contribution to the Correlation Coefficient of the Group of children with PTE	0.066 (15%)	0.102 (30%)	0.074 (15%)	0.11 (29%)
21	Correlation coefficient $\hat{\rho}$	0.431 (100%)	0.336 (100%)	0.48 (100%)	0.385 (100%)

Line 5, when considering the mother-daughter cohort, shows the total contribution of the group of daughters of mothers with no education, and over time this group accounts for a larger part of the total correlation, increasing from 45-47% of the full value over the two waves. Line 10 shows the proportional contribution of the group of daughters with primary educated mothers being negligible, which contrasts with the large effect of the group with mothers that have been either educated at a high school or tertiary level. The contributing effect of those with high school parents drops over time from 39-24%, but of those with tertiary educated mothers, the proportional contribution increases significantly from 15-29% of the full value of the correlation coefficient. Again, we find the decrease being driven by the group with mothers who have attained a high school level of education. The interpretation of the effects of this group on mobility can be interpreted the same as for the men above.

This table shows how highly polarized the intergenerational transmission of education still is in South Africa, despite the economic growth, large strides in educational attainment between generations and educational reforms. As Checchi et al (2013) linked educational attainment and socio-economic conditions, we can see from the above that children growing up in the most disadvantaged families are increasingly likely to remain disadvantaged. While the children from more affluent backgrounds are increasingly likely to retain this privilege and their relative advantages.

While these results speak to the persistence of educational status, they do not explicitly consider the quality of that education, especially at the lower end. The literature (e.g. Leibbrandt et al 2016), identify poor education quality as being a poverty trap. In addition, school dropout in South Africa is driven by students falling behind (i.e. being more than two years older than expected for the current grade). Falling behind is largely determined by school quality and historically black schools are particularly affected in this regard. If children's and parents' ability, as well as school quality were variables in the dataset, that would allow for much clearer insights into the factors contributing to intergenerational persistence, especially at the lower end of the distributions of wealth and education.

CONCLUSION

This paper has used the NIDS dataset to measure the intergenerational mobility of education, over a ten-year period in South Africa. Intergenerational educational persistence, as measured by the regression coefficient of parents' education as a predictor of educational attainment of children, showed large decreases for both father-son and mother-daughter pairs over the generation considered. There was also evidence of a convergence in population group effects, implying race played a smaller role explaining educational status over the period considered, other than for coloured men, and white women whose effects seemed to increase over time. This finding is in line with the literature which speaks to general increases in educational attainment in South Africa in recent years. The descriptive statistics have shown increases in the average levels of education over the period for children, and a large difference between average attainment for parents and children. The convergence in population group effects over the period is also in line with the literature, suggesting differences in educational attainment between groups has decreased.

The estimated correlation coefficient between children's and parents' educational attainment levels mimicked this decreasing movement over time. Based on the decomposition of the intergenerational correlation coefficient, we find that at the lower end of the distribution of both parents' education there seems to be an increase in the proportional contribution this group has on the correlation coefficient. On the higher end of the educational distribution, the contributing effect of tertiary educated parents on the correlation coefficient has increased for both males and females over the same period as well. This implies that while there has been a net decrease in the persistence of educational status in South Africa, this decrease is not uniform and is driven mainly by the increase in mobility of children born to parents with high school levels of educational attainment. Not all cohorts experience this increase in mobility, with an offsetting increase in persistence for the poorest and richest in the sample.

These results have significant policy implications. This lack of change in persistence at the lower end of the distribution of parental education indicates policy has done very little to compensate for the lack of educational inputs from parents. While education levels have generally improved, average education increased in all groups and primary school enrollment is at their highest levels, progressing to the highest levels of educational attainment is still highly dependent on parental educational attainment and family background in South Africa.

BIBLIOGRAPHY

Azam, M. and Bhatt, V., 2015. Like father, like son? Intergenerational educational mobility in India. *Demography*, 52(6), pp.1929-1959.

Beutel, A. & Anderson, K., 2008. RACE AND THE EDUCATIONAL EXPECTATIONS OF PARENTS AND CHILDREN: The Case of South Africa. *The Sociological Quarterly*, Issue 0038-0253, pp. 335-361.

Black, S.E., Devereux, P.J. and Salvanes, K.G., 2005. The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, 120(2), pp.669-700.

Blake, J., 1985. Number of siblings and educational mobility. *American Sociological Review*, pp.84-94.

Blanden, J., Gregg, P. and Machin, S., 2005. Intergenerational mobility in Europe and North America. *Report supported by the Sutton Trust, Centre for Economic Performance, London School of Economics*.

Burns, J. & Keswell, M., 2012. INHERITING THE FUTURE: INTERGENERATIONAL PERSISTENCE OF EDUCATIONAL STATUS IN KWAZULU-NATAL, SOUTH AFRICA. *Economic History of Developing Regions*, 27(1), pp. 150-175

Chevalier, A., Denny, K. and McMahon, D., 2003. *A multi-country study of inter-generational educational mobility* (No. 03/14). Working Paper Series, UCD Centre for Economic Research.

Checchi, D., Fiorio, C.V. and Leonardi, M., 2013. Intergenerational persistence of educational attainment in Italy. *Economics Letters*, 118(1), pp.229-232.

Corak, M., 2013. Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives*, 27(3), pp.79-102.

Finn, A., Leibbrandt, M., Ranchhod, V. (2016). Patterns of persistence: Intergenerational mobility and education in South Africa. Cape Town: SALDRU, University of Cape Town. SALDRU Working Paper Number 175/ NIDS Discussion Paper 2016/2

Hertz, et al 2007. The inheritance of educational inequality: International comparisons and fifty-year trends. *The BE Journal of Economic Analysis & Policy*, 7(2).

Lillard, L.A. and Willis, R.J., 1994. Intergenerational educational mobility: Effects of family and state in Malaysia. *Journal of Human Resources*, pp.1126-1166.

Louw, M., Van der Berg, S. and Yu, D., 2006. Educational attainment and intergenerational social mobility in South Africa. University of Stellenbosch, Stellenbosch Economic Working Paper, 9(06), pp.1-43

Nimubona, A.D. and Vencatachellum, D., 2007. Intergenerational education mobility of black and white South Africans. *Journal of Population Economics*, 20(1), pp.149-182.

Piraino, P., 2015. Intergenerational earnings mobility and equality of opportunity in South Africa. *World Development*, Volume 67, pp. 396-405.

Solon, G., 2002. Cross-country differences in intergenerational earnings mobility. *Journal of Economic Perspectives*, 16(3), pp.59-66.

Thomas, D., 1996. Education Across Generations in South Africa. *The American Economic Review*, 86(2), pp. 330-334.

Van der Berg, S., 2007. Apartheid's enduring legacy: Inequalities in education. *Journal of African economies*, 16(5), pp.849-880.

APPENDIX

Table 5: Sample sizes and descriptive statistics associated with the creation of the analytical sample of males from the National Income Dynamics Study for Waves 1 and 5.

SAMPLE RESTRICTION	WAVE 1				WAVE 5			
	Age 18-64	Age 18-45	Education information	Non-missing father's education	Age 18-64	Age 18-45	Education information	Non-missing father's education
MEAN AGE	35.8	29.3	29.3	30.9	35.5	29.3	29.1	30
CHILD MEAN EDUCATION	8.6	9.3	9.3	9.1	10	10.5	10.5	10.7
EDUCATIONAL CATEGORY:								
NO SCHOOLING	9.3	5.1	5.1	7.6	3.8	1.5	1.5	1.5
PRIMARY SCHOOL	22.9	18.4	18.4	18.8	13.2	8.7	8.7	7.8
HIGH SCHOOL	61.5	70.4	70.4	65.8	67.8	75.5	75.5	74.1
TERTIARY	6.3	6.1	6.1	7.8	15.2	14.3	14.3	16.6
RACE PROPORTIONS:								
AFRICAN	74.9	77.4	77.5	82.6	74.7	78.4	80.2	82.8
COLOURED	16.4	16	15.9	10.1	14.6	13.8	13.7	11
INDIAN/ASIAN	1.8	1.5	1.5	0.9	2.5	2.2	1.9	1.7
WHITE	6.9	5.1	5.1	6.4	8.2	5.6	4.2	4.5
PROPORTION URBAN	55.4	5.3	54.2	48.6	62.2	60.9	59.4	57.6
PROPORTION RURAL	44.6	45.7	45.8	51.4	37.8	39.1	40.6	42.4
MEAN HOUSEHOLD SIZE	4.7	4.9	4.9	4.1	4.8	4.9	4.9	4.4
PARENTS' MEAN EDUCATION	3.9	4.5	4.5	4.5	4.9	5.8	5.8	5.8
PARENTS' EDUCATION STANDARD DEVIATION	4.7	4.8	4.8	4.8	5.1	5.1	5.1	5.1
PARENTS' EDUCATIONAL CATEGORY:								
NO SCHOOLING	52.1	44.9	45	44.9	44.5	35	35	35.1
PRIMARY SCHOOL	19.5	22.8	22.8	22.9	18.1	21.1	21.1	21.1
HIGH SCHOOL	25.1	28.9	28.9	28.9	32.3	38.1	38.1	38.1
TERTIARY	3.3	3.4	3.3	3.3	5.1	5.8	5.8	5.7
NUMBER OF OBSERVATIONS	6599	4889	4850	2031	11814	8891	7624	2808

Table 6: Sample sizes and descriptive statistics associated with the creation of the analytical sample of females from the National Income Dynamics Study for Waves 1 and 5.

SAMPLE RESTRICTION	WAVE 1				WAVE 5			
	Age 18-64	Age 18-45	Education information	Non-missing mothers' education	Age 18-64	Age 18-45	Education information	Non-missing mothers' education
MEAN AGE	37	29.9	29.9	32.2	37.2	29.8	29.8	31.6
CHILD MEAN EDUCATION	8.3	9.5	9.5	9.1	9.7	10.7	10.7	10.7
EDUCATIONAL CATEGORY:								
NO SCHOOLING	11.1	4.5	4.5	6.6	5.6	1.6	1.6	1.9
PRIMARY SCHOOL	23.6	17.7	17.7	21.2	15	7.5	7.5	8.1
HIGH SCHOOL	58.5	70.6	70.6	63.6	62.4	73.2	73.2	70.1
TERTIARY	6.8	7.2	7.2	8.6	17	17.7	17.7	19.9
RACE PROPORTIONS:								
AFRICAN	76.8	78.9	79	82.2	77	79.7	81.5	84.6
COLOURED	15.6	15.1	15	10.7	14	13.6	13.1	10.3
INDIAN/ASIAN	1.8	1.6	1.6	1.4	2.2	1.9	1.7	1.6
WHITE	5.8	4.4	4.4	5.7	6.8	4.8	3.7	3.5
PROPORTION URBAN	51.9	52.2	52.2	50.3	57.6	57.7	56.2	57
PROPORTION RURAL	48.1	47.8	47.8	49.7	42.4	42.3	43.8	43
MEAN HOUSEHOLD SIZE	5.3	5.5	5.5	4.7	5.4	5.6	5.6	4.7
PARENTS' MEAN EDUCATION	3.5	4.3	4.3	4.3	4.4	5.6	5.6	5.6
PARENTS' EDUCATION STANDARD DEVIATION	4.4	4.5	4.5	4.5	4.7	4.8	4.8	4.8
PARENTS' EDUCATIONAL CATEGORY:								
NO SCHOOLING	53.6	44	43.9	44	45.7	32.4	32.4	32.4
PRIMARY SCHOOL	22.1	26.2	26.2	26.1	22.8	26.6	26.6	26.6
HIGH SCHOOL	22	26.9	27	27	27.3	35.6	35.6	35.6
TERTIARY	2.3	2.9	2.9	2.9	4.2	5.4	5.4	5.4
NUMBER OF OBSERVATIONS	8383	5935	5912	2791	14713	10396	9133	3833