Intergenerational Mobility in South Africa: How much has South Africa Changed since the Early 1990s?

Thesis

Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the

School of Economics

Faculty of Commerce

UNIVERSITY OF CAPE TOWN

By: Hagos Fesshaye Ressom

Supervisors: Prof. Martin Wittenberg and Assoc. Prof. Nicola Branson

AUGUST 2024

Abstract

South Africa has been grappling with persistently high rates of poverty, unemployment, and inequality. These unfortunate realities have their roots in the country's colonial and apartheid past, the legacy of which continues to influence the post-apartheid socioeconomic outcomes. This thesis examines changes over time in the relationship between the education and earnings of South African parents and their children once these are adults. The study uses two independent surveys, one collected in the early 1990s and the other in 2017. The timing of the two surveys enables for an examination of changes in mobility during the first 25 years of democracy. The study contributes to the South African literature on inequality by enriching and extending the cross-sectional empirical evidence on trends in educational and earnings inequality to the intergenerational dimension.

The study shows that while intergenerational education mobility has increased considerably, earnings persistence has remained high and almost unchanged between the two periods. Intergenerational relative education mobility, as measured by the intergenerational correlation coefficient, decreased from 0.72 in the first period to 0.43 in the second. On the other hand, intergenerational earnings persistence remained high, with the elasticity estimated to be 0.73 in the first period and 0.70 in the second (among currently employed). These results demonstrate that the well-documented cross-sectional trends in educational and earnings inequality in the post-apartheid period are paralleled in the intergenerational dimension. The study further highlights the continued relevance of parental education to earnings, with the effect of parental education on children's earnings shown to be significant and positive in both periods. The contribution of parental education to children's earnings inequality, however, has been small (relative to the contribution of own education) in both periods.

The thesis concludes by emphasizing on the contrasting trends in intergenerational education and earnings inequality in the post-apartheid period and recommends that fixing the quality of education could go a long way in addressing the persistently high levels of intergenerational earnings persistence.

Declaration

I, *Hagos Fesshaye Ressom*, declare that this thesis is my own original work and that other sources have been acknowledged through referencing. I further declare that the thesis has not been submitted for the award of a PhD degree at any other university.

Signed: Signed by candidate

Date: August 14, 2024

Acknowledgement

I would like to extend my special and sincere gratitude to my supervisors, Prof. Martin Wittenberg and Assoc. Prof. Nicola Branson, for their superb guidance and encouragement. Their thoughtful mentorship has not only kept me engaged with my studies but, more importantly, taught me to think critically. They have also been instrumental in shaping the thesis.

In 2020, when COVID struck and the University had to close its student residences, Professor Martin Wittenberg invited me to stay at his guest house. I am very grateful and would like to thank the Professor and his family for their support and kindness.

I am extremely thankful to the African Economic Research Consortium (AERC) for funding my studies. My doctoral studies would not have been possible without the support of AERC. I also acknowledge additional financial support from UCT Postgraduate funding office and am grateful. I am also grateful to Assoc. Prof. Nicola Branson and Prof. Reza Daniels for giving me the opportunity to work as a research assistant in their research projects while waiting for the results of the examination of the thesis and graduation.

I am thankful to my friends, in particular, Hisseine Bahar, Said Musa, Tedros Hagos, Obrian Ndhlovu, Solomon Aboagye and Zandile Ndebele for their friendship and companionship.

Finally, my special thanks go to my family. To my parents, Fesshaye Ressom and Letom Yemane, for their constant support and encouragement. To my wife, for her unconditional love and understanding during this long absence from home. To my sisters and brothers, for always checking on and encouraging me.

Dedication

To my lovely wife, Sinit Emun, and our son, Yotam.

Table of Contents

List of Figures

List of Tables

Chapter 1

1. Introduction

1.1 Motivation

Intergenerational mobility is concerned with the association between the socioeconomic outcomes of parents and their children as adults. The stronger the association in socioeconomic outcomes across generations, the lower the intergenerational mobility (and the higher the intergenerational persistence). Empirical studies of intergenerational mobility across countries show that the socioeconomic outcomes of parents and their children are positively correlated (e.g., Solon, 1999; Hertz *et al.*, 2008; Blanden, 2009; Corak, 2016). Although it is not possible - nor is it necessarily desirable - to eliminate the correlation between the socioeconomic outcomes of parents and their children, low levels of intergenerational mobility are a cause for concern. If the socioeconomic outcomes of individuals are strongly correlated with their parental background, then the life prospects of individuals are unlikely to be different from that of their parents. This means children from poor parents may not have the opportunity to fully realize their human capital potential, which is a loss to society. Moreover, low levels of intergenerational mobility only serve in perpetuating deep-seated inequality.

This thesis is concerned with changes over time in intergenerational mobility in South Africa. South Africa, with its complex colonial and apartheid history, presents an interesting case. The country has been grappling with persistently high rates of poverty, unemployment, and inequality. These unfortunate realities have their roots in the country's past, the legacy of which continues to influence the post-apartheid socioeconomic outcomes.

The pre-1994 history of South Africa is marked by colonialism, racial segregation, and apartheid (Feinstein, 2005). The system of governance was heavily biased in favour of the white minority at

the expense of the black[1] majority.[2] Blacks were marginalized in all aspects of life and discriminated against from having equal access to opportunities. They were restricted in terms of where they could live, and their mobility was regulated. Their employment opportunities were limited and systematically controlled. Moreover, as we will discuss in chapter 2, the education of blacks was poorly resourced and inferior in quality.

With the fall of apartheid and transition to democracy in 1994, South Africa set out to address the previous course of inequalities in opportunities (Government of South Africa, 1994). Various policies have been enacted to expand access to services and equalize opportunities. However, despite these initiatives, the country continues to be challenged by persistently high levels of economic inequality.

The focus of this thesis is on education and earnings. Earnings (or income) is often the primary focus in economic mobility studies.[3] The study of education as a socioeconomic outcome in mobility studies is also important as education is integrally connected to earnings and plays a critical role in a wide variety of outcomes (Blanden, Doepke & Stuhler, 2023). With regards to the importance of education to the economic outcomes of South Africans, Anderson, Case & Lam (2001) noted that "It is impossible to analyse issues such as racial differences in income, trends in employment, or intergenerational transmission of inequality without looking at the role of education". Since Anderson *et al*.'s remark, the importance of education to the economic outcomes of South Africans has increased even more, as we will discuss below.

Thanks to the rich cross-sectional datasets collected in the post-apartheid period, the trends in educational attainment, labour market earnings and the evolution of the relationship between

[1] The term 'black' here (and elsewhere in the thesis) refers collectively to the African, coloured and Indian/Asian racial groups.

[2] For comparative purposes, the composition of the South African population by race in 1950 was 69.5% African, 8.6% coloureds, 2.7% Indian/Asian, and 19.2% white (Feinstein, 2005).

[3] Socioeconomic outcome can be measured in various ways such as by income, earnings, occupation, education, or any other indicator of socioeconomic status. Economists tend to focus on intergenerational mobility in earnings or income. On the other hand, sociologists typically focus on intergenerational mobility in occupation or class positions.

education and earnings have been extensively studied. Finn, Leibbrandt & Ranchhod (2017) documents the stylized facts of the observed patterns, which include:

(i) An increase in average education along with a decline in educational inequality.

The mean education of the working-age population has been steadily increasing over the years, coupled with a decline in variability. For instance, mean years of completed education increased steadily from 7.7 years in 1993 to 10.5 years in 2018 (Branson & Lam, 2022). At the same time, the coefficient of variation declined steadily from 0.5 in 1994 to just over 0.3 in 2011 (Lam, Finn & Leibbrandt, 2015).

(ii) Persistently high earnings inequality.

The trends in income and earnings inequality in South Africa have been the subject of several studies. The findings show that inequality has been persistently high in the post-apartheid years and has even increased over the years (e.g., Bhorat, van der Westhuizen & Jacobs, 2009; Leibbrandt *et al.*, 2010; Wittenberg, 2017). Inequality in the post-apartheid period may have been exacerbated by structural shifts in the economy. The available evidence indicates that the economy of South Africa has been undergoing a structural shift that stemmed from a move towards de-industrialization and skills-biased technological change (Bhorat, Stanwix & Thornton, 2022). The shift in the economy increased demand for skilled workers. The demand for skilled workers led to higher returns to postsecondary education, which in turn created pressure on inequality.[4] Empirical studies of the contribution of education to earnings inequality show that increasing returns to postsecondary education has been one of the main factors for keeping earnings inequality persistently high even though inequality in education has been declining (Lam, Finn & Leibbrandt, 2015; Lam, 2020).

The above two points suggest that, in the post-apartheid era, while inequality in educational attainment has been declining, inequality in earnings has remained persistently high. These contrasting socioeconomic developments at the cross-sectional level are quite interesting and

[4] The shift towards skills-biased technological change of the economy and its consequences on inequality is not unique to South Africa. This has been one of the main mechanisms for increasing inequality in many developed economies (Gottschalk & Smeeding, 1997; Katz & Autor, 1999).

invite one to look at the intergenerational dimension. This thesis seeks to examine changes over time in the intergenerational inequality of South Africans in education and earnings.

The degree of association between the socioeconomic outcomes of parents and children have been observed to change over time (e.g., Hertz *et al.*, 2008; Chetty *et al.*, 2014b; Lefranc, 2018; Fan *et al.*, 2021; Connolly, Haeck & Lapierre, 2021). Corak (2004) notes that in developed countries the main drivers of change in the degree of association across generations over time have been changes in the returns to education and globalization.

1.2 Research Questions

The thesis aims to enrich the stylized facts of the cross-sectional trends in educational attainment and labour market earnings inequality during the post-apartheid period by looking at the intergenerational aspect of changes in these outcomes. The overarching question we ask is: has the association between the socioeconomic outcomes of parents and their children weakened or strengthened over time? This broad question is addressed in three sets of focused research questions.

(i) How do we measure intergenerational mobility? Given a measure of intergenerational mobility in education, how has the association between the educational attainments of parents and their children evolved? Has the link weakened or strengthened over time?

(ii) Does parental background matter to children's labour market earnings over and above the child's education? If yes, has the importance of parental background weakened or strengthened in the democratic era? Further, has there been a change in the contribution of parental background to children's earnings inequality over time?

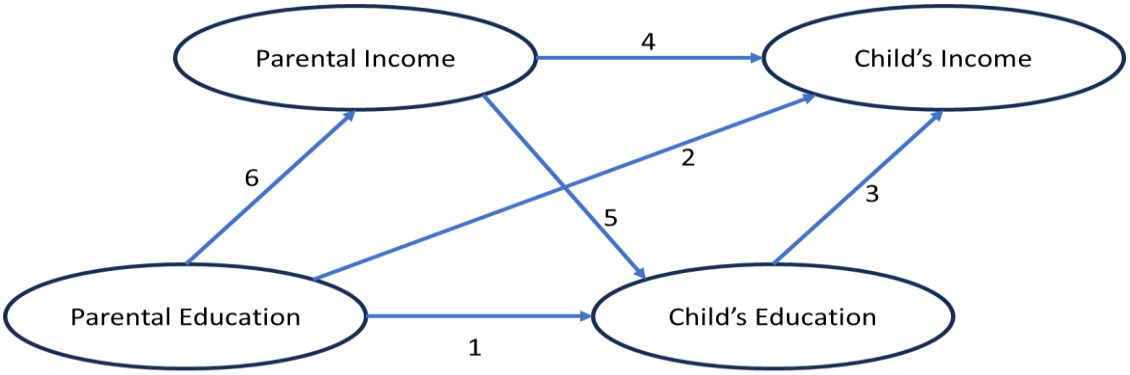
(iii) Has intergenerational persistence in earnings weakened or strengthened in the democratic era? Further, has there been a change in the mediating role of education in the transmission of earnings advantage across generations over time? If yes, was the role of education to increase or decrease earnings persistence?

The thesis addresses these three sets of research questions in three separate chapters.

1.3 Conceptual Approach of the Thesis

It is a well-established fact that parental background matters to the life outcomes of children. In particular, both parental education and income are known to matter to the education and income of their children. Figure 1.1 captures the paths of influence of the education and income of parents on that of their children.

Figure 1.1 Path diagram of the relationship between education and income of parents and their children



The thesis builds on the conceptual framework depicted in Figure 1.1. It will look at the various arrows that run from parents to children, with a focus on how the relationships have changed over time. In chapter 3 we will look at arrow 1. The chapter will examine how the association between the education of parents and the education of children has evolved over time. In chapter 4 we will look at arrow 2, controlling for arrows 1 and 3. The chapter will investigate changes over time in the importance of parents' education to children's earnings. In chapter 5 we will first look at the overall association between the earnings of parents and their children (arrow 4 plus arrows 5 & 3). The chapter here will examine changes over time in the overall association between the earnings of parents and their children. We will then look at the direct (arrow 4) and indirect (through education – arrows 5 & 3) effects of parents' earnings on children's earnings controlling for the confounding effect of parents' education (arrows 1 and 6). The chapter here will examine changes over time in the mediating role of education in the intergenerational transmission of earnings.

1.4 Datasets

The empirical analyses are carried out using two South African datasets. These are the Survey of Socioeconomic Opportunity and Achievement (SSOA) (Treiman, 2017) and the fifth wave of the National Income Dynamics Study (NIDS-w5) (SALDRU, 2018).[5] SSOA was collected between 1991 and 1994. The main survey was carried out in 1991/92. The survey collected an additional sample from rural areas in 1993/94. SSOA has hardly been used in previous studies. However, the dataset offers a wealth of information, as can be gathered from its name. The survey collected extensive information on respondents' socioeconomic background, including their educational, occupational and residential histories, and parental background.

NIDS-w5 is the wave 5 dataset of the National Income Dynamics Study panel. NIDS is a longitudinal household survey that started in 2008. NIDS-w5 was conducted in 2017. In wave 5, the survey undertook a top-up sample to boost its representativeness. NIDS collects data on each household member using the relevant instrument (adult, child, or proxy questionnaire). The survey collects a variety of information including demographics, education, labour market participation, economic activity and parental background.

The special feature of SSOA and NIDS, in the context of intergenerational mobility studies, is that they possess parental information on both coresident and non-coresident parents. Further, the timing of the two datasets allows for an investigation of changes over time in the socioeconomic mobility of South Africans.

We note that the two surveys differ in their degree of representation of the South African population. While NIDS is a nationally representative sample, SSOA is not. SSOA has national geographic coverage but missed farmworkers who were living on commercial farms.[6] We should thus bear in mind this difference between the two surveys as we go into the analysis.

[5] The datasets are publicly available and can be accessed from the website of DataFirst at the University of Cape Town (https://www.datafirst.uct.ac.za).

[6] Farmworkers who were living on commercial farms (in pre-1994 South Africa) are not a clearly identifiable group in the publicly available datasets, and thus it is difficult to get estimates of their exact size. However, we can provide the size of their occupational category in the early 1990s. According to the 1991 Census, the proportion of people in the occupational category '*farming and related occupations*' was about 13.81 percent of the employed in the pre-

Additional information on the datasets and web links to the datasets is given in Appendix 1.A.

1.5 Contribution of the Thesis

The thesis contributes to the South African literature on intergenerational inequality in several ways. First, we examine trends in intergenerational education mobility for cohorts born between 1923 and 1992 using datasets that are collected before and after the fall of apartheid. Empirical analyses to date have focused on post-apartheid datasets. Moreover, the cohorts in the two datasets overlap to some extent which allows for an assessment of measurement differences across the two surveys.

Second, we document the trends of intergenerational education mobility using both absolute and relative measures of mobility. Absolute mobility measures are absent in the studies of intergenerational mobility for South Africa. However, absolute mobility measures provide important summary statistics on intergenerational mobility. For instance, in cases where a country experiences upward educational mobility, say due to educational expansion that benefited everyone, a measure of relative mobility may remain unchanged if the changes in educational attainment for every child are as strongly related to parents' education as before. On the other hand, measures of absolute mobility can capture the change in the overall mobility across generations.

Third, we show how earnings can be adjusted for employment volatility to obtain a measure of the permanent earnings of individuals. Intergenerational earnings mobility studies typically rely on earnings measures that assume full employment. However, the South African labour market features high levels of employment volatility. The high levels of employment volatility make it difficult to get a measure of the permanent earnings of individuals.

Fourth, we highlight the confounding effect of parental education in analyzing the contribution of the mediating role of education to intergenerational earnings persistence. We also show how the confounding effect of parental education can be disentangled. Most studies do not control for the

1994 four provinces of the Republic of South Africa (i.e., the Cape, the Transvaal, Orange Free State and Natal). The size of the farmworkers who were living on commercial farms in the early 1990s is likely to be smaller than this however as it is inclusive of all in this occupational category.

confounding role of parental education in analyzing the mediating role of education in the intergenerational transmission of earnings advantage.

Fifth, we present new empirical evidence on changes over time in the association between the socioeconomic outcomes of parents and their children. Further, the fact that one of the datasets is collected right before the fall of apartheid provides an interesting base for the comparative analysis of changes in the socioeconomic mobility of South Africans in the democratic era.

1.6 Structure of the Thesis

The thesis is structured into six chapters. In the next chapter we provide a brief context for the educational attainments of South Africans. We discuss the nature and the changes over time of inequality in educational opportunity in South Africa. By providing the historical context to the educational provision in South Africa, the chapter sets the stage for the examination of changes over time in the intergenerational mobility of South Africans in the three chapters that follow.

In chapter 3, we examine changes in intergenerational education mobility across birth cohorts born between 1923 and 1992. The analysis is carried out using both absolute and relative measures of mobility.

In chapter 4, we investigate changes in the association between parental background (proxied by the education of parents) and children's earnings. In addition, we examine changes in the contribution of parental education to children's earnings inequality between the two periods represented by our datasets.

In chapter 5, we examine changes in intergenerational earnings persistence between the two periods. Further, we investigate whether there have been changes in the contribution of the mediating role of education to intergenerational earnings persistence between the two periods.

In the final chapter, we bring together the main findings of the thesis. One of the key findings of the thesis is that while intergenerational education mobility has increased considerably, earnings persistence has remained high and almost unchanged between the two periods. We conclude by highlighting the contrasting trends in intergenerational education and earnings inequality in the

post-apartheid period and recommend a way in which education can be instrumental in reducing intergenerational earnings persistence.

Appendix

1.A The Datasets

(i) The Survey of Socioeconomic Opportunity and Achievement (SSOA)

SSOA was collected between 1991 and 1994. The main survey was carried out in 1991/92. The survey collected an additional sample from rural areas in 1993/94. About 90 per cent of the sample was collected in 1991/92. The survey has national geographic coverage. It covered the former Republic of South Africa and the - at the time - nominally independent 'TVBC' states; namely, Transkei, Venda, Bophutatswana, and Ciskei. The survey, however, did not include farm workers who were living on commercial farms.

The sampling of the survey was based on a complex area probability sample, in which sampling groups and the number of cases to be interviewed were determined by the researchers. The survey first stratified the population into race-area groupings (sampling groups).[7] South Africa, up until June 1991, had a policy of residential segregation by race. The sampling within each sampling group followed the same approach. The survey used the administrative structure in place at the time. The country was divided into nine development regions. Each development region was divided into planning regions. Each planning region was in turn divided into magisterial districts. The sample size was allocated to administrative units at each administrative level in proportion to their population sizes, based on the 1985 Census figures.

Each magisterial district was further divided into enumerator areas. The survey randomly selected a prescribed number of enumerator areas within each magisterial district and then households within each selected enumerator area, in each case with probability proportional to size, using the 1985 population figures. Within a selected household, the survey interviewed one randomly selected adult (aged 20 years or older), with males having twice as much probability of being selected into the sample as females. In total the survey collected data from 9,086 individuals.

[7] Race refers to the four racial groups of South Africa (namely, African, coloureds, Indian/Asian and whites), and area refers to urban/rural geographic divide.

In analyzing the data, the calibrated weights that come with the dataset are used as weights.

The source of the dataset is DataFirst of the University of Cape Town. The version of the dataset used is version 1.

More information on SSOA can be found at https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/670 and the documentation therein.

(ii) The National Income Dynamics Study – wave 5 (NIDS-w5)

The National Income Dynamics Study (NIDS) is the first nationally representative panel survey of South Africa. The first wave was collected in 2008 and sampled 28,226 individuals living in 7,296 households across the country. The subsequent waves followed members of the households that were selected in the first wave and any children born to women. Additionally, any other individual who happened to fulfil the criteria of household membership in any subsequent wave is interviewed in that wave but is not tracked (Brophy *et al.*, 2018). The survey was last conducted in 2017 for its fifth time. In wave 5, the survey undertook a top-up sample to boost its representativeness. The survey successfully interviewed 39,434 individuals (in wave 5), of which 2,016 were from the top-up sample (SALDRU, 2018).

The sampling of the survey in wave 1 followed a stratified two-stage cluster sample design in sampling households. The population was stratified by district-council (DC). In the first stage, primary sampling units (PSUs) from Stats SA's 2003 Master sample of PSUs were randomly selected in each stratum.[8] The survey planned for a total of 400 PSUs. The number of PSUs selected per DC was allocated in proportion to the number of households. In the second stage, 24 dwelling units were drawn systematically within each selected PSU. The survey then interviewed all households living at the selected dwelling unit. In each selected household, all resident household members are NIDS sample members. Additionally, non-resident members that were out of scope at the time of the survey are included as NIDS sample members. These are individuals living in institutions (such as boarding school hostels, prisons or hospitals). The survey refers to both the resident and non-resident household members as continuing sample

[8] Stats SA refers to the National statistical service of South Africa.

members (CSMs). The survey interviewed each CSM using the relevant instrument (adult, child or proxy questionnaire).

This thesis uses the wave 5 dataset of NIDS. In analyzing the data, the cross-sectional calibrated weights that come with the dataset are used as weights. Although the wave 5 dataset of NIDS is the main sample, we also use the first 4 waves of NIDS to correct for missing values in some variables in wave 5.

The source of the dataset is DataFirst of the University of Cape Town. More information on NIDS can be found at:

https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/451 - (for NIDS-wave 1), and https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/712 - (for NIDS-wave 5). Further information on NIDS-w5 regarding the top-up sample and weights can be found in Brophy *et al.* (2018), Branson (2019) and Branson & Wittenberg (2019).

(iii) Summary Statistics of Samples

Table 1.1 presents summary statistics of respondents in the two datasets. The summaries are based on the adult samples in each dataset. In SSOA the sample is restricted to those aged 20 years and above, and in NIDS-w5 to those 15 years and above. The table shows that SSOA under-sampled Africans and oversampled whites and is more urban. On the other hand, the NIDS sample features a good degree of representation of the South African population.

Table 1.1 Summary Statistics of the two Samples (unweighted)

	SSOA			NIDS-w5		
	n	mean	st.dev.	n	mean	st.dev.
Race						
African	9,086	0.548	0.498	31,274	0.753	0.431
coloured	9,086	0.096	0.295	31,274	0.139	0.346
Indian	9,086	0.084	0.277	31,274	0.024	0.153
white	9,086	0.273	0.445	31,274	0.083	0.277
Sex						
Female	9,086	0.473	0.499	31,261	0.560	0.496
Male	9,086	0.527	0.499	31,261	0.440	0.496
Urban/rural res						
Urban	9,086	0.840	0.367	29,560	0.581	0.493
Rural	9,086	0.160	0.367	29,560	0.419	0.493
Province						
Cape	8,963	0.220	0.415			
Natal	8,963	0.154	0.361			
Orange Free State	8,963	0.061	0.239			
Transvaal	8,963	0.410	0.492			
Homelands	8,963	0.156	0.362			
Western Cape				31,283	0.047	0.212
Eastern Cape				31,283	0.121	0.326
Northern Cape				31,283	0.101	0.301
Free State				31,283	0.068	0.252
KwaZulu-Natal				31,283	0.051	0.221
North-West				31,283	0.260	0.438
Gauteng				31,283	0.059	0.235
Mpumalanga				31,283	0.147	0.354
Limpopo				31,283	0.068	0.252
N	9,086			31,283		

Notes: (1) The table is based on adult respondents in SSOA (aged 20 years and above) and NIDS-w5 (aged 15 years and above). The table provides unweighted summary statistics of the respondents in the two samples.

Chapter 2

2. Educational Attainment of South Africans in Context

2.1 The racial context

The colonial history of South Africa differs from that of most colonial territories in Africa in that it attracted a large proportion of European settlers (Beinart & Dubow, 1995, p. 1). The historical development of South Africa as a colonial settlement has been the driving force in defining race relations, which in turn defined every aspect of life along racial lines. After the formation of the Union in 1910, the white-controlled government formalized racial segregation in various areas of life. For instance, the Mines and Works Act of 1911 provided for formal racial segregation in employment, the Natives Land Act of 1913 provided for the creation of a system of "reserves" giving white people 87 per cent of the country's land, and the Natives (Urban Areas) Act of 1923 empowered local municipal authorities to racially segregate urban settlements and to control mobility using passes (Beinart & Dubow, 1995, p. 3-4; Beinart, 2001). According to Beinart & Dubow (1995), these legislations (and the practice of segregation at large) were put not only to restrict blacks in terms of where they could live and work but also to secure the socioeconomic position of whites.

Racial segregation was further strengthened with the accession of the apartheid regime to political power in 1948 (Beinart & Dubow, 1995, p. 12; Beinart, 2001). The apartheid regime advanced the concept of 'separate development' of the races and dedicated itself to securing the privileges of the white minority at the expense of the black majority (Beinart, 2001). Several Acts were passed towards this end. For instance, the Prohibition of Mixed Marriages Act of 1949 prohibited marriage between whites and non-whites, the Population Registration Act of 1950 required people to identify and register from birth as belonging to one of four racial groups (namely, African, coloured, Indian or white), the Group Areas Act of 1950 provided for urban areas to be divided into racially segregated zones, and the Bantu Authorities Act of 1951 provided for the establishment of homelands for Natives (Africans) and abolished the Natives Representative Council.

The discriminatory racial laws and practices during the long colonial and apartheid periods privileged whites and trampled down the rights of blacks in all aspects of social, political, and economic life. Under apartheid, while white people enjoyed good educational and employment opportunities, black people lived in a dehumanizing environment in which their educational and employment opportunities were severely circumscribed and were continually reminded of their powerlessness (Fiske & Ladd, 2004, p. 2).

The colonial domination and the apartheid system of governance and their injustices were fought back and vehemently opposed at every turn by the South African populace (e.g., McKenna, 2011). The apartheid regime fell in 1994. With the fall of Apartheid, South Africa transitioned to democracy with the promise of racial equity in all aspects of life.[9]

This chapter intends to provide a brief historical background to the provision of education in South Africa with a particular focus on racial disparity in educational opportunity in the schooling system. The comparative analysis is carried out using the apartheid regime's racial classification of South Africans; namely, African, coloured, Indian/Asian and white.[10]

2.2 Racial Disparity in Educational Opportunity before 1994

2.2.1 Racial Segregation in Education

The educational system of South Africa was racially segregated before the advent of apartheid. Cross & Chisholm (1990) points out that the foundations of segregated schooling go back to the development of a segregationist social policy in the late nineteenth century. Education was racially segregated into four separate and hierarchically different schooling systems (namely, Native education, coloured education, Indian education and white education) during the first two decades of the twentieth century. While white children were brought into a form of mass schooling in well-functioning state schools, black children were not fully incorporated and were largely under voluntary, mission control (Cross & Chisholm, 1990).

[9] Constitution of the Republic of South Africa, p.1
[10] Throughout the thesis, we use these four racial classifications to categorize race.

The apartheid regime, which came to hold political power in 1948, passed several Acts to bring the educational system closer to its policy of 'separate development' of the races (Cross & Chisholm, 1990; Fiske & Ladd, 2004, p. 43). The Acts include the Bantu Education Act of 1953, the Extension of University Education Act of 1959, the coloured Persons Education Act of 1963, the Indian Education Act of 1965, and the National Education Act of 1967. These Acts gave the state large control powers over the education of blacks. The control of the education of blacks was progressively transferred to the central government, with the education of Africans put under the Department of Native Affairs in 1953, the education of coloureds under the Department of Coloured Affairs in 1963, and the education of Indians under the Department of Indian Affairs in 1965 (Horrell, 1968, 1970).

By the 1970s the schooling system of South Africa was highly segregated along racial lines with 'appropriate' curricula for each racial group. Further, the administration of education was split on the basis of race and geography into 19 separate education departments, with each having unequal access to human and financial resources (Donn, 1995). Some of the departments served African, coloured, Indian or white students in the country's urban areas, while others served Africans in the 'self-governing' territories and 'independent' homelands.[11]

The racially segregated educational system had enormous implications to the provision of education. On the one hand, the education of whites was funded generously, and its quality was comparable with the best in the industrialized world. On the other hand, the education of blacks was poorly resourced and overcrowded with insufficient and poorly qualified teachers (Pillay, 1984). The education for blacks was not homogenous either. African education was quite underdeveloped compared to that of coloureds and Indians (Pillay, 1984).

The racial disparity in the provision of education before the fall of apartheid can be seen by looking at the racial differences in schooling inputs and outputs. The following section provides a

[11] The 1959 Promotion of Bantu Self-Government Act, which was an extension of the Bantu Authorities Act of 1951, paved the way for the nominal independence of the homelands. 'Independence' was granted to Transkei (in 1976), Bophuthatswana (in 1977), Venda (in 1979) and Ciskei (in 1981). The remaining homelands remained 'self-governing territories. These were Gazankulu, KaNgwane, KwaNdebele, KwaZulu, Lebowa, and QwaQwa (Beinart, 2001; Fiske & Ladd, 2004, p. 26-7).

comparison of the racial disparity in education in terms of enrolment rates, funding, pupil-teacher ratios, teacher qualifications and matriculation pass rates.

2.2.2 Racial Disparity in the Provision of Education

2.2.2.1 School Enrollment

White children were granted the right to free and compulsory education during the first decade of the twentieth century. On the other hand, during that time while coloureds were given only the right to free education, African children were denied the right to free and compulsory education (Cross & Chisholm, 1990).

The number of Africans attending school was low up to the mid-1950s, but their number started to increase rapidly in subsequent years (Fedderke, de Kadt & Luiz, 2000). The mid-1950s marked the start of the implementation of the Bantu education system, which saw the rise of mass schooling for Africans (Cross & Chisholm, 1990).

Horrell (1968) provides enrollment figures for Africans for selected years between 1925 and 1967 (Table 2.1). The figures show that the number of Africans attending school relative to the African population increased progressively over time. However, a majority of those enrolled were attending in the first four years of schooling. For instance, about 71.6 % of African pupils were attending the first four grades in 1967 (Table 2.1). Fiske & Ladd (2004, p. 48) note that the number of African children attending school reached 3.5 million by 1975, with enrollment in secondary school levels[12] increasing fivefold between 1965 and 1975, to 280,000.

[12]An additional year of study (Form I) was included in the curriculum for Africans in the mid-1950s (Horrell, 1968).

Table 2.1 Enrolment of African students for selected years, 1925-67

Year	Number of African pupils	% of the African population at school	% of African pupils attending Sub A to Std. II
1925	209,049	**na**	83.7
1930	284,250	**na**	83.4
1935	353,044	5.42%	81.2
1940	464,024	6.53%	78.8
1945	587,586	7.60%	75.9
1950	747,026	8.88%	73.5
1955	1,013,910	10.67%	72.7
1960	1,506,034	14.10%	72.8
1965	1,957,836	16.07%	71.8
1967	2,237,397	17.49%	71.6

Source: Horrell (1968, p.51-53)
Note: na=not available

Pi1lay (1984) compares the school enrolment ratios among school-going age children[13] for the four racial groups for selected years between 1920 and 1980. The figures show that the racial gap in school enrolment ratios narrowed over time. In 1920, 78.4 per cent of white children of school-going age were enrolled in school. The corresponding figures for Indians, coloureds and Africans were 14.9, 29.7 and 13.8 per cent, respectively. In 1950, while 85.8 per cent of white children of school-going age were enrolled in school, the corresponding figures for Indians, coloureds and Africans were 49.1, 67.9 and 32.9 per cent, respectively. The racial gap in school enrolment ratio between whites and the other racial groups further narrowed in subsequent years. In 1980, 96.3 per cent of white children of school-going age were at school, and the enrollment ratios for Indians, coloureds and Africans were 94.8, 93.0, and 83.1 per cent, respectively. Pillay notes that the increased enrolment of school-going age children was attributable to compulsory schooling up to age 15 for whites, Indians and coloureds, and four years of schooling for Africans.

Although the increase in enrolment of blacks was impressive, progress through school among blacks was quite low. According to Pillay (1984), the percentage of white pupils who reached standard 10 in 1978 from standard 1 in 1969 was 69 per cent. The corresponding figures for Indians, coloureds and Africans were 34.4, 10.5 and 3.9, respectively. Progress in secondary

[13] The proportion of individuals of school-going age computed as 80 per cent of those in the 5-19 age group (Pillay, 1984).

schooling having reached standard six was slightly better among blacks. The percentage of white pupils who reached standard 10 in 1978 from standard 6 in 1974 was 68.3 per cent. The corresponding figures for Indians, coloureds and Africans were 40.6, 22.3 and 7.8 per cent, respectively. The statistics for the other years around this period were also similar. Pillay notes that high drop-out rates were the main reason for the poor progress through school among blacks.

2.2.2.2 Expenditure on Education

Before the mid-1940s State expenditure on African education was pegged at ZAR340,000 under the 1922 Financial Relations 4th Extension Act. Additional finances for African education were to be obtained from tax through the Natives Taxation and Development Act. The education of coloureds was financed by the State through the provinces on the basis of average pupil attendance (Cross & Chisholm, 1990). Further, Cross & Chisholm note that the financing of the education of Africans and coloureds was boosted in 1945 when the State increased expenditure on education. However, the education of Africans continued to be supported by direct taxation paid by Africans through the Natives Taxation and Development Act (Horrell, 1968).

Fedderke, de Kadt & Luiz (2000) compare the trends in racial per-pupil education expenditure over an extended period. The trend for whites runs from 1917 to 1993, for coloureds & Indians (combined) from 1920 to 1993 and for Africans from 1935 to 1993. The trends reflect large racial differences in per-pupil expenditure, with the white schooling system remaining substantially better resourced over the years. Fedderke, de Kadt & Luiz (2000) show that the per-pupil expenditure for whites remained at least seven times that of Africans, and at least twice that of coloureds & Indians. The gap in per-pupil expenditure between whites and coloureds & Indians, however, narrowed to less than twice during the early 1990s.

Table 2.2 provides per-pupil state expenditure by race as a percentage of white per-pupil expenditure for selected years between 1971 and 1980. The numbers show that the racial gap in per pupil expenditure in comparison to that of whites was quite wide. The per pupil expenditure for Africans, in particular, was quite low and changed little over the years. In 1971-72, the per pupil expenditure on Africans was 5.5 per cent of white per pupil expenditure, and in 1978-79 the per pupil expenditure on Africans (the highest in the 1970s) was 9.8 per cent of white per pupil expenditure.

Table 2.2 Per-Pupil State Expenditure by Race as a percentage of white per-pupil expenditure for selected years, 1971-1980

Year	% of white per-pupil expenditure			white Per-Pupil Expenditure
	Africans	coloureds	Indians	
1971-72	5.5*	20.5	27.0	R 461
1974-75	6.5	20.7	28.3	605
1975-76	6.5	21.7	29.4	644
1976-77	7.4	24.1	33.6	654
1978-79	9.8	31.2	49.3	724
1979-80	7.8	20.0	33.3	1169

Source: Pillay (1984)
Note: (1) white per pupil expenditure is in annual rands.
(2) *Figures for Africans are based on the 'Common Area', i.e. excluding the homelands.

2.2.2.3 Pupil-teacher Ratio

Table 2.3 shows the pupil-teacher ratios in African schools for selected years between 1935 and 1967. The pupil-teacher ratio started to increase after the mid-1950s because the percentage of the African population attending school saw rapid expansion during this period (Horrell, 1968, p. 53).

Table 2.3 Pupil-teacher Ratio in African schools for selected years between 1935 and 1967

Year	1935	1940	1945	1950	1955	1960	1965	1967
Pupil-teacher ratio	43.6	43.4	42.3	40.3	46.1	54.3	56.0	58.5

Source: Horrell (1968, p.52-53)
Note: The figure for 1967 is based on the Republic, i.e. excluding the homelands.

Fedderke, de Kadt & Luiz (2000) compare the trends in pupil-teacher ratios by race over time. They show that while the pupil-teacher ratio for whites never peaked above the mid-20 level over the period 1917-1993, the ratio for Africans never dropped below 31 in the period 1935-1993. The pupil-teacher ratio for Africans rose rapidly after the mid-1950s and stayed generally high throughout the period until 1993. Fedderke *et al.* note that the Bantu education system increased the enrolment of African children but did not provide additional teaching resources at a comparable growth rate. Their results further indicate that the trend of the pupil-teacher ratio for coloureds and Indians (combined) declined consistently over the period 1935-1993, having begun

with a similar ratio as Africans in the mid-1930s to almost closing the gap with whites towards the early 1990s.

Further statistics on pupil-teacher ratio by race for selected years between 1960 and 1982 can be found in Pillay (1984). The numbers show that pupil-teacher ratios in African schools remained generally high.

2.2.2.4 Teacher Qualifications

Although differences in pupil-teacher ratio across the racially segregated schooling systems reflect inequality in schooling opportunities by race, the pupil-teacher ratio in itself does not provide a picture of the differences in the qualification of teachers. Defining a qualified teacher as one with at least a matric and teaching diploma, Pillay (1984) shows large differences in the proportion of qualified teachers by race. In the early 1980s, the percentage of underqualified white teachers was only about 3.36 per cent. On the other hand, the percentage of underqualified Indian, coloureds and African teachers, respectively, were 19.70, 66.14 and 85.00[14] per cent.

Fedderke, de Kadt & Luiz (2000) compares the proportion of super-qualified teachers, which they define as the proportion of teachers possessing a tertiary qualification, between the white and African schooling systems over the period 1960-1993. They found the two schooling systems to differ markedly in this respect over the entire period, with the African schooling system being the disadvantaged one.

These findings suggest that the education system of blacks was constrained not only in the number of teachers but also in their quality.

2.2.2.5 Matriculation Passes

The number of matriculants emerging from the racially segregated schooling systems differed greatly, with matriculants from the white schooling system far exceeding those from the other schooling systems. For instance, among the 1970 matriculants (including exemptions and School leaving certificates) white pupils constituted 85.6 per cent. The corresponding figures for Indian,

[14] The figure for Africans does not include the 'Independent' homelands.

coloured and African pupils, respectively, were 4.5, 3.3 and 6.6 per cent (Pillay, 1984). Although the number of black matriculants increased over time, their number remained far smaller in comparison to that of whites. Among the 1980 matriculants, while white pupils accounted for 57.4 per cent, Indian, coloured and African pupils accounted for 8.3, 5.7 and 28.6 per cent, respectively (Pillay, 1984).

Pillay (1984) also shares some statistics on the number of matriculants who gained university exemption by race for selected years in the 1970s. The figures show that the proportion of blacks with university exemption increased over time, although it remained much smaller compared to that of whites. The percentage of African pupils among those with university exemption increased from 6.7 per cent in 1970 to 16.4 per cent in 1978. The figures for coloureds and Indians also increased during the period, for coloureds from 2.8 to 3.8 per cent and Indians from 2.5 to 6.0 per cent.

The segregation of education by race and the racial gaps in schooling inputs and outputs before the fall of apartheid reflect the racial disparity in educational opportunity that existed during the period. The next section briefly reviews the changes in the provision of education in the post-apartheid period.

2.3 Racial Disparity in Educational Opportunity After 1994

2.3.1 Inequality in Educational Opportunity

South Africa took several measures in the post-apartheid years to dismantle the racially defined and iniquitous education system and replace it with a deracialized, unified and equitable one. The right to education was enshrined in the Constitution of the country, which stipulates that: "Everyone has the right to a basic education, including adult basic education, and to further education, which the state, through reasonable measures, must make progressively available and accessible."[15] Various policy instruments were also enacted. Key legislations include: the South African Qualifications Authority Act of 1995 - which provides for the development and implementation of a national qualifications framework; the South African Schools Act of 1996 -

[15] Constitution of the Republic of South Africa, 1996, Section 29 (1).

which provides for a uniform system in the organization, governance and funding of schools and repeals various apartheid-era enacted laws; and the Further Education and Training Act of 1998 - which provides for a coordinated national higher education and training system.

Education was made compulsory for all South African children up to age 15 or ninth grade, whichever comes first. Further, the various education administrations that existed before the fall of apartheid were brought under a unitary authority in 1995 (Donn, 1995).

The post-apartheid education policy clearly envisions for a break from the race-based education policy of the pre-1994 era. The new education policy promoted for a race-blind education system that caters to all South Africans. However, despite the policy intentions, various studies note that the post-apartheid education system has been largely a continuation of the old system (e.g., van der Berg, 2007; Spaull, 2013a, 2019). These studies argue that while formerly white schools continued to be functional and able to educate students, the schools that formerly served black communities remained largely dysfunctional and failed to educate students.

Fiske & Ladd (2004) evaluate whether the post-apartheid education system has been racially equitable. Their analysis is based on the first eight years following the fall of apartheid. They define racial equity in education around three concepts: namely, equal treatment, equal educational opportunity, and educational adequacy. They conclude that South Africa has made significant progress towards equity in education when equity is seen from the point of equal treatment of persons of all races. However, South Africa was deemed less successful towards equity in education when equity is defined either as equal educational opportunity or as educational adequacy for students of all races. The authors pointed out that equity in education has been elusive due to several factors which include the negative legacy of apartheid and both political and financial challenges the country faced in the post-apartheid years.

Fiske & Ladd (2004) identify four aspects of the legacy of apartheid which have particular relevance to the progress of education in the post-apartheid era. These are “residential segregation and persistent poverty among Africans, inadequate resources and low-quality instruction for black children, low levels of educational attainment among black adults and low student achievement, and the absence of an adequate ‘culture of learning’.”

Spaull (2019) notes that in the post-apartheid period, although race continues to be a major factor in determining educational opportunity of South Africans, other factors such as the wealth of parents and the province of birth also play a role.

2.3.2 Inequality in Educational Attainment

More than two decades into democracy, although racial differences in educational attainment have narrowed over time, there continues to be large racial gaps in educational outcomes (Branson & Lam, 2022). Large racial gaps exist in the rate of secondary school completion. Branson & Lam (2022), using the Community Survey of 2016, show wide racial gaps in secondary school completion. For instance, among the 1991 birth cohort, the proportion of whites completing grade 12 was about 86 per cent. The corresponding proportions for Africans and coloureds were only about 53 per cent. Interestingly, the proportion of Indians completing grade 12 has been increasing rapidly across birth cohorts and has almost closed the gap with whites by the 1991 birth cohort.

Large racial gaps also exist in the rate of postsecondary education completion. Branson & Lam's findings show that the proportion of secondary school graduates who go on to obtain some kind of postsecondary education differs by race. For instance, among the 1990 birth cohort, while the proportion of whites who go on to attain some kind of postsecondary education after completing grade 12 was close to 50 per cent, the corresponding proportion for Africans was only about 20 per cent. Furthermore, they document that the distribution of postsecondary enrollment across colleges and universities differs by race. African and coloured students are more likely than white and Indian students to enroll in colleges[16] than universities.

Branson & Lam (2022) point out that racial differences in educational attainment are not only due to school dropout but also grade repetition. Their findings show that grade repetition varies by race. For instance, even though the rates of enrolment of African and white 18-year-olds are similar, Africans are about 1.1 grades behind whites by age 18. Grade repetition is particularly high at the secondary level and is higher among Africans and coloureds. In another similar study,

[16] 'College' in the South African education system refers to public Technical and Vocational Education and Training (TVET) institutions and private Further Education and Training (FET) colleges (Branson & Lam, 2022).

Branson & Lam (2010) document the presence of high levels of grade repetition among Africans in comparison to whites.

2.4 Concluding Remarks

Education is an important indicator of socioeconomic status and plays an important role in intergenerational mobility. This chapter attempted to provide a brief context to the nature of inequality in educational opportunity and the changes that occurred over time in South Africa. The historical context that underpins the educational attainment of South Africans is useful in understanding changes over time in the intergenerational mobility of South Africans, which we discuss in the subsequent three chapters.

Chapter 3

3. Intergenerational Education Mobility Trends in South Africa

3.1 Introduction

Education has been dubbed the great social equalizer. This can be corroborated by the various studies which show that higher educational attainment is associated with several desirable life outcomes such as higher earnings (e.g., Psacharopoulos & Patrinos, 2018), improved health (e.g., Lleras-Muney, 2005; Lochner, 2011), and better parenting skills (e.g., Oreopoulos & Salvanes, 2011). Education is also a key driver in the process of intergenerational mobility in economic and occupational outcomes.

The focus of this chapter is on the trends of intergenerational education mobility in South Africa. Intergenerational mobility in education provides valuable information on the degree of equality of opportunity in society (Björklund & Salvanes, 2011). The lower the intergenerational mobility, the stronger the association in educational attainment across generations, which suggests that educational outcomes of children are tied to their parental background. On the other hand, the higher the intergenerational mobility, the weaker the association in educational attainment across generations, suggesting that parental background is less important in the educational attainment of their children. Changes in the distribution of education across generations is driven both by public education policy and private investment in education. In South Africa, the involvement of the State in education has been expansive both before and after the fall of apartheid. Under apartheid, as discussed in chapter 2, enrollment rates have increased through the years although the education system was racially segregated and unequal. In the post-apartheid period, the racially segregated education system was brought into a unitary system, public education was expanded, and efforts were made to address racial inequity in education.

The intra-generational trend in educational attainment of South Africans is well documented (see e.g., Thomas, 1996; Lam, 1999; Branson & Lam, 2022). The findings show that average education has been increasing and educational inequality has been declining across birth cohorts. The trend in intergenerational education mobility has also been discussed but gaps remain in the

literature. This chapter enriches the existing empirical literature in several ways. First, the chapter provides an extended view of the trend in intergenerational education mobility using datasets that are collected before and after the fall of apartheid. Empirical analyses to date have only focused on post-apartheid datasets. Second, it documents the trends in intergenerational education mobility using both absolute and relative measures of mobility, providing a comprehensive treatment of trends in education mobility. Third, the chapter uses alternative measures of both absolute and relative mobility, some of which have not been applied in the South African context, providing additional insight into the trends in education mobility.

The rest of the chapter is organized as follows. The next section provides a review of the relevant literature. We briefly discuss the mechanisms of educational persistence across generations. We also discuss measures of mobility and measurement issues that arise in estimating intergenerational education mobility. Further, we present a brief review of the empirical literature on intergenerational education mobility, covering both international and South African studies. Section 3 presents the methods we use to analyse the trends in education mobility. We first discuss the measures of absolute mobility and then the measures of relative mobility. Section 4 identifies and describes the datasets. This is followed by the empirical results. We present estimates of the trends in education mobility across 5-year birth cohorts both for the full and African samples. The last section discusses the main findings of the chapter, with emphasis on the younger birth cohorts. Educational mobility in the younger birth cohorts is seen to depart from the trends of mobility set by the older birth cohorts.

3.2 Literature Review

3.2.1 Economic Perspective on Intergenerational Persistence

Empirical studies, both for developed and developing countries, show that the educational attainments of children and their parents are positively correlated (see e.g., Hertz *et al.*, 2008; van der Weide *et al.*, 2021). Various mechanisms lead to such persistence across generations. Economic theory emphasizes on parental investment behaviour in the human capital of their children within the framework of utility-maximizing parents confronted with allocating their limited resources between their own welfare and the welfare of their children (Becker & Tomes, 1979, 1986). The transmission channels include (Lochner, 2008; Björklund & Salvanes, 2011):

(i) Parental educational attainment: The schooling choices of parents may directly affect the schooling choices of their children. In addition, higher educational attainment of parents may interact with the child's schooling, which increases the child's productivity of going to school.

*(*ii) Parental financial resources: Higher parental education is more likely to generate higher parental incomes. This allows more educated parents to overcome borrowing constraints on investment in their children's education. Higher parental financial resources may also lower the opportunity costs of education.

(iii) Parental cultural resources: parental cultural capital including noncognitive abilities (such as persistence, motivation, temperament, and preferences) could destine children to have the same educational outcomes as their parents. For instance, educated parents may influence the child to pursue higher levels of schooling through their attitudes towards education. More educated parents may also provide a stimulating home environment that improves the child's chances of success in school. Furthermore, educated parents may influence the child's schooling outcome through their selection of residential area and school to attend.

(iv) Genetic transmission: It is well documented that differences in abilities account for a substantial share of the variation in socioeconomic outcomes across individuals (e.g., Cunha *et al*., 2006; Heckman, 2008). Moreover, an important component of the ability of individuals is the genetic transmission of abilities. To the extent that genetic cognitive abilities pass on from parents to children, children of genetically endowed parents may end up acquiring higher levels of schooling.

3.2.2 Mobility Measures and Measurement Issues

The main variable of analysis in measuring intergenerational education mobility is the educational attainments of parents and their children.[17] Educational attainment is often measured by

[17] Note that educational attainment does not capture education as equivalent to human capital. Human capital is a synthesis of various outcomes that include cognitive skills, socioemotional skills and mental wellbeing. Educational

completed years of formal education. Measures of intergenerational education mobility capture either the overall change in educational attainment across generations or the association between the educational attainments of parents and their children given the distribution of education across generations (Torche, 2019). The first is captured using absolute mobility measures and the second using relative mobility measures. There are various measures of absolute and relative mobility in the literature.[18] Commonly used measures of absolute mobility include upward mobility (the proportion of children who surpassed the educational attainment of their parents) and downward mobility (the proportion of children who failed to surpass the educational attainment of their parents). And, commonly used measures of relative mobility include Intergenerational Regression Coefficient (IGRC), Intergenerational Correlation (IGC), and Intergenerational Rank Correlation (IRC).[19]

It is important to note that not all measures of relative mobility are 'pure' relative measures (Jäntti & Jenkins, 2015). Relative education mobility measures provide summary measures of the joint distribution of the education of children and parents. The joint distribution can be decomposed into the joint distribution of the ranks of the education of children and parents and the marginal distributions of the education of children and parents (Nybom. & Stuhler, 2017; Emran & Shilpi, 2019). The joint distribution of the ranks determines mobility across generations, and the marginal distributions determine the extent of inequality within each generation. Pure relative mobility measures are sensitive only to changes in the joint distribution of the ranks and are unaffected by changes in the marginal distributions of education (Jäntti & Jenkins, 2015). Among the three relative measures of mobility mentioned in the last paragraph, the IGRC is based on the joint distribution of the ranks and the marginal distributions. Thus, changes in IGRC could be either due to changes in intergenerational mobility or changes in inequality within each

attainment is one important input in the production of important forms of human capital, especially cognitive development (Behrman, 2019).

[18] A comprehensive discussion on measures of mobility (in the context of income mobility) can be found in Jäntti & Jenkins (2015).

[19] In the literature, there are no commonly used names for these three measures of relative mobility. In this study, we follow the names and acronyms given to these measures by Emran & Shilpi (2021).

generation. On the other hand, the IRC is purely based on the joint distribution of the ranks, thus it is a pure relative mobility measure (both marginal distributions of the ranks have standard deviations of 1).[20] IGC also controls for changes in inequality. The difference between IGC and IRC is that while IRC controls for changes in inequality fully, IGC controls for changes in inequality only in a specific way (i.e. through the standard deviations of the marginal distributions). This can be seen in the relationship between IGC and IGRC.

IGC and IGRC are closely related measures. Their relationship can be expressed as:

$$IGC = IGRC \cdot (\sigma_p / \sigma_c)$$

where σ_p and σ_c denote the standard deviations of the education of parents and children, respectively. IGRC provides only a partial view of variations in the degree of intergenerational persistence as it does not control for changes in inequality across generations. IGC is better than IGRC in this regard, however, IGC also does not fully control for changes in inequality across generations. Lefranc (2018) refers to IGRC as a measure of the extent of transmitted inequality, and IGC (and IRC) as measures of the extent of inherited inequality.

The estimation of intergenerational mobility is sensitive to measurement error and co-residency restrictions in datasets; however, the measures of mobility differ in their robustness to such data limitations (Emran & Shilpi, 2021). Measurement error in education may not be as prevalent as in earnings (or income) data but is bound to occur. For instance, surveys often ask respondents for the education of their parents, which may not be reported with accuracy by everyone.[21] In the context of intergenerational income mobility, evidence shows that the rank coefficient is less affected by measurement error compared to the regression and the correlation coefficients (Nybom & Stuhler, 2017). Intergenerational mobility estimates also suffer from co-residency bias

[20] Purely relative measures of mobility could also be derived from transition matrices. For instance, in a 2x2 transition matrix, the odds ratio completely characterizes relative mobility (see e.g., Long & Ferrie, 2013; Modalsli, 2017). In comparison to the IRC, the transition matrix approaches could be quite useful in obviating any problems with heaping, however, their usefulness is limited to analyzing mobility in a categorical variable.

[21] Respondents also misreport their own education. For instance, misreporting of own education is substantial in South African datasets (e.g., Hertz, 2003; Branson, Hofmeyr & Lam, 2014). Branson, Hofmeyr & Lam (2014) found misreported education for about 33% of respondents.

if appropriate datasets are not used. Co-residency bias may result when the dataset does not provide information on non-coresident parents or children. Most household surveys collect information only on co-residing members of households. The co-residency criteria for household membership results in an unrepresentative sample in the analysis of intergenerational mobility. The measures of relative mobility are generally biased downward in co-resident samples, although they differ in their degree of robustness (Emran, Greene & Shilpi, 2018; Emran & Shilpi, 2018). Emran, Greene & Shilpi (2018) provides evidence which shows that the correlation coefficient is more robust compared to the regression coefficient. Further, Emran & Shilpi (2018) shows that the rank coefficient is more robust compared to the regression coefficient and is comparable in robustness to the correlation coefficient in coresident samples.

Educational attainment as measured by completed years of education may not reflect achievement gaps between individuals with the same level of education or differences in quality of schooling. The use of completed years of education as a proxy for educational attainment thus may introduce measurement errors, which may understate not only the extent of cross-sectional educational inequalities but also intergenerational persistence (Blanden, Doepke & Stuhler, 2023). However, completed years of education has the advantage that it is routinely collected in household surveys.

In the literature, there is no commonly agreed definition of parental education. The alternatives include: taking the highest education of both parents (with the argument that educational opportunities in a family are defined by the parent with the highest education), or taking some combination of the education of both parents[22], or taking the mother's education (with the argument that mothers are more influential in a child's development, especially at an early age), or taking mother's education for daughters and father's education for sons (with the argument that the same gender parent is more influential in a child's education) (Torche, 2019).

[22] For instance, Hertz *et al.* (2008) use the simple average of the education of parents as a measure of parental education.

Empirical studies of intergenerational education mobility often restrict the sample of analysis to respondents who are in their mid-20s and above. This is done to allow enough time for schooling completion, which minimizes the right-censoring of education.[23]

3.2.3 Empirical Literature

3.2.3.1 Intergenerational Absolute Education Mobility

Comparative studies of absolute education mobility across countries show that absolute mobility is higher in high-income countries compared to low-income countries (e.g., van der Weide *et al.*, 2021 for 153 countries around the world; Alesina *et al.*, 2021 for 27 African countries). Absolute education mobility is also documented to vary over time (e.g., van der Weide *et al*., 2021; Neidhöfer, Serrano & Gasparini, 2018). According to van der Weide *et al*. (2021), across birth cohorts, absolute upward mobility has been higher in high-income countries compared to low-income countries. Moreover, they show that the gap in upward mobility between high-income and low-income countries has been narrowing over time, largely driven by a fall in the proportion of children with higher education than their parents in high-income countries. Neidhöfer, Serrano & Gasparini (2018) examines the trend in absolute per capita mobility across birth cohorts among 18 Latin American countries. Their findings show that the trend followed an inverted-U pattern. They attributed the fall in per capita mobility over the younger birth cohorts to the bounded nature of education, where the limited range of education coupled with rising parental education across these birth cohorts reduced the possibility for children to experience an improvement.

Absolute education mobility is also documented to vary across geographic spaces within regions (e.g., Alesina *et al.*, 2021; Asher, Novosad & Rafkin, 2022 for India). Alesina *et al.* (2021) found that across Africa upward mobility is higher in areas with colonial investments in railroads and missionary activities as well as in capital cities and coastal areas. Asher, Novosad & Rafkin

[23] Some studies work with younger respondents who are still in school and co-reside with their parents. In this case, a child's education is operationalized in terms of the number of grades completed per year of age or the gap between the expected and actual schooling since the start of schooling (usually age 6) (e.g., Lam, 1999; Behrman, Birdsall & Székely, 1999).

(2022) found upward education mobility to be higher in urban areas as well as in areas with high average education levels across India.

Intergenerational absolute mobility is associated with a variety of country characteristics. According to van der Weide *et al*. (2021), the relationship between absolute upward mobility and per capita GDP follows an inverted-U pattern across countries. Absolute mobility is low for the poorest and the richest countries, although for different reasons. For the poorest countries, where the threshold for surpassing parental education is low, the low absolute upward mobility reflects a poverty trap. For the richest countries, where the threshold for surpassing parental education is quite high, the low absolute mobility signifies the ceiling effect due to the bounded nature of education (van der Weide *et al*., 2021). Overall, absolute mobility is documented to be positively correlated with variables linked to economic development (e.g., public expenditure in education, school quality and child health) and negatively correlated with proxies for inequality and segregation (van der Weide *et al*., 2021; Neidhöfer, Serrano & Gasparini, 2018).

3.2.3.2 Intergenerational Relative Education Mobility

As with absolute mobility, relative education mobility is higher in high-income countries compared to low-income countries (e.g., Hertz *et al.*, 2008 for 42 countries; van der Weide *et al.*, 2021 for 153 countries around the world). Hertz *et al.*, (2008) found Nordic countries to have the highest levels of relative education mobility, even in comparison to non-Nordic Western European countries. On the other hand, they found Latin American countries to have the lowest levels of relative mobility in comparison to the other four regions in their study: namely, Asia, Africa[24], Eastern Bloc countries and Western Europe & the USA. van der Weide *et al.* (2021), using a much larger sample of countries than Hertz *et al*., also found relative mobility to be lower in developing countries compared to developed countries. Among developing regions, South America, South Asia and parts of Sub-Saharan Africa were found to have some of the lowest levels of relative mobility.

[24] The African group includes Egypt, Ghana, Ethiopia (rural) and South Africa (KwaZulu-Natal province).

Relative education mobility is documented to vary over time (e.g., Hertz *et al.*, 2008; Neidhöfer, Serrano & Gasparini, 2018; van der Weide *et al*., 2021). However, the trend of relative mobility as measured by the correlation coefficient is generally more stable compared to the trend as measured by the regression coefficient. For instance, in Hertz *et al.* (2008), the global trend of the intergenerational regression coefficient across country-cohort estimates of the coefficient declined from 0.8 in the early 1920s to about 0.5 in the early 1980s. On the other hand, the intergenerational correlation coefficient remained almost flat at about 0.43 without displaying any significant trend. As discussed in section 3.2.2, the different trends of relative mobility as measured by the regression and correlation coefficients is attributable to changes in the variances of the schooling distributions of children and parents over time. Neidhöfer, Serrano & Gasparini (2018), using datasets for 18 Latin American countries, also show a more stable trend for the correlation coefficient than the regression coefficient across the birth cohorts.

Intergenerational relative mobility is associated with a variety of country characteristics. According to van der Weide *et al*. (2021), the relationship between relative education mobility (as measured by the correlation coefficient) and per capita GDP follows a U-pattern across countries. Relative mobility is high in the poorest and richest countries, albeit for different reasons. van der Weide *et al.* (2021) argue that since most parents in the poorest countries have no education, parental education varies little across households. In such societies, it matters little to which household one is born into, indicating a high level of relative mobility. In the richest countries, differences in private investment in the education of children across households are partly offset by public interventions that give children from underprivileged parents the opportunity to fulfil their potential, which increases relative mobility. Overall, as with absolute education mobility, relative education mobility is positively correlated with variables linked to economic development (e.g., public expenditure in education, school quality and child health) and negatively correlated with proxies for inequality and segregation (van der Weide *et al*., 2021; Neidhöfer, Serrano & Gasparini, 2018).

We next discuss the empirical literature for South Africa. Before moving on, here is a non-exhaustive list of intergenerational education mobility studies for other countries. Narayan *et al*. (2018) for several countries around the world; Azomahou & Yitbarek (2016) for 9 Sub-Saharan

African countries; Behrman, Gaviria & Székely (2001), Torche (2021), Neidhöfer, Ciaschi & Gasparini (2021) for Latin American countries; Knight, Sicular & Yue (2011), Chen et al. (2015) for China; and Azam & Bhatt (2015) for India. Further, World Bank (2021) provides a comprehensive database of intergenerational education mobility estimates using both absolute and relative measures for countries around the world.

3.2.3.3 Intergenerational Education Mobility in South Africa

Several studies have examined intergenerational education mobility in South Africa. A summary of the studies and their findings is given in Appendix 3.A. Most of the studies are based on relative mobility measures, mainly using intergenerational regression and correlation coefficients.

The studies vary in their methods and target populations. This is not a limitation but renders comparison of mobility estimates across studies difficult. For instance, some studies estimate the independent effect of paternal and maternal education on children's education (e.g., Thomas, 1996; Burns & Keswell, 2012). However, this may not provide an estimate of the overall association between the education of parents and children. Similarly, both Nimubona & Vencatachellum (2007) and Kwenda, Ntuli & Gwatidzo (2015) control for several individual and household characteristics in estimating the association between the education of parents and children.

Some studies are based on coresident samples (e.g., Nimubona & Vencatachellum, 2007; Kwenda, Ntuli & Gwatidzo, 2015). As indicated in section 3.2.2, the co-residency sampling criteria, often used in household surveys, provides unrepresentative samples for the analysis of intergenerational mobility, and leads to biased intergenerational mobility estimates.

The available evidence on trends in relative education mobility across birth cohorts offers mixed results. For instance, Hertz *et al.* (2008) shows the trend in relative education persistence, as measured by the regression coefficient, to have been high and flat across the first half of their birth cohorts, followed by a gentle decline across the remaining half. Kwenda *et al.* (2015) also shows a similar trend – although the estimates of Hertz *et al.* are larger than those of Kwenda *et al.* across comparable birth cohorts. On the other hand, Burns & Keswell (2012), perhaps surprisingly, finds that intergenerational persistence increased between successive generations.

The fact that Hertz *et al.* (2008) and Kwenda *et al.* (2015) provide quantitatively different relative persistence estimates for comparable birth cohorts could be due to differences in methods and samples used. First, as indicated above, Kwenda *et al.* controls for several individual and household characteristics. Second, the studies differ in their target populations. While Kwenda *et al.* is based on the 10% sample drawn from the 2011 Census and considers only the African population group, Hertz *et al.* is based on the KwaZulu-Natal Income Dynamics Study (KIDS), which is a KwaZulu-Natal province sample of Africans and Indians/Asians. Third, Kwenda *et al*. may be affected by sample truncation as their analysis is based on children who co-reside with their parents.

We have noted the contributions of this chapter to the literature in section 3.1. In general, this chapter attempts to enrich and extend the South African literature on education mobility using datasets that are appropriate for the analysis of intergenerational mobility.

3.3 Methods

As indicated in section 3.2.2, there are two sets of summary measures of intergenerational mobility: namely, absolute and relative mobility measures. This section discusses the measures of absolute and relative mobility that we use in this chapter.

3.3.1 Absolute Education Mobility Measures

There are several measures of absolute mobility in the literature. In this chapter we use three measures; namely, upward mobility and two varieties of per capita mobility.

We define u*pward mobility* (UM) as the proportion of children who surpassed the educational attainment of their parents.

$$UM_k = Prob(S_{ijk}^c > S_{jk}^p) \qquad (3.1)$$

where i, j and k identify the i^{th} child in household j and birth cohort k. S^c and S^p denote the attained education of the child and his/her parents, respectively. Increasing values of UM across birth cohorts suggest an expansion of the provision of education.

We use two varieties of per capita mobility.[25] The first is the *mean absolute difference* (M1) and is given by:

$$\mathrm{M1}_k = E\left(\left|S_{ijk}^c - S_{jk}^p\right|\right) \qquad (3.2)$$

where i, j, k, S^c and S^p are as described in (3.1). M1 provides the expected value of the absolute difference between the educational attainments of children and their parents. It includes both those who have outperformed their parents (upwardly mobile) and those who have failed to outperform their parents (downwardly mobile) and hence is a measure of the average of total absolute mobility. Smaller values of M1 indicate that the educational attainment of children is not much different from that of their parents, and thus mobility is low.

The second measure of per capita mobility is the *mean difference* (M2) and is given by:

$$\mathrm{M2}_k = E\left(S_{ijk}^c - S_{jk}^p\right) \qquad (3.3)$$

where i, j, k, S^c and S^p are as described in (3.1). M2 provides the expected value of the difference between the educational attainments of children and their parents. While M1 considers only the magnitude of the difference, M2 considers both the magnitude and direction of the difference. The smaller the difference between M1 and M2, the lower the extent of downward mobility.

3.3.2 Relative Education Mobility Measures

We use three regression-based measures of relative mobility. These are intergenerational regression coefficient (IGRC), intergenerational correlation (IGC) and intergenerational rank correlation (IRC).

The IGRC is obtained from a simple regression model of child's education on parental education.

$$S_{ijk}^c = \beta_{0,k} + \beta_{1,k} S_{jk}^p + \varepsilon_{ijk} \qquad (3.4)$$

where i, j and k identify the i^{th} child in household j and birth cohort k. S^c and S^p denote the attained education of the child and his/her parents, respectively. β_0 and β_1 are the parameters of

[25] Per capita absolute mobility measures are proposed by Fields & Ok (1996)

the model and ε denotes the error term. β_1 is the IGRC. It provides a measure of the association between the educational attainments of parents and their children. It measures 'grade persistence' across generations (Hertz *et al.*, 2008). Larger values of β_1 imply higher levels of grade persistence across generations. $(1-\beta_1)$ is thought of as a measure of intergenerational mobility.

The IGC is obtained from a simple regression of child's education on parental education but with both education variables expressed as a ratio of the standard deviation of their respective distributions for a given birth cohort k. Thus, equation 3.4 modifies to:

$$\frac{S_{ijk}^{c}}{\sigma_k^{c}} = \rho_{0,k} + \rho_{1,k} \cdot \frac{S_{jk}^{p}}{\sigma_k^{p}} + \upsilon_{ijk} \qquad (3.5)$$

where $i, j, k,$ S^{c} and S^p are as described in (3.4). ρ_0 and ρ_1 are the parameters of the model, and υ is the error term. σ^c and σ^p denote the standard deviations of the distribution of education of children and parents, respectively. ρ_1 is the IGC. It provides a measure of the association between the educational attainments of parents and their children. Hertz *et al.* (2008) refer to the IGC as measuring 'standardized persistence' across generations. Larger values of ρ_1 imply higher levels of persistence in educational attainments across generations. $(1-\rho_1)$ is thought of as a measure of intergenerational mobility. As discussed in section 3.2.2, the IGC can also be obtained from the IGRC as the two are related as: $\widehat{\rho_{1k}} = \widehat{\beta_{1k}}.\frac{\sigma_k^{p}}{\sigma_k^{c}}$, for a given birth cohort k.

Our third measure of relative mobility is obtained, once again, from a simple regression model of child's education on parental education but with both education variables expressed as percentile ranks. However, while the percentile ranks of children are relative to their birth cohorts, the percentile ranks of parents are relative to the full sample of parents. Thus, equation 3.4 modifies to:

$$R_{ijk}^{c} = \delta_{0,k} + \delta_{1,k} R_j^{p} + \zeta_{ijk} \qquad (3.6)$$

where R_{ijk}^{c} denote the percentile rank of the i^{th} child from household j among children in birth-cohort k, and R_j^{p} denote the percentile rank of the parent from the j^{th} household among all parents. δ_0 and δ_1 are the parameters of the model and ζ is the error term. δ_1 is the IRC. It provides a measure of the association between the relative positions of parents and their children

in their respective education distributions. Larger values of δ_1 imply higher levels of persistence in educational outcomes across generations. $(1 - \delta_1)$ is thought of as a measure of intergenerational mobility.

As indicated in section 3.2.2, the three regression-based measures differ in their effectiveness in capturing the extent of relative mobility. However, we use all three as they offer alternative views of the trend of relative mobility across birth cohorts. Moreover, as also indicated in section 3.2.2, they differ in their robustness to various data quality issues.

The overall level of mobility for each measure (both the absolute and relative measures) is obtained by taking a simple average of the mobility estimates across the birth cohorts. This approach has an advantage as it gives equal weight to all birth cohorts and thus treats each sample cohort as a representative of its population birth cohort (Hertz *et al.*, 2008). The time trend of the relative education mobility estimates across the birth cohorts is obtained by running a simple regression of the mobility estimates on time.

3.4 Data and Descriptive Statistics

3.4.1 The Datasets

The empirical analysis is based on the two datasets described in chapter 1.[26] These are the Survey of Socioeconomic Opportunity and Achievement (SSOA) and the wave 5 dataset of the National Income Dynamics Study (NIDS-w5). Both datasets possess the required information for the analysis of intergenerational education mobility.

We work with samples of adults who are between 25 and 69 years of age. The lower age limit is set to 25 years to allow time for the completion of schooling. The upper age limit is constrained to 69 years because longevity and education are positively correlated (Hoque *et al*., 2019). Hertz *et al.* (2008) notes that to the extent that appearance in older cohorts and education are correlated, the sampling selection bias in favour of the more educated could produce a downwardly biased estimate of the intergenerational regression coefficient.

[26] Additional information on the datasets is given in Appendix 1.A.

Given the age range from 25 to 69 years, the SSOA sample is born between 1923 and 1967[27], and the NIDS-w5 sample is born between 1948 and 1992. The trend analysis is based on 5-year birth cohorts. The choice of the 5-year interval for the birth-cohorts is rather arbitrary. However, it is deemed that a 5-year interval is long enough to give us an adequate number of observations for a given birth cohort and is short enough to give an adequate number of coefficient-estimates to trace the trend of intergenerational mobility over time.

The educational attainment of both children and parents is measured in terms of their completed years of education (excluding any repeated grades). Appendix 3.B presents the mapping table that we used to convert and harmonize the reported highest education levels into completed years of education across the two datasets. The conversion table is adapted from Branson & Leibbrandt (2013).

We measure parental education as the simple average of mother's and father's education where both are reported. The correlation coefficient between the education of mothers and fathers is quite strong in both datasets: 0.77 in SSOA and 0.73 in NIDS-w5. These strong correlations justify the use of the average education of parents to represent parental education. For cases where the education of only one parent is reported, that educational level is assumed as the parental education.

The mobility analysis is carried out for the full sample as well as the African sample in each dataset. As discussed in chapter 2, the African population group is historically the most disadvantaged racial group and is by far the largest in terms of population size. The analysis is not carried out by gender. The gender gap in educational attainment is quite small in South Africa (Branson & Lam, 2010; Branson & Leibbrandt, 2013).

As indicated in chapter 1, our datasets differ in their degree of representation of the South African population. While NIDS is a nationally representative sample, SSOA is not. SSOA has national geographic coverage but missed farmworkers who were living on commercial farms. The fact that SSOA is not a truly representative sample poses a challenge in analyzing mobility changes

[27] The age range for SSOA assumes that the survey was conducted in 1992.

between the two periods. However, if the direction of the bias in the mobility estimates for SSOA are known, then it is possible to assess the direction of the change in mobility between the two periods. Farmworkers who were living on commercial farms were structurally limited in their ability to accumulate human and physical capital and pass it on to their children. In the pre-1994 era, farmworkers were the most disadvantaged group in terms of schooling opportunities (Christie and Gaganakis, 1989). Moreover, farmworkers who were living on commercial farms were more likely to have higher intergenerational transmission of occupations given the indentured nature of labour in the farms. Thus, we could assume that intergenerational education persistence among farmworkers who were living on commercial farms was high. SSOA, by not including them, is thus likely to provide (relative) persistence estimates that are biased downward.

3.4.2 Descriptive Statistics

The mobility analysis is based on 'children' with non-missing education both for themselves and their parents. The proportions of children with non-missing education both for themselves and their parents are 91.2 and 95.0 per cent in the SSOA and NIDS-w5 samples, respectively.

Table 3.1 presents the average education of the children with respect to some background characteristics. The table shows that the average education in NIDS is higher compared to that in SSOA. This agrees with the well documented finding that the average education of South Africans has been increasing over time. The table also shows that the gap in average education across the categories of gender, race and urban/rural divide have declined over time (between SSOA and NIDS-w5). Narrowing gaps in educational attainment over time by background characteristics has been documented in other studies (e.g., Branson & Leibbrandt (2013) for gender; Thomas (1996) and Branson & Lam (2022) for race).

Table 3.1 Average education of children by background characteristics

	SSOA Birth year range: 1923-67		NIDS-w5 Birth year range: 1948-92	
	Sample size	Ave. Educ	Sample size	Ave. Educ.
All	6,462	7.86 (.139)	18,842	10.09 (0.075)
Gender				
Female	3,039	7.35 (0.168)	10,960	10.03 (0.089)
Male	3,423	8.38 (0.188)	7,874	10.17 (0.083)
Race				
African	3,488	6.50 (0.186)	14,617	9.89 (0.074)
coloured	564	8.17 (0.254)	2,570	9.54 (0.285)
Asian/Indian	609	9.21 (0.206)	374	11.14 (0.357)
white	1,801	12.01 (0.099)	1,280	12.81 (0.155)
Urban/Rural				
Urban	5,438	9.01 (0.199)	10,369	10.70 (0.105)
Rural	1,024	5.55 (0.248)	7,489	8.85 (0.105)

Notes: (1) The table shows the average education of children by background characteristics in SSOA and NIDS-w5. The table is based on respondents between 25 and 69 years of age with non-missing education both for themselves and their parents.
(2) Results are weighted using the calibrated weights provided with the datasets.
(3) Numbers in parentheses are standard errors.

Table 3.2 presents the average education of children and their parents by 5-year birth-cohorts of the children. The table shows that while the average education of children has increased across the birth cohorts in each dataset, the standard errors have decreased. This is especially evident in NIDS-w5. As indicated in the introduction of this chapter, the average education of South Africans has been increasing across birth cohorts coupled with a decline in educational inequality. However, we should also note that the decrease in the standard errors (in Table 3.2) may have been enhanced by the increasing sample size across the birth cohorts.

The table further shows that the average education of both children and their parents across the same birth cohorts is higher in SSOA than in NIDS. The higher average education in SSOA is possibly due to the exclusion of farmworkers who were living on commercial farms in that survey. As indicated in the previous section, farmworkers were the most disadvantaged group in terms of schooling opportunities in the pre-1994 era.

Table 3.2 Average education of children and their parents by 5-year birth cohorts

	SSOA			NIDS-w5		
Birth cohort	Sample Size	Child ave. educ	Parent ave. educ	Sample size	Child ave. educ	Parent ave. educ
1988-92				3,774	11.22 (0.062)	6.98 (0.145)
1983-87				3,199	11.17 (0.076)	6.47 (0.147)
1978-82				2,349	10.83 (0.104)	5.17 (0.196)
1973-77				2,031	10.51 (0.128)	4.12 (0.217)
1968-72				1,894	9.70 (0.175)	4.02 (0.224)
1963-67	1,263	9.12 (0.208)	5.80 (0.218)	1,711	8.70 (0.174)	3.32 (0.190)
1958-62	1,065	9.03 (0.226)	5.26 (0.242)	1,562	8.23 (0.226)	3.36 (0.233)
1953-57	1,014	8.38 (0.211)	5.20 (0.242)	1,303	7.47 (.256)	2.78 (0.225)
1948-52	924	7.39 (0.290)	4.30 (0.257)	1,019	6.43 (0.300)	2.91 (0.280)
1943-47	647	7.33 (0.341)	4.53 (0.331)			
1938-42	524	6.53 (0.323)	4.17 (0.288)			
1933-37	400	6.29 (0.418)	4.04 (0.391)			
1928-32	343	5.60 (0.441)	3.42 (0.348)			
1923-27	282	5.43 (0.410)	3.56 (0.332)			

Notes: (1) The table shows the average education of children and their parents across 5-year birth cohorts of the children in SSOA and NIDS-w5. The table is based on children between 25 and 69 years of age with non-missing education both for themselves and their parents.
(2) Results are weighted using the calibrated weights provided with the datasets.
(3) Numbers in parentheses are standard errors.

3.5 Empirical Results

This section presents the trends in intergenerational education mobility in SSOA and NIDS. We first present the trends based on the absolute mobility measures and then the trends based on the relative mobility measures. The trend analysis is carried out for the full sample as well as the African sample in each dataset. We check the robustness of the trend estimates by redefining parental education as the highest education of parents.

3.5.1 Trends in Absolute Mobility

3.5.1.1 Average Education across Birth-years

Figure 3.1 compares the average education of children and their parents across single-year birth cohorts of children in SSOA and NIDS-w5. In each graph, each dot represents the average education of children born in a given year, and the corresponding point down in the 45-degree line represents the average education of the parents of those children. Panel A is based on the full samples and Panel B is based on the African samples of SSOA and NIDS-w5.

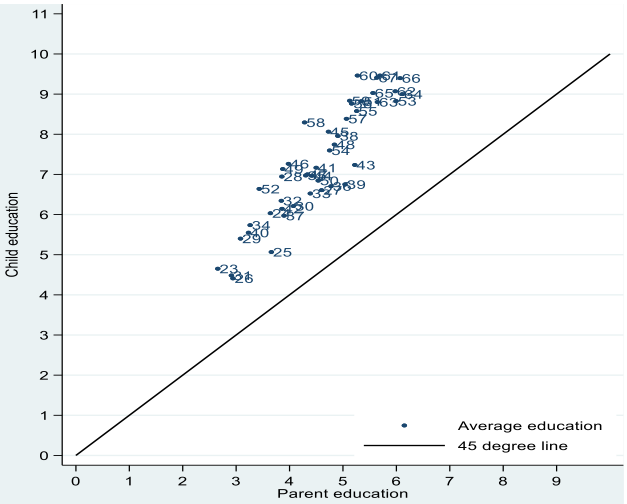
The graphs show that the average education of each single-year birth cohort of children is higher than that of their parents in each sample. The gap between the average education of children and their parents is quite small in the older birth cohorts of SSOA. However, the gap is seen to have increased over successive birth cohorts and is largest for those born in the 1970s (see Figure 3.1).

The figures for SSOA show that the average education of both children and their parents increased across the birth cohorts - both in the full and African samples. The pattern in NIDS, however, is different. In NIDS, the average education of both children and their parents increased across the older birth cohorts. However, in the younger birth cohorts, only the average education of parents kept increasing while the average education of children remained almost stagnant both in the full and African samples (Figure 3.1b&d). The slowdown in the increase of the average education of young South Africans has also been documented in Branson & Lam (2022). We will see in the next section that the stagnation in educational attainment among the younger cohorts has implications to the trend in absolute education mobility.

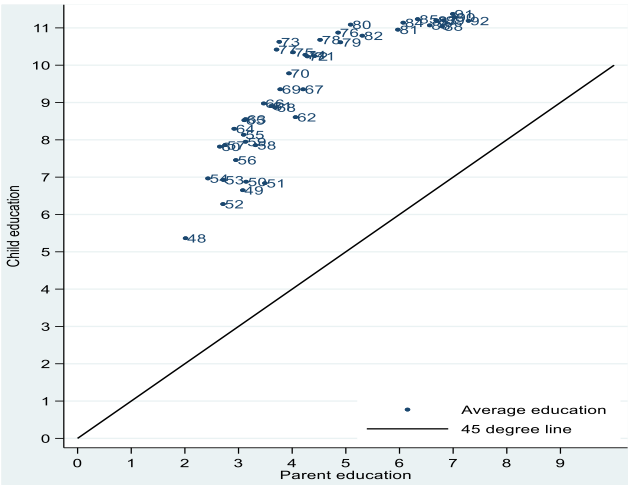
Figure 3.1 Average education of children and parents by birth-year of children

Panel A: Full sample

(a) SSOA

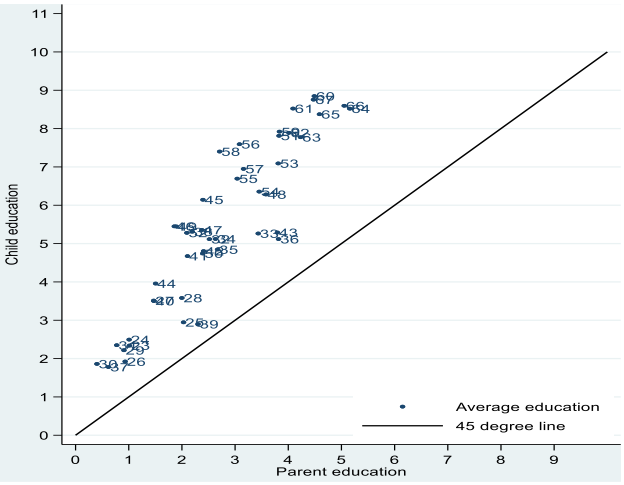


(b) NIDS-w5

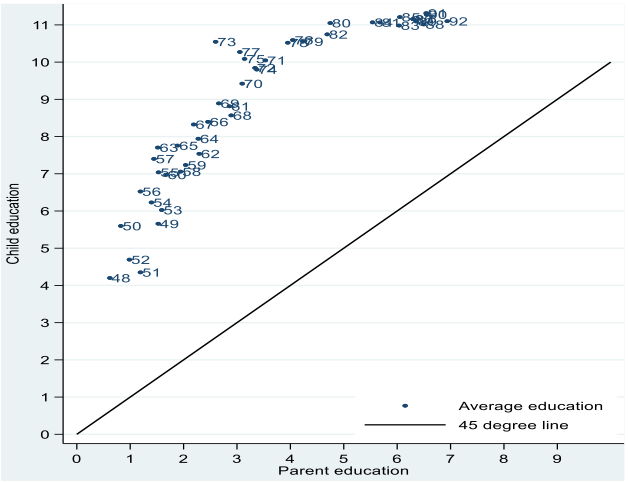


Panel B: African sample

(c) SSOA



(d) NIDS-w5



Notes: (1) The figures present the average education of children and their parents by single-year birth cohorts of children in SSOA and NIDS-w5. In each graph, each dot represents the average education of children born in a given year and the corresponding point down in the 45-degree line represents the average education of their parents. The dots are labelled with the year of birth of children. Panel A is based on the full sample and Panel B is based on the African sample. The samples are based on respondents between 25 and 69 years of age.
(2) SSOA birth-year range is 1923-67, and NIDS-w5 birth-year range is 1948-92.
(3) Parental education is defined as the average education of parents.
(4) Results are weighted using the calibrated weights provided with the datasets

The magnitude and direction of the difference in average education between that of children and their parents across the single-year birth cohorts of children in Figure 3.1 provides a useful insight into the pattern of absolute mobility over time. In the next section, we explore the trends in the educational attainment of children relative to their parents' using our three measures of absolute mobility.

3.5.1.2 Trends in Absolute Mobility

Figure 3.2 presents the trend in education mobility using the three measures of absolute mobility (discussed in the methods section) across the five-year birth cohorts of children.

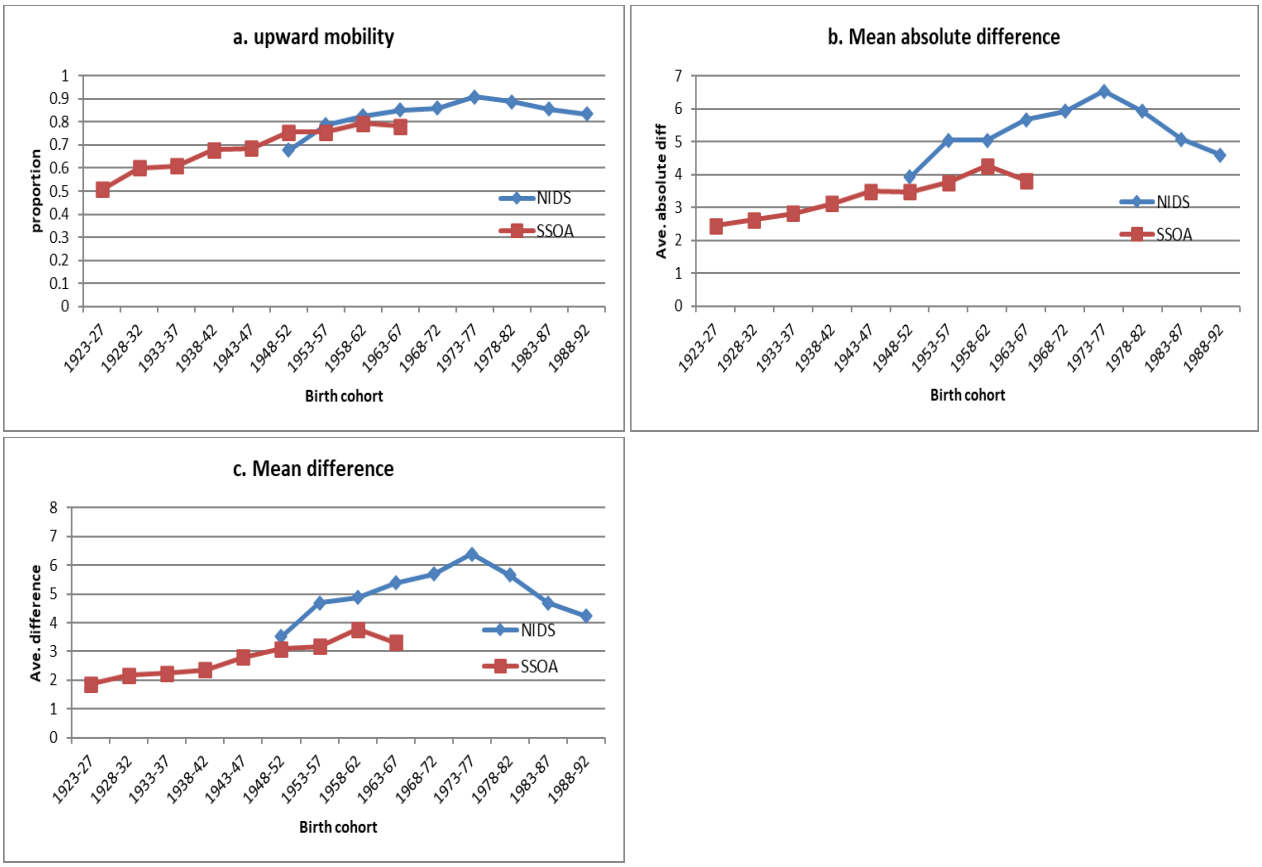
The proportion of children who attain higher education than their parental education, i.e. the upwardly mobile, has been increasing across birth cohorts. Upward mobility is seen to increase across the birth cohorts of SSOA and over many of the initial birth cohorts of NIDS (Figures 3.2a&d). This suggests an expansion in the provision of education over time. Upward mobility peaked for the 1973-77 birth cohort, as evident in NIDS, and gradually declined for birth cohorts born thereafter (Figure 3.2a&d). At its peak, upward mobility was about 91 per cent in the full sample and 93 per cent in the African sample. This means, in the full sample, about 91 per cent of children among the 1973-77 birth cohort attained education levels that are higher than their parental education. In the most recent birth cohort (i.e., the 1988-92 birth cohort), upward mobility stood at 83.5 per cent in the full sample and 85 per cent in the African sample. The estimates for those born in the early 1990s are comparable to upward mobility estimates for South Africa by Alesina *et al.* (2021). Alesina *et al.* estimates upward mobility to have been about 81.4 per cent among children between 14 and 25 years of age, using datasets collected over the 1996-2011 period.[28] The decline in upward mobility in the younger birth cohorts may not mean a reversal in schooling enrolment rates, but rather lower educational attainment of children relative to their parental education among the enrolled. The reason for this can be found in the trends of our two measures of per capita mobility, to which we now turn to.

[28]Alesina *et al.* (2021) defines upward mobility as the proportion of children who completed at least primary education born to parents who have not completed primary education, with parental education defined as the average education of parents.

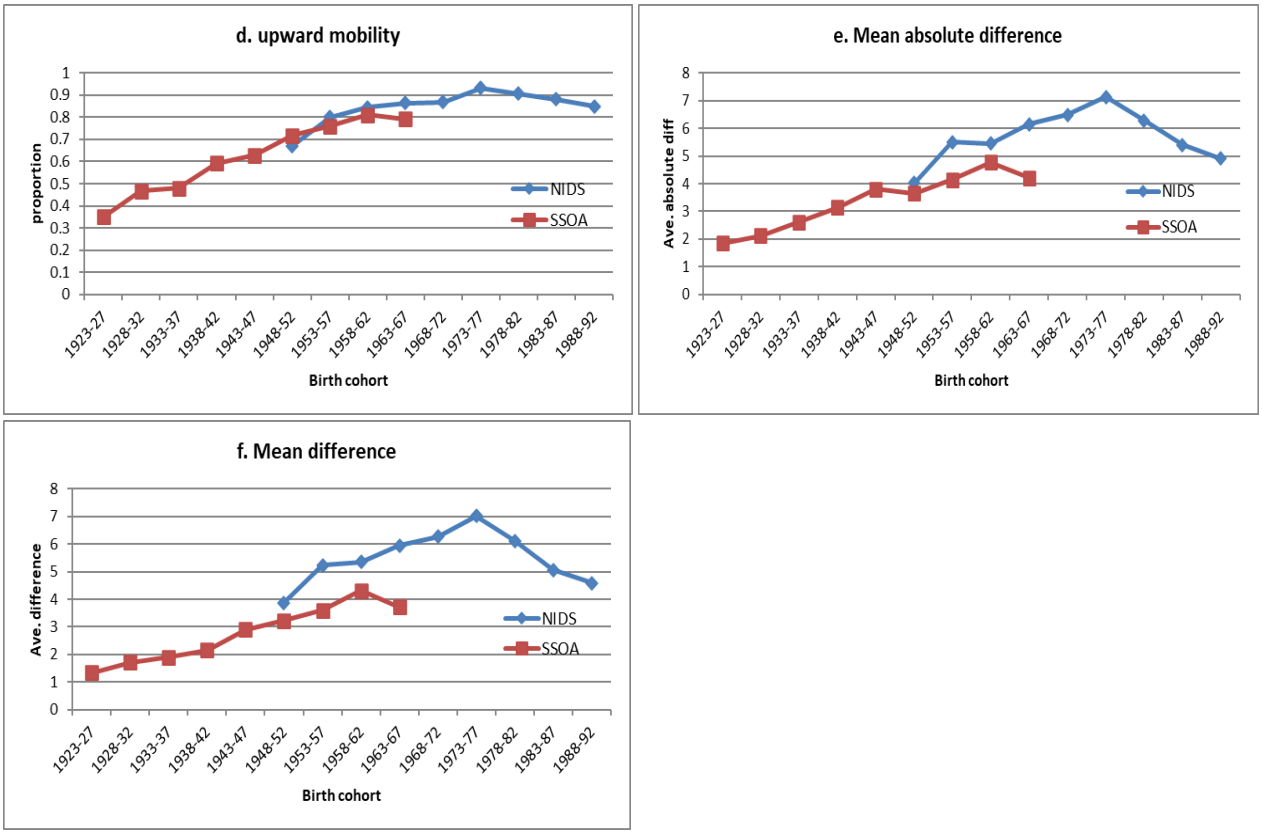
Although upward mobility provides useful information on the extent of absolute mobility, it does not provide information on the magnitude of mobility. Measures of per capita mobility complement upward mobility by providing information on the magnitude of mobility. As indicated in section 3.3.1, we use two measures of per capita mobility; namely, the mean absolute difference and the mean difference.

Figure 3.2 Trends in absolute intergenerational mobility

Panel A: Full sample



Panel B: African sample



Notes: (1) The figures present the trends in absolute mobility across 5-year birth cohorts using three measures of absolute mobility (namely, upward mobility, mean absolute difference and mean difference) in SSOA and NIDS-w5. Panel A is based on the full sample and Panel B is based on the African sample.
(2) Parental education is defined as the average education of both parents.
(3) Results are weighted using the calibrated weights provided with the datasets.

As discussed in section 3.3.1, the mean absolute difference measures the average absolute difference between the educational attainment of children and their parents. The trend in mean absolute difference is shown in Figures 3.2b&e – for the full and African samples. The mean absolute difference is seen to increase across the birth cohorts of SSOA and for many of the initial birth cohorts in NIDS, indicating an increasing gap between the educational attainment of children and their parents over those birth cohorts. In NIDS, the trend is inverted V shaped, peaking for the 1973-77 birth cohort at about 6.5 years and 7.1 years for the full and Africa samples, respectively, and decreasing thereafter.

The mean difference, which takes into consideration the direction of mobility of individuals relative to their parents, further refines the results of the per capita mobility measured by the mean

absolute difference. The trends in mean difference (shown in Figures 3.2c&f) are quite similar to the trends in mean absolute difference. The mean difference between the educational attainment of children and their parents at its peak (in NIDS) is estimated to be about 6.4 years in the full sample and 7.0 years in the African sample.

The mean difference estimates are quite close in magnitude to the mean absolute difference estimates across the older birth cohorts. This suggests that the chances of downward mobility of children in these birth cohorts were quite low. On the other hand, across the younger birth cohorts the mean difference estimates are smaller in magnitude than the mean absolute difference estimates. This suggests that the chances of downward mobility of children have grown higher in these birth cohorts. The decline in upward mobility in the younger birth cohorts thus may not mean a reversal in schooling enrolment rates, but rather increasing lower attainments of children relative to their parental education.

All three measures of absolute mobility show a declining trend among the younger birth cohorts in NIDS. This is further probed by asking whether the decline is prompted by the bounded nature of education, which limits the chances of children from education-rich parents to surpass the educational attainment of their parents. We investigate this by re-coding parental education to grade 12 for children whose parental education is above grade 12. The re-coding provides children from education-rich parents the space to surpass their parental education level and thus the bounded nature of education would not be an issue.[29] The results based on this approach (shown in Appendix 3.C) are quite similar to the trends based on the original parental education; per capita mobility is observed to decline across the younger birth cohorts. This suggests that the decline in per capita mobility in the younger birth cohorts is unlikely to be explained by the bounded nature of education. As we have seen in Figure 3.1, the declining trend of absolute mobility in NIDS corresponds to a period where, across the single-year birth cohorts of children, while the average education of children remained largely unchanged, the average education of parents was increasing. The stagnating educational attainment of children coupled with increasing

[29] Some studies (e.g., van der Weide *et al.*, 2021) have dealt with this issue by restricting their sample to children whose parental education is below tertiary level of education. However, this procedure may introduce sample selection bias.

parental education thus have given rise to a decline in per capita mobility across the younger birth cohorts.

The absolute mobility trends show discrepancies in the overlapping birth cohorts from our two datasets, both for the full and African samples (see Figure 3.2). These discrepancies are likely due to differences in the degree of representation of the two datasets. As indicated in chapter 1, NIDS is a representative sample, but SSOA missed farmworkers who were living on commercial farms. As noted in section 3.4.1, farmworkers were the least educated group, thus SSOA is likely to represent a more educated population, on average. Indeed, the education of parents in SSOA is particularly high as shown in Table 3.2 and this has given rise to smaller per capita mobility estimates in SSOA compared to NIDS across the overlapping birth cohorts.

In sum, the trend in absolute education mobility shows that absolute mobility has been on the rise across the older birth cohorts but seems to be declining among more recent cohorts.

3.5.1.3 Robustness Check

The robustness of the trends in absolute mobility is probed by redefining parental education as the highest education of parents. The results are given in Appendix 3.D. The trends are qualitatively similar to the trends obtained with parental education defined as the average education of parents. However, with the increase in parental education from the average to the highest education of parents, the mobility estimates have decreased across the birth cohorts. One notable change is in the trend for upward mobility. When parental education is defined as the average education of parents, the trend of upward mobility for those born after the mid-1970s is seen to decline gradually. On the other hand, when parental education is defined as the highest education of parents, the trend is seen to decline rather rapidly both in the full and African samples.

3.5.1.4 Summary of the Levels of absolute mobility

Table 3.3 presents a summary of the average levels of absolute mobility in the two datasets - both for the full and African samples. All three measures of absolute mobility indicate that mobility has substantially increased between the early 1990s (in SSOA) and two-and-a-half decades later (in NIDS-w5). For instance, the percentage of children who surpassed their parental education

was about 68.6 per cent in SSOA and 83.2 per cent in NIDS (in the full samples). The table also shows considerable improvement in per capita mobility between the two periods.

Table 3.3 Summary of the levels of absolute mobility

	Upward mobile (proportion)	Mean Absolute Difference	Mean Difference
SSOA			
Full	0.686	3.313	2.757
	(0.032)	(0.202)	(0.210)
African	0.622	3.367	2.766
	(0.054)	(0.334)	(0.343)
NIDS-w5			
Full	0.832	5.303	5.013
	(0.023)	(0.264)	(0.288)
African	0.846	5.706	5.492
	(0.025)	(0.308)	(0.317)

Notes: (1) The table provides a summary of the overall level of absolute mobility based on the three measures of absolute mobility (namely, upward mobility, mean absolute difference and mean difference) in SSOA and NIDS-w5. The summaries are based on the baseline results. The overall level of absolute mobility for each measure is obtained by taking a simple average of the mobility estimates across the birth cohorts. The mobility summaries are for the full and African samples in each dataset.
(2) Parental education is defined as the average education of parents.
(3) Numbers in parentheses are standard errors.

3.5.2 Trends in Relative Mobility

3.5.2.1 Cross Tabulation between the Education of Parents and Children

Table 3.4 presents a cross-tabulation between the education levels of parents and their children in SSOA and NIDS. The entries in each cell are the row percentage and the cross frequency, respectively. The table shows that the row percentages to the left of the main diagonal are generally lower than the percentages to the right of the main diagonal. This indicates upward mobility of children relative to their parents' education.

The table also shows that the educational levels of parents and their children are positively correlated. For instance, the chances of attaining grade 12 or higher are shown to increase with parental level of education in both SSOA and NIDS. In NIDS the chances of attaining postsecondary education increases from about 7.9 per cent for children born to parents with no education to 61.4 per cent for children born to parents with postsecondary education. Further,

although the educational levels of parents and children are positively associated in both datasets, the chances of attaining higher levels of education for children from parents with lower levels of education are generally higher in NIDS than in SSOA. This shows changes in educational association across generations over time.

In the next section, we use regression-based measures of relative mobility to explore changes in the association between the educational attainment of parents and their children over time.

Table 3.4 Cross-tabulation between the education of parents and children

(a) SSOA

parent education		child education					
		None	Primary	Inc. Sec.	Matric	Postsec	Total
None		37.8%	44.5%	13.9%	2.9%	0.9%	100%
		424	685	250	64	24	1447
Primary		5.8%	44.6%	33.5%	12.5%	3.7%	100%
		116	958	869	331	133	2407
Inc. Sec.		0.7%	13.3%	40.6%	31.1%	14.3%	100%
		13	150	575	472	272	1482
Matric		0.3%	3.7%	21.8%	41.4%	32.8%	100%
		1	19	110	244	209	583
Postsec		0.3%	5.0%	9.2%	24.5%	61.1%	100%
		2	8	46	144	343	543
Total		13.1%	33.0%	27.4%	16.3%	10.3%	100%
		556	1820	1850	1255	981	6462

(b) NIDS

parent education		child education					
		None	Primary	Inc. Sec.	Matric	Postsec.	Total
None		12.4%	27.4%	43.5%	8.9%	7.9%	100%
		1182	2098	2499	531	449	6759
Primary		1.1%	14.5%	52.6%	18.7%	13.2%	100%
		77	836	2384	771	591	4659
Inc. Sec.		0.8%	4.9%	45.6%	24.1%	24.6%	100%
		45	297	2137	1054	1015	4548
Matric		0.3%	2.2%	32.0%	28.2%	37.3%	100%
		7	30	463	495	548	1543
Postsec.		0.2%	1.4%	19.2%	17.8%	61.4%	100%
		3	13	254	293	770	1333
Total		4.4%	13.5%	43.0%	18.0%	21.1%	100%
		1314	3274	7737	3144	3373	18842

Notes: (1) The table provides a cross-tabulation between children's and parental education levels in SSOA and NIDS-w5.
(2) The education categories represent: None for 'no schooling'; Primary for grades 1 to 7; Inc. Sec. for grades 8 to 11; Matric for grade 12; and Postsec for above grade 12.
(3) Parental education is defined as the highest education of parents.
(4) Results are weighted using the calibrated weights provided with the datasets.

3.5.2.2 Trends in Relative Mobility

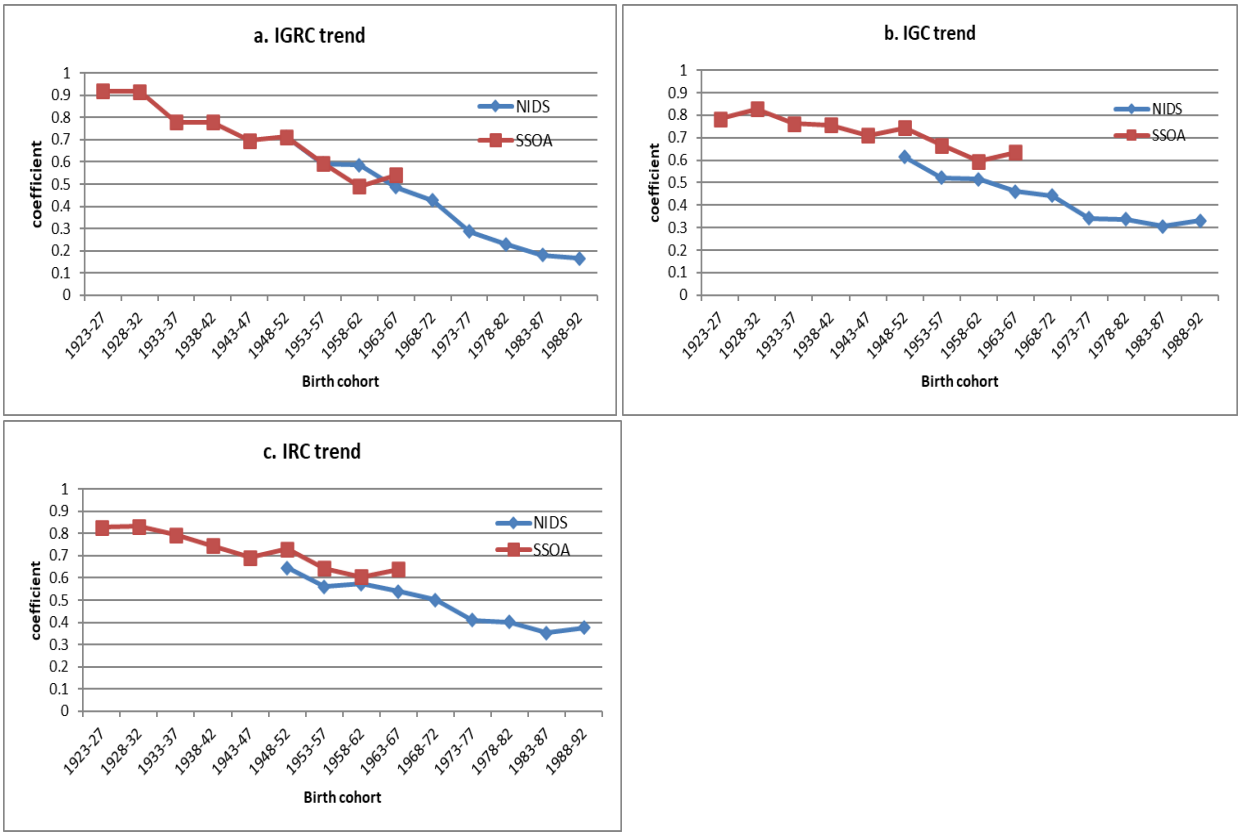
Figure 3.3 presents the trend in educational persistence using the three regression-based measures of relative mobility (discussed in the methods section) across the five-year birth cohorts of children. The trends show that the degree of association in educational attainment across generations has been declining across the birth cohorts.

The trends in the intergenerational regression coefficient (IGRC) show a continually declining pattern of persistence across the birth cohorts of SSOA and NIDS (Figure 3.3a). In the oldest birth cohort of SSOA, an additional year of parental education is associated with an average increase of

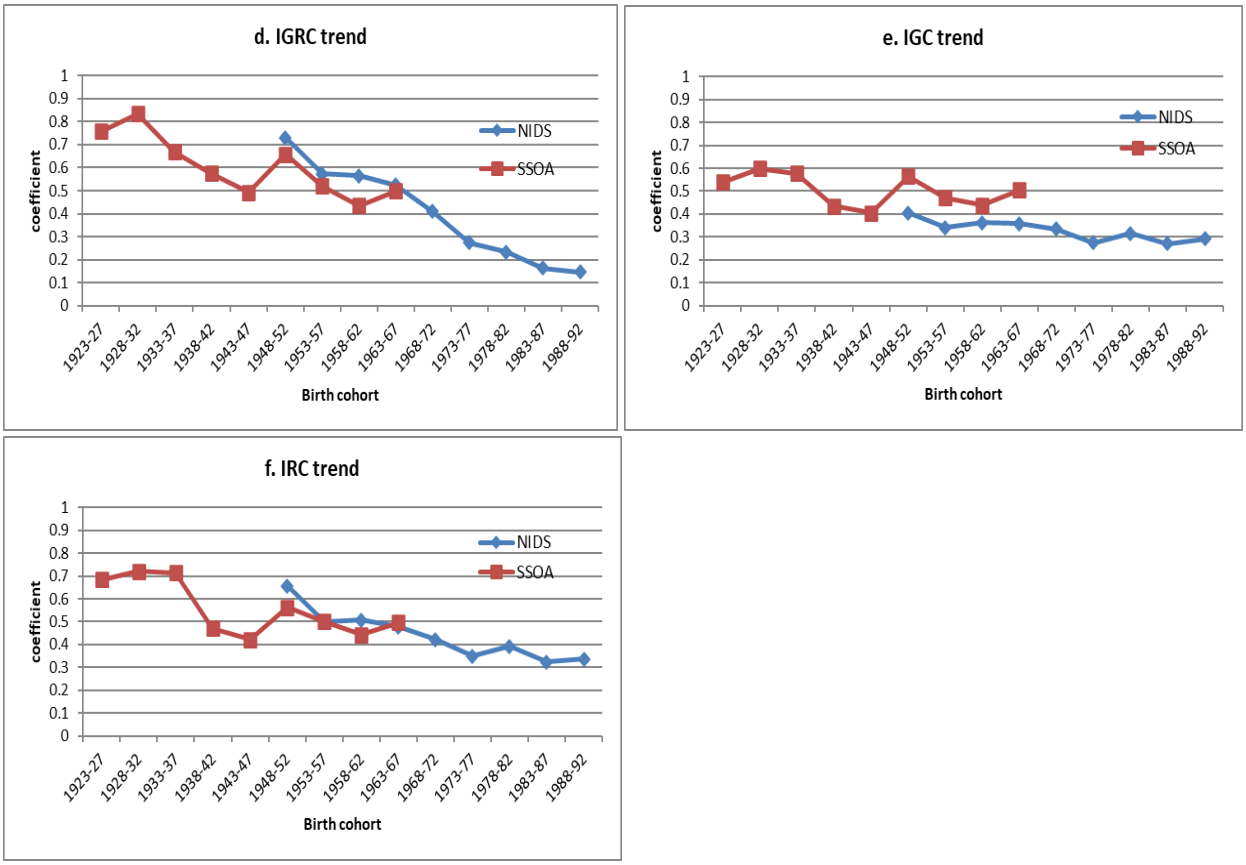
0.92 years in the child's education (that is almost one for one). On the other hand, in the youngest birth cohort of NIDS, the same scenario leads to an average increase of only 0.17 years in the child's education.

Figure 3.3 Trends in relative persistence in SSOA and NIDS-w5

Panel A: Full sample



Panel B: African sample



Notes: (1) The figures present the trends in relative mobility across 5-year birth cohorts using three measures of relative mobility (namely, IGRC, IGC and IRC) in SSOA and NIDS-w5. Panel A is based on the full sample and Panel B is based on the African sample in each dataset.
(2) IGRC, IGC and IRC refer to intergenerational regression coefficient, intergenerational correlation coefficient, and intergenerational rank coefficient, respectively.
(3) Parental education is defined as the average education of parents.
(4) Results are weighted using the calibrated weights provided with the datasets.

The decline of the intergenerational correlation coefficient (IGC) across the birth cohorts is slower compared to the decline of IGRC (compare Figures 3.3a&b). As indicated in section 3.3.2, IGC and IGRC are related as: $IGC = IGRC \cdot (\sigma^p/\sigma^c)$, where σ^p and σ^c denote the standard deviations of the education of parents and children, respectively. In both SSOA and NIDS, the ratio of the standard deviations (σ^p/σ^c) is seen to increase across the birth cohorts (see Appendix 3.E). This is especially true in the younger birth cohorts of NIDS. The slow decline in the IGC across the birth cohorts is thus due to the increase in the ratio of the standard deviations, which offsets the rapid decline in the IGRC. The decline in the IGC for Africans in Figure 3.3e is also slower compared to the decline in IGRC for Africans for the same reasons.

The trends in the intergenerational rank correlation (IRC) are shown in Figures 3.3c&f - for the full and African samples.[30] As discussed in section 3.3.2, the percentile ranks of children are relative to their birth cohorts and the percentile ranks of parents are relative to the full sample of parents.[31] The trends in IRC in Figure 3.3c show a continually declining pattern of relative persistence across the birth cohorts of SSOA and NIDS. In the oldest birth cohort of SSOA, a 10 percentile points increase in parental education rank is associated with an average increase of 8.3 percentile points in the child's education rank. On the other hand, in the youngest birth cohort of NIDS, the same scenario leads to an average increase of only 3.8 percentile points in the child's education rank.

All the three measures of relative mobility show a relatively lower estimate of persistence for the oldest birth cohort of SSOA. This is especially evident in the trends for the African sample. This is likely due to the relatively larger missing education observations in that birth cohort. To the extent that the observations in that birth cohort are unrepresentative of their population birth cohort, the estimate of intergenerational persistence for that cohort is likely to be biased downward (Solon, 1992).

The trend in relative persistence for the African sample in SSOA is not as linear as that for the full sample. All the three measures show persistence levels that suddenly increase for the 1948-52 birth cohort. The timing is rather interesting. The 1948-52 birth cohort can be assumed to have been in primary school during the second half of the 1950s, which corresponds to the early implementation period of the Bantu education system. As indicated in chapter 2, the Bantu education system expanded schooling opportunities for Africans in the mid-1950s. The sudden rise in persistence for the 1948-52 birth cohort could thus be linked to the change in education policy for Africans. The expansion of schooling opportunities for Africans may have resulted in

[30] In the case of ties, individuals are assigned the mean rank of the individuals with that level of education. For instance, if 10 per cent of children in a birth cohort have no education, all the children with no education are assigned a percentile rank of 5 per cent.

[31] In the African sample, the percentile ranks of parents are relative to the full sample of African parents.

educational attainments of African children that systematically differed by parental education, which raises persistence in intergenerational attainment.

As we have seen with the trends for the measures of absolute mobility, the trends for the measures of relative mobility also show discrepancies in the overlapping birth cohorts from our two datasets (see Figure 3.3). We have already indicated in section 3.4.1 that SSOA is likely to provide persistence estimates that are biased downward as it missed farmworkers who were living on commercial farms. In addition, we need to consider the likely effects of changes over time in the composition of the population on education persistence. To the extent that changes in the demographic composition are systematically correlated with education, the persistence estimates for the same birth-cohorts may change over time. For instance, empirical evidence has shown that education significantly reduces the hazard of acquiring HIV in South Africa (Bärnighausen, *et al.*, 2007). This suggests that appearances in the older birth cohorts of NIDS and education are positively correlated. Given that absolute education mobility has been high in the older birth cohorts of NIDS, this selection bias that favours the more educated is thus likely to provide persistence estimates that are smaller than what could have been. Our empirical results (in Figure 3.3) show that the expected downward bias of persistence in SSOA seems to have larger effect on IGRC, especially in the African sample. On the other hand, the downward effect of the change in absolute education mobility on persistence seems to be larger on IGC. Interestingly, IRC provides persistence estimates that show less discrepancies across the overlapping birth cohorts of our two datasets.

A comparison of our trend estimates with other studies shows that the IGRC and IGC trends for NIDS (using the full sample) are comparable to those reported in the GDIM[32] (World Bank, 2021). Moreover, the IGRC and IGC trends for NIDS using the African sample are qualitatively similar to those reported in Kwenda, Ntuli & Gwatidzo (2015). However, Kwenda *et al.* controls for other variables, which makes it difficult to directly compare the magnitudes of the estimates in the two studies. Further, the IGC trend for the African sample in NIDS is similar to that reported

[32] GDIM is an acronym for Global Database on Intergenerational Mobility. The GDIM estimates for South Africa are based on the 4th wave of NIDS.

in Hertz *et al.* (2008) in terms of the invariability of the trend over time. As discussed in the literature review section, Hertz *et al.* used the KwaZulu-Natal Income Dynamics Study (KIDS).

3.5.2.3 Robustness Check

The robustness of the trends in relative mobility is probed by redefining parental education as the highest education of parents. The results are given in Appendix 3.D. The trends are qualitatively similar to the trends obtained with parental education defined as the average education of parents. One notable difference is that when parental education is defined as the highest education of parents the IGRC estimates are lower, but the IGC and IRC estimates remain largely unchanged. This shows that the change in the definition of parental education has impacted the association in educational outcomes more through changes in the variability of parental education than through actual changes in the correlation between the education of parents and children.

3.5.2.4 Summary of the Levels and Trends of relative mobility

Table 3.5 presents a summary of the average levels of relative persistence and the time trend of persistence in the two datasets. The table shows that intergenerational persistence (mobility) in education has decreased (increased) considerably between the early-1990s (SSOA) and two-and-a half decades later (NIDS). All the three measures have decreased between the two periods both for the full and the African samples.

The time trend of persistence shows that relative persistence has been declining over time. The decline in persistence is faster as measured by IGRC than by either IGC or IRC both for the full and African samples in the two datasets. The declining trends of IGC and IRC show that the decrease in persistence over time is not only due to changes in educational inequality but also due to changes in inherited inequality across birth cohorts.

Table 3.5 Summary of the levels of relative persistence and trends across birth-cohorts

	Average level of persistence			Linear trend of persistence		
	IGRC	IGC	IRC	IGRC	IGC	IRC
SSOA						
Full	0.714	0.720	0.723	-0.054***	-0.025***	-0.029***
	(0.051)	(0.025)	(0.028)	(0.005)	(0.004)	(0.004)
African	0.605	0.504	0.557	-0.041***	-0.012	-0.032**
	(0.045)	(0.023)	(0.040)	(0.010)	(0.009)	(0.011)
NIDS						
Full	0.408	0.430	0.484	-0.072***	-0.038***	-0.036***
	(0.067)	(0.036)	(0.034)	(0.005)	(0.004)	(0.004)
African	0.402	0.328	0.441	-0.075***	-0.014***	-0.036***
	(0.069)	(0.015)	(0.036)	(0.005)	(0.003)	(0.006)

Notes: (1) The table presents a summary of the overall level and trend of relative mobility in SSOA and NIDS-w5, based on three measures of relative mobility (namely, IGRC, IGC and IRC). The summaries are based on the baseline results. The overall level of relative mobility for each measure is obtained by taking a simple average of the mobility estimates across the birth cohorts. The trend of mobility for each measure is obtained by running a simple regression of the mobility estimates on time. The mobility summaries are for the full and African samples in each dataset.
(2) Parental education is defined as the average education of both parents.
(3) Numbers in parentheses are standard errors. *** for $p<0.001$, ** for $p<0.05$, * for $p<0.1$

3.6 Discussion of Results

The chapter examined the trends in intergenerational education mobility for South Africa. The analysis was based on two datasets, one collected in the early 1990s and the other in 2017. The datasets enabled us to examine the trends in educational mobility over cohorts born between 1923 and 1992. Moreover, some cohorts in the two datasets overlap, which allowed for an assessment of measurement differences across surveys. The analysis was carried out using both absolute and relative measures of mobility.

The results showed that absolute mobility has been declining in the younger birth cohorts, after having been increasing across the older birth cohorts. The decline in absolute mobility across the younger birth cohorts is more pronounced as measured by per capita mobility than by upward mobility. The trend in upward mobility showed that the proportion of children who surpassed the educational attainment of their parents peaked for those born in the mid-1970s and declined gradually afterwards. The trend in per capita mobility similarly reached its peak for those born in the mid-1970s but was shown to decline rapidly in subsequent birth cohorts.

The increasing trend in upward mobility across the older birth cohorts suggests an expansion of the provision of schooling over time. As discussed in chapter 2, education was made free and compulsory for whites and free (but not compulsory) for coloureds in the early 1900s. In addition, access to schooling for Africans saw a rapid expansion in the mid-1950s.

The gap between the two measures of per capita mobility (i.e., the mean absolute difference and the mean difference) across the older birth cohorts was quite small, which suggests that the chances of downward mobility of individuals were quite low. The decline in upward mobility in the younger birth cohorts may not mean a reversal in schooling enrolment rates, but rather lower attainments of children relative to their parental education. This can be seen by the increasing gap between the mean absolute difference and the mean difference across the younger birth cohorts.

The decline in absolute mobility in the younger birth cohorts holds even when we re-examine the trends after recoding parental education to grade 12 for children whose parental education is above grade 12 so that the bounded nature of education would not be an issue for upward mobility. This suggests that the decline in absolute mobility in the younger birth cohorts may not

be explained by the bounded nature of education. We argued that the decline could be explained by the stagnation of the educational attainment of children in the younger birth cohorts.

Analysis of the association between the educational outcomes of parents and their children showed that the educational attainments of parents and their children are positively associated. However, the degree of association has been declining across birth cohorts. As discussed in chapter 2, in South Africa the involvement of the State in education has been expansive both during apartheid and after, which may have given rise to the decline in educational persistence across birth cohorts. The decline in the degree of association across birth cohorts has been more rapid as measured by IGRC than by IGC or IRC. This indicates that changes in the educational inequalities of children and parents across birth cohorts have played a considerable role in reducing the level of transmitted inequality over time.

Comparisons of the levels of relative mobility between the two surveys as well as with other studies suggests that educational persistence has declined considerably over time in South Africa. For instance, the IGC estimate for the early 1990s is 0.72 (SSOA) and it is 0.43 for the late 2010s (NIDS-w5). Similarly, a comparison of our IGC estimates with IGC estimates for 42 countries in Hertz *et al.* (2008)[33] shows that South Africa would have ranked first as the most persistent country if we used the SSOA estimate and would have ranked right in the middle if we used the NIDS estimate.

Overall, the results indicated that South Africa has experienced substantial increase in education mobility between the two periods both in the overall change in educational attainment across generations (absolute mobility) and in the degree of association between the educational attainment of parents and their children (relative mobility). However, the decline in absolute mobility in the younger birth cohorts is a concerning trend. The intra-generational trend in the educational attainment of children in the younger birth cohorts has been stagnant, leading to an increase in the proportion of young people with below postsecondary levels of education. This is likely due to the poor quality of education offered in many South African schools (see e.g., van der Berg *et al.*, 2011; Spaull, 2015), which hinder learners from pursuing higher levels of

[33] Most of the datasets in Hertz *et al.* (2008) are collected between the mid-1990s and the mid-2000s.

education. The increasing pool of individuals with below postsecondary levels of education in the labour market could induce educated (and/or rich) parents to invest more in the education of their children, which could lead to a reversal of the observed declining pattern in educational persistence.

Appendix

3.A Summary of Previous Studies

Table 3.A.1 Summary of Previous Studies

Study	Data	Children's Education	Method	Mobility measure	Results
Thomas (1996)	Survey of Socioeconomic Opportunity and Achievment (1991); Children age: 20-70yrs; A randomly selected child from a household.	years of schooling	Relative mobility	IGRC	Marginal effect of mother education on child education: African or Asian: 0.3-0.4; whites or coloreds: 0.2
Lam (1999)	October Household Survey 1995; Children age 13-17yrs; Co-resident children.	Grades completed per year of age since age 6	Relative mobility	IGRC	Mother university degree (Vs no educ): 0.22 Father university degree (Vs no educ): 0.25
Nimubona and Vencatachellum (2007)	October Household Surveys 1995-1999; Children age: ≥ 15yrs; Co-resident children.	years of schooling; categories for child education	Relative mobility	IGRC; ordered logit; ordered probit; Pseudopanel estimates	IGRC: Africans 0.23; whites: 0.19
Hertz et al (2008)	KwaZulu Income Dynamics Study (KIDS) 1998; Children age 20 -69; Co-resident and non co-resident children.	years of schooling	Relative mobility: OLS for each cohort	IGRC; IGC	IGRC: 0.69; IGC: 0.44
Girdwood and Leibbrandt (2009)	National Income Dynamics Study (NIDS) wave 1(2008); Children age>=20; Co-resident and non co-resident children.	years of schooling; categories for child education	Absolute mobility	Transition matrix; Ordered logit	upward mobility: Africans 64%; White 43%; coloured 58% (in father-son education comparison)
Burns and Keswell (2012)	KwaZulu-Natal Income Dynamics Study (KIDS), 1993 and 1998; Children's age: ≥ 21 years; Co-resident and non co-resident children.	years of schooling	Relative mobility	IGRC; IGC	IGC; Mother to Son: 0.37; Father to Son: 0.39; Mother to Daughter: 0.49; Father to Daughter: 0.46; Mother to Child: 0.43; Father to Child: 0.42
Kwenda, Ntuli,Gwatidzo (2015)	10% of SA 2011 census; Children age ≥ 18; Co-resident children.	years of schooling; categories for child education	Relative mobility: OLS for each cohort	IGRC; IGC; ordered probit	IGC estimates (simple averages); Mothers to Sons: 0.34; Fathers to Sons: 0.31; Mothers to Daughters: 0.36; Fathers to Daughters: 0.35; Mothers to Child: 0.35; Fathers to Child: 0.33
World Bank (2021)	National Income Dynamics Study (NIDS) wave 4(2014); Children age>=21; Co-resident and non co-resident children.	years of schooling; as well as educational categories for child & parents	Relative mobility: OLS for each cohort	IGC and IGRC	IGC: decreased from 0.74 for 1940s cohort to 0.37 for 1980s cohort; IGRC: decreased from 0.85 for 1940s cohort to 0.23 for the 1980s cohort
			Absolute mobility: based on cohorts	upward mobility rates; Average of difference in years of schooling	Upward mobility: Increased from 61% for the 1940s cohort to 90% for the 1970s cohort, then decreased to 86% for the 1980s cohort

3.B Conversion table of reported highest level of education into years of education

Table 3.B.1 Conversion table of reported education into years of education

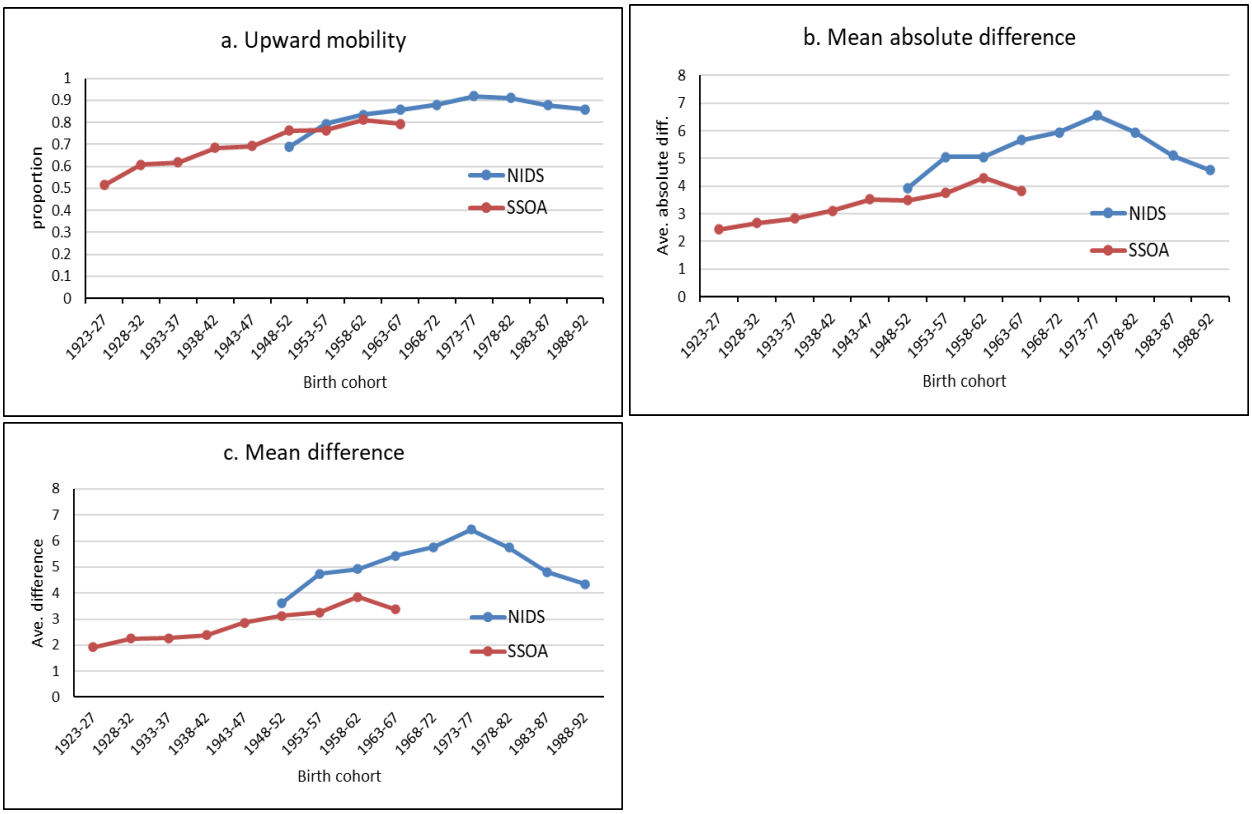
Years of educ.	NIDS	SSOA
0	Grade R/0; No schooling	No schooling; Vocational training course
1	Grade 1 (previously Sub A/ class 1)	sub A/B, Grade 1/2; Literacy course
2	Grade 2 (previously Sub B/ class 2)	Lower primary /some primary
3	Grade 3 (Std. 1)	std 1
4	Grade 4 (Std. 2)	std 2
5	Grade 5 (Std. 3)	Std 3; Higher primary
6	Grade 6 (Std. 4)	Std 4
7	Grade 7 (Std. 5)	Std 5 ; Std 6 elementary;
8	Grade 8 (Std. 6/ Form 1)	Std 6 secondary (Form 1); Some secondary
9	Grade 9 (Std. 7/ Form 2)	Std 7 (Form 2)
10	Grade 10 (Std. 8/ Form 3); NCV 2; N1 (NATED)/ NTC 1	Std 8 (Form 3)
11	Grade 11 (Std. 9/ Form 4); Certificate not requiring Grade 12/ Std. 10; Diploma not requiring Grade 12/ Std. 10; NCV 3; N2 (NATED)/ NTC 2	Std 9 (Form 4); Std 8 or 9 (Form 3 or 4) plus certificate; Std 8 or 9 (Form 3 or 4) plus diploma
12	Grade 12 (Std. 10 /Matric/ Senior Certificate/ Form 5); NCV 4;N3 (NATED)/ NTC 3; N4 (NATED)	Std 10 (Form 5) with school leaving pass (senior certificate); Std 10 (Form 5) with university pass (matric exemption); Completed secondary
13	Certificate requiring Grade 12/ Std. 10; Diploma requiring Grade 12/ Std. 10; N5 (NATED); N6 (NATED)	Post-secondary schooling, but no degree; Post-secondary schooling, with diploma; Some post-secondary
15	Bachelor's degree	Post-secondary schooling, with degree
16	Bachelor's degree and diploma; Honours Degree; Higher Degree (Master's, Doctorate)	Post-graduate schooling beyond a degree
.	Other	Other

3.C Trends in absolute mobility – Robustness Check

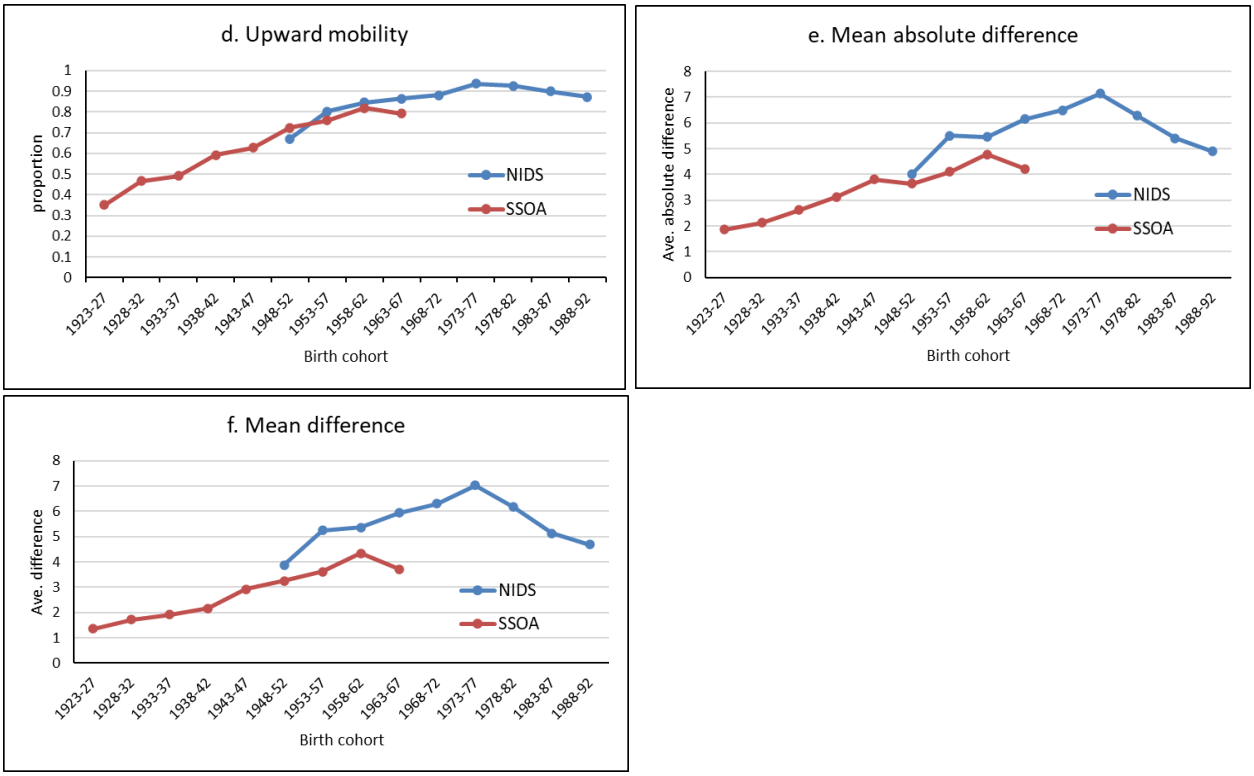
Figure 3.C.1 Trends in absolute mobility - parental education re-coded

Note: Parental education re-coded to grade 12 for children whose parental education is above grade 12.

Panel A: Full sample



Panel B: African sample



Notes: (1) The figures present the trends in absolute mobility across 5-year birth cohorts in SSOA and NIDS-w5 with parental education re-coded to grade 12 for children whose parental education is above grade 12. The trends are based on three measures of absolute mobility (namely, upward mobility, mean absolute difference and mean difference). Panel A is based on the full sample and Panel B is based on the African sample in each dataset. The figures show that absolute mobility increased over much of the older cohorts but is declining over the younger cohorts.

(2) Parental education is defined as the average education of both parents.

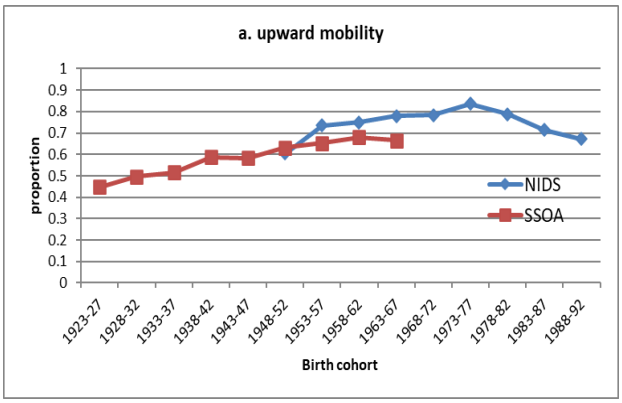
(3) Results are weighted using the calibrated weights provided with the datasets.

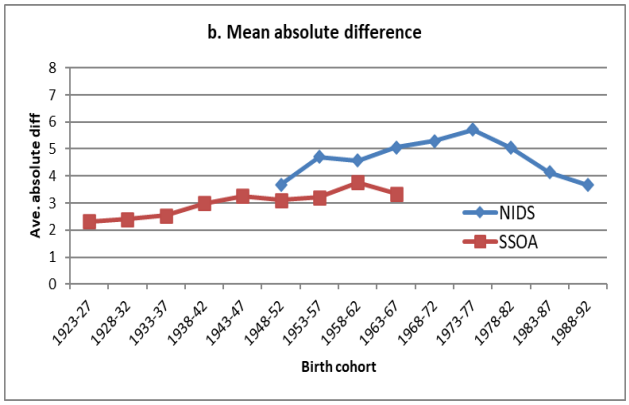
3.D Trends in education mobility – Robustness Check

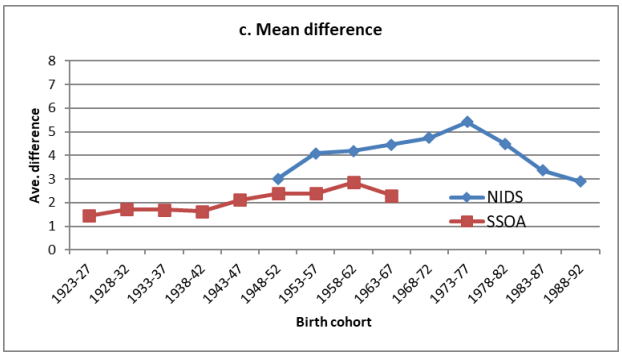
Figure 3.D.1 Trends in absolute mobility - robustness check

Note: Parental education redefined as the highest education of parents.

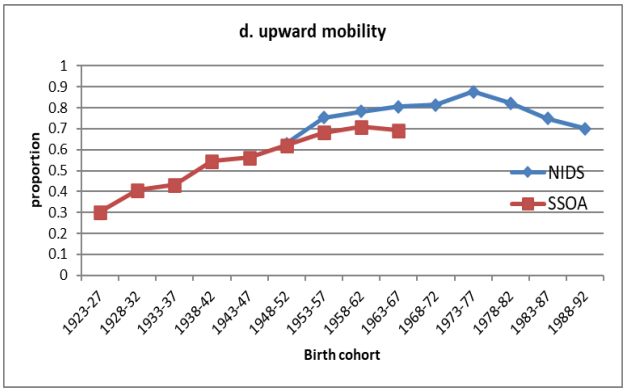
Panel A: Full sample

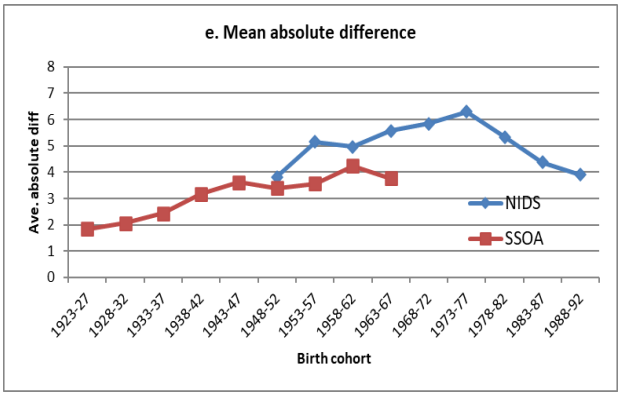


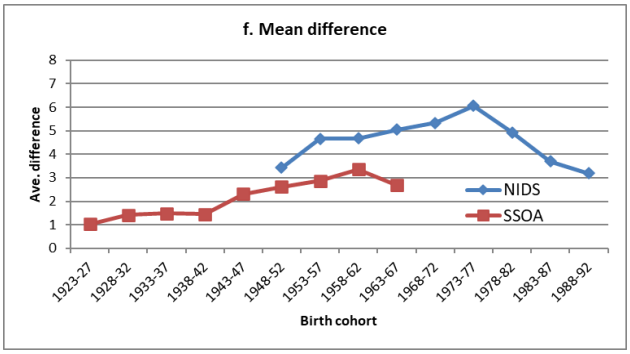




Panel B: African sample





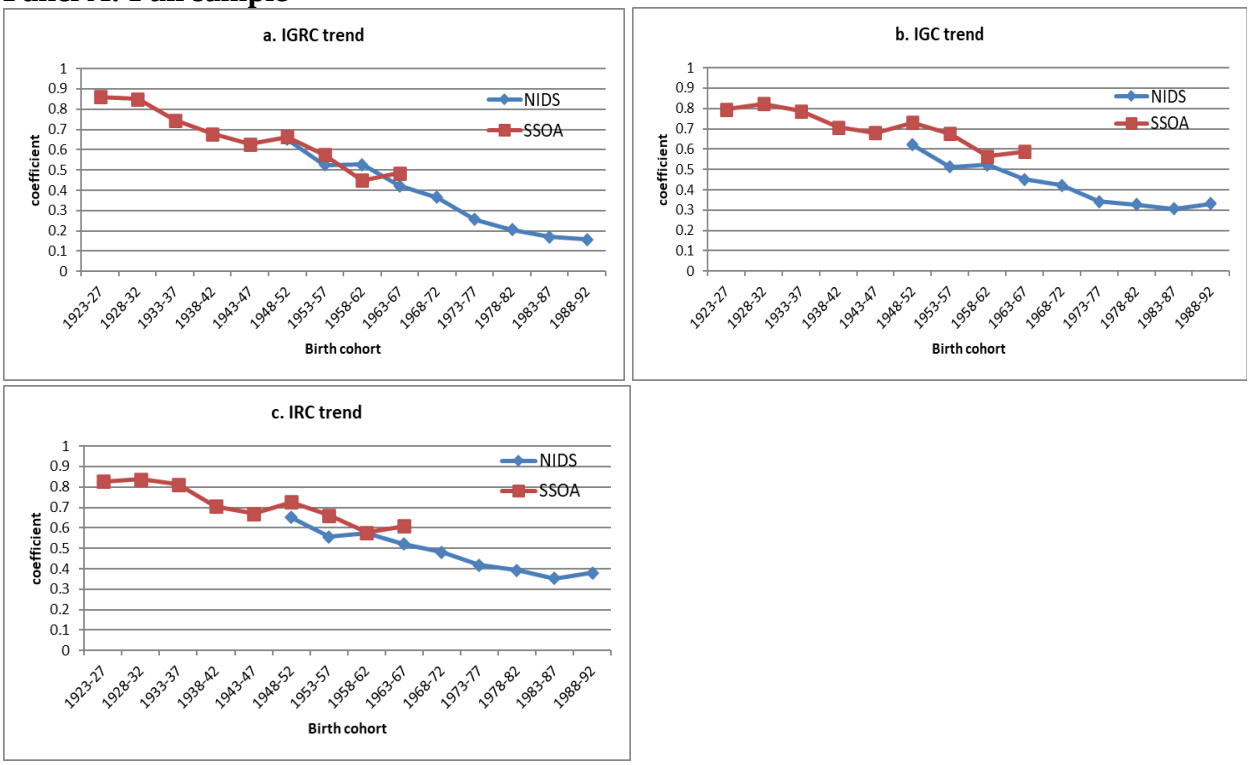


Notes: (1) The figures present the trends in absolute mobility across 5-year birth cohorts using three measures of absolute mobility (namely, upward mobility, mean absolute difference and mean difference) in SSOA and NIDS-w5. Panel A is based on the full sample and Panel B is based on the African sample in each dataset. The figures show that absolute mobility increased over much of the older cohorts but declined over the younger cohorts.
(2) Parental education is defined as the highest education of parents.
(3) Results are weighted using the calibrated weights provided with the datasets.

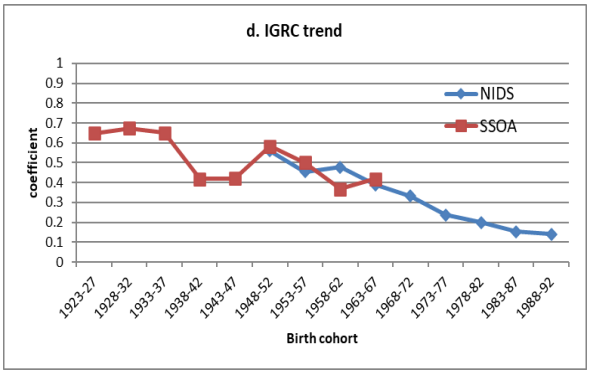
Figure 3.D.2 Trends in relative mobility – robustness check

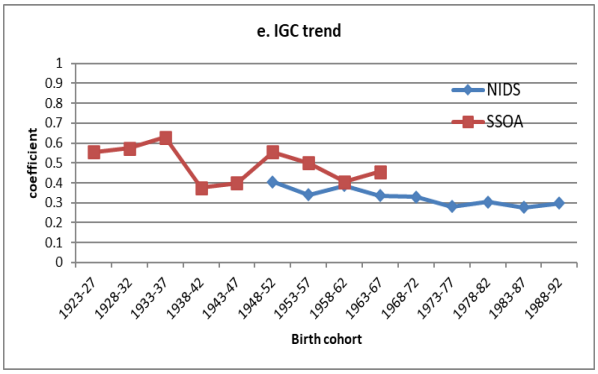
Note: Parental education redefined as the highest education of parents

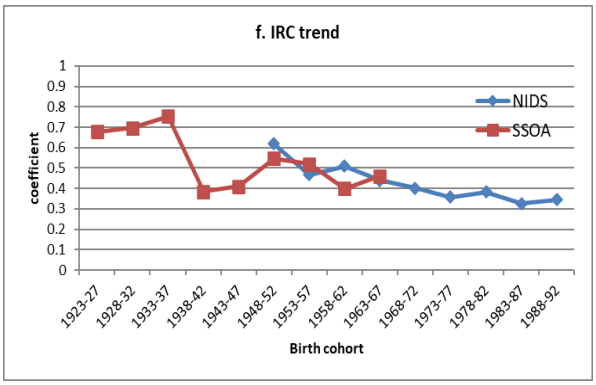
Panel A: Full sample



Panel B: African sample



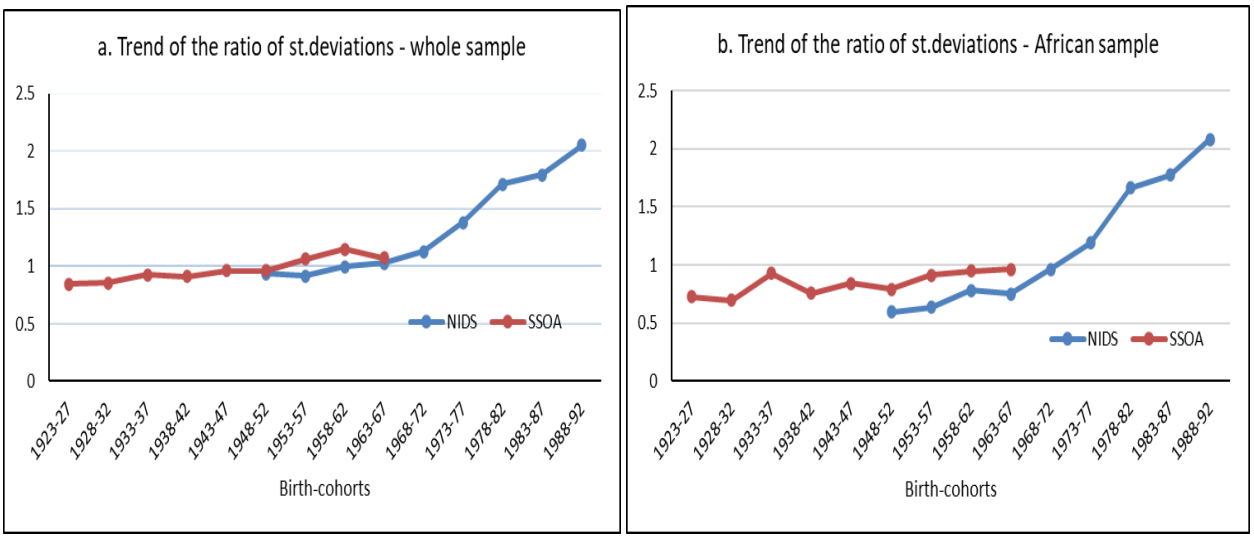




Notes: (1) The figures present the trends in relative mobility across 5-year birth cohorts using three measures of relative mobility (namely, IGRC, IGC and IRC) in SSOA and NIDS-w5. Panel A is based on the full sample and Panel B is based on the African sample in each dataset. The figures show that relative mobility has generally increased across successive birth cohorts, although irregularities exist.
(2) IGRC, IGC and IRC refer to intergenerational regression coefficient, intergenerational correlation coefficient, and intergenerational rank correlation coefficient, respectively.
(3) Parental education is defined as the highest education of parents.
(4) Results are weighted using the calibrated weights provided with the datasets.

3.E Trend in the Ratio of Standard Deviations

Figure 3.E.1 Trends in the ratio of standard deviations



Notes: The figures show the trends in the ratio of the standard deviation of the educaiton of parents to the standard deviation of the education of children across birth cohorts for the full and African samples in SSOA and NIDS-w5. The figures are based on children between 25 and 69 years of age.

Chapter 4

4. Parental Background and Earnings in South Africa

4.1 Introduction

The socioeconomic trajectory of South Africans in the post-apartheid era has been quite intriguing. As indicated in chapter 1, the average education of the working-age population has been increasing coupled with a decline in schooling inequality. Further, as shown in chapter 3, the association between the educational attainments of parents and their children has been declining across birth cohorts. On the other hand, earnings inequality has been persistently high. The dynamics leaves one to wonder, inter alia, what the role of parental background has been in the social process that determines earnings.

This chapter examines changes over time in the importance of parental background to earnings and earnings inequality in the South African labour market. Parental background plays important roles in the socioeconomic outcomes of children. The human capital approach to intergenerational mobility, for instance, emphasizes the importance of parental background in the transmission of economic advantages across generations (e.g., Becker and Tomes, 1979, 1986). Empirical evidence further shows that intergenerational economic mobility is lower in developing countries compared to developed countries with parental background playing a central role in determining earnings (Lam & Schoeni, 1993).

In our empirical analysis, we proxy parental background by the education of parents. Education of parents can capture the long-term influence of parental background on children's life outcomes. The analysis is carried out using our two datasets, namely, SSOA and NIDS-w5. The analysis has two parts. In the first part we investigate changes in the returns to the education of parents between the two surveys. Previous South African studies in this area mainly focused on the returns to own education without paying due attention to the influence of key parental background variables in the relationship. In the second part, we examine changes in the contribution of the education of parents to earnings inequality between the two surveys.

The main analysis is based on employed males between 25 and 59 years of age. We also replicate the analysis for females for comparison purposes. The results based on the main sample show that the addition of education of parents in a Mincerian-type regression reduces the magnitude of the returns to own education in both periods, with about 10 percent reduction in both. In addition, the partial effect of parental education (defined as the average education of both parents) on children's earnings is shown to be positive and statistically significant in both periods. The contribution of parental education to earnings inequality, however, has been small (relative to the contribution of own education) in both periods.

The chapter is organized into 6 sections. The next section provides a review of the relevant literature. We briefly discuss the relationship between parental background and children's earnings in the context of the human capital model and present a review of empirical studies that found significant association between parental background and children's earnings. Further, we present a review of the South African literature on the relationship between education and labour market outcomes of children, with the intention of laying the context for the analysis of the association between parental background and children's earnings in the results section. Section 3 discusses the methods used. We use Mincerian earning regressions to analyse the association between the education of parents and children's earnings and a regression-based decomposition of inequality to investigate the contribution of parental education to earnings inequality. Section 4 presents and describes the datasets. This is followed by the empirical results. We examine the partial association between the education of parents and children's earnings and the contribution of parental education to children's earnings inequality with a focus on changes that occurred over time. The last section discusses the main findings, highlighting the continued relevance of parental background to earnings and earnings inequality.

4.2 Literature Review

4.2.1 Parental Background and Earnings

Standard Mincerian regressions typically model earnings as a function of education and labour market experience.[34] However, these regressions are believed to overestimate the returns to education that would be experienced by a randomly selected individual (e.g., Heckman & Hotz, 1986; Lam & Schoeni, 1993). Theoretical models of the association between parental background and adult economic outcomes of children postulate that the earnings of children are linked to their parents not only through parental investment in the human capital of the child but also through the intergenerational transmission of endowments[35] (Becker & Tomes, 1979, 1986). Thus, to the extent that schooling in the standard Mincerian regressions is positively correlated with endowments acquired from parents and the latter is not accounted for, the returns to education might be biased upward.

Several studies have found substantial correlations between parental background and children's earnings, conditional on child's education. This has been documented to be the case in several developing countries as well as in developed countries with high intergenerational inequality.

Heckman & Hotz (1986) investigated the labour market earnings of Panamanian males using a 1983 dataset. They found that the inclusion of parental background variables (proxied by the education of mother and father) in a Mincerian wage regression reduces the estimated average return to own education from 13.0 to 8.6 per cent. They argued that their result provides evidence that a part of the return to own education in regressions that do not control for parental background is attributable to the influence of parental background on children's earnings. Stelcner, Arriagada & Moock (1988) also found similar results for a sample of males in Peru. Lam & Schoeni (1993) investigated the effect of family background on earnings among married males using a 1982 Brazilian dataset. They found that earnings are significantly positively

[34] Psacharopoulos & Patrinos (2018) review the returns to investment in education for a large number of countries.

[35] Endowments from parents to children refer to genetically transmitted productive traits (such as cognitive ability and race), the learning and life skills acquired through belonging to a particular family culture, as well as social capital acquired through family reputation and connections, as discussed in Becker & Tomes (1979).

associated not only with one's parents' education but also with the education of parents-in-laws, conditional on own education. Moreover, their findings reveal that the returns to own education fall by about one-third when education of parents are accounted for in the earnings equations. Using the same Brazilian dataset, Strauss & Thomas (1991) found parental background (proxied by the education of parents) to play an important role in determining earnings among the urban sample.

Chen & Feng (2011) studied the association between the education of parents and the wages of adult children in China in the context of the human capital model. They found a significant positive correlation between the education of parents and the earnings of children. Liu, Hammitt & Lin (2000) also found similar results for Taiwan as that of Chen & Feng (2011). Armitage & Sabot (1987) used data for Kenya and Tanzania and found that in both countries returns to education rise significantly with parental education.

Franzini, Patriarca & Raitano (2020), using a 2005 Italian sample, found that children of parents with tertiary education earn about 5 per cent higher compared to children of parents with primary education after controlling for child's education. Similarly, Arntz *et al*. (2022) found that conditional on a child's education, children of high-educated parents earned on average about 8 per cent higher compared to children of low-educated parents in Germany in the early 1990s. Other studies from Europe, e.g., Laurison & Friedman (2016) for the UK, and Franzini & Raitano (2009) for 13 European countries, also found substantial earnings differences by parental background.

Schultz (1987) summarizes the various interpretations given to the explanatory role of parental background or parental education on children's earnings. These include:

(i) nepotism and social stratification - where influential parents use their connections to advance the labour market outcomes of their children.

(ii) parents' cultural capital – where more educated parents provide their children with a learning environment at home thereby enhancing the market productivity of education for a given schooling years.

(iii) quality of schooling – where parental education may serve as a proxy for the quality of a child's schooling. More educated parents can afford to vary their market investment in the child's education for given schooling years.

(iv) Genetic transmission of ability – where parental education may serve as a proxy for child's ability. To the extent that genetic abilities are transmitted from parents to children, children of genetically endowed parents may end up earning higher for a given schooling years.

The studies mentioned above provide evidence that parental background has an important role in the labour market outcomes of children. Understanding the role that parental background plays in determining earnings is essential in assessing the importance of the family in transmitting inequality across generations.

4.2.2 Educational attainment and labour market outcomes in South Africa

One of the stylized facts of the South African labour market is that the average education of the working-age population has been steadily increasing, coupled with a decline in schooling inequality (Finn, Leibbrandt & Ranchhod, 2017). Branson & Lam (2022) documents the changes in the distribution of the working-age population by educational level between 1993 and 2018. Their findings show that the proportions of those with no education and between 1 to 7 years of education have seen a steady decline over the years with the former declining from 13 to 3 per cent and the latter declining from 30 to 12 per cent. The proportion of those with incomplete secondary education (grades 8-11) has barely changed over the years, remaining at around 35 per cent. And the proportions of those with grade 12 and tertiary education have seen a steady increase over the years, with the former increasing from 14 to 31 per cent and the latter increasing from 9 to 16 per cent. While the educational attainment dynamics has been remarkable, the quality of education received by majority of South African learners has been poor (van der Berg *et al.*, 2011; Spaull, 2013b and 2015). Spaull (2013b) points to the existence of inequality in educational opportunity by socioeconomic status, wealth, geographic location and language, and notes that a majority of learners (75-80 per cent) are in schools whose performance can only be described as abysmal.

Education is inextricably linked to the labour market outcomes of South Africans. The relationship between education and employment has been investigated by several studies and the findings show that having postsecondary education greatly improves one's employment prospects (e.g., Branson *et al.*, 2012; Branson & Leibbrandt, 2013; Branson & Lam, 2022). Branson & Lam (2022) investigated the relationship between education and employment using the proportions employed across completed grades in 1994/95 and 2017/18. They found that in both periods the probability of employment by education remained almost the same at each grade until grade 12, but sharply rises with education for those with postsecondary education.

The relationship between education and earnings has been observed to follow a convex pattern (Keswell & Poswell, 2004; Branson & Lam, 2022). Again, Branson & Lam (2022) investigated the relationship between education and earnings using datasets collected in 1994/95 and 2017/18. They found that average earnings increase with education in both periods, with the rate of increase being higher at postsecondary levels. Furthermore, the relationship is seen to intensify in its convexity between the two periods.

Studies that dealt with the returns to education in South Africa include Mwabu & Schultz (1996, 2000), Sherer (2000), Chamberlain & van der Berg (2002), Keswell & Poswell (2004), Branson & Leibbrandt (2013), Salisbury (2016) and Depken, Chiseni & Ita (2019).[36] However, none of these studies have explicitly considered the importance of parental background in determining earnings.

The role of education in the evolution of earnings inequality in the post-apartheid period has also been investigated (e.g., Branson *et al.*, 2012; Finn & Leibbrandt, 2018). Branson *et al.* (2012) showed that while increasing returns to postsecondary education had inequality increasing effect, increasing educational attainment of South Africans had the opposite effect, resulting in consistent levels of inequality over time. Finn & Leibbrandt (2018) also found increasing returns to education to have had inequality enhancing effect but did not find increasing educational

[36] Some of these studies have accounted for the nonlinear relationship between education and earnings. For example, Sherer (2000) considered a quadratic specification for education, Mwabu & Schultz (1996, 2000) and Chamberlain & van der Berg (2002) considered a spline function, Keswell & Poswell (2004) considered quadratic and cubic specifications, and Branson & Leibbrandt (2013) allowed education to have different effects by educational level.

attainment to have had inequality reducing effect. Instead, they found declining returns to potential experience to have had inequality reducing effect.

The survey of the international literature for the influence of parental background on children's earnings in section 4.2.1 shows that most of the countries in which parental background is significantly associated with one's earnings are either developing countries or developed countries with high levels of inequality. South Africa is an interesting case in this regard as it is a developing country with high levels of inequality.

4.3 Methods

This section presents the methods used to analyse the association between the education of parents and children's earnings and the contribution of parental education to children's earnings inequality.

4.3.1 Association between parental education and earnings

The standard Mincerian model of human capital expresses earnings as a function of the individual's measured stock of human capital, which includes schooling and labour market experiences (Mincer, 1974). In empirical studies, the standard model is often extended by including variables such as race, gender, region, migration status, health status, tenure on the job or other variables that could affect earnings (Polachek, 2007). In this study, we use the Mincerian earnings model augmented by the addition of the education of parents[37] and other relevant variables.

We write the augmented Mincerian model as:

$$\ln Y_i = \beta_0 + \beta_s S_i + \beta_M M_i + \beta_F F_i + \delta \boldsymbol{X} + \varepsilon_i \qquad (4.1)$$

where the dependent variable lnY is the natural log of earnings, and S, M and F denote completed years of education of respondent, mother, and father, respectively. $\boldsymbol{X}$ is a vector of

[37] Several studies have used education of parents as a proxy for unmeasured ability or family culture to control for the influence of unobserved ability on the relationship between earnings and own education (e.g., Card, 1995; Ashenfelter & Zimmerman, 1997).

relevant control variables including labour market experience (and its square), and ε is a mean zero error term assumed to be uncorrelated with any of the explanatory variables.

As discussed in the previous section, the relationship between education and earnings is non-linear with a convex pattern in the South African labour market. Further, the relationship has remained convex in pattern between the early 1990s and the late 2010s. Thus, in our empirical analysis, we assume a quadratic relationship between education and log earnings.

Our datasets do not provide information on the actual labour market experiences of respondents. We use age as a proxy for labour market experiences. With high levels of grade repetition in the schooling system (Branson, Hofmeyr & Lam, 2014) and employment volatility in the labour market (Banerjee *et al*., 2008; Kerr, 2018), potential experience[38] may not be a good proxy for labour market experience.

The actual explanatory variables are thus education of respondents and its square, education of parents, and age of respondents and its square. In addition, we control for race, marital status (a dummy for whether the respondent is ever married), urban/rural status (a dummy for whether the respondent lives in urban area), and province of respondents.

4.3.2 Contribution of parental education to earnings inequality

The contribution of the education of parents to earnings inequality is investigated using Fields' regression-based decomposition of inequality (Fields, 2003).[39] Fields' approach to inequality decomposition is simple and does not depend on any summary measure of inequality.

The first step in the Fields' decomposition is the estimation of the earnings regression equation. Our earnings equation is the one given in equation 4.1.

[38] Potential experience is often given by: Age-Education-6.

[39] Fields' approach to earnings decomposition improves the approach taken by Morduch & Sicular (2002), and in turn the two studies are extensions of the income decomposition approach of Shorrocks (1982). Fiorio & Jenkins (2007) developed a STATA module for the Fields' regression-based decomposition of inequality.

The second step is to estimate how much of the dependent variable is contributed by each of the explanatory variables in the regression. If we write equation 4.1 in a general form as:

$$\ln Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_K X_{Ki} + \varepsilon_i$$

Then the above equation can be re-written as:

$$\ln Y_i = \beta_0 + f_{1i} + f_{2i} + \ldots + f_{Ki} + \varepsilon_i \qquad (4.2)$$

where each f_k is a 'composite' factor, equal to the product of the regression coefficient and its corresponding variable. Each f_{ki} is the amount of log earnings contributed by the k^{th} explanatory variable for an individual i, with $k = 1,2, \ldots, K$. We estimate the regression coefficients using OLS and re-write the right-hand-side of equation 4.2 as the sum of the estimated contribution of the factors (including the constant and error terms) as:

$$\ln Y_i = \hat{\beta}_0 + \hat{f}_{1i} + \hat{f}_{2i} + \ldots + \hat{f}_{Ki} + \hat{\varepsilon}_i$$

The third step is to estimate the share of total inequality contributed by each of the factors. Having expressed log earning as the sum of the estimated factor contributions, Fields (2003) used a result by Shorrocks (1982), which decomposes total inequality by additive factor components. According to the result, the share of total inequality contributed by the k^{th} factor (s_k) for an inequality index $I(\ln Y)$ defined over a vector of log earnings $lnY = (lnY_1, \ldots, lnY_N)$, is given by:

$$s_k(\ln Y) = \frac{\mathrm{cov}(f_k, \ln Y)}{\sigma^2(\ln Y)} = \frac{\sigma(f_k) * corr(f_k, \ln Y)}{\sigma(\ln Y)} \qquad (4.3)$$

where $s_k(\ln Y)$ denotes the share of total inequality in $\ln Y$ contributed by f_k and is referred to as 'relative factor inequality weight'. σ^2, cov and $corr$ stands for variance, covariance, and correlation, respectively. We can gather from equation 4.3 that $s_k(\ln Y)$ is the slope coefficient in the regression of factor f_k on the dependent variable $\ln Y$.

The above decomposition does not depend on the choice of the measure summarizing inequality as long as some general decomposition conditions, discussed in Shorrocks (1982), are met.[40] Note that the sum of the relative factor inequality weights is equal to 1: $\sum_{k=1}^{K+1} s_k(\ln Y) = 1$. The $(K+1)^{th}$ term is for the contribution of the residual. The constant term is irrelevant in the inequality decomposition calculation as it is the same for every individual. Note also that the contribution to inequality by a factor could be negative. A negative contribution indicates the factor has an inequality-decreasing role. In addition, we can only relate inequality in the dependent variable to contributions from the factors $f's$, not the $X's$.

4.4 Data and Descriptive Statistics

4.4.1 The Datasets

The empirical analysis is based on the Survey of Socioeconomic Opportunity and Achievement (SSOA) and the fifth wave of the National Income Dynamics Study (NIDS-w5). The datasets are discussed in chapter 1.

We obtain the earnings of respondents in SSOA by adding their gross earnings from the primary job or business and secondary economic activities (if any). Both earnings (from primary and secondary jobs) are given in the form of brackets.[41] For those with earnings from both primary and secondary activities, we add corresponding lower and upper bounds of both earnings to obtain their total earnings (in brackets).[42] In NIDS-w5, we obtain the earnings of respondents by adding their gross monthly earnings from all employment-related sources of income. The employment income sources include earnings from regular wages, self-employment, casual work

[40] These conditions are met by almost all inequality indices including the Gini coefficient, the coefficient of variation, the Atkinson index, and the generalized entropy family (Fields, 2003).

[41] The questionnaire provided equivalent ranges of the bracket earnings on weekly, monthly, and yearly bases. Majority of respondents reported their earnings on monthly basis. For example, the percentage of responses for the schedule of earnings from main job are: 16.9% weekly, 71.7% monthly, and 11.4% yearly.

[42] Among the currently employed individuals, about 4.5 per cent reported earnings from secondary economic activities.

as well as agricultural activities. Moreover, any earnings from bonus payments or share of profits related to a wage job are included as earnings.

The main analysis is based on employed males who are between 25 and 59 years of age. The relatively higher lower-age-limit of 25 years allows us to minimize the occurrence of zero earnings, which may introduce sample selection issues when working with log earnings.[43] We replicate the analysis for females in the same age range for comparison purposes.

Earnings are gross monthly earnings and are converted to constant 2017 South African Rands (ZAR).

4.4.2 Descriptive Statistics

Table 4.1 describes the dependent and explanatory variables in our study of the partial association between the education of parents and the earnings of their children. The table further presents summary statistics of the variables based on employed males between 25 and 59 years of age who reported positive earnings. In Appendix 4.A we provide summary statistics of the same variables based on males between 25 and 59 years of age who are in the labour force to serve as a reference.[44] A comparison of the two tables shows that in SSOA those with positive earnings have slightly higher averages both for own education and the education of parents in comparison to the larger sample of labour force participants. Moreover, the racial composition of the sample of positive earners features a higher proportion of whites and a lower proportion of Africans in comparison to the racial composition in the larger sample. On the other hand, in NIDS the averages for both own education and the education of parents in the two samples are quite similar. Further, the racial composition of the two samples is almost the same.

[43] In both SSOA and NIDS, the proportion of employed individuals in the 25-59 years age range who reported zero earnings is quite small. In SSOA, the figures are 0.61% (14 individuals) for males, and 0.59% (8 individuals) for females. In NIDS, 0.13% (6 individuals) for males, and 0.24% (10 individuals) for females. The proportion of employed individuals in the 25-59 years age range with missing earnings, however, is relatively large. In SSOA, 7.06% (162 individuals) for males, and 6.86% (96 individuals for females). In NIDS, 7.44% (302 individuals) for males, and 5.64% (230 individuals) for females. Furthermore, from the above figures, there does not appear to be much difference in paid labour force participation rates between men and women.

[44] The employment rate among males between 25 and 59 years of age is about 80.1 per cent in SSOA and 85.5 per cent in NIDS-w5 (based on the broad definition of unemployment).

Table 4.1 Definition of variables and their sample statistics: Males 25-59 years of age with positive earnings

		SSOA			NIDS-w5		
Variable	**Variable definition**	n	Mean	St.dev.	n	Mean	St.Dev.
ln_earnings	ln of monthly earnings (in 2017 ZAR)	2,534	8.70*	1.02	3,638	8.66	1.28
ch_edu	Completed years of education of child	2,534	8.95	4.13	3,623	10.74	2.94
ch_edu_sq	child education squared	2,534	97.27	64.60	3,623	124.07	53.71
mth_edu	Completed years of education of mother	2,239	5.37	4.56	3,335	5.75	4.93
fth_edu	Completed years of education of father	2,011	6.06	4.69	2,957	5.45	5.15
prnt_edu	Average parental education	2,329	5.64	4.41	3,425	5.62	4.61
married	=1 if ever married, 0 otherwise	2,534	0.84	0.37	3,638	0.50	0.50
urban	=1 if lives in urban area, 0 in rural area	2,534	0.73	0.45	3,638	0.74	0.44
Race							
African	=1 if African, 0 otherwise	2,534	0.58	0.49	3,636	0.81	0.39
coloured	=1 if coloured, 0 otherwise	2,534	0.10	0.30	3,636	0.08	0.28
Indian	=1 if Indian/Asian, 0 otherwise	2,534	0.05	0.22	3,636	0.02	0.15
white	=1 if white, 0 otherwise	2,534	0.27	0.44	3,636	0.08	0.27
Age	Age of respondent	2,534	38.68	9.29	3,638	37.93	9.19
Age_sq	Age squared	2,534	1582.68	758.87	3,638	1523.03	746.16
N		**2,534**			**3,638**		

Notes: (1) The summary statistics are based on employed males between 25 and 59 years of age who reported positive earnings in SSOA and NIDS-w5.
(2) * The figure is based on the predicted earnings from interval regression. The mean real log earnings for the lower and upper bounds of the bracket earnings are 8.40 and 8.82, respectively.
(3) Results are weighted using the calibrated weights provided with the datasets.

A comparison of SSOA and NIDS (in Table 4.1) shows that while the average education for both respondents and mothers is higher in NIDS than in SSOA, the average education for fathers is higher in SSOA. The higher average education of fathers in SSOA is perhaps due to the exclusion of farmworkers in that sample. As we have discussed in chapter 3 (section 3.4.1), farmworkers were the most disadvantaged group in terms of schooling opportunities. Furthermore, as noted above, the racial composition of the SSOA sample features higher proportion of whites relative to their size in the larger sample of labour force participants. As discussed in chapter 2, the educational attainment of whites has been historically high, although the racial gaps have been narrowing over time.

The percentage of missing values for the education of parents is rather large in both datasets. Among the samples of positive earners, mothers' education is missing for about 11.6 per cent in SSOA and 8.3 per cent in NIDS. Fathers' education is missing for about 20.6 per cent in SSOA and 18.7 per cent in NIDS. To the extent that the missingness in the education of parents is systematic, limiting the regression analysis to those with non-missing education of parents will introduce sample selection bias. Thus, in our regression analysis, we define indicator variables for missing education of parents and reset the missing values in each education variable to the mean value of that variable. The indicator variables are then included in the regression model to control for the effect of the missingness. Although this approach is likely to provide a downwardly biased coefficient for the education of parents (see e.g., Jones, 1996; Allison, 2002), it allows us to use all the available information.

Before proceeding to the regression analysis of the association between the education of parents and children's earnings, we first examine descriptively the partial association between the two variables. Table 4.2 presents a cross-tabulation between the educational levels of parents and sons. The first two tables (Panel A) are based on the SSOA sample. The first table is a cross-tabulation between the education levels of mothers and sons, and the second is a cross-tabulation between the education levels of fathers and sons. Panel B shows the same two tables based on the NIDS sample. In each table, the cell entries show the mean log real monthly earnings and the frequency of sons with parents at a given education level. For instance, the entries of the first cell in the first table of Panel A indicate that the mean log real monthly earnings of sons born to mothers with no education and who had no education themselves is 7.76. There are 100 individuals like this in the sample.

Table 4.2 shows that the mean earnings of sons increase with the increase in the level of education of parents in both SSOA and NIDS. This is shown in the column for overall mean. This preliminary result of a positive association between the education of parents and the earnings of children signals the importance of the education of parents to the earnings of their children. For a given education level of sons, the mean earnings of sons vary across the levels of education of their parents. This variation can be observed at each level of education of sons across the education levels of mothers and fathers in both SSOA and NIDS (Table 4.2). The

mean earnings of sons with the same level of education seem to rise with the increase in parents' education level, although there are exceptions.

Table 4.2 Mean log earnings of sons in the cross-tabulation between the education levels of parents and sons

Panel A: SSOA

Mother educ	Child education level					
	0	1	2	3	4	Total
0	7.76	8.22	8.33	8.68	9.35	8.20
	100	249	170	58	31	608
1	7.78	7.94	8.39	8.95	9.47	8.43
	21	194	281	135	69	700
2	7.97	8.14	8.89	9.51	9.83	9.30
	2	21	163	217	182	585
3	.	8.29	9.31	9.49	9.90	9.61
	0	2	26	86	107	221
4	.	.	8.88	9.27	9.90	9.65
	0	0	8	27	90	125
Total	7.77	8.11	8.53	9.24	9.77	8.70
	123	466	648	523	479	2239

Father educ	Child education level					
	0	1	2	3	4	Total
0	7.92	8.10	8.22	8.51	9.56	8.15
	71	172	98	26	17	384
1	7.65	8.18	8.41	8.99	9.47	8.46
	26	191	255	140	66	678
2	7.97	7.87	8.82	9.42	9.76	9.15
	2	23	176	193	151	545
3	11.60	8.10	9.48	9.61	9.83	9.67
	1	3	25	88	84	201
4	.	8.40	9.16	9.37	9.93	9.70
	0	3	8	45	147	203
Total	7.88	8.12	8.53	9.27	9.77	8.73
	100	392	562	492	465	2011

Panel B: NIDS-w5

Mother educ	Child education level					
	0	1	2	3	4	Total
0	7.66	7.85	8.14	8.64	9.31	8.25
	88	248	550	149	140	1175
1	7.92	7.73	8.01	8.92	9.43	8.50
	7	110	431	195	157	900
2	6.92	7.32	8.36	8.77	9.49	8.76
	3	32	328	220	224	807
3	8.29	7.74	8.57	9.35	9.90	9.40
	1	2	62	70	99	234
4	.	.	9.20	9.13	10.05	9.75
	0	0	35	40	132	207
Total	7.65	7.78	8.22	8.89	9.64	8.66
	99	392	1406	674	752	3323

Father educ	Child education level					
	0	1	2	3	4	Total
0	7.54	7.73	8.26	8.55	9.17	8.29
	89	244	567	178	151	1229
1	7.50	7.74	7.98	8.72	9.34	8.38
	3	78	319	137	110	647
2	8.03	7.63	8.54	8.97	9.67	8.94
	3	27	257	157	181	625
3	.	8.22	8.55	9.29	10.02	9.42
	0	3	57	80	116	256
4	.	6.21	8.27	9.34	10.03	9.62
	0	1	33	35	121	190
Total	7.57	7.73	8.29	8.85	9.67	8.69
	95	353	1233	587	679	2947

Notes: (1) The tables are based on employed males between 25 and 59 years of age who reported positive earnings. The tables present the cross-tabulation between the educational levels of parents and their children. In each cell, the first entry is the mean log real earnings of children, and the second entry is their frequency in the cross-tabulation. Education is ordered into 5 categories: 0 stands for 'no education', 1 for 'primary (grades 1 to 7)', 2 for 'Incomplete secondary (grades 8 to 11)', 3 for 'matric (grade 12)', and 4 for 'postsecondary (above grade 12)'.
(2) Results are weighted using the calibrated weights provided with the datasets.

4.5 Empirical Results

This section examines the partial association between the education of parents and earnings of their children as adults in SSOA and NIDS-w5, with a focus on changes in the association between the two surveys. In addition, it assesses the contribution of parental education to children's earnings inequality in the two surveys.

The main results are based on employed males who are between 25 and 59 years of age. We check the robustness of the coefficient estimates by restricting the analysis to employed males between 30 and 49 years of age. This restriction allows us to check whether the results are affected by life-course fluctuations in earnings. The available evidence in the literature suggests that the bias in permanent earnings due to life-course fluctuations is small if earnings are measured between the early 30s and mid-40s (e.g., Haider & Solon, 2006; Nybom & Stuhler, 2016). We further extend the analysis by considering employed females who are between 25 and 59 years of age.

4.5.1 Association between the Education of Parents and Children's Earnings

4.5.1.1 Regression of log real earnings: Males

As discussed in section 4.3.1, we use regression methods to analyse the partial association between the education of parents and children's earnings. Specifically, we use interval regression in SSOA as earnings are given in brackets[45] and OLS in NIDS. Table 4.3 presents the regression output. The first five columns of the regression output are for SSOA, and the next five are for NIDS. In each set, the first column considers all the explanatory variables except the education of parents. The following four columns add different variations of parental education to the first specification: mothers' education, fathers' education, both parents' education separately and finally the average education of parents.

[45] Standard interval regression models assume that the error term is normally distributed: $\varepsilon \sim \mathrm{N}(0, \sigma^2)$. In our sample, histogram graphs (with overlayed normal curves) of the two dependent variables (i.e., the log of the lower bound earnings and the log of the upper bound earnings) showed that their distributions can be approximated to follow normal distribution. Thus, we assumed that the actual earnings follow a log normal distribution.

Table 4. 3 Regression of log real earnings: Males

`	SSOA					NIDS-w5				
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Child education	0.007	0.004	0.009	0.007	0.003	-0.120***	-0.111***	-0.109***	-0.107***	-0.108***
	(0.023)	(0.023)	(0.024)	(0.023)	(0.024)	(0.031)	(0.030)	(0.030)	(0.030)	(0.030)
Child education squared	0.006***	0.005***	0.005***	0.005***	0.005***	0.017***	0.016***	0.016***	0.016***	0.016***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Mother education		0.017		0.006			0.030***		0.025***	
		(0.011)		(0.011)			(0.007)		(0.008)	
Father education			0.022*	0.019				0.024***	0.014*	
			(0.013)	(0.015)				(0.007)	(0.008)	
Parental education					0.028**					0.034***
					(0.013)					(0.007)
Mother education missing		-0.040		0.000			-0.141		-0.080	
		(0.069)		(0.073)			(0.113)		(0.094)	
Father education missing			-0.078	-0.075				-0.107	-0.093	
			(0.061)	(0.065)				(0.129)	(0.131)	
Parental educ. Missing					-0.043					-0.240
					(0.078)					(0.153)
Race										
Coloured	0.396***	0.375***	0.361***	0.357***	0.349***	0.352**	0.332**	0.321**	0.321**	0.322**
	(0.088)	(0.084)	(0.082)	(0.081)	(0.080)	(0.144)	(0.148)	(0.142)	(0.144)	(0.145)
Indian/Asian	0.497***	0.525***	0.484***	0.497***	0.515***	0.220	0.179	0.161	0.156	0.151
	(0.128)	(0.134)	(0.120)	(0.126)	(0.126)	(0.199)	(0.173)	(0.194)	(0.175)	(0.180)
White	1.035***	0.967***	0.954***	0.939***	0.916***	0.910***	0.771***	0.779***	0.720***	0.741***
	(0.080)	(0.076)	(0.068)	(0.071)	(0.069)	(0.120)	(0.122)	(0.127)	(0.126)	(0.125)
Age	0.079***	0.086***	0.090***	0.091***	0.093***	0.018	0.028	0.027	0.032	0.029
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.027)	(0.026)	(0.027)	(0.026)	(0.026)
Age Squared	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Married	0.352***	0.351***	0.346***	0.346***	0.347***	0.465***	0.477***	0.470***	0.476***	0.477***
	(0.077)	(0.074)	(0.073)	(0.073)	(0.072)	(0.070)	(0.069)	(0.070)	(0.068)	(0.069)
Urban	0.175	0.172	0.177	0.176	0.171	0.194**	0.182**	0.187**	0.181**	0.178*
	(0.118)	(0.113)	(0.114)	(0.113)	(0.111)	(0.091)	(0.091)	(0.089)	(0.089)	(0.092)
Province										
Natal	0.070	0.065	0.058	0.057	0.056					

	(0.133)	(0.131)	(0.131)	(0.131)	(0.128)					
Orange Free State	-0.047	-0.064	-0.061	-0.064	-0.081					
	(0.114)	(0.111)	(0.110)	(0.110)	(0.109)					
Transvaal	0.121*	0.126*	0.114*	0.117*	0.119*					
	(0.065)	(0.065)	(0.066)	(0.066)	(0.064)					
Homelands	0.149	0.157	0.152	0.154	0.155					
	(0.178)	(0.172)	(0.172)	(0.171)	(0.170)					
Eastern Cape						-0.006	-0.004	0.003	0.001	-0.008
						(0.123)	(0.122)	(0.122)	(0.121)	(0.119)
Northern Cape						0.038	0.056	0.068	0.070	0.063
						(0.154)	(0.152)	(0.148)	(0.149)	(0.148)
Free State						0.468***	0.454***	0.464***	0.456***	0.452***
						(0.141)	(0.136)	(0.139)	(0.136)	(0.136)
KwaZulu-Natal						0.309*	0.371**	0.355**	0.388**	0.381**
						(0.174)	(0.175)	(0.172)	(0.172)	(0.177)
North-West						0.498***	0.517***	0.522***	0.530***	0.517***
						(0.149)	(0.147)	(0.148)	(0.147)	(0.147)
Gauteng						0.311***	0.326***	0.325***	0.333***	0.327***
						(0.117)	(0.113)	(0.117)	(0.115)	(0.113)
Mpumalanga						0.447***	0.488***	0.484***	0.503***	0.494***
						(0.131)	(0.127)	(0.130)	(0.127)	(0.128)
Limpopo						0.410*	0.439*	0.457**	0.461**	0.455**
						(0.231)	(0.226)	(0.229)	(0.227)	(0.227)
Constant	5.573***	5.395***	5.291***	5.265***	5.228***	6.571***	6.211***	6.273***	6.105***	6.188***
	(0.534)	(0.551)	(0.552)	(0.560)	(0.553)	(0.622)	(0.591)	(0.607)	(0.583)	(0.591)
N	2,093	2,093	2,093	2,093	2,093	3,621	3,621	3,621	3,621	3,621
R-squared	0.487	0.489	0.490	0.490	0.491	0.398	0.407	0.404	0.409	0.408
F-Stat	140.2	128.6	129.9	116.1	133.6	77.57	74.53	70.93	67.52	74.39

Notes: (1) The table presents regression of log real earnings based on employed males who are between 25 & 59 years of age.

(2) The reference category for race is African, and for province the reference category is *The Cape* in the case of SSOA and *Western Cape* in the case of NIDS.

(3) Results are weighted using the calibrated weights provided with the datasets. Parental education refers to the average education of parents.

(4) Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

The quadratic specification for the education of respondents is highly significant in both samples, signifying a strong non-linear (convex) relationship between one's education and earnings (Table 4.3). According to the specification that includes the education of both parents (columns 4 and 8 in Table 4.3), keeping the other explanatory variables constant, in SSOA the return to education increases with education in a convex pattern, and in NIDS the returns to education decrease until 3.5 years of schooling and increase afterwards. Graphs of the predicted log real earnings for given levels of education are given in Appendix 4.B. The graphs suggest that the relationship between education and earnings has intensified in convexity over time, which concurs with the results in Keswell & Poswell (2004) and Branson & Lam (2022).

The addition of the education of parents either individually or simultaneously reduces the magnitude of the returns to own education in comparison to the returns observed in the specifications without the education of parents.[46] For instance, the addition of the education of both parents reduces the returns to own education by about 10.3 per cent in SSOA and 9.9 per cent in NIDS.[47] Likewise, the addition of parental education (i.e. the average education of parents) reduces the returns to own education by about 11.2 per cent in SSOA and 9.0 per cent in NIDS. Furthermore, the addition of the education of parents either individually or simultaneously has no effect on the coefficient of determination. The reduction in the returns to own education after controlling for the education of parents suggests that earnings of children are linked to their parents not only through parental investment in the schooling of the child but also through the intergenerational transmission of endowments (e.g., social capital acquired through family reputation and connections, life skills acquired through belonging to a particular family culture, genetic transmission of ability and other productive characteristics).

Further, the partial effect of the education of parents on sons' earnings is shown to be positive in both samples (Table 4.3). In the regressions that include the education of both parents, in SSOA

[46] We don't account for measurement error in schooling. Further, when schooling is measured with error, the addition of variables that are highly correlated with the observed schooling tends to increase the downward bias of the schooling coefficient due to measurement error (Griliches, 1977).

[47] For each sample, we first calculate the average marginal effects (AMEs) of own education in the two models (with and without the education of parents) and then compute the difference in AMEs between the models.

mothers' and fathers' education are both positively associated with sons' earnings, however, neither is statistically significant (column 4 in Table 4.3). In NIDS mothers' and fathers' education are both significantly positively associated with sons' earnings, with higher significance shown for the education of mothers (column 8 in Table 4.3). A comparison of the returns to the education of parents between the two surveys shows that while fathers' education is strongly associated with sons' earnings in SSOA, mothers' education is strongly associated with sons' earnings in NIDS. Although the regression results indicate that the relative significance of the education of mothers and fathers have changed over time, these changes may not be interpreted in terms of changes in the relative importance of the education of parents. As indicated in chapter 3, the correlation between the education of mothers and fathers is highly strong in both datasets. In the context of assortative mating, the strong correlation between the education of parents implies that the variables are proxies for each other. In such cases, what is more significant is just a function of what is reported with less error (Lubotsky & Wittenberg, 2006). Thus, it may not be appropriate to make conclusions about the relative importance of the education of parents to the earnings of their children based on the regression output alone.

In light of the strong correlation between the education of mothers and fathers, we use the average education of parents as a measure for parental education to analyse the partial correlation between the education of parents and children's earnings. The regression results are given in columns 5 and 10 of Table 4.3 for SSOA and NIDS, respectively. The results indicate that in SSOA, keeping the other variables constant, one additional year of parental education is associated with about 2.8 per cent increase in predicted earnings. In NIDS, the same scenario leads to about 3.4 per cent increase in earnings. As shown in Table 4.3, the coefficient for NIDS is statistically more significant and is more precisely estimated. While the partial effect of parental education on sons' earnings is shown to be stronger in NIDS than in SSOA, analysis of the overall association between the two has revealed that the association is stronger in SSOA than in NIDS. One additional year of parental education is associated with about 14.1 per cent increase in predicted earnings in SSOA, and with about 11.7 per cent increase in earnings in NIDS. The smaller returns to parental education in SSOA (than in NIDS) when we control for children's education indicates that the indirect effect of parental education (that runs through children's education) on children's earnings is larger in SSOA than in NIDS. We have seen in

chapter 3 that the overall association between parental education and children's education is larger in SSOA than in NIDS.

The indicator variables for the missingness of the education of parents in the earnings regressions are not statistically significant in both samples (Table 4.3). This indicates that the earnings of those with missing parent education are not systematically different from the earnings of those with non-missing parent education. We note that the coefficients for the education of parents (using our approach) are slightly lower in magnitude in comparison to those obtained using the default approach in statistical packages (i.e., excluding all observations with missing values in the variables of analysis).

The regression output in Table 4.3 also shows some interesting changes in the extent of association between the other explanatory variables and sons' earnings. The returns to being coloured, Indian/Asian or white relative to being African are all seen to have decreased between SSOA and NIDS. In NIDS, the relative returns are significant for coloureds and whites, but not for Indians/Asians. The insignificance of the coefficient for Indians/Asians may be due to their small size in the sample rather than real changes in their relative returns over time. The returns to being married and residing in an urban area (both dummies) are seen to have increased between the two surveys. The returns to being married are statistically highly significant in both periods. This accords with Casale & Posel (2010) who document a significant male marital premium among African men.

4.5.1.2 Robustness Check

The robustness of the regression coefficient estimates is probed by restricting the analysis to employed sons between 30 and 49 years of age. The regression output is presented in Table 4.4. The reduction in the returns to own education when we control for the education of parents is larger in this sample than in the baseline sample. For instance, the addition of the education of both parents reduces the returns to own education by about 14.1 per cent in SSOA and 11.8 per cent in NIDS relative to the returns observed in the specification without the education of parents (vs. 10.3 per cent in SSOA and 9.9 per cent in NIDS in the baseline sample). Further, the magnitude of the partial effects of parental education on sons' earnings are larger in this sample

than in the baseline. Keeping the other variables constant, one additional year of parental education is associated with earnings increase of about 3.8 per cent in SSOA, and 4.2 per cent in NIDS. These results give further support to the importance of the direct effect of the education of parents on children's earnings.

Table 4.4 Regression of log real earnings: Robustness check

`	SSOA			NIDS-w5		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Child education	0.018	0.019	0.013	-0.164***	-0.149***	-0.149***
	(0.028)	(0.027)	(0.029)	(0.040)	(0.039)	(0.040)
Child education squared	0.005***	0.004***	0.004***	0.020***	0.018***	0.018***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Mother education		0.005			0.035***	
		(0.014)			(0.010)	
Father education		0.031*			0.009	
		(0.017)			(0.010)	
Parental education			0.038**			0.042***
			(0.015)			(0.009)
Mother educ. missing	-	Yes	-	-	Yes	-
Father educ. missing	-	Yes	-	-	Yes	-
Parental educ missing	-	-	Yes	-	-	Yes
Race	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes
Age squared	Yes	Yes	Yes	Yes	Yes	Yes
Married	Yes	Yes	Yes	Yes	Yes	Yes
Urban	Yes	Yes	Yes	Yes	Yes	Yes
Province	Yes	Yes	Yes	Yes	Yes	Yes
Constant	4.487***	4.229***	4.205***	4.311*	3.211	3.314
	(1.590)	(1.542)	(1.541)	(2.214)	(2.115)	(2.198)
N	1,369	1,369	1,369	2,138	2,138	2,138
R-squared	0.486	0.489	0.491	0.401	0.416	0.416
F-Stat	94.36	82.30	94.21	48.57	42.08	47.03

Notes: (1) The table presents OLS regression of log real earnings based on employed males between 30 & 49 years of age. Parental education refers to the average education of parents.
(2) Results are weighted using the calibrated weights provided with the datasets.
(3) Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

The age restriction (of the samples) has increased the bigger coefficients and decreased the smaller coefficients of the education of parents in comparison to the baseline results in the specification that includes the education of both parents (columns 2 &5 of Table 4.4). Thus, in SSOA fathers' education is found to be significantly associated with sons' earnings, which was

not the case in the baseline result. In NIDS only mothers' education is found to be significantly associated with sons' earnings, while fathers' education is made insignificant. These changes probably reflect the multicollinearity between the education of mothers and fathers, and what is made more significant is just a function of what is reported with less error.

4.5.1.3 Regression of log real earnings: Females

The partial association between the education of parents and children's earnings is further analysed by considering a sample of female. The analysis is carried out based on employed daughters who are between 25 and 59 years of age. The regression output is given in Appendix 4.C. The results show that the addition of the education of parents in the regression equations reduces the magnitude of the returns to own education relative to the returns observed in the specification without the education of parents in both SSOA and NIDS. For instance, the addition of the education of both parents reduces the returns to own education by about 11.2 per cent in SSOA and 6.2 per cent in NIDS. In addition, the partial effect of the education of parents on daughters' earnings is shown to be positive and statistically significant in both samples. In SSOA, keeping the other variables constant, one additional year of parental education is associated with about 3.6 per cent increase in predicted earnings. In NIDS, the same scenario leads to about 3.0 per cent increase in earnings. The results for daughters support the results we obtained for sons in terms of the importance of the education of parents to the earnings of their children.

A comparison of the partial association between parental education and earnings for daughters and sons between the two surveys shows that, unlike for sons, daughters' earnings are more strongly associated with parental education in SSOA than in NIDS. We should, however, note that the estimates for daughters are likely to be biased due to sample selection issues that arise not only from employment participation decisions but also from employment opportunity restrictions for women - especially in the first period (SSOA).

4.5.2 Decomposition of earnings inequality

As discussed in section 4.3.1, we use regression-based decomposition of inequality to analyse the contribution of parental background to children's earnings.[48] [49] The decompositions are based on the earnings regressions that use parental education (i.e., the average of the education of parents) for the education of parents.

4.5.2.1 Decomposition of earnings inequality: Males

Table 4.5 presents the decomposition of inequality in log real earnings. The columns for SSOA and NIDS-w5 in the table show the relative factor inequality weights, which are the percentage contribution of the explanatory factors to earnings inequality in the two samples. In SSOA the largest contributor to earnings inequality, after the residual, is race. Race is closely followed by education.[50] On the other hand, in NIDS the largest contributor to earnings inequality, after the residual, is solely education. The contribution of race to inequality in NIDS is quite small.

The changes in the contributions of race and education to inequality between the two surveys is noteworthy. The contribution of education seems to have remained almost unchanged between the two surveys (20.1 per cent in SSOA and 22.6 per cent in NIDS). On the other hand, the contribution of race has decreased between the two surveys (25.5 per cent in SSOA and 5.4 per cent in NIDS). We should, however, be cautious in interpreting the contributions of race and education to inequality in the two surveys. During apartheid, as discussed in chapter 2, education systems were nested within race and differed greatly in content and quality of provision across races. Thus, it may not be possible to fully untangle the contribution of race from that of education in the first survey. It is likely that a part of the contribution of race to inequality in-

[48] In SSOA, log earnings of individuals are the predicted log earnings from interval regression. The predicted earnings are constrained to fall within their brackets. Although the within bracket imputation of log earnings can minimize prediction errors, the fact that we are dealing with predicted values is likely to reduce the variability in log earnings.

[49] Results are obtained using STATA's ineqrbd module (Fiorio & Jenkins, 2007).

[50] The contribution of education to earnings inequality is the sum of the contribution of education and the square of education.

Table 4. 5 Contribution of parental education to earnings inequality: Males

	Percentage contribution of explanatory factor to earnings inequality $[s_k(ln\ Y)]$ in:	
Factor	**SSOA**	**NIDS-w5**
Education	20.11 (0.006)	22.63 (0.006)
Parental education	5.99 (0.002)	4.86 (0.002)
Race	25.54 (0.007)	5.43 (0.002)
Age	1.58 (0.003)	0.95 (0.002)
Marital Status	2.04 (0.003)	4.48 (0.003)
Urban	2.27 (0.002)	1.01 (0.001)
Province	-0.90 (0.001)	1.09 (0.002)
Residual	43.37 (0.011)	59.55 (0.008)
Total	**100.00**	**100.00**

Notes: (1) The table presents the percentage contribution of the explanatory factors to earnings inequality in SSOA and NIDS-w5. Numbers in parenthesis are standard errors (for the percentage contributions in their decimal form). The analysis is based on employed males who are between 25 and 59 years of age.
(2) The results are obtained using STATA's ineqrbd module. In the associated regression of log earnings on the explanatory variables, education and age are entered as polynomials of degree 2, race and province are entered as categorical variables, and marital status and urban are entered as dummy variables.

Table 4.6 Contribution of parental education to inequality: Robustness check

	Percentage contribution of explanatory factor to earnings inequality $[s_k(lnY)]$ in:	
Factor	**SSOA**	**NIDS-w5**
Education	23.70 (0.008)	24.83 (0.008)
Parental education	8.76 (0.003)	6.74 (0.003)
Race	24.18 (0.008)	2.87 (0.002)
Age	0.56 (0.002)	0.48 (0.002)
Marital Status	3.27 (0.004)	4.60 (0.004)
Urban	2.93 (0.003)	0.85 (0.001)
Province	-2.88 (0.003)	1.18 (0.003)
Residual	39.49 (0.014)	58.47 (0.011)
Total	**100.0**	**100.0**

Notes: (1) The table presents the contribution of the explanatory factors to earnings inequality in SSOA and NIDS-w5. Numbers in parenthesis are standard errors (for the percentage contributions in their decimal form). The analysis is based on employed males who are between 30 and 49 years of age.

SSOA is attributable to education. In general, earning returns across racial groups may be different not only by education but also by other productive characteristics (e.g., work experience) in South Africa.[51] Nevertheless, the substantial decrease in the contribution of race between the two surveys may also indicate an actual decline in the importance of race to inequality over time. Seekings & Nattrass (2005) argue that the importance of race in explaining inequality has diminished over time.

The contribution of parental education[52] to earnings inequality is shown to be quite small in both samples. The contribution of parental education is about 6.0 per cent in SSOA and 4.9 per cent in NIDS. These percentages are small in comparison to the contribution of own education. We should, however, note that these contributions are due to the partial effect (of parental education on sons' earnings).

The contribution of the residual to earnings inequality is quite large in the first survey and is even larger in the second survey (43.4 per cent in SSOA and 59.6 per cent in NIDS). The increase in the contribution of the residual suggests that other factors not accounted for by our regression model are gaining importance in their contribution to earnings inequality. Moreover, differences in measurement between the two surveys could be a contributing factor.

4.5.2.2 Robustness Check

The robustness of the results of the decomposition of earnings inequality is probed by restricting the analysis to employed sons who are between 30 and 49 years of age. The decomposition result is presented in Table 4.6. The results are qualitatively similar to those obtained based on males between 25 and 59 years of age. However, the contribution of parental education to earnings inequality is slightly larger in this sample than in the baseline sample in both SSOA and NIDS.

[51] In this respect, a more elaborate regression model could have been one that also includes the interaction terms of race with other characteristics of individuals (e.g. education). However, for our comparative analysis, given the change in the system of governance and the subsequent changes in access to education, employment and other opportunities, we have opted not to include race interactions.

[52] Parental education is defined as the average education of parents. The results for the contribution of the education of parents to inequality remains almost unchanged if both mothers' and fathers' education are used instead.

4.5.2.3 Decomposition of earnings inequality: Females

The decomposition of earnings inequality is further examined by considering a sample of female. The analysis is carried out based on employed daughters who are between 25 and 59 years of age. The decomposition result is presented in Appendix 4.D. The results are somewhat different from those obtained for sons. In SSOA, unlike in the case of sons, education is the only large contributor to earnings inequality, after the residual. The contribution of race to inequality is quite small compared to its contribution in the case of sons. In addition, the contribution of education to inequality for daughters is quite large in both SSOA and NIDS in comparison to those shown for sons. As indicated in section 4.5.1, the different results based on the samples of daughters and sons are likely due to sample selection issues that arise not only from employment participation decisions but also from employment opportunity restrictions for women - especially in the first period (SSOA).

The contribution of parental education to inequality is small in comparison to the contribution of own education and is somewhat larger in SSOA than in NIDS (8.0 per cent in SSOA and 4.7 per cent in NIDS).

4.6 Discussion of Results

The chapter examined changes over time in the association between parental background (proxied by the education of parents) and children's earnings and the importance of parental background as a source of earnings inequality in South Africa. The analysis was carried out using two datasets: namely, the Survey of Socioeconomic Opportunity and Achievement (SSOA) and the 5th wave of the National Income Dynamics Study (NIDS-w5). As indicated in chapter 1, the two datasets differ in their degree of representation of the population. NIDS is a nationally representative sample, but SSOA missed farmworkers who were living on commercial farms. Thus, the objective was not to quantify the change in the association between parental background and earnings or the contribution of parental background to earnings inequality between the two periods but to assess the importance of parental background to earnings and earnings inequality over time. The analysis has revealed several interesting results.

The addition of the education of parents in a Mincerian-type regressions was shown to reduce the magnitude of the returns to own education in both surveys. Moreover, the reduction in the returns to own education is slightly larger in the first survey than in the second. The addition of parental education (i.e., the average education of parents) reduces the returns to own education by about 11.2 per cent in SSOA and 9.0 per cent in NIDS relative to the returns observed in the specification without the education of parents. The reduction in the returns to own education after controlling for the education of parents suggests that earnings of children are linked to their parents not only through parental investment in the schooling of children but also through the intergenerational transmission of endowments. The chapter did not investigate further whether the reduction in the returns to own education when we controlled for parental background represents nepotistic practices in the labour market or unobserved individual characteristics. However, empirical studies for other countries (e.g., Lam & Schoeni (1994) for Brazil and Agnarsson & Carlin (2002) for Sweden) have found evidence that suggests that parental background effects primarily reflect unobserved individual characteristics rather than nepotism in the labour market. In South Africa, the distribution of employment opportunities along racial lines during apartheid could be seen as a type of collective nepotism. This may indicate that the importance of parental background effects that represent nepotistic practices in the labour market were quite strong in the first period. On the other hand, the shift in the economy towards skills-biased technological change in the post-apartheid years may have increased the relevance of parental background effects that represent unobserved individual productive characteristics.

Further, the partial effects of the education of parents on children's earnings were shown to be positive and statistically significant in both periods. For instance, one additional year of parental education is associated with earnings increase of about 2.8 per cent in SSOA and 3.4 per cent in NIDS, for a given level of child's education. The chapter did not attempt to compare the estimates for the two periods, owing to the differences in the degree of representativeness of the two samples. However, the results suggest that the partial effects of parental education on children's earnings have remained relatively large and statistically significant over time.

The contribution of education to earnings inequality was shown to be large in both periods. On the other hand, the contribution of parental education to earnings inequality was shown to be

small in both periods. However, the contribution of parental education to inequality could be economically important as these contributions are above and beyond the contribution of the child's education to earnings inequality.

The chapter started by asking what the role of parental background has been in the social process that determines earnings, given the post-apartheid trends in educational attainment and earnings inequality. Firstly, it is important to note that the racial distribution of the employed may have changed over time. As shown in our descriptive statistics section, in SSOA the racial distribution of the employed featured higher proportion of whites (and lower proportion of Africans) in comparison to their size among the labour force participants. On the other hand, in NIDS the racial distribution of the employed was shown to be similar to that of the labour force participants. Secondly, although the chapter did not statistically compare the returns to parental education between the two periods, the relatively larger magnitude in the second period may indicate an increasing role of parental background in the determination of children's earnings. Overall, the results highlighted the continued relevance of parental background to earnings and earnings inequality in the highly unequal South African labour market.

Appendix

4.A Definition of variables and their sample statistics

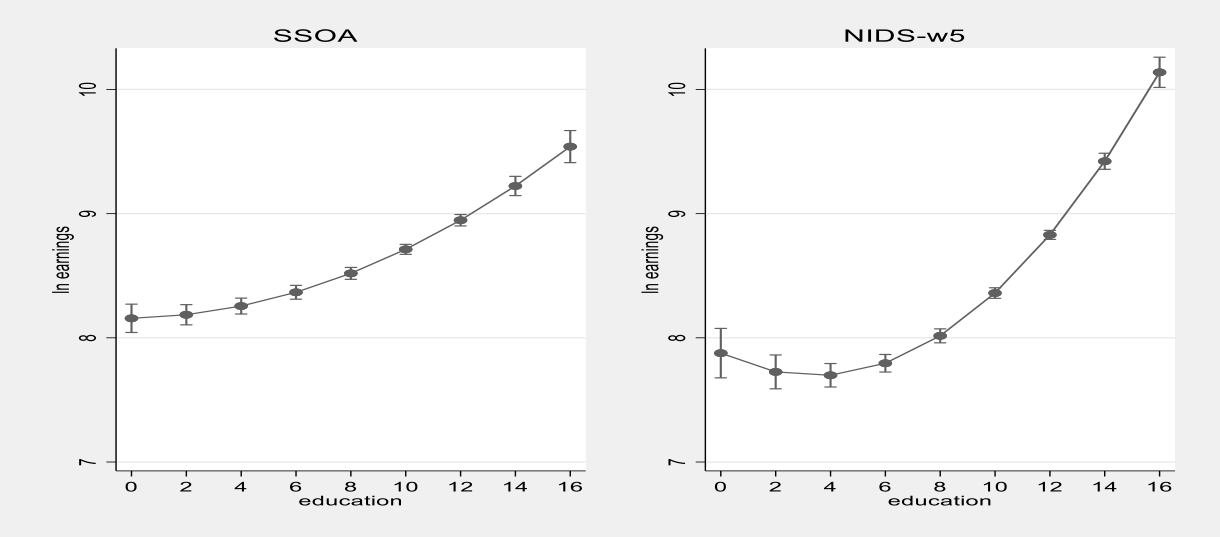
Table 4.A.1 Definition of variables and sample statistics: Labour force participating males

		SSOA			NIDS-w5		
Variable	**Variable definition**	n	Mean	St.dev.	n	Mean	St.Dev.
ch_edu	Completed years of education of child	3,131	8.64	4.09	4,921	10.70	2.86
ch_edu_sq	child education squared	3,131	91.36	63.20	4,921	122.58	52.11
mth_edu	Completed years of education of mother	2,742	5.00	4.44	4,425	5.79	4.86
fth_edu	Completed years of education of father	2,442	5.60	4.60	3,888	5.41	5.07
prnt_edu	Average parental education	2,860	5.20	4.28	4,541	5.64	4.54
married	=1 if ever married, 0 otherwise	3,129	0.79	0.41	4,943	0.48	0.50
urban	=1 if lives in urban area, 0 in rural area	3,131	0.72	0.45	4,939	0.73	0.44
Race							
African	=1 if African, 0 otherwise	3,131	0.65	0.48	4,941	0.81	0.39
coloured	=1 if coloured, 0 otherwise	3,131	0.09	0.28	4,941	0.09	0.28
Indian	=1 if Indian/Asian, 0 otherwise	3,131	0.05	0.21	4,941	0.03	0.16
white	=1 if white, 0 otherwise	3,131	0.21	0.41	4,941	0.08	0.27
Age	Age of respondent	3,131	37.81	9.20	4,943	37.67	9.17
Age_sq	Age squared	3,131	1514.40	742.67	4,943	1503.03	741.83
N		**3,131**			**4,943**		

Notes: (1) The summary stats are based on labour force participating males between 25 and 59 years of age in SSOA and NIDS-w5.
(2) Results are weighted using the calibrated weights provided with the datasets.

4.B Relationship between own education and earnings: Males

Figure 4.B.1 Predicted log real earnings for given levels of education in SSOA and NIDS



Notes: (1) The figures present the predicted log real earnings for given levels of education (based on averages of the predictions over the estimation sample) with the other explanatory variables in the regression models set at their mean values. The figures are based on the regression specifications that include the average education of parents (columns 4 and 8, in Table 4.3). The figures show that the relationship between education and earnings follows a convex pattern in SSOA and NIDS-w5. Moreover, the relationship has intensified in convexity over time.

4.C Regression of log real earnings: Females

Table 4.C.1 Regression of log real earnings: Females 25-59 years of age

`	SSOA			NIDS-w5		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Child education	0.026	0.021	0.018	-0.169***	-0.162***	-0.166***
	(0.043)	(0.044)	(0.044)	(0.024)	(0.023)	(0.023)
Child education squared	0.008***	0.008***	0.008***	0.022***	0.021***	0.021***
	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
Mother education		0.002			0.008	
		(0.016)			(0.007)	
Father education		0.039***			0.025***	
		(0.014)			(0.007)	
Parental education			0.036***			0.030***
			(0.013)			(0.007)
Mother educ. missing	-	Yes	-	-	Yes	-
Father educ. missing	-	Yes	-	-	Yes	-
Parental educ missing	-	-	Yes	-	-	Yes
Race	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes
Age squared	Yes	Yes	Yes	Yes	Yes	Yes
Married	Yes	Yes	Yes	Yes	Yes	Yes
Urban	Yes	Yes	Yes	Yes	Yes	Yes
Province	Yes	Yes	Yes	Yes	Yes	Yes
Constant	6.030***	5.963***	5.877***	5.843***	5.460***	5.466***
	(0.801)	(0.778)	(0.776)	(0.462)	(0.447)	(0.459)
N	1,262	1,262	1,262	3,856	3,856	3,856
R-squared	0.415	0.422	0.426	0.473	0.482	0.481
F-Stat	44.91	37.36	41.23	115	107.3	109.6

Notes: (1) The table presents OLS regression of log real earnings based on employed females who are between 25 & 59 years of age.

(2) Results are weighted using the calibrated weights provided with the datasets.

(3) Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

4.D Contribution of parental education to inequality: Females

Table 4.D.1 Contribution of parental education to earnings inequality: Females 25-59 years of age

	Percentage contribution of explanatory factor to earnings inequality $\{s_k(lnY)]$ in:	
Factor	**SSOA**	**NIDS-w5**
Education	32.13 (0.012)	37.90 (0.007)
Parental education	7.97 (0.004)	4.66 (0.002)
Race	9.01 (0.005)	2.49 (0.001)
Age	-0.55 (0.003)	-0.04 (0.002)
Marital Status	-0.14 (0.002)	0.10 (0.000)
Urban	-0.99 (0.002)	2.03 (0.001)
Province	1.65 (0.002)	0.82 (0.001)
Residual	50.93 (0.015)	52.04 (0.008)
Total	**100.0**	**100.0**

Notes: (1) The table presents the contribution of the explanatory factors to earnings inequality in SSOA and NIDS-w5. Numbers in parenthesis are standard errors (for the percentage contributions in their decimal form). The analysis is based on employed females who are between 25 and 59 years of age.
(2) The results are obtained using STATA's ineqrbd module. In the associated regression of log earnings on the explanatory variables, education and age are entered as polynomials of degree 2, race and province are entered as categorical variables, and marital status and urban are entered as dummy variables.

Chapter 5

5. Intergenerational Earnings Persistence in South Africa: How much has South Africa changed since the early 1990s?

5.1 Introduction

This chapter focuses on changes over time in the economic persistence of South Africans. To situate the chapter within the thesis, in chapter 3 we investigated changes over time in the education mobility of South Africans. The results showed that intergenerational persistence in education has been declining across birth cohorts. In chapter 4 we examined changes over time in the importance of parental background to children's earnings. The results showed that the partial effect of parental education on children's earnings has remained positive and statistically significant over time.

This chapter extends the discussion by investigating changes over time in the association between the earnings of parents and their children once these are adults. The comparative analysis is carried out using SSOA and NIDS-w5. As we will discuss in the next section, given the model of human capital, economic persistence can change over time due to changes in the returns to human capital, the relative investment of rich versus poor parents in the human capital of their children, the returns to genetically transmitted earning-related traits, and public investment in the human capital of children. South Africa, in the post-apartheid period, as indicated in chapter 2, witnessed an expansion of public investment in education and other social services. This has the potential to increase intergenerational economic mobility. On the other hand, South Africa has experienced increasing returns to education (e.g., Branson & Leibbrandt, 2013; Lam, Finn & Leibbrandt, 2015). This has the potential to decrease economic mobility.

The chapter further investigates changes in the mediating role of education in intergenerational earnings persistence between the two periods. Education forms an integral part of children's human capital and plays an important mediating role in the transmission of economic advantages across generations. This is because parents can more directly influence a child's education and thereby increase the human capital and earnings potential of the child.

The empirical evidence on economic inequality in South Africa is rather extensive. However, besides a couple of studies on intergenerational inequality, the focus is mostly on cross-sectional inequality. The findings of this literature show that cross-sectional economic inequality has been persistently high (e.g., Bhorat, van der Westhuizen & Jacobs, 2009; Leibbrandt *et al.*, 2010; Wittenberg, 2017). This chapter seeks to enrich the existing literature by providing empirical evidence on changes over time in economic inequality in the intergenerational dimension.

The chapter adds to the literature in several ways. First, we show how earnings can be adjusted for employment volatility to obtain a measure of the permanent earnings of individuals. With high levels of employment volatility in the South African labour market, obtaining the permanent earnings of individuals can be quite difficult. Second, we highlight and disentangle the confounding effect of parental education in analyzing the contribution of the mediating role of education to intergenerational earnings persistence. Controlling the confounding effect of parental education is especially important when one is interested in changes over time in the contribution of the mediating role of education. Parental education impacts not only parental earnings but also children's education & earnings and these relationships can change over time. Third, we provide empirical evidence of changes in intergenerational earnings persistence between two well-spaced periods where one belonged to the pre-1994 and the other to the post-1994 periods. In addition, we provide evidence of changes in the contribution of the mediating role of education to intergenerational earnings persistence between the two periods.

The chapter is structured into 6 sections. Section 2 provides a review of the relevant literature. We discuss the potential determinants of changes over time in intergenerational persistence within the framework of the human capital model of intergenerational economic mobility. Further, we discuss the methodological and data issues in estimating intergenerational persistence and present a brief review of the empirical literature. Section 3 presents the methods that we use to estimate intergenerational persistence and the mediating role of education in the intergenerational transmission of earning advantages. Section 4 identifies and describes the datasets. This is followed by the empirical results. The empirical analysis is carried out using unadjusted and adjusted earnings for employment volatility. The last section discusses the main

findings of the chapter. It highlights why intergenerational earnings persistence in South Africa has remained high in the post-apartheid period.

5.2 Literature Review

5.2.1 Theoretical Framework

Empirical studies of intergenerational economic persistence largely rely on the human capital model of Becker and Tomes (1979). The model demonstrates mechanisms of intergenerational persistence on the assumption that each family allocates resources in a way that maximizes its well-being and the well-being of its descendants. The model has been adapted and simplified to rationalize the double-log functional form of the regression equation for estimating intergenerational elasticity that involves only two generations (e.g., Solon, 1999, 2004, 2014). In this section, we follow Solon (2004 & 2014) to discuss the potential determinants of the optimal parental investment in the human capital of the child and the mechanisms through which intergenerational persistence may change over time.

The simplified model assumes families that are made up of a single parent of generation $t-1$ and a child of generation t. Each family allocates the parent's lifetime earnings ($Y_{i,t-1}$) between its own consumption ($C_{i,t-1}$) and investment towards the child's human capital ($I_{i,t-1}$). This gives rise to the parent's budget constraint as:

$$Y_{i,t-1} = C_{i,t-1} + I_{i,t-1} \tag{5.1}$$

The amount of human capital the child receives ($H_{i,t}$) is a function of the parent's investment in the child with the functional relationship assumed to be:

$$H_{i,t} = \theta \ln I_{i,t-1} + \mathsf{E}_{i,t} \tag{5.2}$$

where $\theta > 0$ is a measure of the effectiveness of parental investment in the child's human capital and $\mathsf{E}_{i,t}$ represents the human capital endowment of the child that is independent of the level of parental investment. These endowments include the genetic transmission of abilities and other characteristics (e.g., race), family reputations and connections, and life skills learnt through

belonging to a particular family culture (Becker & Tomes, 1979). Solon (2004 & 2014) assumes that $\mathrm{E}_{i,t}$ follows a first-order autoregressive process.

$$\mathrm{E}_{i,t} = \tau + \lambda\, \mathrm{E}_{i,t-1} + \nu_{i,t} \tag{5.3}$$

where λ, which lies between 0 and 1, represents the degree of heritability of endowment from parent to child.

The lifetime earnings of the child ($Y_{i,t}$) depends on its human capital, with the functional relationship assumed to be given by:

$$ln\mathrm{Y}_{\mathrm{i,t}} = \mu + \pi \mathrm{H}_{\mathrm{i,t}} \tag{5.4}$$

where π represents the earning returns to human capital and μ is a constant term. Substituting equation 5.2 into equation 5.4 yields:

$$lnY_{i,t} = \mu + (\pi\theta) lnI_{i,t-1} + \pi E_{i,t} \tag{5.5}$$

where $\pi\theta$ represents earning returns to parental investment in the human capital of the child.

The family's problem is to decide on how much to invest in the child's human capital given its budget constraint (equation 5.1). The Cobb-Douglas utility function of the family is assumed to be given by:

$$U_i = (1-a) lnC_{i,t-1} + a lnY_{i,t} \tag{5.6}$$

where a, which lies between 0 and 1, is a measure of the degree of parental altruism and represents the parent's preference towards the child's future earnings relative to current own consumption. The setup is a typical optimization problem, where a family maximizes its lifetime utility (as described in equation 5.6) with the choice variable being the level of investment in the child's human capital. The solution to the first-order condition for the optimal level of investment is:

$$\mathrm{I}_{i,t-1} = \left\{ \frac{a(\pi\theta)}{1\text{-}a(1\text{-}(\pi\theta))} \right\} \mathrm{Y}_{i,t-1} \tag{5.7}$$

Equation 5.7 shows that the optimal level of parental investment in the human capital of the child positively depends on parental altruism (a), the earning returns to the child's human capital (π), the effectiveness of parental investment in the child's human capital (θ) and parental riches ($Y_{i,t-1}$). Solon (2004) further shows that parental investment negatively depends on public investment (in a model of intergenerational earnings persistence that includes the effect of public investment on parental investment behaviour).

Substituting equation 5.7 for $\mathrm{I}_{i,t-1}$ into equation 5.5, we obtain a regression equation for the intergenerational association between the child's earnings ($Y_{i,t}$) and the parent's earnings ($Y_{i,t-1}$).

$$ln\mathrm{Y}_{\mathrm{i,t}} = \mu^{*} + \pi\theta ln\mathrm{Y}_{\mathrm{i,t\text{-}1}} + \pi\mathrm{E}_{\mathrm{i,t}} \qquad (5.8)$$

where $\mu^{*} = \mu + (\pi\theta)ln\{(a\pi\theta)/[1 - a(1 - \pi\theta)]\}$ is the constant term. The child's endowment ($\mathrm{E}_{\mathrm{i,t}}$) is unobserved. Equation 5.8 is the intergenerational mobility equation that is typically used in empirical research with the last term in the right-hand side incorporated into the error term. The equation does not establish causality between $lnY_{i,t-1}$ and $lnY_{i,t}$, but simply shows the association between the earnings of children and their parents.

In a steady state, where the variances of $lnY_{i,t}$ and $lnY_{i,t-1}$ are the same, the slope coefficient from the regression of $lnY_{i,t}$ on $lnY_{i,t-1}$ and the correlation coefficient between $lnY_{i,t}$ and $lnY_{i,t-1}$ are equal. Through this equivalence, Solon (2014) shows that the steady state intergenerational elasticity (β) is given by: $\beta = \frac{(\pi\theta)+\lambda}{1+(\pi\theta)\lambda}$.

The above expression for β shows that the intergenerational elasticity positively depends on the earning returns to the child's human capital (π), the effectiveness of parental investment in the child's human capital (θ), and on the degree of heritability of endowments (λ). In a model of intergenerational earnings persistence that includes the effect of public investment on parental investment behaviour, Solon (2004) further shows that the elasticity negatively depends on the progressivity of public investment. These results imply that intergenerational persistence can change over time due to changes in any of those factors.

Another family factor, besides parental investment in the human capital of their children, that may influence economic persistence across generations is parental choice of residential neighbourhood. The effects of neighbourhood on intergenerational persistence have been gaining traction. Mogstad and Torsvik (2023) review the channels of transmission. The channels are quite relevant to the South African context, and include:

(i) Quality of local institutions - residential segregation by parental income can affect the quality of schools and other local institutions that are important for human development, which in turn influences intergenerational mobility.

(ii) Community level of human capital – with residential segregation, children of high-income parents are more likely to interact with other advantaged children, and this may in turn increase the returns to parental investments in children's education.

(iii) Peer effects and social interactions – the attitudes, ambitions, beliefs and constraints that underlay life choices in formative years may partly be shaped by peers in the neighbourhood. Furthermore, the ambitions parents have for their children may depend on attitudes and decisions made by peer parents in the neighbourhood.

(iv) Economic inequality - increased economic inequality may lead to more segregated neighbourhoods. With residential segregation by socioeconomic status, children of low-income parents may be disadvantaged in terms of the quality of schooling and the social influences of peers they receive, and this may negatively impact their adult economic outcomes.

Some recent studies have also discussed additional drivers of intergenerational persistence in a developing-country context, which include labour market segmentation, imperfect credit and insurance markets, information frictions (Piraino, 2020), and group discrimination (Funjika & Gisselquist, 2020).

5.2.2 Estimation Issues

Conceptually we are interested in the association between the permanent earnings of parents and their adult children. The main estimation issues are centred around the measurement of permanent earnings. The ideal sources of data are long-running panel data or administrative

records that link earnings across generations. These data sources are hard to come by and that poses several estimation challenges. First, the intergenerational persistence equation is vulnerable to measurement errors in the permanent earnings of both parents and children. The use of current earnings of parents to proxy for their permanent earnings results in inconsistent elasticity estimates. Empirical evidence shows that non-classical measurement errors in current earnings attenuate the OLS estimate of intergenerational persistence (e.g., Haider & Solon, 2006; Böhlmark & Lindquist, 2006).[53] In addition, if you have multiple measures of current earnings (e.g., from panel data), the use of the average of the earnings to proxy for the permanent earnings of parents does not fully resolve the attenuation inconsistency (Mazumder, 2005). Similarly, on the left-hand side of the intergenerational mobility equation, the use of current earnings of children to proxy for their permanent earnings is prone to non-classical measurement errors (i.e. the systematic deviations of current from lifetime earnings over the lifecycle). However, empirical evidence shows that left-hand side measurement error is harmless for consistency if children's earnings are measured between the early 30s and mid-40s (e.g., Haider & Solon, 2006; Nybom & Stuhler, 2016).

Second, household datasets in most countries do not possess earnings information for two generations. Intergenerational elasticity estimation in such cases has been carried out using earnings imputed from different datasets. This is done using the Two-Sample Two-Stage Least Squares (TS2SLS) approach.[54] The TS2SLS approach attempts to overcome the data limitation by combining information from two separate samples of different generations from the same population. These two samples are: a 'main sample' of children with observations on their earnings and their parental characteristics (**Z**), and an 'auxiliary sample' of pseudo-parents with observations on their earnings and the same **Z** variables as in the main sample. The

[53] In theory, non-classical measurement error in the explanatory variable can either amplify or attenuate the OLS estimates (Haider & Solon, 2006)

[54]The TS2SLS estimator is a variant of the Two-Sample Instrumental Variables (TSIV) estimator. The latter is discussed by Angrist & Krueger (1992) and Arellano & Meghir (1992). Unlike the one-sample case where the Instrumental Variable (IV) and the Two-Stage Least Squares (2SLS) estimators are equivalent (in the exactly identified case), in the two-sample case the TSIV and the TS2SLS estimators are numerically distinct (Inoue & Solon, 2010).

characteristics of parents (**Z**) are the instruments that link the two samples.[55] The TS2SLS estimator provides consistent estimates of the intergenerational elasticity if either the parental characteristics have no independent effect on the children's earnings or (for the extremely unrealistic prospect of) getting an R^2 of a 100 per cent in the first stage regression (Solon, 1992). Education and occupation are the typical characteristics of parents that are used for the prediction of their earnings. These variables are likely to have a direct effect on children's earnings other than through the parents' earnings, hence the TS2SLS estimate is generally inconsistent. Björklund & Jäntti (1997), using education and occupation as predictors for the earnings of fathers, found the TS2SLS estimate to be upwardly biased by about 30 per cent.

5.2.3 Empirical Literature

5.2.3.1 Intergenerational Persistence Across Countries

Comparisons of intergenerational economic persistence estimates across countries show that persistence is generally lower in developed countries (see e.g., Blanden, 2009; Corak, 2016).[56] Corak (2016) provides comparable estimates of father-son earning elasticities for 22 countries. His results indicate Nordic countries to have the lowest levels of intergenerational persistence. The few developing countries in his study (Peru, China, Brazil, and Chile) are shown to have high levels of intergenerational persistence. Further, the study shows the presence of a positive relationship between cross-sectional inequality in the parental generation and the level of intergenerational persistence among the countries, a relationship which is referred to as the 'Great Gatsby Curve'.

Empirical studies of intergenerational earnings persistence in developing countries are not as extensive as in developed countries, with the data requirements for estimating intergenerational persistence being the main stumbling block in the former (Emran & Shilpi, 2021). However,

[55]Björklund & Jäntti (1997) is the first to use the TS2SLS technique in the estimation of intergenerational elasticity. Since then, the technique has been applied in several studies. Jerrim, Choi & Simancas (2016) provides a list of 28 studies in which the technique has been used.

[56] Direct comparison of intergenerational earnings elasticity across studies is often not reliable because studies differ in their estimation methods, sample selection rules as well as variable definitions.

there are some country-specific studies. Most of these studies are conducted using cross-sectional datasets and use the TS2SLS technique to estimate the elasticity. For example: Dunn (2007) for Brazil, Núñez and Miranda (2011) for urban Chile, Gong, Leigh & Meng (2012) for urban China, and Doan & Nguyen (2016) and Dang (2020) for Vietnam. There are also a few studies that used panel datasets. For example: Lillard & Kilburn (1995) for Malaysia, and Mohammed (2019) for rural India.

In South Africa, an early study by Hertz (2001) (as cited in Piraino, 2015) used co-residing samples of fathers and sons from the KwaZulu-Natal Income Dynamics Study (KIDS) panel dataset to estimate intergenerational earnings elasticity. Solon (2002) reports an adjusted earnings elasticity of 0.44 for South Africa based on the findings of Hertz (2001). The availability of several waves of the National Income Dynamics Study (NIDS) dataset in recent years has allowed for the estimation of intergenerational earnings persistence at the national level (e.g., Piraino, 2015; Finn, Leibbrandt & Ranchhod, 2017). Piraino (2015) used average earnings of respondents across the first three waves of NIDS to proxy for the permanent earnings of sons and the TS2SLS technique to proxy for the permanent earnings of fathers. He estimates the father-son intergenerational earnings elasticity to be in the range of 0.57-0.67.[57] Finn, Leibbrandt & Ranchhod (2017) followed the same approach as Piraino (2015) but using the first 4 waves of NIDS. Further, Finn *et al.* addresses the issue of sample selection that results from limiting the analysis to those employed. Their preferred elasticity estimates are 0.68 for father-son and 0.72 for mother-son samples. Direct comparison of the elasticity estimates of the two studies may not be appropriate as they differ somewhat in their imputing variables for the prediction of fathers' earnings. Crucially, Piraino's study did not include race as an imputing variable, although race is quite important in determining earnings in the South African context.

[57] Bloise, Brunori and Piraino (2021) have used the dataset of Piraino (2015) in their study of the effectiveness of a machine learning approach in estimating intergenerational income mobility on sub-optimal data.

5.2.3.2 Intergenerational Persistence: Changes over Time and Spatial Variation within a Country

Most of the studies that looked at changes in intergenerational persistence over time are from developed countries. Table 5.1 provides a list of studies that examined intergenerational persistence changes over time. The results of the studies indicate that intergenerational persistence has either increased or remained unchanged for the recent cohorts.

Among the studies for the US in Table 5.1, Davis & Mazumder (2017) found intergenerational income persistence to have increased for those born in the early 1960s compared to those born around 1950. The study links the increase in persistence between the two cohorts to the rise in inequality around1980. Furthermore, it attributes the increase in persistence to increases in the returns to education as well as to strengthened relationship between parental income and probability of being married. For cohorts born between 1971 and 1993, Chetty *et al.* (2014b) found intergenerational income persistence to have remained unchanged across 4-year birth cohorts.

Among the UK studies in Table 5.1, Blanden *et al.* (2004) found higher levels of intergenerational earnings persistence for those born in 1970 compared to those born in 1958. The study attributes the increase in persistence to an expansion of the university system in the late 1980s and early 1990s which benefited children from richer parents.

Intergenerational economic persistence may also vary across regions within a given country. Empirical studies of spatial variation of intergenerational persistence, although few, show heterogeneity of the extent of mobility across regions within a country (e.g., Chetty *et al*. (2014a) for the United States; Corak (2020) for Canada; Michelangeli, Östh & Türk (2022) for Sweden). These studies show that intergenerational persistence differs significantly within a country and point out that mobility is largely associated with the quality of local conditions (e.g. level of residential segregation, income inequality, poverty, and quality of schools).

5.2.3.3 The Mediating Role of Education

The paths of transmission of economic advantage across generations can be seen as either direct or indirect, with the indirect path representing parental investment in the child's human capital. Following Becker & Tomes (1979), the direct path captures factors such as the direct inheritance of property, the use of family-based social networks for occupational placing, assortative mating by social origin, and the transmission of personality traits. The indirect path captures the impact of parents' investment on the child's human capital and the returns to human capital conditional on parents' investment. Education is an important component of the child's human capital. Parents invest in the education of the child thereby increasing the earnings potential of the child.

The role of education in economic persistence is documented to change over time. Table 5.1 presents a summary of country-specific studies that examined changes over time in the mediating role of education in intergenerational persistence. Overall, the results suggest that for recent cohorts while the association between parental income and children's education has declined, the return to education has increased. A couple of studies (namely, Pekkala & Lucas (2007) for Finland and Lefranc (2018) for France) have explicitly examined the contribution of education to intergenerational persistence. Both studies found that the contribution of education to intergenerational persistence has increased for recent cohorts.

Some studies have compared the mediating role of education in intergenerational economic persistence across countries. For instance, Gregg *et al.* (2017) compares the mediating role of education in intergenerational persistence in the United States, Great Britain and Sweden. The study found that the high level of persistence in the United States and Great Britain is almost entirely accounted for by the part of the parent-child association that is not mediated by education. Parental income was found to be more important for earnings at a given level of education in the United States and Great Britain than in Sweden.

Table 5.1 Empirical Studies of Intergenerational persistence trends and Mediating Role of Education

Study	Country	Income measure	Cohorts of children compared	Method	Results	
					Intergenerational Persistence Trend	Mediating Role of Education
Levine and Mazumder (2002)	US	earnings of sons and parental income	Compares two periods: 1980 and 1993	OLS	IGE increased from 0.235 to 0.414 between the two periods.	The return to son's education rose from 5 to 10%, and the influence of parental income on son's education declined between the two periods.
Aaronson and Mazumder (2008)	US	earnings of sons; parental income	decennial Censuses from 1940 to 2000	TS2SLS	IGE exhibited a large fall between 1950 and 1980, and a sharp rise after 1980.	Including child's education in the IGE equation reduces the IGE between 30 and 45 percent over the period, without affecting the time pattern of the IGE.
Chetty et al. (2014b)	US	income of parents and children	4-year-birth-cohorts, 1971-93	OLS	rank-rank slope remained stable.	
Davis and Mazumder (2017)	US	income of parents and children	cohorts born in 1948-53 and 1961-64	OLS	IGE increased from 0.28 to 0.45 between the two cohorts. Rank-rank slope increased from 0.25 to 0.36.	The association between parent income and children's education remained relatively constant, the return to education increased between the two cohorts.
Connolly et al. (2021)	Canada	income of parents and children	4-year-birth-cohorts, 1963-85	OLS	rank-rank slope steadily increased from 0.189 for earliest cohort to 0.234 for recent cohort.	
Blanden *et al*. (2004)	UK	earnings of children and parental income	1958 and 1970	OLS	IGC rose from 0.121 to .253 between the two periods. Trend the same for sons and daughters.	Including child's education in the IGE equation, it accountes for 17% for sons and 28% for daughters of the fall in mobility between the two periods.
Nicoletti and Ermisch (2008)	UK	earnings of fathers and sons	singly-year-birth-cohorts, 1950 -72	TS2SLS	IGE and IGC did not change much over the period	
Lefranc (2018)	France	earnings of fathers and sons	5-year-birth-cohorts from 1931 to 1975	TS2SLS	IGE followed V-pattern, falling from 0.6 for those born in the 1930s to about 0.4 for those born in the 1950s and then rose to a level seen at the beginning.	Contribution of education to IGE declined gradually over the 1930s to 1950s cohorts, then remained flat, and increased for recent cohorts.
Pekkarinen *et al*. (2017	Norway	earnings of sons and income of fathers	5-year birth-cohorts, 1932-74	OLS	IGE declined from 0.28 for those born in the early 1930s to 0.19 for those born in early 1940s, after which the elasticity remained constant.	For cohorts between 1930 and 1940s, the association between log parental income and education declined by 18% and return to education fall by 31%. For later cohorts the return to education increased.
Pekkala and Lucas (2007)	Finland	earnings of children and parental income	3 year birth-cohorts, 1930 -1970	OLS	A significant decline in IGE from the 1930 birth cohort until the cohorts of early 1950s, then gradually increased.	The mediating role of education declined from the earliest cohorts until those born in early 1950s, thereafter it gradually increased.
Fan et al. (2021)	China	income of parents and children	cohorts born in 1970–80 and 1981–88	OLS	IGE increased from 0.390 to 0.442 between the two periods	
Chu and Lin (2020)	Taiwan	earnings of fathers and sons	Compares two periods: 1990–94 and 2005–10	TS2SLS	IGE remained stable at around 0.37 between the two periods	

Note: IGE refers to intergenerational elasticity, and rank-rank slope refers to Intergenerational rank correlation.

5.3 Methods

This section presents the methods used to estimate intergenerational earnings persistence[58] [59] and analyse the mediating role of education in the intergenerational persistence. In addition, it discusses how earnings are adjusted for employment volatility to obtain a measure of the permanent earnings of individuals.

The main focus of the empirical analysis is on the association between the earnings of fathers and sons.

5.3.1 Measuring Intergenerational Persistence

We are interested in estimating β in the following regression model:

$$lnY_{ij}^{s} = \alpha + \beta lnY_{j}^{f} + \varepsilon_{ij} \quad (5.9)$$

where lnY_{ij}^{s} denotes the (natural) log of permanent earnings of the i^{th} son in family j and lnY_{j}^{f} denotes the log of permanent earnings of the father in family j, and ε_{ij} is an error term. β, often referred to as the intergenerational elasticity, is the parameter of interest and provides a measure of the share of parental earnings advantage that gets transmitted to children.

Our datasets do not possess permanent earnings of sons or fathers. In relation to the data requirements for estimating the intergenerational persistence equation, the datasets possess current earnings of respondents and information on the characteristics of fathers. To circumvent the data limitation, we proxy the permanent earnings of sons by their current earnings and use the TS2SLS technique to find a proxy for the permanent earnings of fathers.

Denoting $lnY_{ij,t}^{s}$ the log earnings in year t and lnY_{ij}^{s} the permanent log earnings of the i^{th} son in family j, we can relate the observed log earnings of sons to their permanent log earnings as:

[58] As discussed in chapter 3, intergenerational mobility measures can be identified as either absolute or relative measures. The focus of this chapter is on relative persistence, and measures persistence using the intergenerational regression coefficient (or elasticity).

[59] A framework for thinking about how the diverse measures of intergenerational mobility (both absolute and relative measures) relate to one another can be found in Deutscher & Mazumder (2023).

$$ln\mathrm{Y}^{s}_{ij,t} = ln\mathrm{Y}^{s}_{ij} + \upsilon^{s}_{ij,t} \quad (5.10)$$

where the error term $\upsilon_{ij,t}$ is a transitory fluctuation due to actual transitory changes as well as random measurement errors. Because the respondents are of different ages, we incorporate age controls into equation 5.10.

$$ln\mathrm{Y}^{s}_{ij,t} = ln\mathrm{Y}^{s}_{ij} + \varphi^{s}\mathrm{A}^{s}_{ij,t} + \gamma^{s}\left(\mathrm{A}^{s}_{ij,t}\right)^{2} + \upsilon^{s}_{ij,t}$$

where $\mathrm{A}^{s}_{ij,t}$ denotes the age of the i^{th} son in family j in year t and $\left(\mathrm{A}^{s}_{ij,t}\right)^{2}$ is the square of age. Solving the above equation for the permanent earnings of the son ($ln\mathrm{Y}^{s}_{ij}$), then substituting that into equation 5.9, and rearranging terms yields:

$$ln\mathrm{Y}^{s}_{ij,t} = \alpha + \beta\, lnY_{j}^{f} + \varphi^{s}\mathrm{A}^{s}_{ij,t} + \gamma^{s}\left(\mathrm{A}^{s}_{ij,t}\right)^{2} + \varepsilon_{ij} + \upsilon^{s}_{ij,t} \quad (5.11)$$

Equation 5.11 allows us to work with the available current earnings of sons. We note that if the classical measurement error assumption were to hold, the measurement error in the dependent variable (i.e., sons' current earnings) would not bias the estimate of β but leads to a loss of precision in the estimate. However, as discussed in section 5.2.2, the use of current earnings to proxy for permanent earnings introduces non-classical measurement errors, which bias the elasticity estimate. Empirical evidence shows that the bias is small if earnings are measured between the early 30s and mid-40s (e.g., Haider & Solon, 2006; Nybom & Stuhler, 2016).

To find a proxy for the permanent earnings of fathers, we use the TS2SLS approach. In the TS2SLS technique, observed characteristics of fathers are used to impute for their earnings. As noted in section 5.2.2, the technique employs two datasets drawn from the same population: the main sample and a previously collected 'auxiliary' sample. The earnings of fathers are not observed in the main sample, but the sample possesses characteristics of fathers $(\mathbf{Z})$ that have influence on their earnings (such as education and occupation). The auxiliary sample possesses earnings and the same $\mathbf{Z}$ variables, as in the main sample, of respondents (pseudo-fathers). The $\mathbf{Z}$ variables link the two samples and are used to impute the earnings of the actual fathers in the main sample in conjunction with estimated coefficients from an earnings equation of the pseudo-fathers in the auxiliary sample.

The first stage regression in the TS2SLS can be written as:

$$lnY_k^f = \boldsymbol{Z\delta} + \varphi^f A_k^f + \gamma^f (A_k^f)^2 + v_k^f$$

where $\ln Y_k^f$ denotes the log earnings of pseudo-fathers, $\mathbf{Z}$ is a vector of the imputing variables, $\boldsymbol{\delta}$ is a vector of the regression coefficients, A denotes age, and v is the error term. Age profile of respondents is incorporated to control for lifecycle bias. The estimated coefficients from the above equation are used to impute the log of permanent earnings of fathers in the main sample. The prediction equation for the permanent log earnings of fathers in the main sample is:

$$\widehat{lnY_j^f} = \boldsymbol{Z}_j^f \widehat{\boldsymbol{\delta}} + \hat{\varphi}(A) + \hat{\gamma}(A^2) \qquad (5.12)$$

where $\widehat{\boldsymbol{\delta}}$ represents the vector of coefficients estimated in the first stage regression, and $\boldsymbol{Z}$ is the vector of the imputing variables in the main sample. In our empirical analysis, we set age (A) at 45. Empirical studies show that current earnings are closest to permanent earnings when current earnings are measured around mid-life (e.g., Haider & Solon, 2006).

In the second stage of the TS2SLS, we estimate the intergenerational persistence equation (equation 5.11) using the main sample, replacing the unobserved fathers' permanent earnings (lnY_j^f) with their prediction ($\widehat{lnY_j^f}$). In our empirical analysis, we estimate the intergenerational persistence equation using five different sets of imputing variables for fathers' earnings. The first set has only education; the second – education and race; the third – education, race and occupation; the fourth – education, race, and province; and the fifth set has – education, race, occupation and province. These sets are the same as those used by Finn, Leibbrandt & Ranchhod (2017).

Note that the earnings of fathers are imputed variables. Imputed regressors are measured with sampling error, so hypothesis testing based on the estimated covariance matrix of the second step estimator are biased (Murphy & Topel, 1985). Thus, the standard errors should be corrected to account for sampling variation in the prediction of fathers' earnings. In our empirical analysis, we use a bootstrapping technique to obtain asymptotically correct standard errors of the regression coefficients. The bootstrapping procedure follows Björklund & Jäntti (1997). We first

draw a bootstrap sample from the auxiliary sample, run the earnings regression and obtain the parameter estimates that we use to generate the earnings of fathers in the main sample. Then we draw a bootstrap sample of sons from the main sample, impute the earnings of their fathers (using the first stage regression coefficients), obtain the intergenerational elasticity and save the estimate. After repeating the process 500 times, we estimate the standard error of the elasticity estimate by the standard deviation of the bootstrap estimates of the elasticity.

Models of intergenerational persistence generally assume a sample of families each consisting of a parent and a child. One of our datasets (namely, NIDS), however, has fathers with multiple sons. In terms of parental investment in the human capital of children, investment per child may be lower in households with multiple children. The elasticity coefficient may thus be biased towards fathers with more sons if the investment coefficient differs by household size. In this study we assume that there is no dilution of parental investment regardless of how many sons a father has. The presence of multiple sons per parent, moreover, generates error terms that are correlated within the household. The structure of the error term in equation 5.11 thus consists of not only transitory errors associated with the fluctuation of current earnings from permanent earnings and measurement errors associated with permanent earnings, but also the correlation of errors within households. These errors generate both heteroscedasticity and autocorrelation. Given the autocorrelation of error terms, we obtain the standard errors of the parameter estimates (using a bootstrapping technique) with clustering at the primary sampling unit level.

5.3.2 Adjusting Earnings for Employment Volatility

The South African labour market exhibits substantial levels of employment volatility (Banerjee *et al*., 2008; Kerr, 2018). Kerr (2018) estimates the rates of churning[60] to have been at 53 per cent per year between 2011 and 2014. This suggests that a substantial percentage of South African workers work intermittently and thus have highly variable employment probability. The high levels of employment volatility make it difficult to get a measure of the permanent earnings of individuals. In the presence of employment volatility, permanent earnings are determined not only by the amount one earns when employed but also by how often one is employed. To see this

[60] Defined as the sum of hires and separations expressed as percentage of firm level employment.

let the binary variable z_i indicate whether a person is currently employed. We can then write current earnings as:

$$Y_i = z_i Y_i^1 + (1 - z_i) Y_i^0$$

where Y_i^1 and Y_i^0 are the earnings if employed or not employed, respectively. Making the straightforward assumption that $Y_i^0 = 0$, and taking expectations we obtain:

$$E(Y_i|\boldsymbol{X}_i) = E(Y_i|\boldsymbol{X}_i, z_i = 1)E(z_i|\boldsymbol{X}_i)$$

$$E(Y_i|\boldsymbol{X}_i) = E(Y_i|\boldsymbol{X}_i, z_i = 1)\Pr(z_i = 1|\boldsymbol{X}_i) \quad (5.13)$$

Equation 5.13 provides a measure of the permanent earnings of individuals. In most studies, the implicit assumption is that $\Pr(z_i = 1|\boldsymbol{X}_i) \cong 1$ for the employed. However, in the case of South Africa this assumption is not reasonable due to high rates of employment volatility. Permanent earnings will be lower than expected earnings while employed, with the gap increasing as the probability of employment decreases.

Given the high rates of employment volatility, in our empirical analysis we adjust earnings for employment volatility. The procedure for adjusting the earnings follows Wittenberg (2022).[61] The basic assumption in Wittenberg (2022) is that

$$Y_i^1 = E[Y_i|\boldsymbol{X}_i, z_i = 1](1 + \delta_i), \quad (5.14)$$

with $\delta_i = \frac{Y_i^1 - E[Y_i|\boldsymbol{X}_i, z_i = 1]}{E[Y_i|\boldsymbol{X}_i, z_i = 1]}$ (i.e. δ_i is the proportional deviation from the conditional mean).

Solving for $E[Y_i|\boldsymbol{X}_i, z_i = 1]$ in equation 5.14, substituting that into equation 5.13, then taking logs of the resulting expression, followed by taking the expected values, Wittenberg shows that the log of permanent earnings can be given by:

$$\ln(E[Y_i|\boldsymbol{X}_i]) = E[\ln(Y_i^1)|\boldsymbol{X}_i] - E[\ln(1 + \delta_i)|\boldsymbol{X}_i] + \ln[\Pr(z_i = 1|\boldsymbol{X}_i)] \quad (5.15)$$

[61] Unpublished mimeo.

From equation 5.15, an approximately unbiased estimator of the log of permanent earnings in the case of the employed[62] can be given by:

$$\ln[\widehat{E(Y_i|\boldsymbol{X}_i)}] = lnY_i^1 + \ln\,[\Pr(\widehat{z_i = 1}|\boldsymbol{X}_i)] \qquad (5.16a)$$

Equation 5.15 can also be used to impute for the log of permanent earnings of the unemployed. In the case of the unemployed[63], an approximately unbiased estimator of the log of permanent earnings can be given by:

$$ln[\widehat{E(Y_i)}|\boldsymbol{X}_i] = E[\ln\widehat{(Y_i^1)}|\boldsymbol{X}_i] + \frac{1}{2}\widehat{\sigma_i^2} + \ln[\Pr(\widehat{z_i = 1}|\boldsymbol{X}_i)] \qquad (5.16b)$$

where σ_i^2 is the conditional variance of lnY_i given $\boldsymbol{X}_i$ for Y_i that follows a log-normal distribution. The first term in equation 5.16b can be estimated by the predicted earnings from a Mincerian earnings regression. The variance in the second term of equation 5.16b can be estimated from the standard deviation of the residuals in the Mincerian earnings regression, and the third term can be obtained from a probit model.

In our empirical analysis, we use Mincerian regression models to predict the log of current earnings and a probit model for estimating the probability of being employed. Specifically, we regress log earnings on age, age squared, education and race.

$$\ln Y_i = o + \varphi A_i + \gamma A_i^2 + \sum_{i=1}^{6}\phi_i S_i + \sum_{i=1}^{4}\psi_i R_i + \varsigma_i$$

where S and R denote education and race, respectively. Both are in categories. Our probit model is given by:

$$\Pr(z_i = 1|\boldsymbol{X}_i) = \Phi(\boldsymbol{\Gamma}'\boldsymbol{X}_i)$$

[62] Taking logs of equation 5.14, and recognizing that for the currently employed

$\ln[E(Y_i|\boldsymbol{X}_i, z_i = 1)] = \ln(Y_i|\boldsymbol{X}_i, z_i = 1) = \ln(Y_i^1)$, results in $\ln(1 + \delta_i) = 0$.

[63] Take logs of equation 5.14 and assume a log-normal distribution for earnings (with parameters μ_i and σ_i^2, and thus $E[\log(Y_i)] = \mu_i$ and $E[Y_i] = e^{\mu_i + \frac{1}{2}\sigma_i^2}$) to simplify $\ln(1 + \delta_i)$ (Wittenberg, 2022). Further, no selection effects are assumed.

where z_i is a dummy for being employed, $\boldsymbol{X}$ is a vector of explanatory variables, Φ is the standard normal cumulative density function, and $\boldsymbol{\Gamma}$ is a vector of coefficients. The explanatory variables are age, age squared, education (in categories), race, marital status, and urban residence.

5.3.3 Decomposition of the Intergenerational Elasticity and the Mediating Role of Education

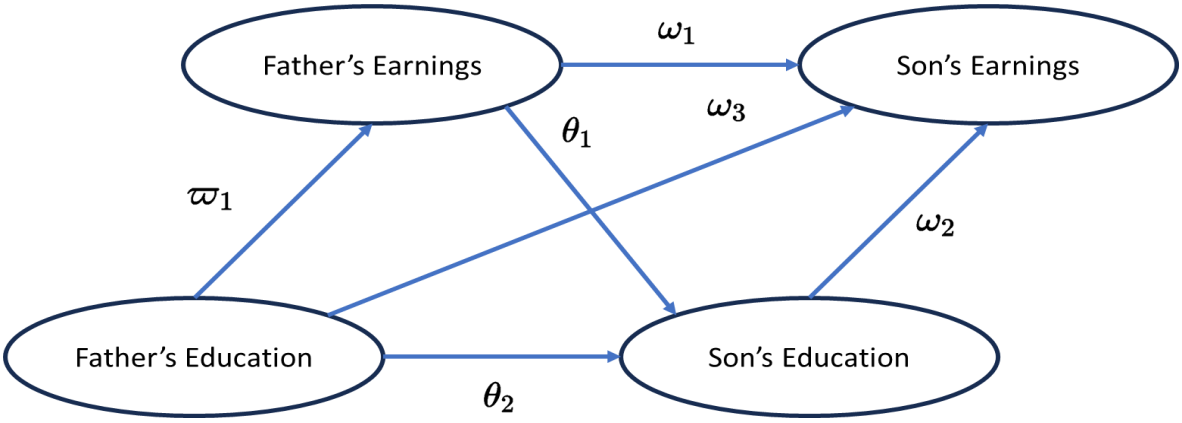
The mediating role of education in earnings persistence is analysed by decomposing the intergenerational elasticity into the direct and indirect (through education) effects of the earnings of fathers on that of sons.[64]

In studying the mediating role of education in intergenerational economic persistence, most studies do not control for the confounding effects of parental education. However, controlling for the confounding effects of parental education could be important as parental education is closely related to parental earnings as well as to children's education and earnings. Empirical evidence shows that parental education causally impacts children's education (Black & Devereux 2011; Björklund & Salvanes, 2011). Further, as discussed in the previous chapter, various studies have shown that parental education is strongly associated with children's earnings, controlling for child's education. In the case of South Africa, as found in the previous two chapters, parental education has not only been associated with children's education and earnings, but its degree of association with these variables have also changed over time. Thus, in our empirical analysis, we control for the confounding effects of fathers' education in analyzing the changes over time in the mediating role of education in intergenerational earnings persistence.

Figure 5.1 provides a diagrammatic scheme of the paths of influence of fathers' earnings on sons' earnings and how fathers' education might confound the relationship.

[64] Some decomposition studies of intergenerational persistence are interested in accounting for the contributions of various pathways, irrespective of whether the pathways are direct or indirect. Nonetheless, the method of decomposition of the elasticity is the same as the approach we use here (see e.g., Bowles & Gintis, 2002; Blanden, Gregg & Macmillan, 2007).

Figure 5.1 Confounding role of fathers' education in the relationship between the earnings of fathers and sons



Note: The labels of the arrows represent regression coefficients.

Given Figure 5.1, the relationship between the earnings of fathers and sons can be captured using three regression models. First, sons' earnings can be modelled as a linear function of their education as well as fathers' education and earnings.

$$lnY_{ij}^{s} = \omega_0 + \omega_1 lnY_j^f + \omega_2 S_{ij}^{s} + \omega_3 S_j^f + \zeta_{ij} \quad (5.17)$$

where lnY_{ij}^{s} and lnY_j^f refer to the (natural) log earnings of the i^{th}son and his father in the j^{th} family, respectively. Similarly, S_{ij}^{s} and S_j^f denote the education of the son and his father, and ζ_{ij} is an idiosyncratic error term. ω_0, ω_1, ω_2 and ω_3 are parameters to be estimated. ω_1 captures the association between the earnings of fathers and sons given the education of sons and fathers, ω_2 captures sons' earnings returns to education controlling for fathers' education and earnings, and ω_3 captures the association between the education of fathers and earnings of sons given the earnings of fathers and education of sons.

Second, sons' education can be modelled as a linear function of fathers' education and earnings.

$$S_{ij}^{s} = \theta_0 + \theta_1 lnY_j^f + \theta_2 S_j^f + \xi_{ij} \quad (5.18)$$

where θ_0, θ_1 and θ_2 are parameters of the model and ξ_{ij} is an idiosyncratic error term. θ_1 captures the association between fathers' earnings and sons' education given fathers' education,

and θ_2 captures the association between the education of fathers and sons controlling for fathers' earnings.

Third, fathers' earnings can be modelled as a linear function of their own education.

$$ln\, Y_{ij}^{f} = \varpi_0 + \varpi_1 S_{ij}^{f} + \nu_{ij} \quad (5.19)$$

where ϖ_0 and ϖ_1 are parameters of the model and ν_{ij} is an idiosyncratic error term. ϖ_1 captures the association between the education and earnings of fathers.

Given the relationships expressed in equations 5.17, 5.18 and 5.19, the regression coefficient of the linear projection of the earnings of sons on the earnings of fathers can be given by:

$$plim\beta = \omega_1 + \theta_1\omega_2 + \omega_3\varpi_1{}' + \theta_2\omega_2\varpi_1{}' \quad (5.20)$$

where $\varpi_1{}' = R^2/\varpi_1$ with $\varpi_1{}'$ representing the regression coefficient in the reverse regression of the education of fathers on the log earnings of fathers and R^2 denoting the coefficient of determination of that regression. Equation 5.20 shows that the intergenerational elasticity decomposes into four terms. The first term is the direct effect of fathers' earnings on sons' earnings (ω_1), controlling for fathers' education and sons' education. The second term is the indirect effect of fathers' earnings on sons' earnings that runs through sons' education ($\theta_1\omega_2$). The last two are omitted variable bias terms that quantify the confounding effect of fathers' education in the relationship between the earnings of fathers and sons. In particular, the third term captures that portion of the direct effect of fathers' education on sons' earnings that is correlated with fathers' earnings ($\omega_3\varpi_1{}'$). The last term captures that portion of the indirect effect of fathers' education on sons' earnings that is correlated with fathers' earnings but runs through sons' education ($\theta_2\omega_2\varpi_1{}'$). The last two terms are like spurious effects which misattribute effects to fathers' earnings that actually belong to fathers' education. The contribution of the mediating role of education to intergenerational earnings persistence, controlling for the confounding effects of fathers' education, is thus given by: $(\theta_1\omega_2)/(\omega_1 + \theta_1\omega_2)$.

5.4 Data and Descriptive Statistics

5.4.1 The Datasets

Once again, we use the Survey of Socioeconomic Opportunity and Achievement (SSOA) and the fifth wave of the National Income Dynamics Study (NIDS-w5) in our analysis.[65] Both datasets meet the requirements for the estimation of the intergenerational persistence equation. Both are surveys with a broad range of labour market and education questions and collect parental information regardless of residency status.

The datasets possess information on various characteristics of fathers. However, neither of them has earnings information for all fathers. We use the TS2SLS approach to impute the earnings of fathers from their reported characteristics. As discussed in section 5.3.1, the TS2SLS technique requires an auxiliary sample of pseudo-fathers. We use the 1980 Census for SSOA, and the Project for Statistics on Living Standards and Development (PSLSD) survey for NIDS-w5 as the auxiliary samples.

The 1980 census (Department of Statistics, 1985) covered the former four provinces of the Republic of South Africa (namely, the Cape, Orange Free State, Transvaal, and Natal), as well as the so-called National States of Ciskei, KwaZulu, Gazankulu, Lebowa, QwaQwa, KaNgwane, and KwaNdebele. The Census did not cover the "independent states" of Transkei, Venda, and Bophuthatswana, but census data for Bophuthatswana was released with the final 1980 census dataset. In our analysis, we combine both the 1980 census and the Bophuthatswana census and refer to them jointly as 'the census'. The census collected socio-demographic information on all household members including their education, labour market participation, occupation, and annual incomes. The census asks individuals for their annual market incomes.[66] The census's income definition is broader than the earnings definition used in SSOA (i.e., its corresponding main sample in the TS2SLS). Nonetheless, we use the reported income of individuals as a proxy

[65] The datasets are discussed in chapter 1.

[66] The census defines income as consisting of gross salary, profits from business, estimated cash value of fringe benefits, and any other regular income such as interest, dividend, and net rent from fixed property.

for their earnings with the assumption that labour earnings constitute the main source of income for most employed individuals.

The PSLSD survey (SALDRU, 1995) is a nationally representative household survey. It was conducted in 1993. The survey collected demographic and socioeconomic information on all household members. In the PSLSD, we obtain earnings of individuals by adding their gross monthly earnings from wages, bonuses and profit shares related to the wage job, and casual employment.

All earnings are converted to annual earnings.[67] The estimation of the intergenerational persistence equations is carried out using both unadjusted and adjusted earnings for employment volatility. Earnings are adjusted for employment volatility following the procedure presented in section 5.3.2. This is done for SSOA, NIDS-w5 and PSLSD, but not for the census. In the census, respondents were asked for their annual income. Based on how the question is asked, we assume that respondents would provide income figures that reflect the amount of time they worked during the year (which means, the incomes are implicitly adjusted for employment volatility).

The main empirical analysis is based on males between 25 and 44 years of age who are in the labour force.[68] Permanent earnings for the unemployed are imputed based on their expected earnings and adjusted for volatility using the procedure described in section 5.3.2.

The sample of pseudo-fathers (from the auxiliary datasets) is restricted to employed males between 30 and 59 years of age. This means that the main sample in SSOA were aged between 12 and 31 years and those in NIDS-w5 were between 1 and 20 years when their corresponding auxiliary samples were collected. The gap between the 1980 census and its corresponding main sample, SSOA, is about 12 years, which is not a large enough gap to assume that those in their early 30s in 1980 (the time the census was conducted) to be fathers with adult children in the early 1990s (the time the SSOA was collected). The ideal dataset would have been one that is

[67] Earnings of individuals in the main samples are discussed in chapter 4.

[68] We use the broad definition of unemployment. The broad definition includes both those who are not working but are actively searching for work and those who are willing to work but have given up searching.

collected about 20 years earlier to SSOA (as is the case between the PSLSD and NIDS-w5). However, this may not be a major issue in our analysis as the main use of the auxiliary samples is to estimate relative returns to the characteristics of fathers.

5.4.2 Descriptive Statistics

This section presents summary statistics of sons (and of their fathers) in the main samples with respect to some background characteristics and examines the effect of the adjustment of earnings for employment volatility on the distribution of earnings.

5.4.2.1 Descriptive Statistics of Main Samples

We have already seen some descriptive statistics of SSOA and NIDS-w5 based on certain qualifications in the previous two chapters. In chapter 3 the analysis was based on all respondents between 25 and 69 years of age. In chapter 4 the main analysis was based on employed males between 25 and 59 years of age. In this chapter, the main analysis is based on males between 25 and 44 years of age who are in the labour force.

Table 5.2 presents summary statistics of the respondents (and of their fathers) with respect to age, proportion employed, race, education, occupation, and province.[69] [70]

A comparison of the distribution of education of sons in SSOA and NIDS-w5 in Table 5.2 shows that in NIDS a smaller percentage of the sons are in the lower education categories compared to SSOA. This reflects the increase in educational attainment over time. We have seen in chapter 3, and it is a well-documented finding, that the educational attainment of South Africans has been increasing over time.

[69] For comparative purposes, in Appendix 5.A we provide summary statistics of all males between 25 and 44 years of age (without restriction to labour force participation) with respect to the same background characteristics. The distribution of respondents in the main samples is quite comparable to the distribution of respondents in the bigger samples along all the background characteristics considered.

[70] We assume that the race of respondents reflects the race of their fathers. Girdwood & Leibbrandt (2009) report of 'insignificant' number of mixed marriages in NIDS.

Table 5.2 Summary statistics of respondents (and of their fathers) in the main samples

	SSOA				NIDS-w5			
	Sons		Fathers		Sons		Fathers	
	mean	sd	mean	sd	mean	sd	mean	sd
Age	34.0	5.603			33.6	5.516		
Employed	0.833	0.373			0.891	0.311		
Race								
African	0.669	0.471			0.857	0.351		
Coloured	0.089	0.285			0.078	0.268		
Indian/Asian	0.052	0.223			0.016	0.127		
White	0.189	0.392			0.049	0.216		
Education								
None	0.065	0.246	0.229	0.420	0.008	0.088	0.364	0.481
Primary	0.298	0.458	0.408	0.492	0.064	0.245	0.205	0.404
Incomplete secondary	0.313	0.464	0.250	0.433	0.474	0.499	0.244	0.430
Matric	0.204	0.403	0.052	0.223	0.226	0.418	0.119	0.324
Post-secondary	0.120	0.325	0.061	0.240	0.228	0.420	0.068	0.252
Occupation								
Professional/Manager	0.142	0.349	0.114	0.318	0.111	0.314	0.107	0.309
Operator/Semi-skilled	0.297	0.457	0.331	0.471	0.250	0.433	0.205	0.404
Craft/trade	0.223	0.416	0.191	0.394	0.251	0.434	0.129	0.335
Clerk/sales	0.187	0.390	0.128	0.334	0.201	0.401	0.096	0.294
Elementary	0.151	0.358	0.236	0.425	0.187	0.390	0.464	0.499
Province								
Cape	0.194	0.395	0.170	0.376				
Natal	0.147	0.354	0.159	0.366				
Transvaal	0.304	0.460	0.311	0.463				
Orange Free State	0.031	0.174	0.036	0.186				
Homelands	0.324	0.468	0.324	0.468				
Western Cape					0.109	0.312	0.069	0.253
Northern Cape					0.029	0.168	0.023	0.149
Eastern Cape					0.069	0.254	0.160	0.367
KwaZulu-Natal					0.165	0.371	0.182	0.386
Free State					0.057	0.232	0.063	0.244
Mpumalanga					0.098	0.297	0.109	0.312
Limpopo					0.072	0.259	0.144	0.351
North-West					0.060	0.238	0.083	0.275
Gauteng					0.340	0.474	0.168	0.374
N	**2,308**				**3,637**			
# of sons with father's education	1,824				2,834			
# of sons with father's occupation	1,998				2979			
# of sons with fathers' province	1,932				2,977			

Notes: (1) The table presents summary statistics of respondents (and of their fathers) in SSOA and NIDS-w5 by some background characteristics. The table is based on labour force participating males between 25 and 44 years of age in the two samples.
(2) *Includes those who were reported as having never worked and recoded to elementary occupations.
(3) Results are weighted using the calibrated weights provided with the datasets.

Occupation is categorized into five categories: namely, elementary, clerk/sales, craft/trade, operator/semi-skilled and professional/manager occupations. Occupation is categorized in the above way to accommodate the PSLSD occupational classification. The occupational classification in the PSLSD dataset is a collapsed version of the standard one-digit occupational classification. Conforming to the PSLSD classification allows us to use the same occupational classifications across all the datasets.[71] The distribution of sons by occupation in the two samples shows that their distributions are quite comparable (see Table 5.2).

Province is constructed as consisting of five categories in SSOA and the census. These categories are (i) the Cape, (ii) Natal, (iii) Transvaal, (iv) Orange Free State, and (v) the Homelands (consisting of KwaZulu, Gazankulu, Lebowa, QwaQwa, KaNgwane, KwaNdebele, Ciskei, and Bophutatswana). In NIDS and PSLSD, province has the nine post-apartheid provinces. Province of father is not explicitly asked in either SSOA or NIDS. We use the province where the son was living at a young age as the province of the father. In particular, we use the province where the son was living at age 14 as the province of the father in SSOA, and the province where the son was living in 1994 as the province of the father in NIDS.

We note that it may not be appropriate to directly compare the distributions of the characteristics of fathers between the two samples (in Table 5.2) as the samples of fathers are not random in themselves. This is especially true in the case of NIDS. NIDS considers all household members, thus fathers with more sons are overrepresented. We further note that the proportions of missing values for the characteristics of fathers are rather large (as can be deduced from the figures for the number of sons with non-missing father's characteristics at the bottom rows of Table 5.2). Education of fathers is missing for about 21.0 per cent of sons in SSOA and for about 22.1 per cent of sons in NIDS. Similarly, occupation of fathers is missing for about 13.4 per cent in SSOA[72] and for about 18.1 per cent in NIDS. Further, in NIDS about 28.2 per cent of sons

[71] In NIDS, occupation is given in the standard one-digit occupational classification. In SSOA and the Census, occupation is available in forms that can readily be coded into the one-digit occupational classification.

[72] In SSOA, for respondents whose father's occupation is missing but economic activity is known (about 6.7 per cent of respondents), we predict fathers' occupation on the basis of their economic activity. This is accomplished by

reported that their fathers never worked. In our analysis, we treat those fathers who are reported to have never worked as having elementary occupations. We assume that it is less likely for those fathers to never have jobs, but their jobs may have been irregular and/or low-skilled jobs, which can be classified as elementary occupations. Note that the missingness in education, occupation and province of fathers impacts the working sample sizes when estimating the intergenerational earnings elasticities as these variables are used in imputing the earnings of fathers.

Appendix 5.B presents summary statistics of respondents in the auxiliary samples by the same background characteristics as in Table 5.2.

5.4.2.2 Adjusting Earnings for Employment Volatility

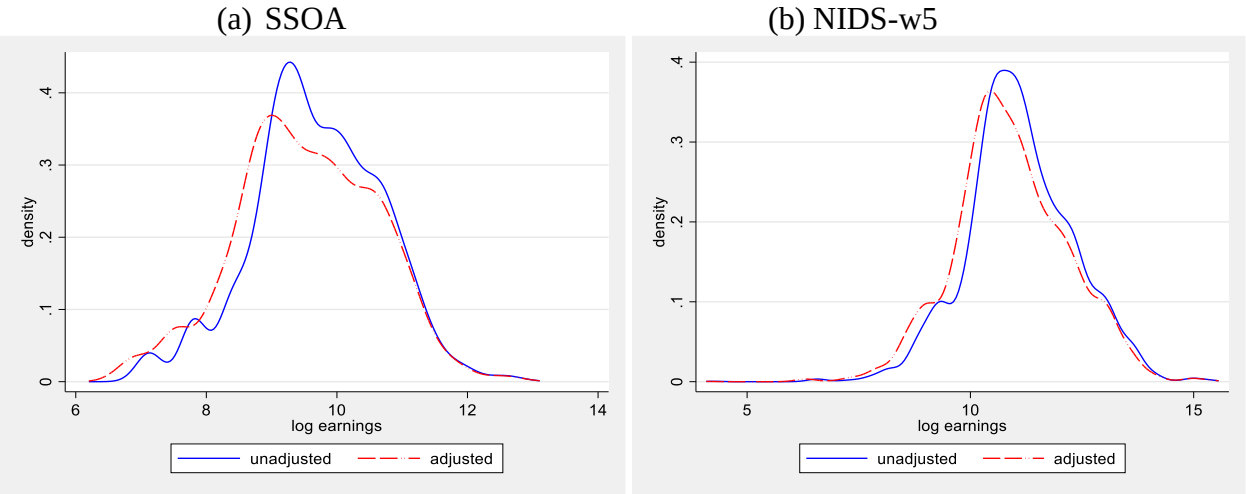
Figure 5.2 presents the kernel densities of the unadjusted and adjusted log annual earnings of the employed in SSOA and NIDS-w5. The adjustment of the earnings for employment volatility has given rise to earnings distributions that are pushed to the left in comparison to the distribution of the unadjusted earnings, with the gap between the two generally widening as we move towards the left, in both samples. This is as expected. The permanent earnings of individuals are expected to be lower than their current earnings in the face of employment volatility, with the gap increasing as the probability of employment decreases. The adjustment of the earnings has also resulted in distributions with wider spread. The spread seems to be wider in the case of SSOA than NIDS. The results of the probit model that we used to adjust the earnings for employment volatility show that individual characteristics (especially, race and education) were more important in determining employment in SSOA than in NIDS-w5.[73] This has produced a distribution with more variability relative to the distribution of the unadjusted earnings in the former. We will see in the empirical results section that these different patterns of earnings

running multinomial logistic regression of occupation on economic activity, race, and education of fathers, and assigning the missing occupation to occupational category with the highest predicted probability.

[73] The results of the probit models for the probability of being employed and the earnings regressions conditional on being employed for SSOA and NIDS-w5 are given in Appendix 5.C.

adjustments have huge consequences to the estimated intergenerational elasticities in the two samples in comparison to the elasticities based on the unadjusted earnings.

Figure 5.2 Kernel densities of unadjusted and adjusted log annual earnings



Notes: (1) The kernel densities compare the distribution of adjusted log earnings (adjusted for employment volatility) with unadjusted log earnings in SSOA and in NIDS-w5. The figures are based on employed males between 25 and 44 years of age in both samples.
(2) The Kernel density estimates are based on the Gaussian Kernel function with bandwidth set at 200.
(3) Results are weighted using the calibrated weights provided with the datasets.

We next examine how the distribution of the adjusted log earnings of the employed changes when the sample of the unemployed (with their predicted permanent log earnings) is added to the sample of the employed.

The expected earnings of the unemployed are predicted based on their characteristics. Table 5.3 presents the distribution of the employed and the unemployed in the main samples with respect to race and education. The table shows that the distribution of the employed and the unemployed with respect to these two characteristics are more similar in NIDS-w5 than in SSOA.

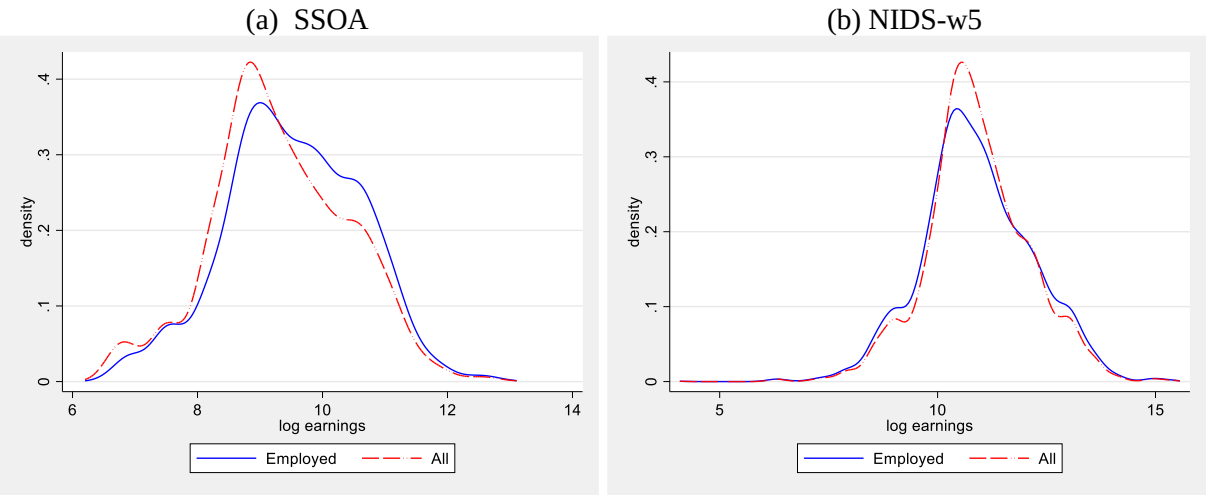
Table 5.3 Percentage distribution of the employed and unemployed sons in main samples

	SSOA		NIDS-w5	
	Employed	**Unemployed**	**Employed**	**Unemployed**
Race				
African	60.5	92.3	83.5	88.5
Coloured	10.5	4.5	7.6	8.9
Indian/Asian	5.1	1.1	2.6	0.8
White	23.9	2.1	6.4	1.8
Education				
None	5.8	8.0	0.8	0.5
Primary	23.8	44.1	5.9	9.2
Incomplete Secondary	31.2	32.1	45.7	52.6
Matric	23.6	14.1	23.0	26.6
Postsecondary	15.6	1.7	24.7	11.0
N	**1979**	**329**	**2914**	**723**

Notes: (1) The table presents the distribution of the employed and the unemployed with respect to race and education in SSOA and NIDS-w5. The table is based on labour force participating males who are between 25 and 44 years of age.
(2) The employment status of individuals is based on the broad definition of unemployment.
(3) Results are weighted using the calibrated weights provided with the datasets.

Figure 5.3 presents the kernel densities of the adjusted log annual earnings of the employed and labour force participants in the main samples of SSOA and NIDS-w5. The addition of the unemployed to the sample of the employed seems to have larger effects in the middle parts of the kernel densities than in the tails, with the changes more pronounced in the case of SSOA. As we have seen in Table 5.3, the distribution of the unemployed and the employed with respect to both race and education are more similar in NIDS than in SSOA. Further, the addition of the unemployed to the sample of the employed seems to decrease the spread of the earnings distribution in both samples. We will see in the next section that the addition of the unemployed to the sample of the employed slightly lowers the estimated elasticities in both SSOA and NIDS.

Figure 5.3 Kernel densities of adjusted log annual earnings in SSOA and NIDS



Notes: (1) The kernel densities compare the distribution of the adjusted earnings of the employed and the labour force participants (using the predicted permanent earnings for the unemployed) in SSOA and in NIDS-w5. Earnings are adjusted for employment volatility. The figures are based on males between 25 and 44 years of age in the two samples. ***Employed*** refers to the currently employed, and ***All*** refers to labour force participants (employed as well as the unemployed (broad definition)).
(2) The Kernel density estimates are based on the Gaussian Kernel function with bandwidth set at 200.
(3) Results are weighted using the calibrated weights provided with the datasets.

5.5 Empirical Results

This section examines intergenerational earnings persistence in SSOA and NIDS-w5 and compares the persistence between the two surveys. Further, it examines the mediating role of education in the intergenerational earnings persistence in each survey and evaluates changes in the mediating role of education between the two surveys.

The empirical analysis is carried out first using the unadjusted earnings and then using the adjusted earnings for employment volatility. The main analysis is based on males between 25 and 44 years of age who are in the labour force, with a focus on the employed subsample. We further extend the analysis by considering employed females who are between 25 and 44 years of age.

5.5.1 Intergenerational Earnings Elasticity

As discussed in the methods section, the estimation of the intergenerational elasticity is carried out using the TS2SLS approach, and the earnings of fathers are imputed using five alternative sets of their characteristics. The characteristics of fathers that are used to impute for their earnings are education, race, occupation, and province.

5.5.1.1 Intergenerational Earnings Elasticity: Unadjusted Earnings

Table 5.4 presents the intergenerational elasticity estimates using the unadjusted earnings. The estimates are based on the employed samples in SSOA and NIDS-w5. We have five elasticity estimates for each sample. These correspond to the five sets of imputing variables for the earnings of fathers. We refer to them as specifications (1) - (5).

The elasticity estimates are remarkably consistent across all the five specifications in SSOA and from specification (2) to (5) in NIDS. The elasticity estimates suggest that in SSOA, using specification (5) as an example, when a father earns 10 per cent more than the mean earnings of his generation, his son's earnings are approximately 6.5 per cent higher than the generational mean. In NIDS, a similar scenario leads to the son's earnings being about 7.1 per cent higher than the generational mean.

The elasticity estimate based on specification (1) for NIDS is smaller in magnitude compared to the other four specifications. The relatively lower elasticity estimate based on specification (1) - where only education is used to impute for fathers' earnings – is likely due to the decrease in inequality and increase in intergenerational mobility of educational attainment of South Africans over time. We have indicated in chapter 1 that inequality in educational attainment has been declining. Further, we have shown in chapter 3 that intergenerational mobility in educational attainment has been increasing over time. Note that the addition of race and other characteristics of fathers to education in specifications (2) to (5) have resulted in estimates displaying higher levels of intergenerational earnings persistence.

Having estimated the intergenerational earnings elasticities for the two surveys, we can discuss changes over time in intergenerational persistence between the two periods. We note in advance that the comparative results should be read with the following two points in mind. The first point relates to differences in the degree of representation of the two surveys. As indicated in chapter 1, NIDS is a nationally representative sample, but SSOA is not as it missed farm workers who were living on commercial farms. As discussed in chapter 3 (section 3.4.1), farmworkers who were living on commercial farms were structurally limited in their ability to accumulate human and physical capital and pass it on to their children. Thus, we could assume that intergenerational persistence in earnings among this group was high. SSOA, by not including farmworkers who were living on commercial farms, is thus likely to provide elasticity estimates that are biased downward. The second point relates to measurement issues and sample selection effects. Differences in survey measurement of variables between the surveys and the missingness in the characteristics of fathers (i.e., education, occupation, and province) might affect the comparative results.

With the above two caveats in mind, we compare the elasticity estimates between the two surveys to test whether the estimates for NIDS-w5 are lower than those for SSOA. We use a Z-statistic that compares regression coefficients from two different regressions.[74] The results show

[74] The test of equality of the intergenerational earnings elasticity used a Z-statistic for comparing regression coefficients from two different regressions: $Z_{calculated} = \frac{\widehat{\beta}_{NIDS} - \widehat{\beta}_{SSOA}}{\sqrt{(SE\widehat{\beta}_{NIDS})^2 + (SE\widehat{\beta}_{SSOA})^2}}$ (Clogg, Petkova & Haritou, 1995).

that the elasticity estimates are not statistically significantly different across all the five specifications (see Table 5.4, at the lower section). We are interested in what these results might tell us regarding the direction of change in intergenerational earnings persistence between the two periods. Before answering the question, let's re-estimate the elasticities using earnings that are adjusted for employment volatility.

5.5.1.2 Intergenerational Earnings Elasticity: Adjusted Earnings

We have indicated in section 5.3.2 that the South African labour market exhibits substantial levels of employment volatility and that we need to correct earnings for employment volatility to get a measure of the permanent earnings of individuals. Table 5.5 presents the intergenerational elasticity estimates using earnings that are corrected for employment volatility. Panel A is based on the employed sample, and Panel B is based on labour force participants (both employed and unemployed pooled).

The adjustment of the earnings for employment volatility has a larger effect on the estimated elasticities of SSOA than NIDS. A comparison of elasticity estimates in Panel A of Table 5.5 and Table 5.4 (both for the employed samples) shows that in SSOA the elasticity estimates based on the adjusted earnings are larger compared to their corresponding elasticity estimates based on the unadjusted earnings across all the five specifications. On the other hand, in NIDS the elasticity estimates based on the adjusted earnings are not much different from those based on the unadjusted earnings across all the five specifications. For instance, the elasticity estimate based on the adjusted earnings obtained using specification (5) in SSOA shows an increase of about 12.4 per cent relative to its corresponding elasticity estimate based on the unadjusted earnings. The same comparison in NIDS shows a decrease of about 0.7 per cent. As discussed in section 5.4.2, the earnings adjustment, which takes into consideration the probability of being employed,

Table 5.4 Intergenerational earnings elasticity: unadjusted earnings

		Variables used for imputing earnings of fathers (Z)				
		Education	Education & race	Education, race & occupation	Education, race & province	Education, race, occupation & province
		(1)	(2)	(3)	(4)	(5)
SSOA	β	0.649	0.664	0.681	0.642	0.646
	s.e.	(0.046)	(0.037)	(0.040)	(0.037)	(0.040)
	N	[1238]	[1238]	[1184]	[1013]	[969]
NIDS-w5	β	0.560	0.720	0.685	0.716	0.706
	s.e.	(0.057)	(0.064)	(0.062)	(0.069)	(0.066)
	N	[2090]	[2089]	[1909]	[1854]	[1719]
Test of: β of NIDS < β of SSOA						
Z-calculated		-1.215	0.758	0.054	0.945	0.777
p-value		0.111	0.776	0.520	0.829	0.782

Note: (1) The table shows intergenerational elasticity estimates for father-son pairs in SSOA and NIDS-w5 with fathers' earnings predicted using five alternative specifications. Earnings are unadjusted for employment volatility. The table is based on employed males between 25 and 44 years of age in both samples.[75]
(2) Numbers in parenthesis beneath the regression coefficients are cluster bootstrapped standard errors[76], and below them in brackets are sample sizes.

[75] The 25-44 age range for sons is chosen to facilitate comparability of our estimates with previous South African studies. However, as indicated in the literature review section, lifecycle bias (in children's earnings) is relatively small when earnings are measured between the early 30s and mid-40s. Re-analysis of the sample by restricting sons' age to those between 30 and 44 years of age have shown that the elasticity estimates do not significantly change (results not shown here but can be obtained from author upon request).

[76] As indicated in the methods section, the standard errors of the elasticity estimates are obtained using a bootstrapping method. The bootstrapping assumes clustering at the primary sampling unit level in both the auxiliary and main samples.

Table 5.5 Intergenerational earnings elasticity: Adjusted earnings

		Variables used for imputing earnings of fathers (Z)				
		Education	Education & race	Education, race & occupation	Education, race & province	Education, race, occupation & province
		(1)	(2)	(3)	(4)	(5)
Panel A: Employed						
SSOA	β	0.726	0.748	0.766	0.725	0.726
	s.e.	(0.052)	(0.039)	(0.043)	(0.041)	(0.044)
	N	[1237]	[1237]	[1183]	[1012]	[968]
NIDS-w5	β	0.556	0.724	0.689	0.707	0.701
	s.e.	(0.060)	(0.065)	(0.064)	(0.066)	(0.065)
	N	[2083]	[2083]	[1904]	[1852]	[1718]
Test of : β of NIDS < β of SSOA						
Z-calculated		-2.141	-0.317	-0.999	-0.232	-0.319
p-value		0.016	0.375	0.159	0.409	0.375
Panel B: All (Labour force participants)						
SSOA	β	0.674	0.731	0.756	0.697	0.702
	s.e.	(0.045)	(0.031)	(0.034)	(0.033)	(0.036)
	N	[1464]	[1464]	[1392]	[1216]	[1154]
NIDS-w5	β	0.522	0.687	0.652	0.671	0.663
	s.e.	(0.056)	(0.062)	(0.060)	(0.062)	(0.063)
	N	[2653]	[2653]	[2378]	[2313]	[2112]
Test of : β of NIDS < β of SSOA						
Z-calculated		-2.116	-0.635	-1.508	-0.370	-0.537
p-value		0.017	0.261	0.066	0.356	0.295

Note: (1) The table presents intergenerational elasticity estimates using earnings that are adjusted for employment volatility for father-son pairs in SSOA and NIDS-w5. The estimates in Panel A are based on employed males between 25 and 44 years of age and the estimates in Panel B are based on labour force participating males between 25 and 44 years of age. Fathers' earnings are predicted using five alternative specifications.

(2) Numbers in parenthesis beneath the regression coefficients are cluster bootstrapped standard errors, and below them in brackets are sample sizes.

has different effects for different individuals with the effects being larger in SSOA than in NIDS. This has resulted in a distribution with a wider spread in SSOA and has given rise to larger elasticity estimates in that sample.

A comparison of the elasticity estimates based on the adjusted earnings of the employed between the two surveys shows that the elasticity estimates are larger in SSOA than in NIDS across all the five specifications. However, except for specification (1), none of the differences are statistically significant. The summary of the differences between the elasticities is given in the lower section of Panel A in Table 5.5.

When the unemployed (with their predicted permanent earnings) are added to the employed sample, the elasticity estimates for the combined sample are shown to be lower than those for the employed sample across all the five specifications in both SSOA and NIDS (see Table 5.5, Panels A&B). We have seen in section 5.4.2 that the addition of the unemployed sample to the sample of the employed has narrowed the spread of the distribution of the adjusted earnings of the employed sample in both SSOA and NIDS. The decrease in the variability of earnings when we combine the employed and unemployed samples has resulted in lower elasticity estimates in both SSOA and NIDS. A comparison of the elasticity estimates for the combined samples (employed and unemployed combined) between the two surveys shows that the elasticity estimates are larger in SSOA than in NIDS across all the five specifications. The summary of the differences between the elasticities is given in the lower section of Panel B in Table 5.5.

Reverting to the main question of the change in intergenerational persistence between the two periods, the comparative analysis using the unadjusted earnings shows that the elasticity estimates are generally comparable between the two surveys. On the other hand, the analysis using the adjusted earnings shows that the elasticity estimates for SSOA are somewhat larger in magnitude than those for NIDS. Given the high levels of employment volatility in the South African labour market, the elasticity estimates based on the adjusted earnings seem more sensible results. Overall, taking the caveats mentioned in section 5.5.1.1 into consideration, the results suggest that intergenerational earnings persistence were probably higher in the first period than the second.

A comparison of our results with other South African studies (in particular, Finn, Leibbrandt & Ranchhod, 2017) shows that the elasticity estimates based on the adjusted earnings of labour force participants in NIDS are slightly lower than those in Finn *et al.* based on the different subsamples with a Heckman correction for selection into employment. As indicated in our literature review section, Finn, Leibbrandt & Ranchhod (2017) used the average earnings of respondents across the first four waves of NIDS to proxy for the permanent earnings of sons and the TS2SLS technique to proxy for the permanent earnings of fathers with the PSLSD as an auxiliary sample.

5.5.1.3 Intergenerational Earnings Elasticity: Females

The estimation of intergenerational earnings elasticity for the father-daughter pairs is conducted following the same procedure as that for the father-son pairs. The analysis is based on employed daughters who are between 25 and 44 years of age and is carried out using both the unadjusted and adjusted earnings for employment volatility. The elasticity estimates are shown in Appendix 5.E. The elasticity estimates for daughters using both the unadjusted and adjusted earnings are generally larger in SSOA and smaller in NIDS compared to their corresponding elasticity estimates for employed sons. Again, as indicated in chapter 4 (section 4.5.1), the different results based on the samples of daughters and sons are likely due to sample selection issues that arise not only from employment participation decisions but also from employment opportunity restrictions for women - especially in the first period (SSOA).

A comparison of the elasticity estimates between SSOA and NIDS (for daughters) shows that the elasticity estimates based on the unadjusted earnings are comparable in the two surveys, except for specification (1). The elasticity estimate based on specification (1) for NIDS is smaller in magnitude compared to the estimates based on the other four specifications in NIDS and is also much smaller in comparison to its corresponding estimate in SSOA. As indicated in section 5.5.1.1, the relatively lower elasticity estimate based on specification (1) - where only education is used to impute for fathers' earnings – is likely due to the decrease in inequality and increase in intergenerational mobility of educational attainment in South Africa over time.

When adjusted earnings are used, the elasticity estimates for both SSOA and NIDS increase across all the 5 specifications relative to their elasticity estimates based on the unadjusted earnings, with larger increases seen in the case of SSOA. As a result, the elasticity estimates for NIDS are shown to be significantly smaller compared to those for SSOA across all the 5 specifications. The results suggest that the direction of earnings persistence between the two periods for daughters are the same as that for sons.

5.5.2 The Mediating Role of Education in Earnings Persistence

As discussed in the methods section, the mediating role of education in intergenerational earnings persistence is examined by decomposing the overall association between the earnings of fathers and sons into the direct and the indirect (through education) components, controlling for the confounding effect of fathers' education.

For the earnings of fathers, we consider specification 5 (where education, race, occupation and province of fathers are used to impute for their earnings).

5.5.2.1 Mediating Role of Education: Males

Table 5.6 presents the results of a decomposition of the elasticity of the association between the earnings of fathers and sons. Panel A is based on the unadjusted earnings and Panel B is based on the adjusted earnings (for employment volatility). The contribution of the mediating role of education to earnings persistence is computed controlling for confounding effects of fathers' education.

The results show that the contribution of the mediating role of education to earnings persistence is about the same in the two surveys, whether one uses unadjusted or adjusted earnings. In both surveys the contribution is estimated to be in the range of 19-20 per cent (Table 5.6). Stated differently, controlling for the confounding effects of fathers' education, about 80 per cent of the association between the earnings of fathers and sons is due to the part that is not mediated by the child's education. If we do not control for the confounding effects of fathers' education, the contribution of the mediating role of education to persistence is about 16 per cent in SSOA and 20 per cent in NIDS (not shown in Table 5.6). This shows that the contribution of the mediating

role of education to earnings persistence changes depending on whether we control for the confounding effects of fathers' education or not. These results highlight the importance of controlling for parental education when examining the contribution of the mediating role of education to earnings persistence. This is especially important when the analysis is focused on changes over time in the contribution of the mediating role of education to earning persistence as the confounding effects of parental education may change with time. We have seen in the previous two chapters that in South Africa the relationship of parental education with both children's education and children's earnings have changed over time.

The mediating role of education in earnings persistence can be further examined by looking at the changes in the relationship between the earnings of fathers and sons over time. We have indicated in the methods section that the association between the earnings of fathers and sons, controlling for the confounding effect of fathers' education, can be captured using three regression models. The regression models along with a summary of the estimated regression coefficients are given in Table 5.6.

The results of the first regression, which relates sons' education with fathers' earnings and fathers' education, shows that sons' education is significantly positively associated with fathers' earnings in both SSOA and NIDS (see Table 5.6). The coefficient ($\boldsymbol{\theta_1}$), however, is smaller in NIDS. Similarly, the association between fathers' education and sons' education in the relationship is significantly positive in both SSOA and NIDS, and the size of the coefficient ($\boldsymbol{\theta_2}$) is smaller in NIDS. The smaller size of the coefficients for fathers' earnings and fathers' education in the second survey is possibly due to the expanded schooling opportunity over time (discussed in chapter 2), which weakens the association between parental economic status and children's educational attainment.

The results of the second regression, which relates sons' earnings with sons' education and fathers' earnings and fathers' education, shows that sons' earnings are significantly positively associated with both sons' education and fathers' earnings in SSOA and NIDS (Table 5.6). The returns to both fathers' earnings ($\boldsymbol{\omega_1}$) and own education ($\boldsymbol{\omega_2}$) is larger in NIDS. The association between fathers' education and sons' earnings in the relationship is not significant in both SSOA and NIDS. The coefficient for fathers' education ($\boldsymbol{\omega_3}$) is close to zero in the relationship. This is-

Table 5.6 Mediating role of education: Sons

The estimated regression models are:

(1) $S_{ij}^{s} = \theta_0 + \theta_1 lnY_j^f + \theta_2 S_j^f + \theta_3 A_{ij}^s + \theta_4 (A_{ij}^s)^2 + \xi_{ij}$

(2) $lnY_{ij}^{s} = \omega_0 + \omega_1 lnY_j^f + \omega_2 S_{ij}^s + \omega_3 S_j^f + \omega_4 A_{ij}^s + \omega_5 (A_{ij}^s)^2 + \zeta_{ij}$

(3) $ln\, Y_{ij}^f = \varpi_0 + \varpi_1 S_{ij}^f + \nu_{ij}$

(4) $lnY_{ij}^s = \alpha + \beta lnY_j^f + \varphi A_{ij}^s + \gamma (A_{ij}^s)^2 + \varepsilon_{ij}$

Summary of regression coefficients

Panel A: Decomposition of intergenerational elasticity (*unadjusted earnings*)

	$\boldsymbol{\theta_1}$	$\boldsymbol{\theta_2}$	$\boldsymbol{\omega_1}$	$\boldsymbol{\omega_2}$	$\boldsymbol{\omega_3}$	$\boldsymbol{\varpi_1}$	$\boldsymbol{\varpi_1}'$	$\boldsymbol{\beta}$	**Indirect effect** $(\boldsymbol{\theta_1\omega_2})$	**% due to indirect effect** $[(\theta_1\omega_2)/(\omega_1+\theta_1\omega_2)]\cdot 100$
SSOA	1.286	0.209	0.399	0.081	0.027	0.175	3.419	0.657	0.104	20.73
	(.269)	(.067)	(.050)	(.015)	(.016)	(.006)		(0.040)		
NIDS	0.715	0.087	0.559	0.201	-.017	0.124	5.808	0.706	0.144	20.47
	(.157)	(.023)	(.101)	(.017)	(.014)	(.003)		(0.049)		

Change in the indirect path between SSOA and NIDS=+0.04; change in the direct path= +0.16.

Panel B: Decomposition of intergenerational elasticity (*adjusted earnings*)

	$\boldsymbol{\theta_1}$	$\boldsymbol{\theta_2}$	$\boldsymbol{\omega_1}$	$\boldsymbol{\omega_2}$	$\boldsymbol{\omega_3}$	$\boldsymbol{\varpi_1}$	$\boldsymbol{\varpi_1}'$	$\boldsymbol{\beta}$	**Indirect effect** $(\boldsymbol{\theta_1\omega_2})$	**% due to indirect effect** $[(\theta_1\omega_2)/(\omega_1+\theta_1\omega_2)]\cdot 100$
SSOA	1.286	0.209	0.463	0.091	0.026	0.175	3.419	0.737	0.117	20.16
	(.269)	(.067)	(.052)	(.017)	(.017)	(.006)		(0.041)		
NIDS	0.615	0.092	0.561	0.225	-.021	0.137	5.217	0.701	0.138	19.75
	(.139)	(.023)	(.093)	(.018)	(.015)	(.004)		(0.046)		

Change in the indirect path between SSOA and NIDS=+0.021; change in the direct path= +0.098.

Note: (1) The tables present the contribution of the mediating role of education to earnings persistence in SSOA and NIDS-w5. The tables are based on employed males between 25 and 44 years of age in both samples. Panel A is based on the unadjusted earnings and Panel B is based on the adjusted earnings (for employment volatility). Fathers' earnings are imputed based on their education, race, occupation, and province (Specification 5).
(2) Numbers in parenthesis beneath the regression coefficients are cluster bootstrapped standard errors.

because much of the father-son association in the relationship is captured by the association between fathers' earnings and sons' earnings.

The changes in the relationship between the education and earnings of fathers and sons between the two surveys have implications for changes over time in the mediating role of education in intergenerational earnings persistence. Table 5.6 shows that the association between fathers' earnings and sons' education (θ_1) has declined between the two surveys. On the other hand, the association between sons' education and sons' earnings (ω_2) has increased between the two surveys. Thus, while the influence of fathers' earnings on sons' education has decreased, the returns to sons' education has increased between the surveys. These two opposing changes have resulted in a small positive net change in the mediating role of education in the intergenerational persistence between the two surveys. Further, the change in the direct effect of fathers' earnings on sons' earnings (ω_1) between the two surveys is shown to be positive. These results indicate that both the direct and indirect effects of fathers' earnings on sons' earnings tended to increase persistence between the two surveys. Note that the confounding effects of fathers' education in the relationship also changes between the two surveys.

As shown in the methods section, the overall level of persistence is the sum of the direct and indirect effects of fathers' earnings on sons' earnings and the confounding effects of fathers' education in the relationship.

5.5.2.2 Mediating Role of Education: Females

The mediating role of education in the intergenerational earnings persistence for the father-daughter pairs is carried out following the same procedure as that for father-son pairs. The result of the decomposition is presented in Appendix 5.F. Again, the contribution of the mediating role of education to earnings persistence is computed controlling for the confounding effects of fathers' education. The results show that the contribution of the mediating role of education to earnings persistence is larger in NIDS than in SSOA (31 per cent in NIDS and 23 per cent in SSOA, using the adjusted earnings). The higher contribution of the mediating role of education in NIDS is due to higher returns to education in that survey. If we do not control for the confounding effects of fathers' education, the contribution of the mediating role of education to persistence is about 14 per cent in SSOA and 20 per cent in NIDS (not shown in the Table).

These are comparable to those reported for sons. The decline in the contribution of the mediating role of education to persistence in both surveys when we do not control for the confounding effects of fathers' education is attributable to a strong positive association of fathers' education with both fathers' earnings, daughters' education and daughters' earnings in the surveys. These results starkly show the importance of controlling for the confounding effects of parental education when analyzing changes over time in the contribution of the mediating role of education to intergenerational earnings persistence.

5.6 Discussion of Results

The chapter examined changes over time in intergenerational earnings persistence and the mediating role of education for South Africa. The chapter is based on two independent datasets, with one collected in the early 1990s and the other in 2017. The analysis followed the same procedure in estimating the intergenerational elasticities in the two samples. However, as noted in the empirical results section, the comparative results should be read with the following caveats in mind. First, the datasets differ in their degree of representation of the South African population. NIDS is a representative sample of the population, but SSOA is not as it missed farmworkers who were living on commercial farms. SSOA, by not including farmworkers who were living on commercial farms, is likely to provide elasticity estimates that are biased downward. Second, differences in survey measurement of variables across the surveys and the missingness in the characteristics of fathers (i.e., education, occupation, and province – as these variables were used in imputing the earnings of fathers) might affect the comparative results.

The chapter paid special attention to the challenge posed by the high levels of employment volatility in the South African labour market in obtaining a measure of the permanent earnings of individuals. This was accomplished by adjusting earnings for the probability of being employed. The use of the adjusted earnings provided elasticity estimates that are generally larger than those using the unadjusted earnings. The increase in the elasticity estimates when earnings are adjusted for employment volatility suggests that transitory components in earnings underestimate the elasticity estimates (Solon, 1992).

Comparing the two surveys, the elasticity estimates using the adjusted earnings were marginally higher in SSOA than in NIDS. In SSOA, when a father earns 10 per cent more than the mean earnings of his generation, his son's earnings are approximately 7.3 per cent higher than the generational mean. In NIDS, a similar scenario leads to the son's earnings being about 7.0 per cent higher than the generational mean (among the employed samples). The results suggest that regression towards the mean has been quite slow. Assuming that the elasticity estimates for SSOA are biased downward (because SSOA missed farmworkers who were living on commercial farms), South Africa's level of earnings persistence was probably higher in the first period than the second. However, even if persistence were to decline between the two periods, the elasticity estimates for the second period remain high (by international standards, see e.g., Corak, 2016). Thus, the decline may not be large enough to be seen as economically significant.

The next logical question to ask is why intergenerational earnings persistence has remained high. As noted in the introduction of the thesis, cross-sectional earnings inequality has been persistently high in South Africa. The forces that kept earnings inequality high are more likely the main reasons for the high intergenerational persistence in earnings, which broadly include the legacies of apartheid and the post-apartheid structural changes in the economy. The legacies of apartheid have continued to define the opportunities available to most South Africans (National Planning Commission, 2012). One key legacy in this respect is apartheid's policy of residential segregation. Both theoretical works (e.g., Durlauf, 1996; Fernandez & Rogerson, 1997) and empirical studies (e.g., Chetty et al. 2014a; Corak 2020) show that residential neighbourhood matters to intergenerational mobility. For instance, in a study of intergenerational income mobility across areas within the United States, Chetty *et al.* (2014a) found that high-mobility areas have less residential segregation, less inequality, better school quality, higher social capital and family stability. The findings of Chetty *et al*. could well relate to the South African context given the historical legacy and institutional arrangements that apartheid has left behind. Furthermore, in the South African context, the persistently high economic inequality in the post-apartheid era may have helped in maintaining residential segregation by socioeconomic status, leading to reduced intergenerational mobility.

In the post-apartheid years, the shift in the structure of the economy towards de-industrialization and skills-biased technological change have had differential impact in the labour market by education. Increased demand for skilled workers has led to increasing returns to postsecondary education. This is likely to have increased the gap in children's earnings by parental socioeconomic status, which raises intergenerational earnings persistence. It should also be noted that, as indicated in the introduction of the thesis, increasing returns to education has been one of the main factors preventing economic mobility across countries.

The analysis of changes over time in the contribution of the mediating role of education to intergenerational earnings persistence highlighted the importance of controlling for the confounding effects of parental education. The contribution of the mediating role of education to persistence changes depending on whether the confounding effects of parental education are controlled for or not. For instance, the contribution of the mediating role of education to persistence when the confounding effects of fathers' education are not controlled for is estimated to be about 16 per cent in SSOA and 20 per cent in NIDS. On the other hand, when the confounding effects of fathers' education are controlled for, it is estimated to be about 20 per cent in both surveys. As the confounding effects of parental education are likely to change over time, controlling for parental education is important to properly compare changes over time in the contribution of the mediating role of education to earnings persistence. We have seen in the previous two chapters that in South Africa the relationship of parental education with both children's education and earnings have changed over time.

In sum, the high intergenerational earnings persistence estimates we found both for the early 1990s and two-and-a-half decades into the democratic era suggests that the persistently high cross-sectional earnings inequality in the post-apartheid period is paralleled in the intergenerational dimension.

Appendix

5.A Summary statistics of Main Samples

Table 5.A.1 Summary statistics of all males 25-44 years of age

	SSOA				NIDS-w5			
	Sons		Fathers		Sons		Fathers	
	mean	sd	mean	sd	mean	sd	mean	sd
Age	34.0	5.614			33.6	5.542		
Race								
African	0.665	0.472			0.859	0.348		
Coloured	0.096	0.295			0.081	0.272		
Indian/Asian	0.052	0.223			0.016	0.125		
White	0.187	0.390			0.045	0.207		
Education								
None	0.063	0.243	0.234	0.424	0.010	0.098	0.370	0.483
Primary	0.303	0.460	0.412	0.492	0.071	0.257	0.208	0.406
Inc. Sec.	0.311	0.463	0.242	0.428	0.476	0.499	0.239	0.427
Matric	0.206	0.405	0.052	0.222	0.223	0.417	0.121	0.326
Postsecondary	0.117	0.321	0.060	0.238	0.220	0.414	0.062	0.241
Occupation								
Professional/Manager	0.140	0.347	0.110	0.313	0.104	0.306	0.102	0.303
Operator/Semi-skilled	0.292	0.455	0.329	0.470	0.240	0.427	0.203	0.403
Craft/trade	0.223	0.417	0.188	0.391	0.256	0.436	0.128	0.335
Clerk/sales	0.187	0.390	0.131	0.338	0.210	0.407	0.092	0.289
Elementary	0.158	0.364	0.242	0.429	0.190	0.392	0.474	0.499
Province								
Cape	0.199	0.399	0.167	0.373				
Natal	0.148	0.356	0.159	0.366				
Transvaal	0.305	0.461	0.302	0.459				
Orange Free State	0.033	0.180	0.040	0.196				
Homelands	0.314	0.464	0.332	0.471				
Western Cape					0.112	0.316	0.076	0.265
Northern Cape					0.029	0.168	0.022	0.146
Eastern Cape					0.077	0.267	0.169	0.375
KwaZulu-Natal					0.165	0.371	0.184	0.387
Free State					0.059	0.235	0.065	0.247
Mpumalanga					0.096	0.294	0.103	0.304
Limpopo					0.071	0.256	0.137	0.344
North-West					0.061	0.239	0.081	0.273
Gauteng					0.331	0.471	0.163	0.369
N	**2,443**				**5,738**			

Notes: (1) The table presents summary statistics of respondents (and of their fathers) in SSOA and NIDS-w5 by some background characteristics. The table is based on all males between 25 and 44 years old (whether labour force participants or not).
(2) *Incudes those who were reported as having never worked and coded to elementary occupations.
(3) Results are weighted using the calibrated weights provided with the datasets.

5.B Summary Statistics of Auxiliary Samples

Table 5.B.1 Summary statistics of auxiliary samples: Employed males 30-59 years of age

	Census 1980		PSLSD	
	mean	sd	mean	sd
Age	41.3	8.195	41.2	8.003
Race				
African	0.639	0.480	0.605	0.489
Coloureds	0.085	0.279	0.101	0.302
Indian/Asian	0.033	0.179	0.037	0.189
White	0.243	0.429	0.257	0.437
Education				
None	0.274	0.446	0.118	0.323
Primary	0.288	0.453	0.270	0.444
Inc. Sec.	0.295	0.456	0.333	0.471
Matric	0.078	0.269	0.139	0.346
Postsecondary	0.065	0.246	0.139	0.346
Occupation				
Professional/Manager	0.145	0.352	0.211	0.408
Operator/Semi-skilled	0.214	0.410	0.253	0.435
Craft/trade	0.182	0.386	0.158	0.365
Clerk/sales	0.168	0.374	0.167	0.373
Elementary	0.290	0.454	0.210	0.408
Province				
Cape	0.209	0.407		
Natal	0.112	0.315		
Transvaal	0.451	0.498		
Orange Free State	0.087	0.282		
Homelands	0.141	0.348		
Western Cape			0.135	0.341
Northern Cape			0.015	0.120
Eastern Cape			0.079	0.270
KwaZulu-Natal			0.166	0.372
Free State			0.084	0.278
Mpumalanga			0.095	0.293
Limpopo			0.044	0.206
North-West			0.093	0.290
Gauteng			0.289	0.453
N	**3,182,799**		**3,662**	

Notes: (1) The table presents summary statistics of respondents with respect to some background characteristics in the auxiliary samples (Census 1980 and PSLSD). The table is based on employed males between 30 and 59 years of age in the two datasets.
(2) Categories for education: no education, primary education (grades 1 to 6), lower secondary (grades 7 to 9), upper secondary (grades 10 to 11), matric (grade 12), and postsecondary (above grade 12).
(3) PSLSD results are weighted using the calibrated weights provided with the datasets.

5.C Probit and Earnings Regression Models

Table 5.C.1 Probit model for the probability of being employed

Variables	SSOA	NIDS-w5
Age	0.187**	0.119**
	(0.0840)	(0.0582)
Age squared	-0.00236*	-0.00165*
	(0.00124)	(0.000865)
Education		
Primary	-0.287*	-0.327
	(0.161)	(0.236)
Inc. Sec.	-0.134	-0.226
	(0.163)	(0.227)
Matric	0.265	0.0659
	(0.177)	(0.232)
Postsec.	0.486**	0.235
	(0.205)	(0.235)
Race		
Coloured	0.456***	0.0436
	(0.132)	(0.0740)
Indian/Asian	0.478***	0.199
	(0.148)	(0.220)
White	0.727***	0.398**
	(0.126)	(0.192)
Married	0.490***	0.631***
	(0.0839)	(0.0741)
Urban	0.102	0.310***
	(0.116)	(0.0523)
Constant	-3.031**	-1.457
	(1.386)	(0.990)
N	2,307	3,616

Notes: (1) The table is based on males between 25 and 44 years of age who are in the labour force in SSOA and NIDS-w5.
(2) Reference category for education is 'no education', reference for race is African. Married and Urban are binary variables (with reference never married and rural, respectively).
(3) Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 5.C.2 Earnings Regression Model

Variables	SSOA	NIDS-w5
Age	0.156***	0.0781*
	(0.0449)	(0.0440)
Age squared	-0.00190***	-0.000784
	(0.000654)	(0.000646)
Education		
Primary	0.430***	0.0515
	(0.0986)	(0.176)
Inc. Sec.	0.794***	0.532***
	(0.0967)	(0.165)
Matric	1.248***	1.017***
	(0.101)	(0.168)
Postsec.	1.521***	1.799***
	(0.104)	(0.168)
Race		
Coloured	0.253***	0.177***
	(0.0676)	(0.0521)
Indian/Asian	0.493***	0.590***
	(0.0687)	(0.137)
White	0.930***	1.210***
	(0.0504)	(0.0935)
Constant	5.526***	8.279***
	(0.757)	(0.756)
N	1,847	2,601
R-squared	0.460	0.329

Notes: (1) The table is based on employed males between 25 and 44 years of age in SSOA and NIDS-w5.
(2) Reference category for education is 'no education', reference for race is African. Married and Urban are binary variables (with reference categories never married and rural, respectively).
(3) Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

5.D First stage regressions in the TS2SLS

Table 5.D.1 First stage regressions in SSOA

	Variables used for imputing earnings of fathers (Z)				
Variables	**Education**	**Education & Race**	**Education, Race & Occupation**	**Education, Race & Province**	**Education, Race, Occupation & Province**
Education	Yes	Yes	Yes	Yes	Yes
Race	-	Yes	Yes	Yes	Yes
Occupation	-	-	Yes	-	Yes
Province	-	-	-	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes
Age squared	Yes	Yes	Yes	Yes	Yes
Constant	5.176***	5.446***	6.050***	5.352***	5.992***
	(0.0250)	(0.0238)	(0.0228)	(0.0239)	(0.0229)
N	3,114,240	3,114,240	3,042,120	3,114,240	3,042,120
R-squared	0.264	0.331	0.357	0.333	0.362

Notes: The auxiliary sample for SSOA is the 1980 Census. The regressions are based on employed males who are between 30 and 59 years of age.

Table 5.D.2 First stage regressions in NIDS-w5

	Variables used for imputing earnings of fathers (Z)				
Variables	**Education**	**Education & Race**	**Education, Race & Occupation**	**Education, Race & Province**	**Education, Race, Occupation & Province**
Education	Yes	Yes	Yes	Yes	Yes
Race	-	Yes	Yes	Yes	Yes
Occupation	-	-	Yes	-	Yes
Province	-	-	-	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes
Age squared	Yes	Yes	Yes	Yes	Yes
Constant	0.899*	1.427***	2.176***	1.553***	2.241***
	(0.519)	(0.458)	(0.448)	(0.451)	(0.445)
N	2,908	2,908	2,906	2,908	2,906
R-squared	0.466	0.587	0.614	0.605	0.625

Notes: The auxiliary sample for NIDS-w5 is PSLSD. The regressions are based on employed males who are between 30 and 59 years of age.

5.E Intergenerational earnings elasticity: Females

Table 5.E.1 Intergenerational earnings elasticity: Females

		Variables used for imputing earnings of fathers (Z)				
		Education	Education & race	Education, race & occupation	Education, race & province	Education, race, occupation & province
		(1)	(2)	(3)	(4)	(5)
Panel A: Unadjusted Earnings						
SSOA	β	0.738	0.673	0.696	0.652	0.643
	s.e.	(0.071)	(0.058)	(0.059)	(0.062)	(0.065)
	N	[741]	[741]	[703]	[593]	[567]
NIDS-w5	β	0.578	0.687	0.658	0.666	0.663
	s.e.	(0.051)	(0.053)	(0.057)	(0.057)	(0.062)
	N	[2040]	[2040]	[1837]	[1846]	[1689]
Test of: β of NIDS < β of SSOA						
Z-calculated		-1.830	0.178	-0.463	0.166	0.223
p-value		0.034	0.571	0.323	0.568	0.587
Panel B: Adjusted Earnings						
SSOA	β	0.859	0.807	0.829	0.782	0.769
	s.e.	(0.069)	(0.058)	(0.060)	(0.063)	(0.067)
	N	[735]	[735]	[697]	[587]	[561]
NIDS-w5	β	0.586	0.706	0.680	0.684	0.685
	s.e.	(0.049)	(0.050)	(0.054)	(0.054)	(0.051)
	N	[2034]	[2034]	[1831]	[1841]	[1684]
Test of: β of NIDS < β of SSOA						
Z-calculated		-3.226	-1.319	-1.846	-1.181	-0.998
p-value		0.001	0.093	0.032	0.119	0.159

Note: (1) The table presents intergenerational elasticity estimates for father-daughter pairs in SSOA and NIDS-w5 with fathers' earnings predicted using five alternative specifications. The analysis is based on employed females between 25 and 44 years of age in the two samples. The estimates in Panel A are using earnings that are unadjusted for employment volatility, and those in Panel B are using the adjusted earnings.

(2) Numbers in parenthesis beneath the regression coefficients are cluster bootstrapped standard errors, and below them in brackets are sample sizes.

5.F Mediating role of education: Females

Table 5.F.1 Mediating role of education: Females

The estimated regression models are:

(1) $S_{ij}^{s} = \theta_0 + \theta_1 lnY_j^f + \theta_2 S_j^f + \theta_3 A_{ij}^s + \theta_4 (A_{ij}^s)^2 + \xi_{ij}$

(2) $lnY_{ij}^{s} = \omega_0 + \omega_1 lnY_j^f + \omega_2 S_{ij}^s + \omega_3 S_j^f + \omega_4 A_{ij}^s + \omega_5 (A_{ij}^s)^2 + \zeta_{ij}$

(3) $ln\, Y_{ij}^f = \varpi_0 + \varpi_1 S_{ij}^f + \nu_{ij}$

(4) $lnY_{ij}^s = \alpha + \beta lnY_j^f + \varphi A_{ij}^s + \gamma (A_{ij}^s)^2 + \varepsilon_{ij}$

Summary of regression coefficients

Panel A: Decomposition of intergenerational elasticity (*unadjusted earnings*)

	θ_1	θ_2	ω_1	ω_2	ω_3	ϖ_1	ϖ_1'	β	**Indirect effect** $(\theta_1\omega_2)$	**% due to indirect effect** $[(\theta_1\omega_2)/(\omega_1+\theta_1\omega_2)]\cdot 100$
SSOA	0.743	0.408	0.260	0.126	0.035	0.166	3.531	0.671	0.094	26.55
	(.246)	(.063)	(.071)	(.026)	(.021)	(.008)		(0.063)		
NIDS	0.551	0.091	0.252	0.244	0.024	0.119	5.923	0.663	0.134	34.74
	(.169)	(.023)	(.095)	(.019)	(.015)	(.004)		(0.061)		

Change in the indirect path between SSOA and NIDS=+0.040; change in the direct path= -0.008.

Panel B: Decomposition of intergenerational elasticity (*adjusted earnings*)

	θ_1	θ_2	ω_1	ω_2	ω_3	ϖ_1	ϖ_1'	β	**Indirect effect** $(\theta_1\omega_2)$	**% due to indirect effect** $[(\theta_1\omega_2)/(\omega_1+\theta_1\omega_2)]\cdot 100$
SSOA	0.743	0.408	0.362	0.144	0.030	0.166	3.531	0.796	0.107	22.87
	(.246)	(.063)	(.074)	(.026)	(.021)	(.008)		(0.065)		
NIDS	0.488	0.092	0.305	0.280	0.020	0.131	5.316	0.685	0.137	30.92
	(.151)	(.023)	(.085)	(.020)	(.015)	(.004)		(0.056)		

Change in the indirect path between SSOA and NIDS=+0.030; change in the direct path= -0.057.

Note: (1) The tables present the contribution of the mediating role of education to earnings persistence in SSOA and NIDS-w5. The tables are based on employed females between 25 and 44 years of age in both samples. Panel A is based on the unadjusted earnings (for employment volatility) and Panel B is based on the adjusted earnings. Fathers' earnings are imputed based on their education, race, occupation, and province (Specification 5).
(2) Numbers in parenthesis beneath the regression coefficients are cluster bootstrapped standard errors.

Chapter 6

6. Summary of Results and Conclusion

The thesis examined changes over time in the relationship between the educational and earning outcomes of parents and their adult children in South Africa. South Africa, with its complex colonial and apartheid history, presents an interesting case. The country has been grappling with persistently high rates of poverty, unemployment, and inequality. These challenges are rooted in the country's past, the legacy of which continues to influence the post-apartheid socioeconomic outcomes.

The study made use of two datasets. These are the Survey of Socioeconomic Opportunity and Achievement (SSOA) and the wave 5 dataset of the National Income Dynamics Study (NIDS-w5). The first dataset was collected right before the fall of apartheid and the second was collected after more than two decades into the democratic era. The timing of the two surveys allowed for an examination of changes in mobility during the first 25 years of democracy.

The cross-sectional trends in educational and earnings inequality are extensively studied. There are also a number of studies on the trends in intergenerational education inequality and a couple of studies on the levels of intergenerational earnings inequality. This study enriches and extends the literature on the trends in intergenerational inequality in several ways. First, it provided an extended view of trends in intergenerational education mobility using datasets that are collected before and after the fall of apartheid. Empirical analyses to date have only focused on post-apartheid datasets. Second, it provided new evidence of changes in intergenerational earnings persistence between two well-spaced periods where one belonged to the pre-1994 and the other to the post-1994 periods. In addition, it investigated changes in the mediating role of education in intergenerational earnings persistence between the two periods. Third, it examined changes in the importance of parental background (proxied by parental education) to children's earnings and earnings inequality between the two periods.

Summary of main findings:

- ***Absolute education mobility has declined across the younger birth cohorts.***

The trend in absolute education mobility was analysed using three measures of absolute mobility. These are upward mobility and two varieties of per capita mobility (namely, *mean absolute difference* and *mean difference*). All three measures showed that absolute mobility declined in the younger birth cohorts, after having been increasing across the older birth cohorts. The trend in upward mobility showed that the proportion of children who surpassed the educational attainment of their parents reached its peak for those born in the mid-1970s at about 91 per cent, and started a gradual decline afterwards and was about 83 per cent for those born in the early 1990s. The per capita mobility measures were also seen to decline for the younger birth cohorts. The trend in *mean absolute difference*, for example, reached its peak for those born in the mid-1970s at about 6.5 years and is shown to decline rapidly in subsequent birth cohorts such that it was about 4.6 years for those born in the early 1990s. The *mean difference* estimates are shown to be quite close in magnitude to the *mean absolute difference* estimates across the older birth cohorts. This suggests that chances of downward mobility in the older birth cohorts were quite low. However, across the younger birth cohorts, the *mean difference* estimates are shown to be smaller than *the mean absolute difference* estimates, suggesting an increase in the chances of downward mobility in the younger birth cohorts.

The decline in absolute mobility in the younger birth cohorts holds even when we re-examine the trends by re-coding parental education to grade 12 for children whose parental education is above grade 12 so that the bounded nature of education would not be an issue. This suggests that the decline in absolute mobility in the younger birth cohorts may not be explained by the bounded nature of education. We argued that the decline resulted due to the stagnation of the educational attainment of children in the younger birth cohorts.

- ***Relative education persistence (mobility) has declined (increased) across birth cohorts.***

The trend in relative education persistence was analysed using three measures of relative mobility. These are the intergenerational regression coefficient (IGRC), the intergenerational correlation coefficient (IGC) and the intergenerational rank coefficient (IRC). Relative education persistence was shown to decline more rapidly across the birth cohorts when measured by IGRC

than by IGC or IRC. This indicates that changes in educational inequality of children and their parents across the birth cohorts played a considerable role in reducing the level of transmitted inequality over time. However, the decline in IGC and IRC across the birth cohorts also indicates that the decrease in persistence was not only due to changes in educational inequality but also due to changes in inherited inequality.

The summary of the trends in education mobility showed that South Africa has experienced substantial increases in education mobility between the two periods (SSOA and NIDS-w5) both in the overall change in educational attainment across generations (absolute mobility) and in the degree of association between the educational attainment of parents and their children (relative mobility).

- ***The partial effect of parental education on children's earnings has been positive and statistically significant in the early 1990s and remained so some 25 years later. However, the contribution of parental education to earnings inequality has been small (in comparison to the contribution of own education) in both periods.***

The addition of parents' education to Mincerian earning regressions reduced the magnitude of the returns to own education in both periods. This suggests that parental background has a direct influence on the earnings of their children, over and above the child's own education. In addition, the partial effect of parental education on children's earnings was positive and statistically significant in the early 1990s and remained so some 25 years later. Among a sample of employed males who are between 25 and 59 years of age, one additional year of parental education is shown to be associated with earning increase of about 2.8 per cent in SSOA and 3.4 per cent in NIDS, for a given level of child's education.

The contribution of parental education to children's earnings inequality was quite small in both periods (about 6.0 per cent in SSOA and 4.9 per cent in NIDS) in comparison to the contribution of own education (20.1 per cent in SSOA and 22.6 per cent in NIDS). However, the contribution of parental education to inequality could be economically important as these are above and beyond the contribution of own education to earnings inequality.

- ***Intergenerational earnings persistence has been high in the early 1990s and remained almost unchanged two-and-a-half decades later.***

Intergenerational elasticity measured using earnings adjusted for employment volatility showed that earnings persistence was marginally lower in the second period than the first. However, even if persistence were higher in the first period, the elasticity estimates for the second period remain high by international standards. According to our preferred estimates, in SSOA when a father earns 10 per cent more than the mean earnings of his generation, his son's earnings were approximately 7.3 per cent higher than the generational mean. In NIDS a similar scenario leads to the son's earnings being about 7.0 per cent higher than the generational mean.

The analysis of changes over time in the contribution of the mediating role of education to intergenerational earnings persistence highlighted the importance of controlling for the confounding effects of parental education. The contribution of the mediating role of education to persistence changes depending on whether we control for the confounding effects of parental education or not. For instance, among a sample of employed males who are between 25 and 44 years of age, the contribution of the mediating role of education to earnings persistence when we control for the confounding effects of parental education was about 20 per cent both in SSOA and NIDS. However, when we do not control for the confounding effects of parental education, it was about 16 per cent in SSOA and 20 per cent in NIDS.

Overall, the study showed that while intergenerational education persistence has decreased considerably, earnings persistence has remained high and almost unchanged between the two periods. These contrasting trends in intergenerational education and earnings inequality mirror the post-apartheid cross-sectional trends in education and earnings inequality. The same reasons as to why increased educational attainment of South Africans has not been associated with a decline in earnings inequality might also be the reasons as to why increased intergenerational education mobility of South Africans is not associated with a decline in earnings persistence. South African studies of the cross-sectional trends in educational and earnings inequality have shown that while increasing educational attainment had inequality decreasing effects, increasing returns to postsecondary education had the opposite effects, leading to persistently high earnings inequality in the post-apartheid era. Further, although the average education of the working-age

population has been increasing over the years, the intra-generational trend in average education among the younger birth cohorts has been stagnant. The stagnation in educational attainment could be due to the poor quality of education in many South African schools, which hinders learners from continuing to higher levels of education (esp. postsecondary education). These observations, together with the findings of this study, suggest that while individuals are attaining higher levels of education than their parents (educational mobility), their attained level of education is not high enough for them to benefit from the rewards of the labour market, which leads to intergenerational persistence in earning outcomes.

Education is key to the labour market outcomes of South Africans. Its importance has also been increasing over time, given the shifts towards de-industrialization and skills-biased technological change in the economy. However, the educational attainment among the younger birth cohorts has been stagnant, leading to an increase in the proportion of young people with below postsecondary levels of education in the labour market. This is a concerning trend. The increasing pool of individuals with below postsecondary levels of education coupled with the rise in the returns to postsecondary education in the labour market may motivate rich parents to invest more in the education of their children. The increase in the relative investment of rich (versus poor) parents in the education of their children may lead to increased intergenerational persistence in education, which in turn may lead to higher persistence in earnings.

Given the increasing importance of education to the economic outcomes of South Africans, education remains the main driver in the process of intergenerational economic mobility. Thus, fixing the quality of education could go a long way in addressing both the stagnation in educational attainment and the high levels of intergenerational persistence in earnings.

The thesis focused on describing changes in mobility over time, and relied on the Becker & Tomes (1979, 1986) human capital model of intergenerational mobility for its theoretical underpinnings. The study did not investigate the correlates of mobility or the channels through which economic outcomes persist across generations in the South African context. In this respect, with the available South African datasets, future studies can investigate the correlates of (education) mobility, and which post-apartheid interventions have been effective.

References

Aaronson, D. & Mazumder, B. 2008. Intergenerational Economic Mobility in the U.S., 1940 to 2000, *The Journal of Human Resources*. 43(1): 139-172.

Agnarsson, S. & Carlin, P. S. 2002. Family Background and the Estimated Return to Schooling: Swedish Evidence. *The Journal of Human Resources*. 37 (3): 680-692.

Alesina, A., Hohmann, S., Michalopoulos, S. & Papaioannou, E. 2021. Intergenerational Mobility in Africa. *Econometrica*. 89(1): 1-35.

Allsion, P.D. 2002. Missing Data. Sage Publications, Inc.

Anderson, K. G., Case, A. & Lam, D. 2001. Causes and consequences of schooling outcomes in South Africa: Evidence from survey data. *Social Dynamics.* 27(1): 37-59.

Angrist, J. D. & Krueger, A. B. 1992. The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples. *Journal of the American Statistical Association*. 87(418): 328-336.

Arellano M. & Meghir C. 1992. Female labour supply and on-the-job search: an empirical model estimated using complementary data sets. *The Review of Economic Studies*. 59(3): 537-559.

Armitage, J., & Sabot, R. 1987. Socioeconomic Background and the Returns to Schooling in Two Low-Income Economies. *Economica*, *54*(213): 103–108.

Arntz, M., Lipowski, C., Neidhöfer, G. & Zierahn-Weilage, U. 2022. Computers as stepping stones? Technological change and equality of labor market opportunities. Working Papers 617, ECINEQ, Society for the Study of Economic Inequality.

Ashenfelter, O. & Zimmerman, D. J. 1997. Estimates of the Returns to Schooling from Sibling Data: Fathers, Sons, and Brothers. *The Review of Economics and Statistics*. *79*(1): 1–9.

Asher, S., Novosad, P. & Rafkin, C. 2022. Intergenerational Mobility in India: New Methods and Estimates across Time, Space, and Communities. G²LM|LIC Working Paper No. 66.

Azam, M. & Bhatt, V. 2015. Like Father, Like Son? Intergenerational Educational Mobility in India. *Demography*. 52(6): 1929–59.

Azomahou, T. T. & Yitbarek, E. A. 2016. Intergenerational education mobility in Africa: has progress been inclusive? Policy Research Working Paper Series 7843, The World Bank.

Banerjee, A., Galiani, S., Levinsohn, J., McLaren, Z. & Woolard, I. 2008. Why has unemployment risen in the new South Africa? *Economics of Transition*. 16(4): 715-740.

Bärnighausen, T., Hosegood, V., Timaeus, I.M. & Newell, M.L. 2007. The socioeconomic determinants of HIV incidence: Evidence from a longitudinal, population-based study in rural South Africa. AIDS. 21(Suppl 7): S29-S38. [http://dx.doi.org/10.1097/01.aids.0000300533.59483.95]

Becker, G. S. & Tomes, N. 1979. An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy*. 87 (6): 1153–1189.

Becker, G. S. & Tomes, N. 1986. Human capital and the rise and fall of families. *Journal of Labor Economics*. 4(3): S1-S39.

Behrman, J. 2019. Human capital and social mobility in low-and middle-income countries. No wp-2019-85, WIDER Working Paper Series, World Institute for Development Economic Research (UNU-WIDER).

Behrman, J.R., Birdsall, N. & Székely, M. 1999. Intergenerational Mobility in Latin America: Deeper Markets and Better Schools Make a Difference. In *New Markets, New Opportunities? Economic and Social Mobility in a Changing World*. N. Birdsall and C. Graham, Eds. Washington, DC: Carnegie Endowment for International Peace.

Behrman, J. R., Gaviria, A., Székely, M. 2001. Intergenerational mobility in Latin America. Working Paper, No. 452, Inter-American Development Bank, Research Department, Washington, DC.

Beinart, W. 2001. Twentieth-Century South Africa. Oxford University Press.

Beinart, W. & Dubow, S. 1995. Segregation and Apartheid in Twentieth-Century South Africa. London: Routledge.

Bhorat, H., Stanwix, B. & Thornton, A. 2022. Changing dynamics in the South African labour market. In: *The Oxford Handbook of the South African Economy.* A. Oqubay, F. Tregenna & I. Valodia, Eds. Chapter 31, pp. 673-689. Oxford University Press.

Bhorat, H., van der Westhuizen, C. & Jacobs, T. 2009. Income and Non-Income Inequality in Post-Apartheid South Africa: What are the Drivers and Possible Policy Interventions? DPRU Working Paper 09/138.

Björklund, A. & Jäntti, M. 1997. Intergenerational income mobility in Sweden compared to the United States. *The American Economic Review*, 87(5): 1009–1018.

Björklund, A. & Salvanes, K. G. 2011. Education and Family Background: Mechanisms and Policies. In: *Handbook of the Economics of Education*. Erik Hanushek, Stephen Machin & Ludger Woessmann, Eds. Elsevier, edition 1, volume 3, chapter 3, pages 201-247.

Black, S. E. & Devereux, P. J. 2011. Recent Developments in Intergenerational Mobility. In: *Handbook of Labor Economics*. O. Ashenfelter & D. Card, Eds. Elsevier, edition 1, volume 4, chapter 16. Pages 1487-1541.

Blanden, J. 2009. How Much Can We Learn from International Comparisons of Intergenerational Mobility? CEE Discussion Papers 0111, Centre for the Economics of Education, LSE.

Blanden, J., Doepke, M. & Stuhler, J. 2023. Educational inequality. In: *Handbook of the Economics of Education.* Eric A. Hanushek, Stephen Machin & Ludger Woessmann, Eds. Elsevier, Volume 6, Chapter 6. 405-497.

Blanden, J., Goodman, A., Gregg, P. & Machin, S. 2004. Changes in intergenerational mobility in Britain. In *Generational income mobility in North America and Europe.* M. Corak, Eds. Cambridge, Cambridge University Press.122–146.

Blanden, J., Gregg, P. & Macmillan, L. 2007. Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education, *The Economic Journal*. 117 (519): C43–C60.

Bloise, F., Brunori, P. & Piraino, P. 2021. Estimating intergenerational income mobility on sub-optimal data: a machine learning approach. *The Journal of Economic Inequality*. 19: 643-665.

Böhlmark, A. & Lindquist, M. 2006. Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden. *Journal of Labor Economics*. University of Chicago Press. 24(4): 879–900.

Bowles, S. & Gintis, H. 2002. The inheritance of inequality. *Journal of Economic Perspectives*. 16(3): 3-30.

Branson, N. 2019. Adding a top-up sample to the National Income Dynamics Study in South Africa. Technical paper number 8, South Africa Labour and Development Research Unit.

Branson, N. & Wittenberg, M. 2019. Longitudinal and Cross-Sectional Weights in the NIDS Data 1-5. NIDS technical paper number 9.

Branson, N., Garlick, J., Lam, D. & Leibbrandt, M. 2012. Education and Inequality: The South African Case. A Southern Africa Labour and Development Research Unit Working Paper Number 75. Cape Town: SALDRU, University of Cape Town.

Branson, N., Hofmeyr, C. & Lam, D. 2014. Progress through school and the determinants of school dropout in South Africa. *Development Southern Africa.* Taylor & Francis Journals, 31(1): 106-126.

Branson, N. & Lam, D. 2010. Education Inequality in South Africa: Evidence from the National Income Dynamics Study. *Studies in Economics and Econometrics.* 34(3): 85-109.

Branson, N. & Lam, D. 2022. The economics of education in South Africa. In *The Oxford Handbook of the South African Economy.* A. Oqubay, F. Tregenna & I. Valodia, Eds. Chapter 33. pp. 707-734. Oxford University Press.

Branson, N. & Leibbrandt, M. 2013. Educational Attainment and Labour Market Outcomes in South Africa, 1994–2010. OECD Economics Department Working Paper No. 1022. Paris: OECD.

Brophy, T., Branson, N., Daniels, R.C., Leibbrandt, M., Mlatsheni, C. & Woolard, I. 2018. National Income Dynamics Study panel user manual. Release 2018. Version 1. Cape Town: Southern Africa Labour and Development Research Unit.

Burns, J. & Keswell, M. 2012. Inheriting the future: Intergenerational persistence of education status in KwaZulu-Natal. *Economic History of Developing Regions*. 27(1), 150–175.

Card, D. 1995. Using Geographic Variation in College Proximity to Estimate the Return to Schooling. In *Aspects of Labour Market Behavior: Essays in Honor of John Vandercamp.* L. N. Christofides, E. K. Grant & R. Swidinsky (Eds.). Toronto: University of Toronto Press, pp. 201–22.

Casale, D. & Posel, D. 2010. The Male Marital Earnings Premium in the Context of Bride Wealth Payments: Evidence from South Africa. *Economic Development and Cultural Change*, *58*(2), 211–230.

Chamberlain, D. & van der Berg, S. 2002. Earnings Functions, Labor Market Discrimination and Quality of Education in South Africa. Working Papers 02/2002, Stellenbosch University, Department of Economics.

Chen, Y. & Feng, S. 2011. Parental education and wages: Evidence from China. *Frontiers of Economics in China*. 6(4): 568–591

Chen, Y., Naidu, S., Yu, T. & Yuchtman, N. 2015. Intergenerational mobility and institutional change in 20th century China. *Explorations in Economic History*. Elsevier, 58(C): 44-73.

Chetty, R., Hendren, N., Kline, P. & Saez, E. 2014a. Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics*, 129 (4): 1553-1623.

Chetty, R., Hendren, N., Kline, P., Saez, E. & Turner, N. 2014b. Is the United States still a land of opportunity? Recent trends in intergenerational mobility. *The American Economic Review*. 104(5): 141-147.

Christie, P. & Gaganakis, M. 1989. Farm Schools in South Africa: The Face of Rural Apartheid. *Comparative Education Review*. 33 (1): 77-92.

Chu, Y. L. & Lin, M. 2020. Intergenerational earnings mobility in Taiwan: 1990-2010. *Empirical Economic.* 59(1): 11-45. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2835881

Clogg, C. C., Petkova, E. & Haritou, A. 1995. Statistical methods for comparing regression coefficients between models. *American Journal of Sociology*. 100(5): 1261-1293.

Connolly, M., Haeck, C. & Lapierre, D. 2021. Trends In Intergenerational Income Mobility and Income Inequality in Canada. Statistics Canada, Analytical Studies Branch research paper series, 1205-9153, no. 458.

Corak, M. 2004. Generational income mobility in North America and Europe: An introduction. In: *Generational Income Mobility in North America and Europe.* M. Corak, Ed. pp. 1-37. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511492549.002

Corak, M. 2016. Inequality from Generation to Generation: The United States in Comparison. IZA Discussion Paper No. 9929. Institute of Labor Economics (IZA).

Corak, M. 2020. The Canadian Geography of Intergenerational Income Mobility. *The Economic Journal*. 130(631): 2134–2174.

Cross, M. & Chisholm, L. 1990. The Roots of Segregated Schooling in Twentieth-Century South Africa. In *Pedagogy of Domination: Toward a Democratic Education in South Africa*. M. Nkomo, Ed. Trenton, NJ: Africa World Press, pp.43-76.

Cunha, F., Heckman, J. J., Lochner, L. & Masterov, D. V. 2006. Interpreting the Evidence on Life Cycle Skill Formation. In: *Handbook of the Economics of Education*. E. Hanushek and F. Welch, Eds. Elsevier, Volume 1, Chapter 12, pp 697-812.

Dang, T. 2020. Intergenerational Earnings and Income Mobility in Vietnam. *LABOUR*. 34 (1): 113–136.

Davis, J. & Mazumder, B. 2017. The decline in intergenerational mobility after 1980. *Working Paper*, No. 2017-05 (Chicago, Federal Reserve Bank of Chicago).

Department of Statistics (now Statistics South Africa). 1985. South African Census 1980 [dataset]. Version 1.3. Pretoria: Department of Statistics [producer], 1985. Cape Town: DataFirst [distributor], 2020. DOI: https://doi.org/10.25828/zytt-pn96

Depken, C., Chiseni, C. & Ita, E. 2019. Returns to education in South Africa: Evidence from the National Income Dynamics Study. *Zagreb International Review of Economics & Business*. 22(1): 1-12.

Deutscher, N., & Mazumder, B. 2023. Measuring Intergenerational Income Mobility: A Synthesis of Approaches. *Journal of Economic Literature*, 61 (3): 988–1036**.**

Doan, Q. & Nguyen, N. 2016. Intergenerational income mobility in Vietnam. MPRA Paper 70603, University Library of Munich.

Donn, G. 1995. Education Policy in the New South Africa: Centre – Province Relations. Occasional Papers No. 56. Centre of African Studies. Edinburgh University.

Dunn, C. E. 2007. The intergenerational transmission of lifetime earnings: Evidence from Brazil. *The B.E. Journal of Economic Analysis & Policy.* 7(2): 1-42.

Durlauf, S.N. 1996. A theory of persistent income inequality. *Journal of Economic Growth*: 1(1), 75–93.

Emran, M. S., Greene, W. & Shilpi, F. 2018. When Measure Matters: Co-residency, Truncation Bias, and Intergenerational Mobility in Developing Countries. *Journal of Human Resources*. 53(3): 589-607.

Emran, M. S. & Shilpi, F. 2018. Estimating intergenerational mobility with incomplete data: Co-residency and truncation bias in rank-based relative and absolute mobility measures. Policy Research Working Paper Series 8414. The World Bank.

Emran, M. S. & Shilpi, F. 2021. Economic approach to intergenerational mobility: Measures, methods, and challenges in developing countries. In: *Social Mobility in Developing Countries: Concepts, Methods, and Determinants*. Vegard Iversen, Anirudh Krishna, and Kunal Sen, Eds. Oxford University Press. pp. 197–220.

Fan, Y., Yi, J. & Zhang, J. 2021. Rising Intergenerational Income Persistence in China. *American Economic Journal: Economic Policy*. 13 (1): 202-30.

Fedderke, J., de Kadt, R. & Luiz, J. 2000. Uneducating South Africa: The failure to address the 1910–1993 legacy. *International Review of Education*. 46(3/4): 257-281.

Feinstein, C.H. 2005. *An Economic History of South Africa: Conquest, discrimination, and development.* Cambridge University Press, Cambridge.

Fernandez, R. & Rogerson, R. 1997. Keeping people out: income distribution, zoning, and the quality of public education. *International Economic Review*, 23–42.

Fields, G.S. 2003. Accounting for Income Inequality and Its Change: A New Method, with Application to the Distribution of Earnings in the United States. In: *Worker Well-Being and Public Policy* (*Research in Labor Economics, Vol. 22*). S.W. Polachek, ed. Emerald Group Publishing Limited, Bingley, pp. 1-38.

Fields, G. S. & Ok, E. A. 1996. The Meaning and Measurement of Income Mobility. *Journal of Economic Theory*. 71(2): 349-377.

Finn, A., Leibbrandt, M. & Ranchhod, V. 2017. Patterns of persistence: Intergenerational mobility and education in South Africa. Version 3. Cape Town: SALDRU, UCT. (SALDRU Working Paper Number 175/ NIDS Discussion Paper 2016/2).

Fiorio, C.V. & Jenkins, S.P. 2007. ineqrbd: Regression-based inequality decomposition, following Fields (2003). UKSUG, Available at: https://www.stata.com/meeting/13uk/fiorio_ineqrbd_UKSUG07.pdf

Fiske, E. & Ladd, H. 2004. Elusive Equity: Education Reform in Post-apartheid South Africa. Washington, DC: Brookings Institution Press.

Franzini, M. & Raitano, M. 2009. Persistence of inequality in Europe: the role of family economic conditions. *International Review of Applied Economics*. 23(3): 345–366.

Franzini, M., Patriarca, F. & Raitano, M. 2020. Market competition and parental background wage premium: the role of human and relational capital. *Journal of Economic Inequality*. 18: 291–317.

Funjika, P. & Gisselquist, R. M. 2020. Social mobility and inequality between groups. WIDER Working Paper Series wp-2020-12, World Institute for Development Economic Research (UNU-WIDER).

Girdwood, S. & Leibbrandt, M. 2009. Intergenerational mobility: Analysis of the NIDS wave 1 dataset. NIDS discussion paper 15. SALDRU, UCT.

Gottschalk, P. & Smeeding, T. M. 1997. Cross-National Comparisons of Earnings and Income Inequality. *Journal of Economic Literature*. 35(2): 633–687.

Gong, H., Leigh, A. & Meng, X. 2012. Intergenerational Income Mobility in Urban China. *Review of Income and Wealth*. 58(3): 481–503.

Government of South Africa, 1994. White paper on reconstruction and development, Government Gazette, Republic of South Africa. Cape Town, 23 NOVEMBER 1994.

Gregg, P., Jonsson, J., Macmillan, L. & Mood, C. 2017. The Role of Education for Intergenerational Income Mobility: A comparison of the United States, Great Britain, and Sweden. *Social Forces*. 96(1): 121-152.

Griliches, Z. 1977. Estimating the Returns to Schooling: Some Econometric Problems. *Econometrica, 45*(1), 1–22.

Haider, S. & Solon, G. 2006. Life-cycle variation in the association between current and lifetime earnings. *The American Economic Review*. 96(4): 1308–1320.

Heckman, J. J. 2008. Role of Income and Family Influence on Child Outcomes. *Annals of the New York Academy of Sciences*. 1136. 307-23.

Heckman, J.J. & Hotz, V.J. 1986. An investigation of the labor market earnings of Panamanian males: evaluating the sources of inequality. *Journal of Human Resources.* 21(4): 507–542.

Hertz, T. 2001. *Education, inequality and economic mobility in South Africa* (Ph.D. dissertation). University of Massachusetts.

Hertz, T. 2003. Upward bias in the estimated returns to education: Evidence from South Africa. *The American Economic Review*. 93(4): 1354– 68.

Hertz, T., Tamara, J., Piraino, P., Selcuk, S., Smith, N., & Verashchagina, A. 2008. The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends. *The B.E. Journal of Economic Analysis & Policy*. 7(2):1-48.

Hoque, M.M., King, E.M., Montenegro, C.E. & Orazem, P. F. 2019. Revisiting the relationship between longevity and lifetime education: global evidence from 919 surveys. *Journal of Population Economics.* 32(2): 551–589.

Horrell, M. 1968. Bantu education to 1968. South African Institute of Race Relations, Johannesburg.

Horrell, M. 1970. The Education of the Coloured Community in South Africa 1652-1970. South African Institute of Race Relations, Johannesburg.

Inoue, A. & Solon, G. 2010. Two-sample instrumental variables estimators. *The Review of Economics and Statistics*. 92(3): 557–561.

Jäntti, M., & Jenkins, S. P. 2015. Chapter 10 - Income Mobility. In: *Handbook of Income Distribution*. Anthony B. Atkinson and Francois Bourguignon, Eds. Version 2, Elsevier.

Jerrim, J., Choi, Á. & Simancas, R. 2016. Two-Sample Two-Stage Least Squares (TSTSLS) estimates of earnings mobility: how consistent are they? *Survey Research Methods*, 10(2), 85-101.

Jones, M. P. 1996. Indicator and Stratification Methods for Missing Explanatory Variables in Multiple Linear Regression. *Journal of the American Statistical Association*, *91*(433), 222–230.

Katz, L. F. & Autor, D. H. 1999. Changes in the Wage Structure and Earnings Inequality. In *Handbook of Labor Economics*. Orley C. Ashenfelter and David Card, Eds. Volume 3A. Amsterdam: North-Holland.

Kerr, A. 2018. Job flows, worker flows and churning in South Africa. *South African Journal of Economics*, 86(S1), 141–166.

Keswell, M. & 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.

Knight, J., Sicular, T. & Yue, X. 2011. Educational Inequality in China: The Intergenerational Dimension. University of Western Ontario, Centre for Human Capital and Productivity (CHCP) Working Papers 2011-13.

Kwenda, P., Ntuli, M. & Gwatidzo, T. 2015. Temporal Developments in Intergenerational Transmission of Education: Case for Black South Africans. *Research in Social Stratification and Mobility.* 42: 96–113.

Lam, D. 1999. Generating Extreme Inequality: Schooling, Earnings, and Intergenerational Transmission of Human Capital in South Africa and Brazil. Population Studies Center, University of Michigan, Report 99-439, 1999.

Lam, D. 2020. Why Has Income Inequality Increased while Education Inequality Has Decreased in Many Developing Countries? In: *The Political Economy of Inequality: U.S. and Global Dimensions*, Sisay Asefa and WeiChiao Huang (Eds). Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, pp. 79-116. https://doi.org/10.17848/9780880996730.Ch5

Lam, D. Finn, A. & Leibbrandt, M. 2015. Schooling inequality, returns to schooling, and earnings inequality: Evidence from Brazil and South Africa. WIDER Working Paper 2015/050, United Nations University World Institute for Development Economics Research.

Lam, D. & Schoeni, R. F. 1993. Effects of Family Background on Earnings and Returns to Schooling: Evidence from Brazil. *Journal of Political Economy*. 101(4): 710–740.

Lam, D. & Schoeni, R. F. 1994. Family Ties and Labor Markets in the United States and Brazil. *The Journal of Human Resources*. *29*(4): 1235–1258.

Laurison, D. & Friedman, S. 2016. The Class Pay Gap in Higher Professional and Managerial Occupations. *American Sociological Review*. *81*(4): 668–695.

Lefranc, A. 2018. Intergenerational earnings persistence and economic inequality in the long run: Evidence from French cohorts, 1931–75. *Economica*. 85(340): 808–845.

Leibbrandt, M., Woolard, I., Finn, A. & Argent, J. 2010. Trends in South African income distribution and poverty since the fall of Apartheid. OECD Social, Employment and Migration Working Papers, No. 101, OECD Publishing. Available at: http://dx.doi.org/10.1787/5kmms0t7p1ms-en

Levine, D. & Mazumder, B. 2002. Choosing the right parents: Changes in the intergenerational transmission of inequality between 1980 and the early 1990s. Working Paper Series WP-02-08, Federal Reserve Bank of Chicago.

Lillard, L. & Kilburn, M. R. 1995. Intergenerational Earnings Links: Sons and Daughters. https://www.rand.org/pubs/drafts/DRU1125.readonline.html.

Liu, J-T, Hammitt, J. K. & Lin, C. J. 2000. Family background and returns to schooling in Taiwan. *Economics of Education Review*. 19(1): 113-125.

Lleras-Muney, A. 2005. The relationship between education and adult mortality in the United States. *The Review of Economic Studies*. 72 (1), 189–221.

Lochner, L. 2008. Intergenerational Transmission. In: *The New Palgrave Dictionary of Economics*. S.N. Durlauf and L.E. Blume, Eds. 2nd Edition, New York: Palgrave-Macmillan.

Lochner, L., 2011. Non-production benefits of education: crime, health, and good citizenship. In: *Handbook of the Economics of Education*. E.A. Hanushek, S. Machin, L. Woessman, Eds. Elsevier, Amsterdam, vol. 4, pp. 183–282.

Long, J. & Ferrie, J. 2013. Intergenerational Occupational Mobility in Great Britain and the United States since 1850. *American Economic Review*, 103 (4): 1109–37.

Lubotsky, D. & Wittenberg, M. 2006. Interpretation of Regressions with Multiple Proxies. *The Review of Economics and Statistics*, *88*(3), 549–562.

Mazumder, B. 2005. Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data. *The Review of Economics and Statistics*. 87(2): 235–255.

McKenna, A. 2011. The history of Southern Africa (The Britannica Guide to Africa). Britannica Educational Publishing, Rosen Educational Services, New York.

Michelangeli, A., Östh, J. & Türk, U. 2022. Intergenerational income mobility in Sweden: A look at the spatial disparities across municipalities. *Regional Science Policy & Practice*, 14(4), 981–1004.

Mincer, J. 1974. *Schooling, Experience and Earnings*. New York: National Bureau of Economic Research.

Modalsli, J. 2017. Intergenerational Mobility in Norway, 1865-2011. *The Scandinavian Journal of Economics*, *119*(1), 34–71.

Mohammed, A.R.S. 2019. Does a Good Father Now Have to be Rich? Intergenerational Income Mobility in Rural India. *Labour Economics*, 60(C): 99–114.

Mogstad, M. & Torsvik, G. 2023. Chapter 6 - Family background, neighborhoods, and intergenerational mobility. In: *Handbook of the Economics of the Family*. Lundberg, S. & Voena, A, Eds. North-Holland, vol. 1 (1), pp. 327-387.

Morduch, J., & Sicular, T. 2002. Rethinking Inequality Decomposition, with Evidence from Rural China. *The Economic Journal*, *112*(476), 93–106.

Murphy, K. M., & Topel, R. H. 1985. Estimation and Inference in Two-Step Econometric Models. *Journal of Business & Economic Statistics*, *3*(4), 370–379.

Mwabu, G. & Schultz, T.P. 1996. Education returns across quintiles of the wage function: alternative explanations for returns to education by race in South Africa. *The American Economic Review.* 86 (2): 335–339.

Mwabu, G. & Schultz, T. P. 2000. Wage Premiums for Education and Location of South African Workers, by Gender and Race. *Economic Development and Cultural Change*, *48*(2), 307–334.

Narayan, A., van der Weide, R., Cojocaru, A., Lakner, C., Redaelli, S., Mahler, D. Ramasubbaiah, R. & Thewissen, S. 2018. Fair Progress? Economic Mobility Across Generations Around the World. World Bank.

National Planning Commission. 2012. National Development Plan 2030: Our Future – make it work. Republic of South Africa.

Neidhöfer, G., Ciaschi, M. & Gasparini, L. 2021. Intergenerational mobility in education in Latin America.CAF – Working Paper No. 2021/14.

Neidhöfer, G., Serrano, J. & Gasparini, L. 2018. Educational inequality and intergenerational mobility in Latin America: A new database. *Journal of Development Economics*, 134: 329–349.

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

Nuñez, J. & Miranda, L. 2011. Intergenerational income and educational mobility in urban Chile. *Estudios de Economia*, 38(1), 195–221.

Nybom, M. & Stuhler, J. 2016. Heterogeneous income profiles and lifecycle bias in intergenerational mobility estimation. *The Journal of Human Resources*, 51(1): 239–268.

Nybom, M. & Stuhler, J. 2017. Biases in standard measures of intergenerational income dependence. *The Journal of Human Resources*. 52(3): 800-825.

Oreopoulos, P. & Salvanes, K. G. 2011. Priceless: The Nonpecuniary Benefits of Schooling. *Journal of Economic Perspectives*. 25 (1): 159-84.

Pekkala, S. & LUCAS, R. E. B. 2007. Differences across cohorts in Finnish intergenerational income mobility. *Industrial Relations: A Journal of Economy and Society.* 46(1): 81–111.

Pekkarinen, T., Salvanes, K. G. & Sarvimäki, M. 2017. The evolution of social mobility: Norway during the twentieth century. *The Scandinavian Journal of Economics*. 119(1): 5–33.

Pillay, P. 1984. The development and underdevelopment of education in South Africa. Second Carnegie Inquiry into Poverty and Development in Southern Africa. Conference Paper no. 95. Cape Town.

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

Piraino, P. 2020. Drivers of mobility. *WIDER Working Paper* Series wp-2020-6, World Institute for Development Economic Research (UNU-WIDER).

Polachek, S. 2007. Earnings over the Lifecycle: The Mincer Earnings Function and Its Applications. IZA Discussion Papers 3181, Institute of Labor Economics (IZA).

Psacharopoulos, G. & Patrinos, H. 2018. Returns to investment in education: a decennial review of the global literature. *Education Economics*. 26 (5): 445-458.

SALDRU. 1995. Southern Africa Labour and Development Research Unit. Project for Statistics on Living Standards and Development 1993, Merged [dataset]. Version 1.1. Cape Town: SALDRU [producer], 1995. Cape Town: DataFirst [distributor], 2013. DOI: https://doi.org/10.25828/rgzn-x098

SALDRU. 2018. Southern Africa Labour and Development Research Unit. National Income Dynamics Study 2017, Wave 5. Version 1.0.0 Pretoria: Department of Planning, Monitoring, and Evaluation

[funding agency]. Cape Town: Southern Africa Labour and Development Research Unit [implementer], 2018. Cape Town: DataFirst [distributor], 2018. DOI: https://doi.org/10.25828/fw3h-v708

Salisbury, T. 2016. Education and inequality in South Africa: Returns to schooling in the post-apartheid era. *International Journal of Educational Development*. 46:.43-52.

Schultz, T.P. 1987. Education investments and returns in economic development. Economic Growth Center Discussion Papers. 536. Yale University, New Haven, CT. Available at: https://elischolar.library.yale.edu/egcenter-discussion-paper-series/536

Seekings, J. & Nattrass, N. 2005. Class, race, and inequality in South Africa. Yale University Press, New Haven and London.

Sherer, G. 2000. Intergroup economic inequality in South Africa: the post-apartheid era. *The American Economic Review*. 90 (2): 317–321.

Shorrocks, A. F. 1982. Inequality Decomposition by Factor Components. *Econometrica*. 50(1): 193-211.

Solon, G. 1992. Intergenerational Income Mobility in the United States. *The American Economic Review*. 82(3): 393– 408.

Solon, G. 1999. Intergenerational Mobility in the Labor Market. Handbook of Labor Economics. In *Handbook of Labor Economics*. O. Ashenfelter & D. Card, Eds. Elsevier, edition 1, volume 3, chapter 29, pages 1761-1800.

Solon, G. 2002. Cross-Country Differences in Intergenerational Earnings Mobility. *The Journal of Economic Perspectives. 16*(3): 59–66.

Solon, G. 2004. A model of intergenerational mobility variation over time and place. In *Generational income mobility in North America and Europe*. Miles Corak (Ed.), Cambridge, UK: Cambridge University Press.

Solon, G. 2014. Theoretical Models of Inequality Transmission across Multiple Generations. *Research in Social Stratification and Mobility*. 35: 13–18.

Spaull, N. 2013a. Poverty & privilege: Primary school inequality in South Africa. *International Journal of Educational Development*. 33(5): 436-447.

Spaull, N. 2013b. South Africa's education crisis: The quality of education in South Africa 1994–2011. Report commissioned by the Centre for Development and Enterprise, Johannesburg.

Spaull, N. 2015. Schooling in South Africa: How low-quality education becomes a poverty trap. In *The South African Child Gauge*. A. de Lannoy, S. Swartz, L. Lake & C. Smith (Eds). Cape Town: University of CapeTown, Children's Institute, pp. 34–41.

Spaull N. 2019. Equity: A price too high to pay. In *South African Schooling: The Enigma of Inequality–A study of present situation and future possibilities*. N. Spaull and J. Janssen, Eds. Cham: Springer.

Stelcner, M., Arriagada, A-M. & Moock, P. 1988. Wage determinants and school attainment among men in Peru. Living standards measurement study (LSMS) working paper no. LSM 38 Washington, D.C.: World Bank.

Strauss, J. & Thomas, D. 1991. Wages, schooling, and background: investments in men and women in urban Brazil. Economic growth center, Yale University. Available at: https://elischolar.library.yale.edu/egcenter-discussion-paper-series/657

Thomas, D. 1996. Education across generations in South Africa. *American Economic Review*. 86(2): 330-334.

Torche, F. 2019. Educational mobility in developing countries. WIDER Working Paper, No. 2019/88. The United Nations University World Institute for Development Economics Research (UNU-WIDER).

Torche, F. 2021. Intergenerational Mobility in Latin America in Comparative Perspective. UNDP LAC Working Paper No. 02. Background Paper for the UNDP LAC 2021 Regional Human Development Report.

Treiman, D. 2017. Survey of Socioeconomic Opportunity and Achievement 1991-1994 [dataset]. Version 1. Los Angeles: UCLA, Cape Town: DataFirst [producers], 2017. Cape Town: DataFirst [distributors], 2017. DOI: https://doi.org/10.25828/kcx2-kn30

van der Berg, S. 2007. Apartheid's Enduring Legacy: Inequalities in Education. *Journal of African Economies*. 16(5): 849–880.

van der Berg, S., Burger, C., Burger, R., de Vos, M., du Rand, G., Gustafsson, M., Moses, E., Shepherd, D., Spaull, N., Taylor, S., van Broekhuizen, H. & von Fintel, D. 2011. Low quality education as a poverty trap. Stellenbosch Economic Working Papers: 25/11, University of Stellenbosch.

van der Weide, R., Lakner, C., Mahler, D. G., Narayan, A. & Ramasubbaiah, R. 2021. Intergenerational Mobility around the World. Policy Research Working Paper Series 9707. The World Bank.

Wittenberg, M. 2017. Wages and Wage Inequality in South Africa 1994–2011: PART 2 – Inequality Measurement and Trends. *South African Journal of Economics.* 85(2): 298-318.

Wittenberg, M. 2022. Unbiased estimators of permanent earnings and log of permanent earnings. Unpublished Mimeo. University of Cape Town, Cape Town.

World Bank. 2021. Global Database on Intergenerational Mobility. Washington, D.C.: World Bank Group.

Datasets Used

NIDS Wave 1:

Southern Africa Labour and Development Research Unit. National Income Dynamics Study (NIDS) Wave 1, 2008 [dataset]. Version 7.0.0. Pretoria: SA Presidency [funding agency]. Cape Town: Southern Africa Labour and Development Research Unit [implementer], 2018. Cape Town: DataFirst [distributor], 2018. https://doi.org/10.25828/e7w9-m033

NIDS wave 2:

Southern Africa Labour and Development Research Unit. National Income Dynamics Study Wave 2, 2010-2011 [dataset]. Version 4.0.0. Pretoria: SA Presidency [funding agency]. Cape Town: Southern Africa Labour and Development Research Unit [implementer], 2018. Cape Town: DataFirst [distributor], 2018. https://doi.org/10.25828/j1h1-5m16

NIDS wave 3:

Southern Africa Labour and Development Research Unit. National Income Dynamics Study Wave 3, 2012 [dataset]. Version 3.0.0. Pretoria: SA Presidency [funding agency]. Cape Town: Southern Africa Labour and Development Research Unit [implementer], 2018. Cape Town: DataFirst [distributor], 2018. https://doi.org/10.25828/7pgq-q106

NIDS wave 4:

Southern Africa Labour and Development Research Unit. National Income Dynamics Study 2014-2015, Wave 4 [dataset]. Version 2.0.0. Pretoria: Department of Planning, Monitoring, and Evaluation [funding agency]. Cape Town: Southern Africa Labour and Development Research Unit [implementer], 2018. Cape Town: DataFirst [distributor], 2018. https://doi.org/10.25828/f4ws-8a78

NIDS wave 5:

Southern Africa Labour and Development Research Unit. National Income Dynamics Study 2017, Wave 5 [dataset]. Version 1.0.0 Pretoria: Department of Planning, Monitoring, and Evaluation [funding agency]. Cape Town: Southern Africa Labour and Development Research Unit [implementer], 2018. Cape Town: DataFirst [distributor], 2018. https://doi.org/10.25828/fw3h-v708

SSOA:

Treiman, D. Survey of Socioeconomic Opportunity and Achievement 1991-1994 [dataset]. Version 1. Los Angeles: UCLA, Cape Town: DataFirst [producers], 2017. Cape Town: DataFirst [distributors], 2017. DOI: https://doi.org/10.25828/kcx2-kn30

Project for Statistics on Living Standards and Development (PSLSD) 1993:

Southern Africa Labour and Development Research Unit. Project for Statistics on Living Standards and Development 1993, Merged [dataset]. Version 1.1. Cape Town: SALDRU [producer], 1995. Cape Town: DataFirst [distributor], 2013. DOI: https://doi.org/10.25828/rgzn-x098

South African Census 1980:

Department of Statistics (now Statistics South Africa). 1985. South African Census 1980 [dataset]. Version 1.3. Pretoria: Department of Statistics [producer], 1985. Cape Town: DataFirst [distributor], 2020. DOI: https://doi.org/10.25828/zytt-pn96