



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

UNIVERSITY OF PRETORIA
Department of Economics

Education and Labour Market Outcomes

By

OBREIN MUINE SAMAHIIYA

A dissertation submitted to the University of Pretoria in fulfillment of the requirements of the degree of PHILOSOPHIAE DOCTOR in ECONOMICS in the Faculty of Economic and Management Sciences.

SUPERVISOR: PROFESSOR YOSEPH Y. GETACHEW

Co-SUPERVISOR: PROFESSOR STEVEN F. KOCH

JULY 2025

Abstract

Unemployment is widespread in Africa, particularly in South Africa, where high levels contribute to poverty and inequality, correlating with poor educational outcomes. This thesis explores the impact of education on labour market outcomes in South Africa using data from the South African General Household Survey (GHS). It is comprised of three stand-alone papers each focussing on a specific topic, which corresponds to chapters 2, 3 and 4.

Chapter 2 estimates the effects of education on the probability of employment. I use the compulsory schooling reform of 1997 to identify the effect of education on employment outcomes. The mandatory schooling policy required all children aged 7 and older to be in school up until they reached the age of 15 or Grade 9. Using the GHS data for the period 2014-2018 and the two-stage least squares technique, the results indicate that compulsory schooling increased average educational attainment by between 0.04 and 0.18 years. However, the effect differs by gender: an additional year of compulsory schooling raised education by 0.04 to 0.20 years for males and by 0.09 to 0.17 years for females.

On average, an additional year of education increases the probability of employment by 7.3 to 28.4 percentage points, depending on the survey year. When disaggregated by gender, education is associated with a 3.0 to 18.0 percentage point increase in employment probability for women and a 3.0 to 13.9 percentage point increase for men. The results further reveal substantial variation in the effect of education on employment across race and gender. Black African and White males employment gains from education are larger than for females while among individuals of mixed race, there is no clear gender pattern. These results emphasize the essential contribution of education to improving employment opportunities in South Africa. In addition, they underscore the importance of targeted education policies that addresses gender and racial disparities in employment.

In Chapter 3 I explore whether the 1997 Compulsory Schooling Law helps explain gender-based differences in wage distributions and whether labour market selection effects influence this relationship. The analysis employs conditional quantile regression and a selection-adjusted

approach that accounts for non-random selection into employment. While compulsory schooling laws aim to retain children in school, based on the assumption that increased schooling enhances employment prospects and earnings, South Africa's education system under performs relative to many other countries, which may weaken the intended impact of compulsory schooling reforms on labour market outcomes.

Using the GHS data for 2018 the study reveals that returns to education are generally higher for women than for men, that the compulsory schooling law raised these returns, and that returns vary across both the wage distribution and levels of education — suggesting that schooling contributes to wage inequality. Additionally, adjusting for non-random selection into employment lowers the estimated returns for women but does not change the relative pattern of returns before and after the compulsory schooling policy was implemented. Expanding access to higher education for previously disadvantaged groups and enhancing the quality of skills training through targeted policies could contribute to narrowing labour market earnings disparities.

Chapter 4 examines how education impacts labour market earnings among Black South Africans. According to the country's compulsory schooling law, children must start Grade R or Grade 1 in the year they turn six or seven, provided their birthday falls on or before June 30 of the admission year. I take advantage of the natural variation in educational attainment created by this school entry cut-off to estimate the returns to education using a regression discontinuity design, drawing on data from the GHS for the years 2014 to 2018. This study adds to the sparse literature on the effects of school entry policies in Sub-Saharan Africa.

The results reveal signs of birth date manipulation around January 1st, possibly to benefit from school admission rules. To minimize bias from this potential manipulation, I exclude individuals born on January 1st from the analysis sample. This approach helps reduce distortions in the estimated treatment effects. The evidence suggests that starting school at a younger age is linked to fewer years of schooling completed and lower labour market earnings. However, among those who began school early, education has a statistically significant positive effect on earnings. These results highlight key policy concerns about the optimal school starting age and emphasize the importance of providing additional support to younger entrants who might face academic

disadvantages.

Dedication

This thesis is dedicated to my wife, Johanna, and my children who endured numerous hours of loneliness and uncertainty during this rewarding but difficult journey. I also dedicate this thesis to my departed parents who raised and taught me from a young age.

Acknowledgments

This thesis would not have been possible without the support of many people. Many thanks to my supervisors, Professors Yoseph Y. Getachew and Steven F. Koch, who read my numerous revisions and helped make some sense of the confusion. Also thanks to the University of Pretoria and my employer, the University of Namibia for providing funding for my studies. Moreover, many thanks to Professors who taught me during the first year of my studies. I thank also my classmates, with whom I worked together as we submitted assignments and work that was required.

Declaration

I, Obrein Muine Samahiya, declare that this thesis titled, "Education and Labour Market Outcomes" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help. I further acknowledge the use of Chat-GPT for editing some parts of the thesis.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

TABLE OF CONTENTS

Abstract	i
Dedication	iv
Acknowledgments	v
Declaration	vi
	Page
List of Tables	x
List of Figures	xiii
1 Introduction	1
1.1 Background	1
1.2 Impact of labour laws	6
1.3 Thesis Outline and Contribution	9
1.3.1 Race, Gender, Education and Employment Outcomes	11
1.3.2 Impact of Compulsory Schooling on Returns to Education	12
1.3.3 Too Early to Benefit? School Entry Laws and Returns to Education	14
2 Race, gender, education and employment outcomes	18
2.1 Introduction	19
2.2 Literature Review	22
2.3 Identification Strategy	23
2.3.1 Instrument	24
2.3.2 Model Specification	25
2.3.3 Data	26

TABLE OF CONTENTS

2.4	Results and Discussion	29
2.4.1	Descriptive statistics	29
2.4.2	Regression results	34
2.4.3	Sensitivity analysis	39
2.4.4	Discussion	41
2.5	Conclusion	43
	Appendix for Chapter 2	46
3	Impact of compulsory schooling on returns to education	73
3.1	Introduction	74
3.2	Estimation Strategy	79
3.2.1	Quantile regression	79
3.2.2	Quantile selection model	80
3.2.3	Instrument	83
3.2.4	Data	84
3.3	Results and Discussion	86
3.3.1	Descriptive statistics	86
3.3.2	Quantile regression results	89
3.3.3	Quantile selection results	93
3.3.4	Discussion	95
3.4	Conclusion	98
	Appendix for Chapter 3	100
4	Too Early to Benefit? School Entry Laws and Returns to Education	108
4.1	Introduction	109
4.2	Literature review	112
4.3	Estimation strategy	116
4.3.1	Fuzzy RD	116
4.3.2	Data	118
4.4	Results	120

TABLE OF CONTENTS

4.4.1	Descriptive statistics	120
4.4.2	Evidence of manipulation	122
4.4.3	Impact on Education	125
4.4.4	Impact on earnings	132
4.4.5	Robustness checks	138
4.4.6	Sensitivity analysis	142
4.5	Conclusion	150
	Appendix for Chapter 4	152
5	Conclusion	156
	Bibliography	162

LIST OF TABLES

TABLE	Page
2.1 Summary statistics for the overall, female and male samples	30
2.2 OLS and 2SLS estimates of the effect of education on employment probability	36
2.3 2SLS estimates of the effect of education on employment probability of Black Africans	37
2.4 2SLS estimates of the effect of education on employment probability of Coloureds . .	38
2.5 2SLS estimates of the effect of education employment probability of Whites	38
2.6 2SLS estimates of the effect of grade completion on employment probability	40
A1 Years of education derived from highest education level	46
A2 Summary of the literature on the effect of education on employment probability . . .	47
A3 OLS estimates of the effect of education on employment probability for the overall sample	48
A4 OLS estimates of the effect of education on employment probability for the female sample	49
A5 OLS estimates of the effect of education on employment for the male sample	50
A6 2SLS estimates of the effect of education on employment for the overall sample: second stage	51
A7 2SLS estimates of the effect of education on employment probability for the the overall sample: first stage	52
A8 2SLS estimates of the effect of education on employment probability for the female sample: second stage	53
A9 2SLS estimates of the effect of education on employment probability for the female sample: first stage	54

LIST OF TABLES

A10	2SLS estimates of the effect of education on employment probability for the male sample: second stage	55
A11	2SLS estimates of the effect of education on employment probability for the male sample: first stage	56
A12	2SLS estimates of the effect of education on the likelihood of employment for black women: second stage	57
A13	2SLS estimates of the effect of education on the likelihood of employment for black women: first stage	57
A14	2SLS estimates of the effect of education on the likelihood of employment for black men: second stage	58
A15	2SLS estimates of the effect of education on the likelihood of employment for black men: first stage	58
A16	2SLS estimates of the effect of education on the likelihood of employment for coloured women: second stage	59
A17	2SLS estimates of the effect of education on the likelihood of employment for coloured women: first stage	59
A18	2SLS estimates of the effect of education on the likelihood of employment for coloured men: second stage	60
A19	2SLS estimates of the effect of education on the likelihood of employment for coloured men: first stage	60
A20	2SLS estimates of the effect of education on the likelihood of employment for white women: second stage	61
A21	2SLS estimates of the effect of education on the likelihood of employment for white women: first stage	61
A22	2SLS estimates of the effect of education on the likelihood of employment for white men: second stage	62
A23	2SLS estimates of the effect of education on the likelihood of employment for white men: first stage	62

LIST OF TABLES

A24	Impact of grade 9 completion on employment outcomes: Second stage estimates for the overall sample	63
A25	Impact of grade 9 completion on employment outcomes: First stage estimates for the overall sample	64
A26	Impact of grade 10 completion on employment outcomes: Second stage estimates for the overall sample	65
A27	Impact of grade 10 completion on employment outcomes: First stage estimates for the overall sample	66
A28	Impact of grade 11 completion on employment outcomes: Second stage estimates for the overall sample	67
A29	Impact of grade 11 completion on employment outcomes: First stage estimates for the overall sample	68
A30	Impact of grade 12 completion on employment outcomes: Second stage estimates for the overall sample	69
A31	Impact of grade 12 completion on employment outcomes: First stage estimates for the overall sample	70
A32	Impact of tertiary education on employment outcomes: Second stage estimates for the overall sample	71
A33	Impact of tertiary education on employment outcomes: First stage estimates for the overall sample	72
3.1	Summary statistics for women and men by compulsory schooling treatment exposure	88
3.2	Returns to education for women and men estimated via OLS and quantile regression	90
3.3	Returns to grade completion for women and men estimated via OLS and quantile regression	92
3.4	Returns to education for women estimated via the quantile selection model	93
B1	Returns to education for women estimated via the quantile regression	100
B2	Returns to education for men estimated via the quantile regression	101
B3	Returns to grade completion for women estimated via the quantile regression	102

B4	Returns to grade completion for men estimated via the quantile regression	103
B5	Returns to education for women estimated via the quantile selection model	104
B6	Returns to education for men estimated via the quantile selection model	105
B7	Returns to grade completion for women estimated via the quantile selection models .	106
B8	Returns to grade completion for men estimated via the quantile selection models . . .	107
4.1	Summary statistics of key variables for all individuals, women and men	121
4.2	McCrary manipulation test	124
4.3	Impact on years education and logarithm of monthly earnings for the overall sample	128
4.4	Impact on years of education and the logarithm of monthly earnings for women . . .	131
4.5	Impact on years of education and logarithm of monthly earnings for men	133
4.6	Impact of education on the logarithm of monthly earnings at different placebo cut-offs for the overall sample	139
4.7	Covariate balance: assessing impact of school entry age on language of interview for the overall and female samples	141
4.8	Covariate balance: assessing impact of school entry age on language of interview for the male sample	142
4.9	Covariate balance: assessing impact of school entry age on relation to head of house- hold for the overall sample	143
4.10	Impact of years of education on the logarithm of monthly earnings at different band- widths	144
4.11	Impact of school entry age on years of education at different donut holes for overall, female and male samples	146
4.12	Impact of education on the logarithm of monthly earnings at different donut holes for overall, female and male samples	148
4.13	Impact of education on logarithm of monthly earnings at different polynomial orders for the overall, female and male samples	149

LIST OF FIGURES

FIGURE	Page
1.1 Education completion rate (for persons aged 20 years or older) and unemployment rate in South Africa	2
1.2 Proportion of individuals with a certain level of educational attainment by race	4
1.3 Unemployment and labour force participation rates for women and men	5
2.1 Proportion of individuals with a certain level of education by gender and year of birth	31
2.2 Proportion of individuals attaining a certain level of education by race and year of birth	33
2.3 Fraction of individuals that are employed by level of education and year of birth . . .	34
2.4 Fraction of individuals that are employed by race and year of birth	35
3.1 Distribution of log monthly earnings by gender and treatment status	87
4.1 Birth date heaping around January 1	123
4.2 Manipulation test for birthdays: 2014-2018	124
4.3 Impact of school entry age on years of completed education for the overall sample . .	126
4.4 Impact of school entry age on years of completed education for women	130
4.5 Impact of school entry age on years of completed education for men	132
4.6 Impact of school entry age on the logarithm of monthly earnings for the overall sample	134
4.7 Impact of school entry age on the logarithm of monthly earnings for women	136
4.8 Impact of school entry age on the logarithm of monthly earnings for men	137
C1 McCrary manipulation test for women: 2014-2018	152
C2 McCrary manipulation test for men: 2014-2018	153
C3 Covariate balance tests: gender, marital status and age group	154
C4 Covariate balance tests: province of residence at the time of the survey	155

CHAPTER 1

INTRODUCTION

1.1 Background

The human capital theory and the signalling and screening theories link education to labour market outcomes. The former posits that education is an investment in human capital, which enhances skills and productivity, leading to better employment opportunities and higher wages (Becker, 1962; Mincer, 1958; Schultz, 1961, 1960). On the contrary, the latter paradigm, views education as a “screening, filtering and sorting device” used by employers to hire workers, rather than a productivity enhancer (Arrow et al., 1973; Spence, 1973; Stiglitz, 1975). Empirical studies consistently find a positive correlation between education levels and earnings. Each additional year of schooling typically raises earnings by 8-10% on average (Patrinos et al., 2021).

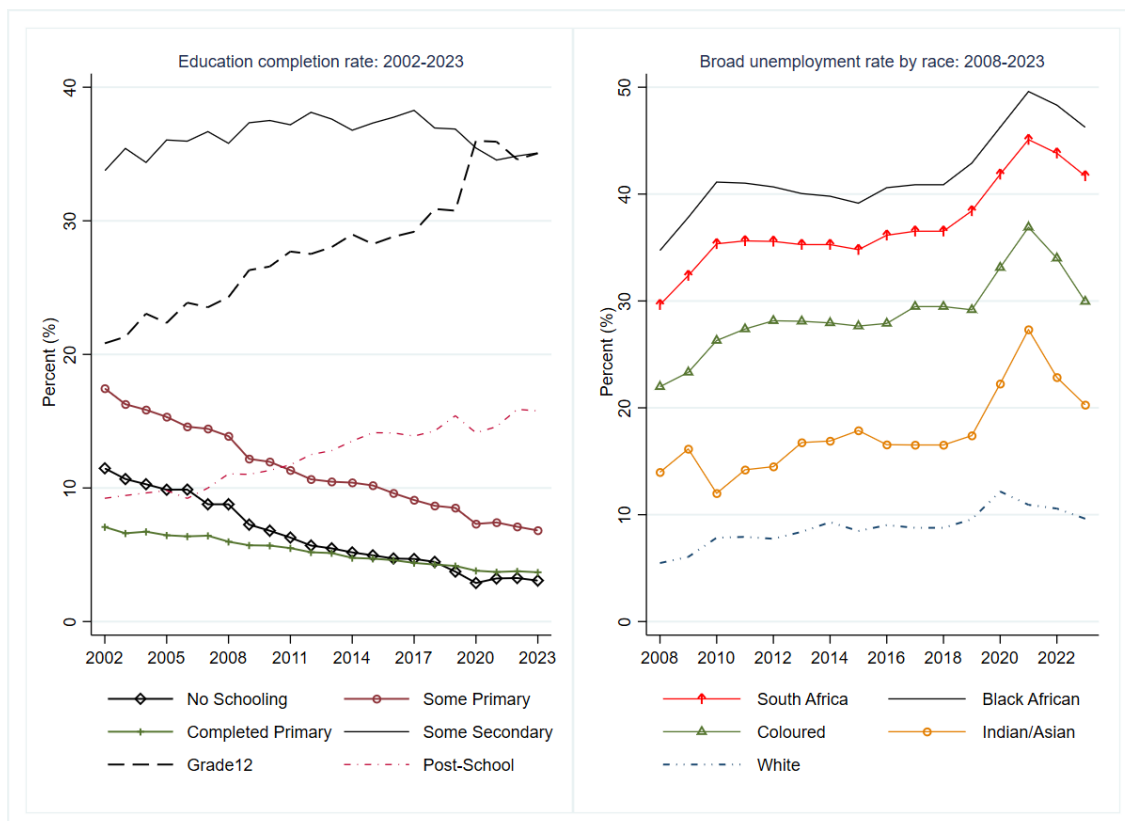
It is imperative to understand the impact of education on labour market outcomes. For example, estimating the employment effects of education provides policymakers with important information on how to address socioeconomic challenges such as income inequality and poverty. Moreover, the returns to education offer a valuable summary of the economic advantages of investing in human capital. This information is essential for policymakers when determining resource allocation. Individuals, too, can use these insights to inform their own choices about investing in their

CHAPTER 1. INTRODUCTION

education (Patrinos et al., 2021; Psacharopoulos, 2024).

This thesis examines the impact of education on labour market outcomes in South Africa. It seeks to achieve three specific objectives: (i) investigate the impact of education on employment outcomes by gender and race, (ii) explore whether the effect of education on labour market earnings varies by compulsory schooling exposure, gender and along the distribution of wages, and (iii) using the school entry rule for identification, estimate the returns to education among Black South Africans. The South African context is characterized by stark socioeconomic inequalities,

Figure 1.1: Education completion rate (for persons aged 20 years or older) and unemployment rate in South Africa



Source: Author's calculations based on data from General Household Surveys: 2002-2023 and the Quarterly Labour Force Surveys: 2008-2023.

with more than 40% of the labour force unemployed (see Figure 1.1) and more than half of the population living in poverty (Francis and Webster, 2019). These disparities have their roots in the country's history of racial discrimination and inequality, particularly in the education sector. The apartheid government created a racially segregated education system with vastly different levels of funding and curriculum quality, resulting in superior education for Whites and inadequate

schooling for non-Whites (Msila, 2007; Burger et al., 2010; Landman et al., 2020; Festus et al., 2016; Thobejane, 2013). The legacy of apartheid has resulted in more than half of the South African population receiving education from a historically disadvantaged school system, which continues to impact their prospects in the labour market (Spaull, 2013, 2015a).

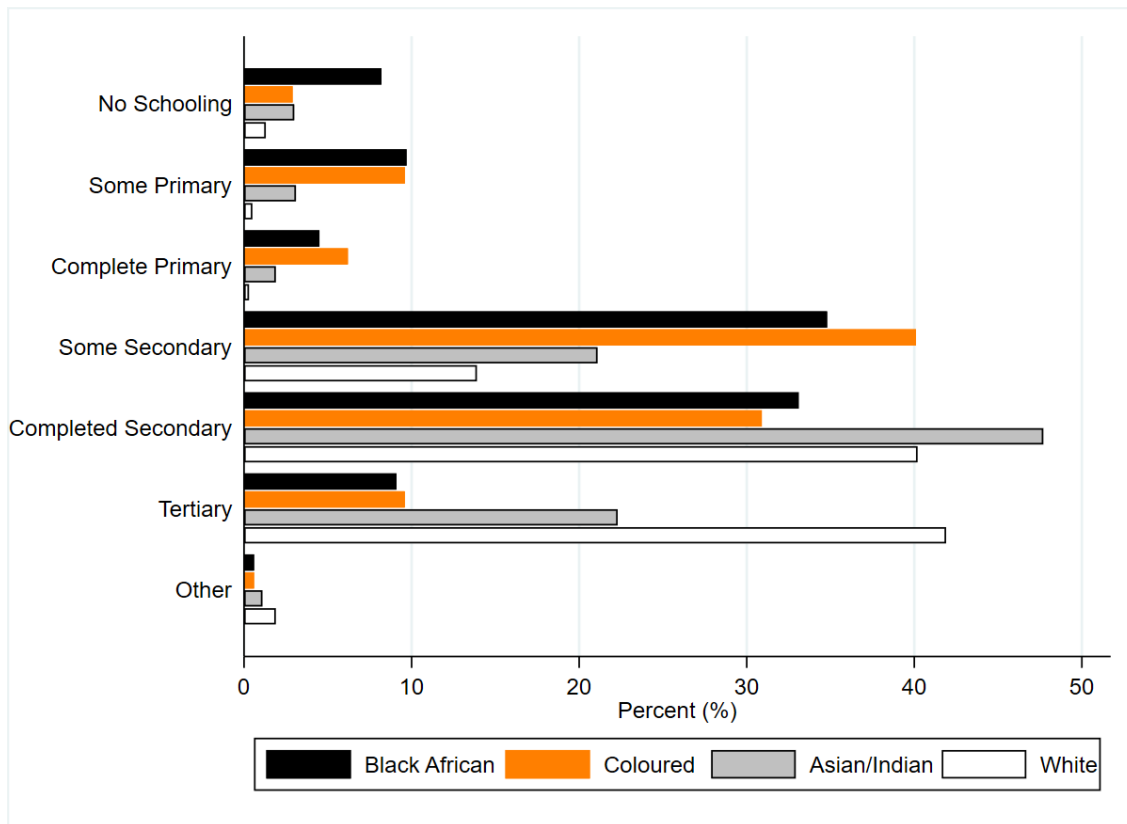
Following the democratic transition in 1994, the new constitution prohibited all forms of racial discrimination and established education as a fundamental right for all citizens (Republic of South Africa, 1996a). In 1996, the government introduced a non-racial national education curriculum and made school attendance compulsory for children aged 7 to 15 years (Republic of South Africa, 1996b). According to UNESCO (2020), from 2002 to 2017, school attendance increased notably across various age groups: by 48 percentage points for 5-year-olds, 20.8 points for 6-year-olds, 2.6 points for those aged 7-15, and 3.4 points for the 16-18 age group. Likewise, attendance in post-secondary institutions saw growth among young people aged 19-23, with rates rising from 0.3% to 6% at TVET/college institutions and from 3% to 6% at universities. As such, the ratio of people without schooling dropped from 11.5% in 2002 to 3.1% in 2023. Similarly, the percentage of people with incomplete primary schooling dropped from 17.5% to 6.8% during the same period. Further, the completion rate of secondary schooling (Grade 12) increased from 20.8% to 35.1% while the completion rate for tertiary education rose from 9.2% to 15.8% during the corresponding period (see Figure 1.1).

Despite these achievements, the education system still grapples with high levels of inequality, dropout rates, and sub-optimal outcomes (Spaull, 2013; Lam et al., 2011; Branson et al., 2014). Spaull (2015a) reports that only 16% of Grade 3 learners achieve an appropriate level (50% or higher) in Mathematics, with the majority performing below grade expectations. Moreover, among the poorest 10%, only 38% of young people who start Grade 1 reach Grade 12, contrasting sharply with 87% from the wealthiest 10% (Spaull, 2015a). Figure 1.2, which shows the proportion of individuals with a certain level of educational attainment, suggests that education outcomes are still correlated with race more than 30 years after the end of apartheid. Accordingly, by 2022, 42% of Whites had a tertiary qualification compared to only 9% Black Africans and 9.6% of Coloureds.

Additionally, the proportion of young people dropping out of school after the compulsory schooling

CHAPTER 1. INTRODUCTION

Figure 1.2: Proportion of individuals with a certain level of educational attainment by race



Source: Author's calculations based on data from the 2022 South African Census. The sample includes individuals aged 34 to 64 years.
Note: Some tables in the Census data did not specify the age range.

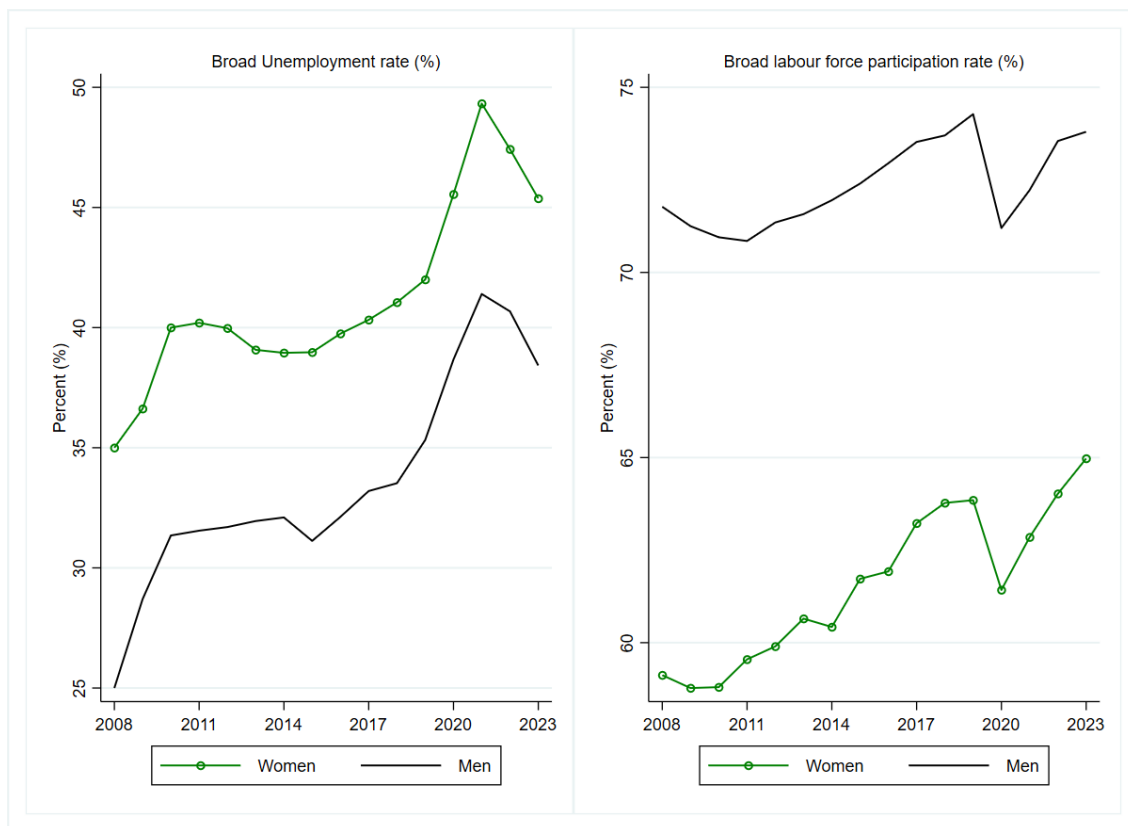
phase has remained high throughout the past two decades (see Figure 1.1), leaving them without necessary skills for the labour market (Lam et al., 2009). Branson et al. (2014) notes that only about 30% of Grade 9 males in 2008 had advanced to Grade 11 in 2010. Figure 1.1 shows that 33.8% of individuals that were surveyed in 2002 had left school without completing secondary schooling, a proportion which rose to 35.1% in 2023. Several factors contribute to this situation, including the poor socioeconomic status of many children (Branson and Leibbrandt, 2013a) and inequality (Branson and Leibbrandt, 2013a; Spaul, 2015a; Timæus et al., 2013).

South Africa is one of the most unequal countries in the world, with a Gini coefficient of 0.69 (Anand et al., 2016). The labour market is a major contributor to income inequality. The unemployment rate has consistently remained high in the last 16 years, with the broad measure of unemployment rising from 29.7% in 2008 to 41.7% in 2023. The unemployment rate is highest among Black Africans at 46.3%, compared to Whites at 9.6% (see Figure 1.1). Unemployment

CHAPTER 1. INTRODUCTION

is closely linked to poverty and inequality, and often correlated with crime, violence, and social unrest. In 2015, only 4.1% of White South Africans were considered poor, compared to 70.8% of Black South Africans, reflecting racial disparities in employment described previously. Moreover, a 10% reduction in the unemployment rate leads to a 3% drop in the Gini coefficient (Anand et al., 2016), underscoring the importance of increasing job opportunities for those historically disadvantaged under apartheid to address the socioeconomic challenges in the country. Gender disparities

Figure 1.3: Unemployment and labour force participation rates for women and men



Source: Author's calculations based on data from the Quarterly Labour Force Surveys.

also exist, with Black African and Coloured women being under-represented in high-skilled jobs (Espinoza et al., 2019). Labour force participation (LFP) trends show a substantial and persistent gender gap over time (see Figure 1.3). The LFP gender gap has narrowed in recent years, being 10.6 in 2023 compared to 15.3 points in 2008. Nevertheless, women's LFP rate has been rising since the end of apartheid due to factors such as improvement in educational attainment and increasing returns to education among others (Ntuli and Wittenberg, 2013). Casale (2004) reports that between 1995 and 2001, the broad labour force participation rate rose from 47.8 to 64.4%, a

rise of 16.6 percentage points in just 7 years. By the end of 2023 it had risen further to 65% (see Figure 1.3).

However, the increase in women's labour force participation rate has led to higher female unemployment, as employment growth has not been sufficient to absorb all the new entrants into the labour market (Casale, 2004; Mosomi, 2019b). Hence, during the 2008-2023 period, the broad unemployment rate among women increased by 10.4 percentage points, reaching 45.4% in 2023 (see Figure 1.3). According to Mosomi (2019b), the gender gap in employment is more pronounced at younger than at older ages, with the disparity decreasing over the life cycle, suggesting that younger women are particularly vulnerable to high unemployment rates.

1.2 Impact of labour laws

Following the end of apartheid in 1994, the democratic government introduced several laws that may have impacted the structure of the labour market. These labour laws aim to dismantle the racial discriminatory policies and practices that existed during the apartheid era. Back then the labour market was configured in favour of the White minority and to the disadvantage of the majority non-White population such that better employment opportunities and conditions were reserved for Whites (Burger et al., 2010; Landman et al., 2020; Festus et al., 2016; Horwitz and Jain, 2011).

The legal instruments that were introduced include the Labour Relations (LR) (Republic of South Africa, 1995), the Basic Conditions of Employment (BCE) (Republic of South Africa, 1997), the Employment Equity (EE) (Republic of South Africa, 1998), and the Broad-Based Black Economic Empowerment (B-BBEE) (Republic of South Africa, 2004) Acts. The LR Act was introduced to give workers several rights including collective bargaining, union membership and the right to strike.

Before 2019, South Africa did not have a single national minimum wage. Only sector-specific minimum wages were legislated (Piek et al., 2023). Besides, making provision for equitable basic employment conditions such as annual leave, sick leave, leave gratuity, and severance pay, the

BCE Act also makes provision for sector-specific minimum wages. Minimum wage laws were set for the most vulnerable sectors, including those with low union density or where market earnings are relatively low. The industries for which minimum wages are established include domestic work, contract cleaning, private security, wholesale and retail, farm work, forestry, taxi, learnership, hospitality, civil engineering, and children's work in advertising, creative, and cultural endeavours (Millea et al., 2017; Nattrass, 2016). Minimum wages are determined either through the sector bargaining councils or the sector determinations by the Employment Conditions Commission (ECC). The national minimum wage law was only introduced in 2019. Therefore, most of the empirical work on minimum wages before the introduction of the national minimum wage focus on specific sectors or sector.

Evidence on the impact of sector-specific minimum wages on employment outcomes is mixed. Bhorat et al. (2013) examines the retail, domestic work, forestry, security, and taxi sectors, finding that minimum wage laws had no statistically significant effect on employment in those sectors with a minimum wage law. Conradie (2003), Conradie (2005), and Murray and Van Walbeek (2007) find no impact of minimum wage laws on farm employment probability, Bhorat, Kanbur and Stanwix (2014) find that minimum wage laws reduce the probability of being in farm employment. In related work, Piek et al. (2023) argue that the impact of agricultural minimum wage laws was limited to non-seasonal workers, finding a large decrease in the employment growth rate of non-seasonal workers.

Similarly, the impact of minimum wage laws on wages remains inconclusive. For example, Bhorat et al. (2013) find that minimum wage laws increased real hourly wages in the retail, domestic work, taxi and security sectors, but had statistically insignificant effect in the forestry sector. In the retail and security sectors, the estimates show that the rise in real hourly wages more than offset any reductions in hours worked, leading to significant increases in total monthly earnings. In the domestic work sector, although the minimum wage had no significant impact on hours worked, real monthly wages still rose significantly in the post-law period, leaving workers better off overall. In contrast, in the forestry sector, the law's impact on real hourly wages was fully offset by reductions in hours worked, resulting in no significant change in real monthly wages.

Moreover, [Millea et al. \(2017\)](#) find that minimum wage laws had impact on wages of Black Africans and Coloureds only, but not Whites. [Bhorat, Kanbur and Stanwix \(2014\)](#) find that the agricultural minimum wage law increased wages of farm workers by an average of 7%. [Bhorat et al. \(2021\)](#) find mixed impacts of the national minimum wage (NMW) law on wages. Whilst in their cross-sectional analysis they find that the NMW had no statistically significant effect on wages, in the panel analysis they find a statistically significant effect of 15.9%. However, they caution that the estimates should be interpreted with care as the results might be influenced by a relatively short panel and the non-compliance by employers.

The EE Act introduced affirmative action (AA) in the workplace to enhance employment opportunities for previously disadvantaged groups such as Black Africans, Coloureds, Indians/Asians, women, and people living with disabilities. To achieve this, employers are required to formulate employment equity plans with numerical targets and measures to eliminate discriminatory barriers and foster workplace diversity ([Burger et al., 2010](#)). The policy applies to firms employing at least 50 employees or firms employing less than 50 employees, but whose income exceed a pre-determined threshold.¹

However, the success of the AA policy appears less than optimal. The commission for employment equity report of the Department of Employment and Labour for the 2022/23 reporting period shows a disproportionate representation of the various population groups. Whilst Black Africans constitute 80.0% of the economically active population (EAP), they occupy only 16.9% of top management positions, both in the private and public sector. This is in sharp contrast with Whites who occupy 62.9% of top management positions whilst their EAP is only 8% ([Department of Employment and Labour, 2023](#)).

Moreover, research by [Espí et al. \(2019\)](#) indicates that Black African and Coloured women, in particular, remain underrepresented in high-skilled and managerial roles, predominantly working in low-paid and low-skilled positions. Another study by [Burger et al. \(2010\)](#) using the Oaxaca and Blinder decomposition method suggests that the impact of the AA policy on labour

¹According to Schedule 4 of the Employment Equity Amendment Act, 2013 the thresholds per sector ranges from R6 million for the agriculture sector to R75 million for the wholesale trade, commercial agents and allied services ([Republic of South Africa, 2014](#)).

market outcomes is minimal compared to education.

The B-BBEE Act was introduced to promote Black economic empowerment. The Act has several targets for empowering Black people² including ownership and management of enterprises, skills development for employees, enterprise and supplier development for Black owned businesses, and socioeconomic development. The B-BBEE regulations set specific targets for Black economic empowerment in each of the above-mentioned areas. For example, the B-BBEE regulations mandates that 60% of senior management positions in designated enterprises should be held by Blacks of which 30% should be reserved for Black women (Baker and McKenzie, 2023). The work by Klasen and Minasyan (2021) who investigated the impact of the B-BBEE Act on the employment of Black women in senior roles indicate that, an average Black woman has a 0.5 to over 1 percentage point probability of being employed in a senior management position compared to other groups (Black men, White men and women).

Despite the commendable efforts by government, the overall evidence suggests limited progress in addressing racial and gender inequities in the labour market. This indicates that labour laws have not been entirely successful in restructuring the labour market, and they appear to have had a limited impact on the labour market.

1.3 Thesis Outline and Contribution

This thesis comprises three main chapters. Chapter 2 examines the impact of education on employment outcomes across gender and race groups using the compulsory schooling law (CSL) of 1997 for identification. Chapter 3 investigates whether exposure to the mandatory schooling reform yields higher returns to education along the wage distribution. Although several developed countries use exogenous variations from compulsory schooling laws (CSLs) to determine the causal effects of education on labour market outcomes (Angrist and Keueger, 1991; Harmon and Walker, 1995; Acemoglu and Angrist, 2000; Oreopoulos, 2006a; Devereux and Hart, 2010; Grenet, 2013; Alam and Mamun, 2016), few do so in developing countries (Torun, 2018; Aydemir and Kirdar, 2017).

²According to the Act Black people is a generic name meaning Africans, Coloureds and Indians.

CSLs offer exogenous variation in schooling as individuals do not choose their birth date and, by extension, their exposure to the law. Moreover, CSLs, which affects individuals during childhood, are unlikely to influence adult labour market outcomes ([Acemoglu and Angrist, 2000](#); [Riddell and Song, 2011](#)). Crucially, the CSL in South Africa likely compelled some learners to remain in school longer than they otherwise would have since there is presence of compliers in the data. The reform is believed to have had the most significant effect on Black South Africans, particularly those with lower levels of educational attainment. The results presented in this thesis align with the literature in indicating that the instrument is sufficiently strong ([Korwatanasakul, 2023](#); [Clay et al., 2021](#); [Oyelere, 2010](#)).

In Chapter 4, I estimate the return to education using a regression discontinuity design. The study exploits South Africa's school entry rule, which requires children who turn seven years old on or before 30 June of the admission year to begin Grade 1, respectively. This work contributes to the relatively small body of literature examining the long-term effects of early school entry. Additionally, it adds to the literature that cautions against relying solely on date of birth as an instrument for identifying the causal effects of education on labour market outcomes.

In contrast to existing studies, this thesis not only estimates average employment and earnings returns to education, but also assesses heterogeneous returns corrected for sample selection bias at various points in the earnings distribution ([Balestra and Backes-Gellner, 2017](#); [Wang, 2013](#)). Although several analyses have explored the education-labour market relationship in South Africa ([Salisbury, 2016](#); [Branson and Leibbrandt, 2013b](#); [Kingdon and Knight, 2004](#); [Festus et al., 2016](#); [Bhorat and Kimani, 2017](#); [Baldry, 2016](#); [Anand et al., 2016](#); [Moll, 1992, 1998](#); [Mwabu and Schultz, 1996, 2000](#); [Bhorat, 2000](#); [Keswell and Poswell, 2002, 2004](#); [Fryer and Vencatachellum, 2005](#); [Bhorat and Goga, 2012](#); [Branson and Leibbrandt, 2013b](#); [Branson et al., 2013](#)), few provide causal estimates ([Depken et al., 2019](#); [Biyase and Zwane, 2015](#); [Zwane, 2020](#); [Nikolov et al., 2020](#); [Hertz, 2003](#)), and the heterogeneity of education's effects on the distribution of earnings has rarely been addressed ([Mwabu and Schultz, 1996](#)).

However, the studies by [Hertz \(2003\)](#) and [Nikolov et al. \(2020\)](#) are not nationally representative. [Hertz \(2003\)](#) examines a small sample of Black individuals in the KwaZulu-Natal province,

while [Nikolov et al. \(2020\)](#) focuses on the Cape Town. In contrast, the analyses conducted by [Depken et al. \(2019\)](#), [Biyase and Zwane \(2015\)](#), and [Zwane \(2020\)](#) rely on identification strategies such as parental education and lagged value of household head's education. Although parental education is correlated with a child's education, it may also be directly linked to wages through channels such as parental social capital. Moreover, using the lagged value of the household head's education based on a relatively short panel will only differ from the initial value if the household head pursues further education during the panel period, potentially undermining the assumption of exogeneity.

This thesis aims to fill these gaps in the literature by evaluating the implications of education on the labour market in South Africa, a developing nation marked by significant inequality and poverty. The significance of this work lies in two key areas. First, understanding rates of return to education is essential for shaping education policy ([Dolton and Sandi, 2017](#)). These rates inform public policy regarding investments in education ([Patrinos, 2016](#)). Thus, establishing the causal average return to education is vital for policy development. Second, recognizing that education has heterogeneous effects on earnings is important for analysing distributional impacts. For instance, if returns are greater in the upper earnings distribution, education may increase, rather than reduce, inequality.

1.3.1 Race, Gender, Education and Employment Outcomes

As noted previously, increasing job opportunities for formerly disadvantaged groups is critical to overcoming South Africa's socioeconomic challenges. Chapter 2 investigates whether education serves as a pathway to improve employment in South Africa, focusing on differences across gender and race. Using the 1997 compulsory schooling law (CSL) as an instrumental variable (IV) for education, this chapter applies the two-stage least squares (2SLS) regression approach to the cross-sectional data from five General Household Surveys (2014-2018).

While the relationship between education and employment has been studied before, few have addressed the potential endogeneity of education ([Alam and Mamun, 2016](#); [Somani, 2021](#); [Bhorat and Kimani, 2017](#); [Festus et al., 2016](#)). Moreover, the broader literature presents mixed results

with some studies showing that education positively (Abekah-Nkrumah et al., 2019; Boahen et al., 2021; Drucker et al., 2022) or negatively (Berhe, 2021; Alam and Mamun, 2016) affects employment, while others find little to no impact (Liwiński, 2020; Torun, 2018; Hall, 2016). These variations may arise from differences in context, data, methods, and identification strategies (Alam and Mamun, 2016; Liwiński, 2020; Torun, 2018). Additionally, the diversity in education systems and labour markets globally underscores the need for country-specific research.

Given South Africa's apartheid history, it provides a unique case for examining the effect of education on employment (Mtapuri and Tinarwo, 2021). As in the international literature, findings from South African studies are conflicting. Some studies show that education reduces the likelihood of unemployment or increases employment probability (Branson and Leibbrandt, 2013b; Festus et al., 2016; Bhorat and Kimani, 2017), while others find little to no effect (Baldry, 2016; Anand et al., 2016).

The findings indicate that, on average, an additional year of education increases the probability of employment by between 7.3 and 28.4 percentage points, depending on the survey year. The results reveal substantial heterogeneity by gender and race, both in the impact of compulsory schooling on educational attainment and in the effect of education on employment outcomes. When disaggregated by gender, the estimated employment gains from education range between 3.0 and 18.0 percentage points for women and 3.0 and 13.9 percentage points for men.

1.3.2 Impact of Compulsory Schooling on Returns to Education

The underlying assumption of compulsory schooling laws (CSLs) is that increased education resulting from CSLs yields favourable labour market outcomes, such as improved earnings (Harmon, 2017). While the South African mandatory schooling policy seems sound in theory, evidence shows that the relationship between education and labour market outcomes is more complex than anticipated. As observed previously, despite the CSL's intentions, the education system is characterised by poor outcomes. This raises concerns about whether the CSL has achieved its intended outcomes, especially in light of persistently low education completion levels.

Chapter 3 investigates the policy's efficacy by examining the impact of the reform on returns to

education, accounting for non-random selection into employment. By using a quantile selection model, recently developed by [Arellano and Bonhomme \(2017\)](#), the analysis provides nuanced insights into how CSL-affected individuals fare across the wage distribution, focusing on Black South African men and women. This method not only allows for the assessment of differences in returns to education across various wage levels but also corrects for potential selection bias, crucial in South Africa, where the unemployment rate is high (see Figure 1.1).

This research fills several gaps in the existing literature. First, it provides an updated comparison of the returns to education between individuals affected by the CSL and those who were not, thereby allowing for inferences regarding the policy's effectiveness. Second, the study employs a novel technique that adjusts for selection bias in quantile regressions, offering a more refined analysis of the returns to education across the wage spectrum. Given that previous research has largely focused on mean returns, this quantile-based approach offers a deeper understanding of the returns at different points of the wage distribution, providing a more comprehensive view of labour market inequality. Third, this study offers the first update in nearly three decades on returns to education along the wage distribution in South Africa, a critical contribution to understanding labour market dynamics.

The focus on Black Africans is particularly important given their historic marginalization under apartheid. The CSL reform is believed to have had the greatest impact on this group, especially those at the lower end of the educational attainment spectrum. Despite the end of apartheid, Black South Africans continue to experience lower educational outcomes and greater barriers to employment compared to other racial groups, making this analysis especially pertinent.

The literature suggests that returns to education often vary along the wage distribution and between different demographic groups. [Buchinsky \(1994\)](#) found that higher returns to education tend to occur at the upper quantiles of the wage distribution, a phenomenon attributed to the complementarity between education and ability ([Patrinos and Sakellariou, 2006](#)). In lower-income countries, however, returns to education tend to decline at higher wage levels, while in higher-income countries, they tend to rise ([Patrinos et al., 2006](#)). Gender differences also play a significant role in this dynamic, with several studies showing that women often receive higher

returns to education than men ([Dougherty, 2005](#)).

This study confirms that returns to education are indeed heterogeneous, varying by exposure to CSL, gender, and educational level. Notably, women experience higher returns to education than men, particularly those affected by the CSL reform. Men, however, see significant returns only at the upper end of the wage distribution, while women benefit across most of the distribution. The results also highlight the presence of sample selection bias, particularly for women. After correcting for this bias, the returns to education decrease by between 0.9 and 4.4 percentage points, depending on the wage quantile.

Importantly, the findings show that returns to education tend to increase with higher educational levels, particularly at the upper end of the wage distribution, which can contribute to widening income inequality. While CSL has successfully increased educational attainment, the benefits in the labour market may not be evenly distributed, potentially reinforcing existing inequalities by providing greater advantages to those who are already better positioned.

In conclusion, Chapter 3 provides a detailed analysis of the impact of CSL on labour market outcomes in South Africa, highlighting both the policy's strengths and limitations. While compulsory schooling has improved educational attainment and increased earnings particularly among women, significant gaps remain. These findings emphasize the importance of targeted interventions aimed at addressing persistent disparities in education and labour market outcomes, especially for marginalized groups such as Black South Africans.

1.3.3 Too Early to Benefit? School Entry Laws and Returns to Education

This chapter investigates the economic returns to education among Black South Africans using data from the General Household Survey (GHS) covering the period 2014–2018 ([Statistics South Africa, 2017a, 2016, 2017b, 2018, 2019](#)). Identifying the true effect of education on earnings is a complex task, primarily due to omitted variable bias. Unobservable traits such as cognitive ability, ambition, or socio-economic background often influence both educational attainment and labour market outcomes, potentially distorting estimates ([Card, 2001](#); [Patrinos et al., 2021](#)). To address this challenge, the analysis draws on the methodology employed in a growing body of

literature such as [Solli \(2017\)](#), [Oosterbeek et al. \(2021\)](#), [Guo et al. \(2023\)](#), and [Balestra et al. \(2020\)](#), which leverages school entry laws as a source of exogenous variation in education using regression discontinuity (RD) designs.

In the South African context, children begin Grade R and Grade 1 in the year they turn six and seven, respectively, as regulated by the compulsory schooling policy outlined in the South African Schools Act ([Republic of South Africa, 1996b](#)) and the Basic Education Amendment Laws Act ([Republic of South Africa, 2024](#)). Furthermore, the school entry policy requires children to reach the age threshold before 30 June of the admission year. Those born after this date typically delay school entry by one year. This cut-off creates a natural discontinuity in educational exposure based solely on birth timing, allowing researchers to isolate the causal effect of schooling by comparing outcomes for individuals born just before and just after the deadline. Compared to alternative strategies such as instrumental variables (IV) or difference-in-differences (DiD), RD designs offer stronger internal validity when properly implemented ([Guo et al., 2023](#)). This study uses local polynomial regressions, a widely accepted approach in RD analysis, to estimate treatment effects ([Gelman and Imbens, 2019](#); [Clark, 2023](#); [Liwiński, 2020](#); [Dolton and Sandi, 2017](#)).

The idea of using school entry laws to study causal effects of education dates back to the influential work of [Angrist and Krueger \(1992\)](#). While this strategy has been applied in a variety of contexts, evidence on the long-run impact of school starting age (SSA) on earnings remains mixed ([Dobkin and Ferreira, 2010](#); [Kawaguchi, 2011](#); [Fredriksson and Öckert, 2014](#); [Larsen and Solli, 2017](#); [Solli, 2017](#); [Görlitz et al., 2025](#)). For instance, studies in countries like China ([Guo et al., 2023](#)), Brazil ([Matta et al., 2016](#); [Peña, 2017](#)), and Mexico ([Peña, 2017](#)) suggest that older school entrants tend to earn more. In contrast, findings from the Netherlands ([Oosterbeek et al., 2021](#)) indicate that younger students may perform better in the labour market. Elsewhere, including the United States ([Dobkin and Ferreira, 2010](#)), Finland ([Pehkonen et al., 2015](#)), and Norway ([Larsen and Solli, 2017](#)), researchers find no significant effect.

Despite the wide application of RD designs in education economics, there is a notable gap in research focused on Sub-Saharan Africa. To date, no study has used this method to estimate the

effect of school starting age on adult wages in the region. Existing research from Nigeria (Oyelere, 2010) and Uganda (Sakaue et al., 2024) has relied on policy changes such as the introduction of universal primary education, to instrument for educational attainment, but these rely on different identification strategies. This chapter fills an important void by applying a regression discontinuity approach to estimate the causal impact of education on earnings in South Africa, a country with persistently high income inequality and labour market segmentation (Wittenberg, 2017b; Mosomi et al., 2020).

A crucial validity concern in RD designs is the possibility of manipulation around the cut-off. Parents may adjust the timing of childbirth or school to ensure that their children fall on the preferred side of the threshold. This type of strategic behaviour has been observed in studies from China (Huang et al., 2020), Argentina (González and Dip, 2024), and Norway (Beck, 2025), and it threatens the integrity of RD estimates by violating the assumption of continuity in potential outcomes. To address this issue, the study employs a donut RD specification, which excludes observations close to the cut-off date where manipulation is most likely to occur (Barreca et al., 2011; Fan et al., 2018). Further robustness checks are provided in the empirical analysis that follows.

South Africa provides a unique context for this research. The country remains one of the most unequal societies in the world as previously discussed (Wittenberg, 2017b; Mosomi et al., 2020). Historically, apartheid-era policies confined Black African men and women to low-paying jobs. While labour laws introduced post-1994 sought to address this, unemployment among Black Africans remain high and Black African women remain underrepresented in high-paying, skilled jobs (Espí et al., 2019; Mosomi, 2019b).

Previous studies on the returns to education in South Africa have largely employed ordinary least squares³ or two-stage least squares (Depken et al., 2019; Biyase and Zwane, 2015; Zwane, 2020) methods, using lagged education (Biyase and Zwane, 2015; Zwane, 2020) and parental education (Depken et al., 2019) as instruments to measure educational impacts. However, these methods

³(Moll, 1992, 1998; Mwabu and Schultz, 2000; Bhorat, 2000; Keswell and Poswell, 2002, 2004; Branson and Leibbrandt, 2013b; Branson et al., 2013; Salisbury, 2016)

CHAPTER 1. INTRODUCTION

may not fully address potential endogeneity bias or fully explore the heterogeneity. Moreover, the current study is different since its focus is on assessing the impact of the school entry rule using a regression discontinuity design.

Results show that early school entry leads to lower educational attainment: individuals who started school early completed, on average, 9 to 12 months less schooling than their peers. This educational shortfall has significant labour market implications, as early entrants earn approximately 16% to 25% less per month, on average. Despite these disadvantages, early entrants still experience sizable returns to education. For this group, each additional year of schooling is associated with a 20% to 30% increase in monthly earnings, underscoring the strong economic value of education.

The impact of education on earnings also varies across survey years and by gender. The highest returns are observed in 2014, followed by 2016 and 2017, while the effects in the remaining years are not statistically significant. When disaggregated by gender, the only significant return is found among women in 2017, where an additional year of education is associated with a 23% increase in earnings. For men, the estimated returns are consistently positive but not statistically significant across all years.

CHAPTER

RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Unemployment is extensive on the African continent and is especially high in South Africa. Such high levels of unemployment are likely to contribute to both poverty and inequality, and are correlated with poor educational outcomes. This study attempts to establish the causal contribution of education to the probability of employment in South Africa. To address the endogeneity of education, I employ the compulsory schooling law as an instrumental variable for education within a two-stage least squares (2SLS) model. The analysis makes use of South Africa's General Household Survey for the years 2014-2018. On average, an additional year of education increases the probability of employment by between 7.3 and 28.4 percentage points, depending on the survey year. I find heterogeneity across gender and ethnicity, both in the effect of compulsory schooling on completed education and the effect of education on employment. When disaggregated by gender, the estimated employment gains due to education are 3.0 to 18.0 percentage points for women and 3.0 to 13.9 percentage points for men. These findings highlight the pivotal role of education in improving employment prospects in South Africa. Expanding access to quality education — particularly at higher levels — represents a key policy instrument for enhancing labour market outcomes and reducing unemployment.

2.1 Introduction

The relationship between education and employment outcomes is explained by two main theoretical frameworks: the human capital theory and signalling (or screening) theories. The human capital theory suggests that education represents an investment that enhances skills and productivity, thereby improving employment prospects (Becker, 1962; Mincer, 1958; Schultz, 1961, 1960). By contrast, signalling and screening theories contend that education primarily acts as a “screening, filtering, and sorting device” for employers, serving as a signal of inherent abilities rather than a direct contributor to productivity (Arrow et al., 1973; Spence, 1973; Stiglitz, 1975).

Despite differing theoretical perspectives, understanding the role of education in shaping labour market outcomes remains a key area of inquiry for both researchers and policymakers (Giotis et al., 2025). While education enhances skills, productivity, and provides credentials, employment offers financial security, opportunities for career progression, and a pathway to social mobility (Giotis et al., 2025). However, unemployment remains a persistent challenge in many countries, particularly in developing economies where structural barriers and labour market rigidities constrain job creation. In 2023, the unemployment rate stood at 7.5% in North America and 6.3% in Northern, Southern, and Western Europe (ILO, 2024), compared to a significantly higher 41.7% broad unemployment rate in South Africa (see Figure 1.1).

The current study considers whether or not education offers a channel through which employment can be affected in South Africa. I also consider heterogeneity in that effect by gender and ethnicity. The analysis applies a two-stage least squares (2SLS) regression using the compulsory schooling law (CSL) of 1997 as an instrumental variable (IV) for education. Given that the educational reform influenced people when they were children, using it for identification has the benefit of being exogenous to adult employment outcomes (Acemoglu and Angrist, 2000). The analysis relies on cross-sectional data from five General Household Surveys (GHS) for the years 2014-2018 (Statistics South Africa, 2017a, 2016, 2017b, 2018, 2019).

Since the democratic transition, unemployment in South Africa has remained stubbornly high—rising from 20% in 1994 to 26.7% in 2000 (Francis and Webster, 2019), and further to 41.7% in 2023.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

These persistently high unemployment levels have been a key driver of the country's severe income inequality, with labour market factors contributing approximately 91% to overall inequality (Finn and Leibbrandt, 2018; Anand et al., 2016). South Africa's Gini coefficient of 0.69 places it among the most unequal societies globally (Mosomi et al., 2020).

Increasing access to employment is essential to addressing both inequality and poverty. According to Finn (2015), 83% of households with no employed members experience poverty, as opposed to 50% of those with at least one wage employee. Anand et al. (2016) estimates that reducing unemployment by 10% lowers the Gini coefficient by 3%.

In addition to broad inequality and poverty, the labour market is associated with other socioeconomic challenges in South Africa (Francis and Webster, 2019). In 2023, the labour force participation (LFP) rate for women was 65.0% compared to 73.8% for men (see Figure 1.2), creating a gender gap in labour force participation of about 13.5%. Meanwhile, the gender gap in the unemployment rate was 18.1%¹ in favour of men during the same year. Women, especially Black African and Coloured women, continue to be underrepresented in high-skilled and management positions (Espi et al., 2019). Thus, increasing employment opportunities for women is also essential in attempts to address socioeconomic challenges.

Furthermore, South Africa's past and current racial divisions continue to be reflected in the country's poverty levels. Figure 1.1 shows that the unemployment rate is highest among the Black African population at 46.3%, followed by mixed-race individuals (30.0%), Indians/Asians (20%), and Whites (9.6%). Additionally, Festus et al. (2016) finds that Black Africans are less likely to be employed than other races. As implied above, unemployment is correlated with poverty; Finn (2015) finds that just 4.1% of White South Africans were considered impoverished in 2015, compared to 20.5% of Indians/Asians, 56.8% of mixed-race, and 70.8% of Black South Africans, which correlates with the racial differences in employment described previously. Thus, increasing employment opportunities for those discriminated against under apartheid is also essential to addressing socioeconomic challenges that remain in the country.

¹Calculated from Figure 1.2 as follows: Gender gap in unemployment rate = $\frac{45.4 - 38.4}{38.4} * 100 = 18.1\%$.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Although I am not the first to examine the relationship between education and employment in South Africa, few have attempted to deal with the likely endogeneity of education. Regardless, the wider literature presents conflicting results. Some analyses find positive or negative effects of education on employment (Alam and Mamun, 2016; Bhorat and Kimani, 2017), while many others find little to no effect (Baldry, 2016; Anand et al., 2016; Liwiński, 2020; Torun, 2018). These differences in results likely arise from heterogeneity in context, data, methods and identification strategies (Alam and Mamun, 2016; Liwiński, 2020; Torun, 2018). Moreover, institutional diversity in both education systems and labour markets suggests that the causal effect of education on employment cannot easily be ascertained from research in other countries. Thus, country-specific research is warranted.

Given its apartheid past, South Africa offers an intriguing and unique setting for estimating the effect of education on employment. Similar to the international literature, the South African literature offers contrasting findings. Although Branson and Leibbrandt (2013b); Kingdon and Knight (2004); Festus et al. (2016); Bhorat and Kimani (2017) find that education reduces the likelihood of unemployment, Baldry (2016) and Anand et al. (2016) find little to no effect.

As suggested above, although there are many studies (Branson and Leibbrandt, 2013b; Kingdon and Knight, 2004; Festus et al., 2016; Bhorat and Kimani, 2017; Baldry, 2016; Anand et al., 2016), only Festus et al. (2016) attempt to address causality. Festus et al. (2016) used several exclusion restrictions, including the number of children, the number of adult males, adult females and elderly members in the household. They find that a post-secondary certificate/diploma increases the probability of employment by 4 to 8%, whilst a university degree increases the employment probability by 8 to 10%. Hence, further research is needed. In this analysis, I consider an instrument that is more commonly applied in the literature, a compulsory schooling law.

I find evidence that education is endogenous, which suggests that endogeneity bias is likely present in the majority of previous South African studies. The results show that compulsory schooling raised average educational attainment by between 0.04 and 0.18 years. However, the impact varies by gender: an additional year of compulsory schooling increased education by 0.04 to 0.20 years for males and 0.09 to 0.17 years for females.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

The analysis reveals that, on average, the marginal effect of an additional year of education increases the overall probability of employment by 7.3 to 28.4 percentage points, depending on the survey year. When disaggregated by gender, education is associated with a marginal effect of 3.0 to 18 percentage point increase in employment probability for women and a 3.0 to 13.9 percentage point increase for men. These results suggest that failing to account for unobserved individual heterogeneity may bias estimates of the returns to education on employment outcomes.

The analysis further indicates that the impact of education on employment outcomes varies across race and gender. Black African males' employment probabilities exceed those of their female counterparts, with the marginal effect of an additional year of education increasing male employment by 8.7 to 11 percentage points, compared to a smaller marginal effect of -3.7 to 8.5 percentage points for females. Similarly, for Whites, education tends to boost male employment probabilities (18.8 to 23.0 percentage points) more than it does for females (16.2 percentage points). On the other hand, for mixed race, the pattern is unclear.

The remaining sections are described as follows: The literature is reviewed in section 2.2, before the estimation strategy is discussed in section 2.3; the findings are presented and discussed in section 2.4 whilst the conclusion of the chapter is undertaken in section 2.5.

2.2 Literature Review

The empirical literature on the education-employment nexus is diverse and far from being unified. Whilst some studies find that education increases the probability of employment ([Castel et al., 2010](#); [Diawara et al., 2011](#); [Kingdon and Knight, 2004](#); [Berhe, 2021](#)), others find the opposite ([Baah-Boateng, 2015](#); [Kuepié and Nordman, 2016](#); [Baah-Boateng, 2013](#)) and, yet, some find no effect at all ([Baldry, 2016](#); [Liwiński, 2020](#); [Torun, 2018](#)). Such differences likely arise from differences in context, data, methodologies and identification strategies ([Alam and Mamun, 2016](#); [Liwiński, 2020](#); [Torun, 2018](#)). I provide a brief summary of this literature in Table 2.1.

The literature summarised in this table suggests that the effect of education on employment probability is heterogeneous across a number of dimensions, including the country, the level

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

of schooling considered and gender. [Abekah-Nkrumah et al. \(2019\)](#) find that females with an additional year of senior high school in Ghana were 6.6-8.5 percentage points more likely to be employed compared to only 3.8-4.3 percentage points for males. However, for paid employment [Abekah-Nkrumah et al. \(2019\)](#) find that the effect of education on the employment probability is statistically insignificant for the overall sample, and for both males and females, with estimates ranging from 1.5 to 2.5 percentage points for males and -0.3 to 1.1 percentage points for females.

Similarly, [Torun \(2018\)](#) finds statistically insignificant results for Turkey, with estimates of 1.4 and 0.6 percentage points for females and males, respectively. [Chuanchuan and Jingwen \(2021\)](#) examine the impact gender roles on women employment in China. They find that the more conventional/traditional the concept of gender roles, the lower the probability that women engage in employment.

In summary, a consensus is far from being realised. Several studies find contrasting estimates, both in South Africa and beyond. It is also evident that the majority of these studies do not control for the potential endogeneity of education, which may bias the reported estimates. While some of them attempt to control for endogeneity, they use different methodologies, identification strategies and data, which have yielded contrasting results; finally, some of the exclusion restrictions applied in the South African literature may not be appropriate. Hence, the causal link between education and employment outcomes is still unclear in many settings, including South Africa, which is my focus.

2.3 Identification Strategy

Estimating the impact of education on employment outcomes is challenging due to unobserved factors such as individual ability and motivation that may be correlated with both education and employment. Consequently, the estimated coefficient on education is likely to be biased and inconsistent ([Harmon et al., 2003](#)). To address this issue, I follow the existing literature by employing an instrumental variable technique ([Alam and Mamun, 2016](#); [Liwinski, 2020](#); [Torun, 2018](#)).

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

2.3.1 Instrument

An instrumental variable must satisfy two key properties: relevance and exogeneity (Schlotter et al., 2011). The relevance condition requires that the instrument be correlated with the treatment (endogenous) variable. This condition is crucial to avoid issues related to weak instruments (Bound et al., 1995). The exogeneity condition requires that the instrument affects the outcome only indirectly—through its effect on the endogenous variable. Empirically, this property is much harder to verify, so researchers often rely on theoretical arguments to justify the validity of an instrument.

I use the 1997 Compulsory Schooling Law (CSL) as an instrument for education. This law made schooling mandatory for all learners from the age of seven until the last school day of the year in which they turn fifteen or complete Grade 9, whichever occurs first (Republic of South Africa, 1996b). Furthermore, the Basic Education Laws Amendment Act (BELA) of 2024 extended compulsory education by making Grade R (the reception year) mandatory for six-year-old children (Republic of South Africa, 2024). Nevertheless, this study focuses on the 1997 law, since it is the policy change that directly impacted the individuals in the sample.

The CSL came into effect on 1 January 1997 and was implemented nationwide. As a result, all individuals who were 15 years old or younger on that date were subject to compulsory schooling, while those older than 15 were not. This implies that individuals born on or after 1 January 1982 were exposed to the law, whereas those born before that date were not.

Meanwhile individuals who were 7 years old and younger were exposed to 9 years of compulsory schooling while those who were between 8 and 15 years old on the implementation date were exposed to some years of compulsory schooling, with the actual number depending on their age at the time. Accordingly, the instrumental variable is constructed as the number of years of exposure to the CSL. In the sensitivity analysis, I define education to be a binary measure. Thus, for example, I consider the effect of Grade 9 completion on employment probabilities, using the same indicator of exposure to the CSL.

Importantly, the law plausibly pushed students to stay in school longer than they might have,

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

otherwise, i.e., there are compliers in the data. Several studies ([Korwatanasakul, 2023](#); [Clay et al., 2021](#); [Oyelere, 2010](#)) find a positive relationship between compulsory schooling and years of education, suggesting that there is correlation between the law and schooling completion. The results, below, support this literature, and suggest that the instrument is not weak.

In addition to compliers, the data most likely also includes always takers (those who would have completed Grade 9 and beyond, regardless of the law). The data might also include never takers (those who would not continue with education through Grade 9 or age 15). While the existence of defiers (individuals who reduce their schooling in response to the law) is theoretically possible under the LATE framework, it is highly implausible in the context of compulsory schooling laws, which are legally binding and designed to enforce school attendance. In the sensitivity analysis, I offer some insight into these possibilities.

Exogenous variation in schooling can arise from exposure to the CSL, as individuals do not choose their date of birth and, consequently, cannot self-select into exposure to the law. Moreover, since the CSL affected individuals during childhood, it is unlikely to have a direct impact on adult employment outcomes ([Acemoglu and Angrist, 2000](#); [Riddell and Song, 2011](#)), suggesting that the exclusion restriction is plausibly satisfied. However, given the numerous institutional and legislative changes that accompanied South Africa's transition to democracy (see Chapter 1), I acknowledge that the CSL may not be entirely exogenous.

2.3.2 Model Specification

The goal of this research is to investigate the causal relationship between education and employment, focusing on South Africa. Hence, I do not include labour force non-participants in the analysis. To do so, I employ 2SLS. Identification in my model is underpinned by conditionally exogenous variation in education that is created by the CSL. Thus, I assume that the law does not directly influence employment (exclusion restriction); however, the law is expected to impact education, the latter of which is assumed to impact employment.

The equation of interest is equation (2.1), where Y_i is the dependent variable denoting employment status (= 1, if employed and 0, otherwise), S_i denotes the years of completed education, and

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

X_i represents observed confounders, i.e, variables that are observed and correlated with both schooling and employment. These include age cohort (based on the year of birth), gender, race, marital status, province of residence, survey year, and the number of children under the age of 6 in the household. In some follow-up regressions, I disaggregate by gender and/or race.

$$Y_i = X_i' \delta + S_i \gamma + \epsilon_i \quad (2.1)$$

However, I am concerned with likely endogeneity in schooling completion – there may be unobserved factors that influence both schooling and employment, such as ability, punctuality or work ethic. [Alam and Mamun \(2016\)](#) show that education and employment simultaneously cause each other, finding that high unemployment lowers the opportunity cost of education, thereby encouraging individuals to stay longer in school. Thus, I need an instrument for education, which I project apart from its correlation with the preceding potential confounders, as in equation (2.2), where the previous variables are as noted above and Z is the instrument, which I discuss below.

$$S_i = Z_i \pi + X_i' \Psi + v_i \quad (2.2)$$

2.3.3 Data

The analysis of the effect of education on the probability of employment is based on data obtained from the South African General Household Survey (GHS) for the years 2014-2018 ([Statistics South Africa, 2017a, 2016, 2017b, 2018, 2019](#)). The GHS is a comprehensive survey conducted by Statistics South Africa (Stats SA) to collect data on various socioeconomic aspects of individuals and households in South Africa. It is conducted annually and is designed to gather data representative of the entire country and covers both urban and rural areas.

The GHS data is collected by trained interviewers through face-to-face interviews using structured questionnaires. In this analysis five cross-sections of the survey (2014-2018) are used to form a large enough dataset to allow for reliable estimates across subgroups. The survey covers a wide

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

range of topics, including information about household composition, age, gender, educational attainment, employment status, labour market earnings, and poverty indicators.

There are two reasons for using the GHS over other surveys such as the Quarterly Labour Force Survey (QLFS) and the National Income Dynamic Study (NIDS). The first reason is that the QLFS does not have information on date of birth, which is necessary for this analysis. The second is that there is some evidence of attrition bias in the NIDS dataset, particularly when using the labour market data for men and Africans (Baigrie and Eyal, 2014).

The dependent variable, employment status, is a binary variable defined to include all individuals aged 21 to 65 who were employed in the 30 days prior to the survey whilst unemployment includes all jobless (whether actively looking or not) individuals aged 21 to 65 years in the 30 days prior to the survey. This is a common definition of employment in the literature (Baah-Boateng, 2015). The variable comes from the survey question, which asked individuals to state whether in the last 30 days before the survey they were employed or not. Similar binary outcomes have been used by Alam and Mamun (2016); Berhe (2021); Baah-Boateng (2015), to name a few.

In terms of education, several studies, both local and international, use years of education as a measure of education. These include Bhorat and Kimani (2017) and Liwiński (2020) for South Africa and Poland, respectively. According to human capital and signalling (screening) theories, education is expected to have a positive effect on employment probability. However, as noted above, some empirical studies find a negative relationship between years of education and employment outcomes, while others find statistically insignificant effects (Baah-Boateng, 2015; Torun, 2018).

Gender, a binary variable (= 1, for males, = 0 for females), is included to capture differences in labour market outcomes that have been observed previously. For example, Bhorat and Kimani (2017) find that males are 9 percentage points more likely to be employed than females. Race is another important variable in the context of South Africa (Lam, 2016), with Kingdon and Knight (2004) finding that a married Black African male is 35 percentage points more likely to enter unemployment than his White counterpart. Therefore, the analysis includes dummies for all four

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

race groups in South Africa, namely, Black African, Coloured (mixed race), Indian/Asian and White. In the sensitivity analysis, I split the samples for each group.

Several studies, both local and international, find a positive relationship between marriage and employment outcomes ([Abekah-Nkrumah et al., 2019](#); [Bardasi and Taylor, 2008](#); [Depken et al., 2019](#)). Thus, I include it in the model, capturing two categories, legally married (1) and not legally married (0).

Geographic location might also influence labour market outcomes. [Kingdon and Knight \(2004\)](#) find that province of residency and urban/rural status are correlated with unemployment incidence in South Africa. Similarly, [Baah-Boateng \(2015\)](#) finds a positive correlation between urbanization and unemployment risk in Ghana. Thus, the analysis includes province of residency.

The literature also suggests a relationship between family structure and unemployment ([Naude and Serumaga-Zake, 2001](#); [Viljoen and Dunga, 2013](#)), while [Markussen and Strøm \(2022\)](#) find that the number of children reduces the labour market earnings and labour supply of women, while increasing the earnings and labour supply of men. Similarly, [Vu et al. \(2021\)](#) find that the number of children has a negative impact on maternal participation in the labour market and on working hours in Vietnam. Therefore, I include the number of children under the age of 6 in my regressions.

Finally, youth employment outcomes in South Africa have been weak. [Branson et al. \(2013\)](#), for example, find that younger cohorts face worse labour market conditions than the older generations. Presumably, much of this arises from differences in economic conditions following on from the global financial crisis – South Africa’s economy has not, yet, recovered from that ([Bosch and Koch, 2020](#)) – as well as the labour market environment associated with the end of apartheid and the difficulties associated with absorbing workers who were not previously allowed to work in many parts of the economy. I address this by including dummy variables for each birth-year. I choose this over some other functional form, due to the non-linearities in the relationships – see Figures 2.1 to 2.2, which illustrate smoothed relationships across birth year cohorts. [Torun \(2018\)](#) for Turkey, [Boockmann and Steiner \(2006\)](#) for West Germany, and [Aakvik et al. \(2010\)](#) for

Norway, amongst others, similarly include birth cohorts.

2.4 Results and Discussion

I present results descriptively, with a specific focus on the age pattern of education and employment. I then turn to standard OLS and 2SLS regression results, where the instrument is the number of years exposed to the CSL. I complete my analysis by examining the sensitivity of the results to both the nature of the instrument and measure of completed schooling. Specifically, I use exposure to the CSL, which was designed to influence Grade 9 completion rates, as the instrument and examine the returns to Grade 9 completion. I extend that to further years of grade completion; even though the policy was designed for Grade 9, it is plausible that completing Grade 9 further encouraged some learners to stay in school longer. For both the years of exposure and binary exposure instrument, the analysis is split by race and gender, to examine heterogeneity across these dimensions.

2.4.1 Descriptive statistics

I begin with a brief description of the data; see Table 2.1. It is presented for the overall sample and by gender. The sample covers 5 years of data and includes 231,630 individuals of which 53.2% are females and 46.8% are males. Employed individuals have significantly more years of education, with an average difference of 1.41 years, which is statistically significant. For women and men, the differences are 1.55 years and 1.33 years, respectively. Higher educational attainment, including having completed Grade 9, 10, 11, 12, or tertiary education, is more common among the employed. However, employed individuals have lower exposure to CSL, on average, which is expected given that CSL predominantly affected younger cohorts.

Employed individuals are also significantly older than unemployed individuals, by about 7 to 8 years on average. This pattern holds for both men (+8.65 years) and women (+6.44 years). Younger cohorts (born after 1980) are overrepresented among the unemployed, while older cohorts (born before 1980) are more likely to be employed, reflecting typical life-cycle employment trends.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table 2.1: Summary statistics for the overall, female and male samples

	All individuals			Women			Men		
	Unempl	Empl	Diff	Unempl	Empl	Diff	Unempl	Empl	Diff
	Mean	Mean	b	Mean	Mean	b	Mean	Mean	b
Age in years	32.40	39.60	-7.20***	33.81	40.25	-6.44***	30.40	39.05	-8.65***
Education (years)	9.24	10.65	-1.41***	9.30	10.85	-1.55***	9.15	10.48	-1.33***
Compulsory schooling	5.28	2.50	2.78***	4.86	2.31	2.56***	5.87	2.66	3.20***
At Least Grade 9	0.70	0.81	-0.11***	0.71	0.82	-0.11***	0.68	0.80	-0.12***
At Least Grade 10	0.56	0.74	-0.18***	0.59	0.76	-0.18***	0.53	0.72	-0.19***
At Least Grade 11	0.41	0.64	-0.23***	0.43	0.67	-0.23***	0.37	0.61	-0.23***
At Least Grade 12	0.26	0.51	-0.25***	0.28	0.54	-0.27***	0.24	0.48	-0.25***
Tertiary	0.04	0.18	-0.14***	0.05	0.21	-0.16***	0.04	0.16	-0.12***
African/Black	0.86	0.76	0.10***	0.85	0.76	0.09***	0.87	0.76	0.10***
Coloured	0.09	0.11	-0.02***	0.10	0.12	-0.02***	0.09	0.11	-0.02***
Indian/Asian	0.02	0.03	-0.01***	0.02	0.02	-0.01***	0.01	0.03	-0.02***
White	0.03	0.09	-0.06***	0.03	0.10	-0.06***	0.03	0.09	-0.07***
Female	0.59	0.46	0.13***	1.00	1.00	0.00	0.00	0.00	0.00
Male	0.41	0.54	-0.13***	0.00	0.00	0.00	1.00	1.00	0.00
Legally married	0.17	0.37	-0.20***	0.21	0.33	-0.12***	0.12	0.40	-0.28***
Metropolitan	0.31	0.44	-0.12***	0.31	0.43	-0.12***	0.31	0.44	-0.12***
Cohort 1950–59	0.10	0.08	0.03***	0.11	0.08	0.04***	0.09	0.07	0.01***
Cohort 1960–69	0.11	0.21	-0.10***	0.12	0.23	-0.11***	0.08	0.19	-0.11***
Cohort 1970–79	0.12	0.28	-0.16***	0.13	0.28	-0.15***	0.10	0.27	-0.17***
Cohort 1980–89	0.20	0.32	-0.12***	0.21	0.30	-0.09***	0.18	0.33	-0.15***
Cohort 1990–99	0.40	0.13	0.28***	0.36	0.11	0.25***	0.46	0.14	0.32***
Cohort 2000–03	0.08	0.00	0.07***	0.06	0.00	0.06***	0.09	0.00	0.09***
Children Under 6	0.70	0.44	0.26***	0.81	0.51	0.30***	0.53	0.38	0.15***
Observations	130712	100918	231630	76717	46469	123186	53995	54449	108444

Notes: Unempl stands for unemployed while Empl denote employed. Diff is difference, which is the difference in means between the unemployed and employed groups. b is the coefficient of the t test. ***, **, and * indicate significance at the 1, 5, and 10% critical level. Compulsory schooling represents years of exposure to the compulsory schooling law.

In terms of race, African/Black individuals make up a larger share of the unemployed (86%) compared to the employed (76%), as shown in Chapter 1 (Figure 1.1). Coloured and White individuals are slightly overrepresented among the employed, with minor differences observed for the Indian/Asian group. Females are more likely to be unemployed (59% of the unemployed sample), as highlighted in Chapter 1 (Figure 1.3).

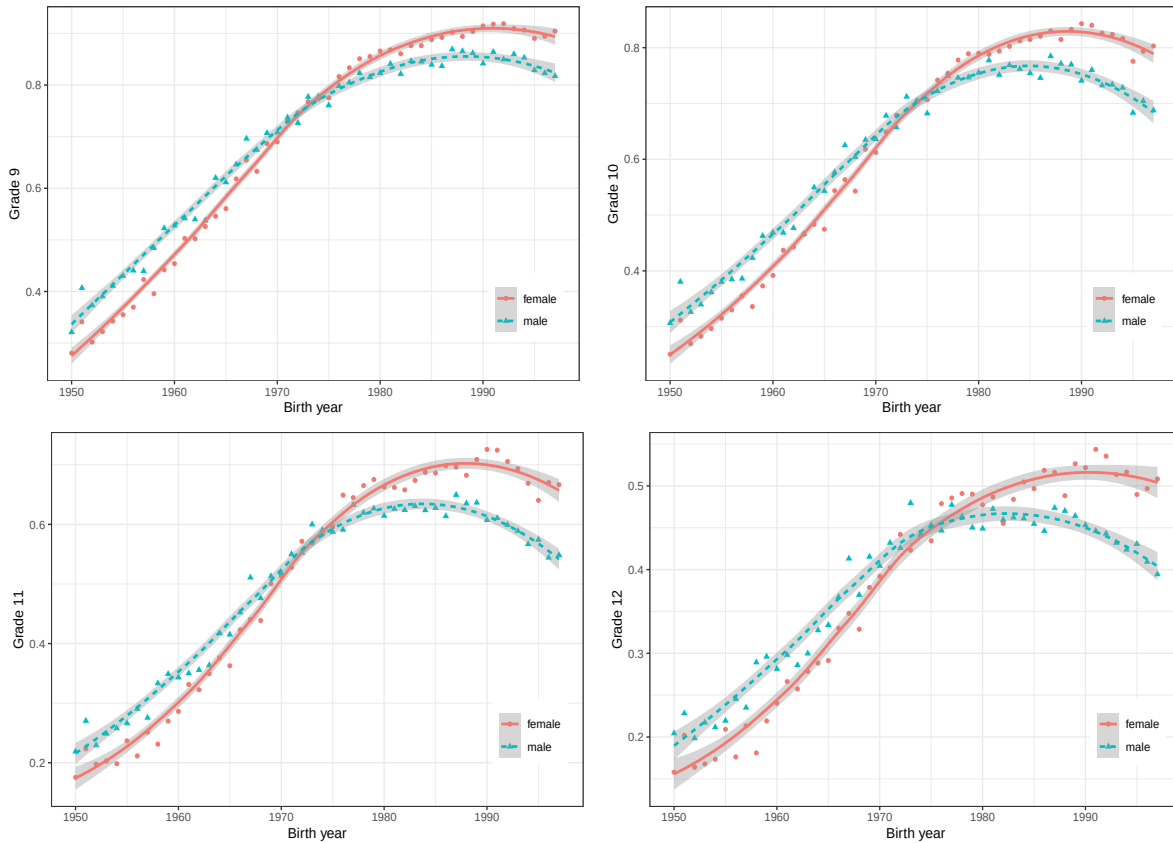
Being legally married and living in a metropolitan area is positively associated with employment. The employment difference by marital status is +20 percentage points overall, +12 percentage points for women, and +28 percentage points for men. The difference for metropolitan residence is +12 percentage points. Having young children is negatively associated with employment for both men and women.

Overall, employment is positively associated with age, education, legal marriage, and metropolitan residence, while it is negatively associated with younger age cohorts, exposure to CSL, and having young children. These patterns are broadly consistent across both men and women,

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

though the magnitudes differ slightly.

Figure 2.1: Proportion of individuals with a certain level of education by gender and year of birth



Estimates indicate that between 58% and 60% of young people drop out of the schooling system after the compulsory schooling phase (Branson and Leibbrandt, 2013b), leading to low educational attainment (Case and Deaton, 1999; Case and Yogo, 1999; Branson and Leibbrandt, 2013a). Low attainment might also arise from the government's over-age policy (Branson and Leibbrandt, 2013b; Burger et al., 2009). Specifically, learners that are two years or more than their appropriate grade-age were not guaranteed entry. Burger et al. (2015) estimate that the number of learners affected by this policy between 1999 and 2003 could have exceeded 800,000.

Despite these issues, educational attainment improved across birth-year cohorts, on average, up until the early to mid-1990s cohorts, as depicted in Figure 2.1. This Figure outlines completion rates across the different grade levels from the 1950 to 1997 birth-year cohorts. The proportion of males who have attained at least a Grade 9 schooling level has increased from 32% among the

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

oldest (born 1950) to 87% among the youngest (born 1997) cohort. The corresponding figures for females are 28% and 93%. For all grade levels, there has been a similar reversing of the education completion gap. For cohorts born between 1970 and 1975, completion rates for women began to exceed that for men, and that gap has widened further, since.

To further examine heterogeneity in the general improvements in educational attainment discussed above and depicted in Figure 2.1, I look at educational attainment across population groups in Figure 2.2. The results remind us of the advantages associated with apartheid. Educational attainment amongst Whites has remained fairly stable across birth cohorts. Education completion amongst Indians/Asians has reached par with that of Whites, and happened about the same time as the gender reversal in education completion rate gaps. However, education completion rates for Black African and mixed race (Coloured) individuals is well below that of the other two population groups, despite extensive improvement up until the 1990 cohort for grades 9-10 and approximately the 1985 cohort for grades 11-12. Amongst Black African and Coloured population groups, respectively, Grade 9 completion rates increased from 16% and 33% for the 1950 cohort to its peak of 91% and 90% for the 1996 cohort.

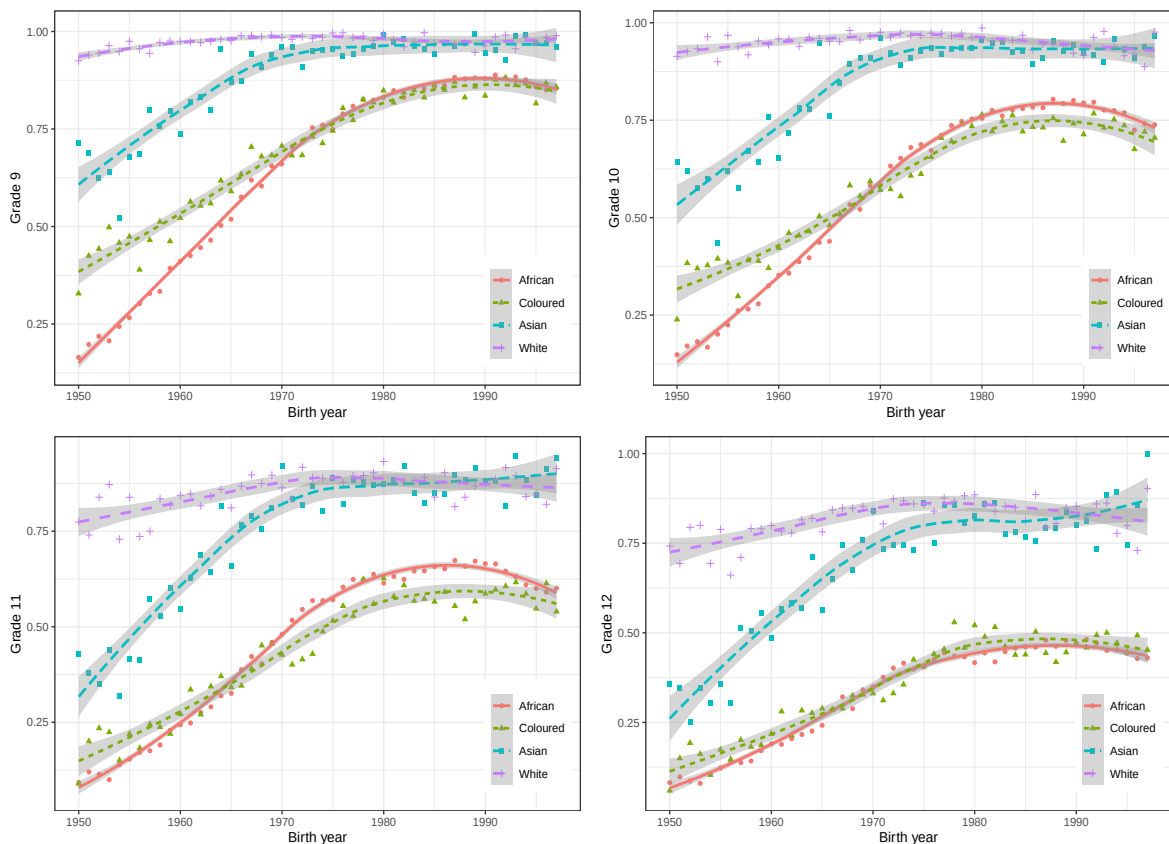
Given the improvements for Black Africans and the stability amongst Whites, it is not surprising that the White-Black educational attainment gap has narrowed over time. The 1950 cohort Grade 9 completion rate gap was 76 percentage points, narrowing to 10 percentage points for the 1997 cohort.

Educational attainment is positively correlated with employment, as expected and depicted in Figure 2.3. Similarly, employment follows the life-cycle, being highest amongst those born between 1960 and 1980, aged approximately 35 to 55, and is lower for the youth and the elderly.² For example, employment rates for the 1950 cohort are 11%, 15%, 38%, and 41% for primary/less, incomplete secondary, secondary, and tertiary education, respectively. For the 1997 cohort, employment rates are 30%, 12%, 19%, and 22%. We also see that the gap in employment rates between education levels is narrower for younger than older cohorts; similarly, employment

²Given that the data is from 2014-2018, at least some share of the 1950 cohort is beyond South Africa's mandatory retirement age. Such individuals are not included in the analysis.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Figure 2.2: Proportion of individuals attaining a certain level of education by race and year of birth

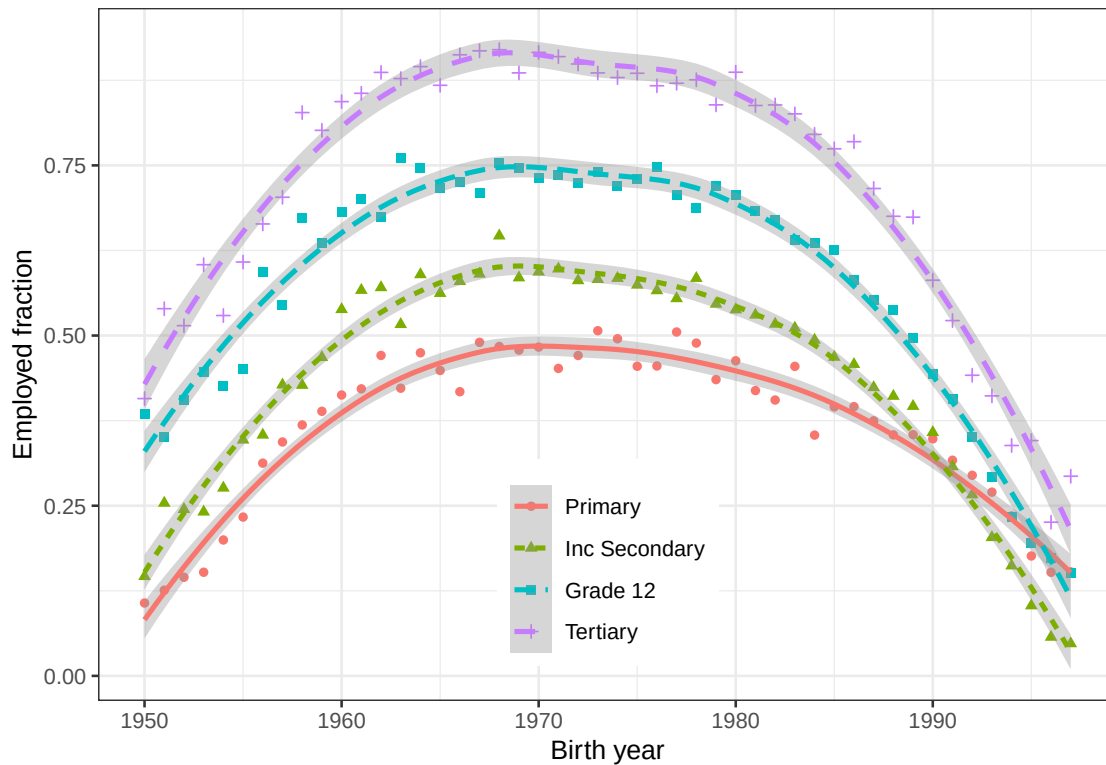


rates amongst the poorest educated are not too dissimilar to those who have completed primary and some secondary schooling.

The life-cycle pattern is observed across race groups, as well. See Figure 2.4. Similar to what we see for population group education gaps, there is an observable population group employment gap. Whites, elderly Whites in particular, the most advantaged under apartheid have higher rates of employment, relative to the other population groups. That advantage narrows across population groups, for the younger cohorts; however, young Black Africans have lower rates of employment than the other three population groups. For the 1950 cohort, employment rates for Black Africans is 14%, while that for Whites is 39%, representing a 25 point employment gap; for the 1997 cohort, the gap has increased to 35 points; White employment rates for the 1997 cohort stand at 48%, while that for Black Africans stands at only 13%. Regardless of whether we look at the level of education or population group, there is strong evidence that labour market

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Figure 2.3: Fraction of individuals that are employed by level of education and year of birth



performance is relatively poor for the youngest cohorts. As we will see in the analysis, since it is the younger cohorts that were exposed to the CSL, the law's employment impact is muted.

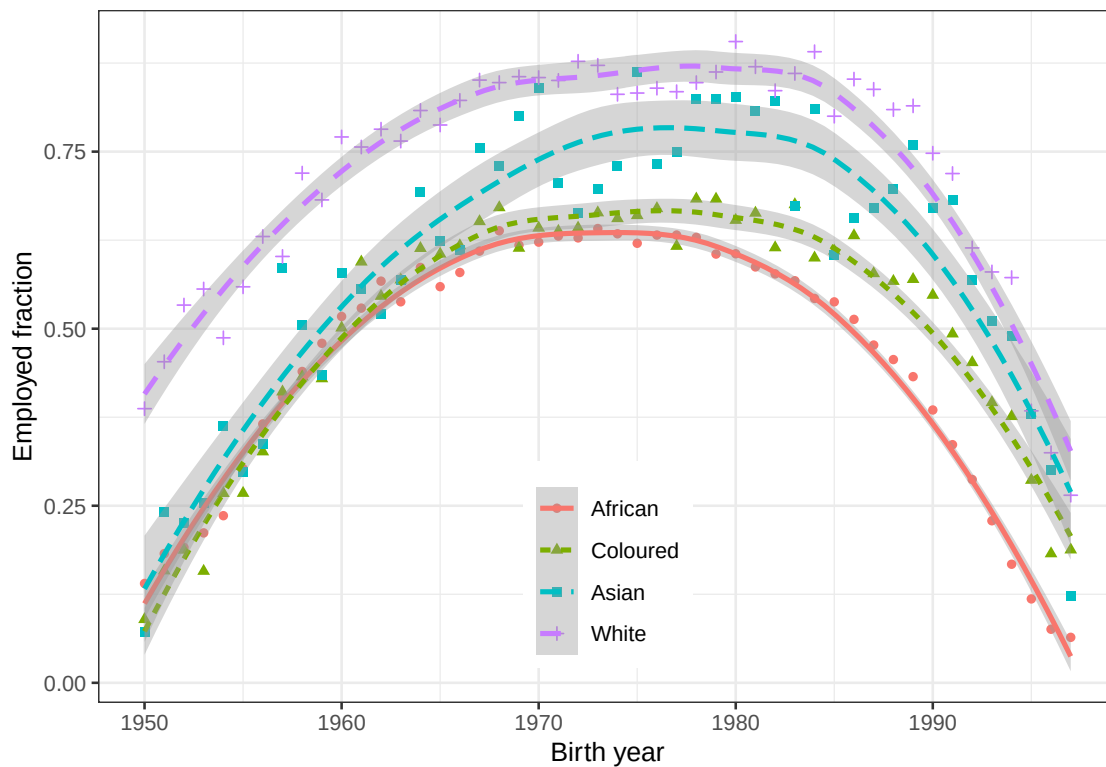
2.4.2 Regression results

The main results are presented in Table 2.2, which is separated into OLS and 2SLS. As expected – see Figure 2.3 and the OLS results – education is positively correlated with an increased employment probability. Further, that correlation is heterogeneous across gender, being larger for females than for males. However, these estimates are likely confounded by unobserved individual information. The Wu-Hausman test rejects the hypothesis that education is exogenous, and, therefore, the 2SLS results control for that possible endogeneity.

As previously mentioned, the years of exposure to the CSL is used as an instrument to control for the endogeneity of education. The coefficient for compulsory schooling is positive and statistically significant for the overall, female and male samples throughout the study period except in 2018

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Figure 2.4: Fraction of individuals that are employed by race and year of birth



for females and males. Thus, each year of exposure to the CSL increases education completion by between 0.035 and 0.196 years. Further statistical tests suggest that the instrument is not weak for most of the years (First stage F test in Table 2.2).

Having established that compulsory schooling is a relevant and strong instrument for years of completed education, the analysis proceeds to the second stage, where I estimate the causal effect of education on employment. The results suggest that education increases the probability of employment except for the overall and female samples in 2014 and male sample in 2017 and 2018 where the coefficient on education is statistically insignificant or where the instrument is weak (males in 2017). For the overall sample, an additional year of education increases the probability of employment by between 7.3 and 28.4 percentage points, on average. Among women, an additional year of education increases employment while for men an extra year of education raises the probability of employment by between 3.0 and 13.9 percentage points. [Abekah-Nkrumah et al. \(2019\)](#), [Boahen et al. \(2021\)](#) and [Drucker et al. \(2022\)](#) find similar results.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table 2.2: OLS and 2SLS estimates of the effect of education on employment probability

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): OLS Estimates						
Education	0.023*** (0.001)	0.023*** (0.001)	0.023*** (0.001)	0.022*** (0.001)	0.022*** (0.001)	0.022*** (0.001)
Female	0.029*** (0.001)	0.030*** (0.001)	0.031*** (0.001)	0.030*** (0.001)	0.031*** (0.001)	0.030*** (0.000)
Male	0.017*** (0.001)	0.017*** (0.001)	0.016*** (0.001)	0.016*** (0.001)	0.015*** (0.001)	0.016*** (0.001)
Panel (a): 2SLS Estimates for the Overall Sample						
Second stage Education	0.009 (0.008)	0.073*** (0.012)	0.113*** (0.020)	0.209*** (0.043)	0.284*** (0.089)	0.064*** (0.006)
First stage Compulsory schooling	0.184***	0.131***	0.088***	0.059***	0.035***	0.111***
F test	382.64 [0.000]	166.30 [0.000]	75.53 [0.000]	31.33 [0.000]	11.42 [0.001]	655.21 [0.000]
Wu-Hausman endogeneity test	4.011 [0.045]	17.525 [0.000]	25.308 [0.000]	48.711 [0.000]	35.561 [0.001]	46.203 [0.000]
Panel (b): 2SLS Estimates for the Female Sample						
Second stage Education	-0.012 (0.012)	0.030* (0.016)	0.101*** (0.025)	0.076*** (0.027)	0.180*** (0.059)	0.035*** (0.008)
First stage Compulsory schooling	0.173***	0.131***	0.097***	0.085***	0.052	0.119***
F test	179.29 [0.000]	86.33 [0.000]	47.21 [0.000]	34.26 [0.000]	13.34 [0.000]	393.36 [0.000]
Wu-Hausman endogeneity test	12.887 [0.000]	0.002 [0.963]	9.594 [0.002]	3.118 [0.077]	12.252 [0.001]	0.464 [0.496]
Panel (c): 2SLS Estimates for the Male Sample						
Second stage Education	0.030*** (0.010)	0.121*** (0.020)	0.139*** (0.035)	0.536** (0.228)	0.598 (0.443)	0.104*** (0.010)
First stage Compulsory schooling	0.196***	0.132***	0.083***	0.037**	0.021	0.106***
F test	205.55 [0.000]	82.18 [0.000]	31.71 [0.000]	5.61 [0.018]	1.84 [0.175]	281.05 [0.000]
Wu-Hausman endogeneity test	1.478 [0.224]	40.813 [0.000]	21.380 [0.000]	72.288 [0.000]	29.155 [0.000]	101.749 [0.000]

Notes: Reported coefficients represent marginal effects from linear probability models. All equations include the following control variables: age, age square, number of children in the household that are under 6 years old, sex where applicable, race (Black African, Coloured, Asian/Indian, and White with Black African serving as the base category), quarter of birth, birth cohort, metropolitan status, and marital status. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

I further estimate the effect of education on employment probability across race and gender; see Tables 2.3, 2.4 and 2.5. I consider only three population groups, due to the small sample size for Indians/Asians. As was the case for the preceding analysis, I find that education is endogenous across race and gender. Similar to the previous analysis, the CSL is found to increase schooling completion among Black Africans across all years for females, and in all years except 2018 for males. Compulsory schooling had a greater impact on males than on females in 2014 and 2015. However, this pattern reverses in 2016 and 2017, with females experiencing larger gains in those years. Nevertheless, the instrument is weak for males in 2017 and both males and females in 2018, as indicated by an F-statistic below the conventional threshold of 10.

For females, education increased the likelihood of employment by 8.5 and 5.0 percentage points

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table 2.3: 2SLS estimates of the effect of education on employment probability of Black Africans

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): Female Sample						
Second stage						
Education	-0.037*** (0.014)	0.010 (0.019)	0.085*** (0.029)	0.050* (0.029)	0.210** (0.082)	0.020** (0.009)
First stage						
Compulsory schooling	0.167*** (0.014)	0.122*** (0.015)	0.086*** (0.016)	0.083*** (0.016)	0.045*** (0.016)	0.114*** (0.007)
F test	135.35 [0.000]	62.38 [0.000]	30.19 [0.000]	27.30 [0.000]	8.18 [0.004]	299.58 [0.000]
Wu-Hausman endogeneity test	26.836 [0.000]	0.971 [0.324]	4.106 [0.043]	0.545 [0.461]	11.318 [0.001]	1.038 [0.308]
Panel (b): Male Sample						
Second stage						
Education	0.005 (0.011)	0.087*** (0.021)	0.110*** (0.038)	0.380** (0.159)	0.528 (0.416)	0.071*** (0.010)
First stage						
Compulsory schooling	0.203*** (0.015)	0.127*** (0.016)	0.075*** (0.016)	0.041** (0.017)	0.022 (0.017)	0.107*** (0.007)
F test	177.96 [0.000]	61.53 [0.000]	21.47 [0.000]	6.10 [0.014]	1.66 [0.198]	238.32 [0.000]
Wu-Hausman endogeneity test	1.089 [0.297]	14.26 [0.000]	8.710 [0.003]	37.902 [0.000]	20.387 [0.000]	35.294 [0.000]

Notes: Reported coefficients represent marginal effects from linear probability models. All equations include the following control variables: age, age square, number of children in the household that are under 6 years old, quarter of birth, birth cohort, metropolitan status, and marital status. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

in 2016 and 2017, respectively. However, in 2014 it reduced the probability of employment by 3.7 percentage points. Similarly, an additional year of schooling increased the employment probability of men by 8.7 and 11.0 percentage points in 2015 and 2016, respectively. Among the mixed race (Coloureds) compulsory schooling increased the years of education among females in 2015. However, in 2014 the exogeneity hypothesis is not reject while in the rest of the years, compulsory schooling is weak. Hence, in 2015, an additional year of education increased the employment probability of Coloured women by 20.4 percentage points. In contrast, among Coloured males, the instrument is weak across all years.

For Whites, the first stage results are not entirely surprising. As we see in Figure 2.2, there is evidence that education for Whites has been fairly constant particularly since the 1970s. Thus, compulsory schooling increases education only in 2014 and 2016 for females and 2014-2016 for males. However, the instrument is weak in 2015 for males and 2016 for females. The effect of education on employment is also heterogeneous, increasing employment in 2014 by 16.2 and 18.8 percentage points for White females and males, respectively. In 2016, the employment probability of males further increased to 23.0 percentage points.

The results suggest that the effect of education on employment probability differs within and

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table 2.4: 2SLS estimates of the effect of education on employment probability of Coloureds

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): Female Sample						
Second stage						
Education	0.106** (0.043)	0.204*** (0.069)	0.678 (0.844)	0.230 (0.203)	0.264 (0.345)	0.162*** (0.042)
First stage						
Compulsory schooling	0.154*** (0.035)	0.148*** (0.042)	0.031 (0.040)	0.057 (0.045)	0.036 (0.043)	0.090*** (0.017)
F test	18.77 [0.000]	11.75 [0.000]	0.62 [0.432]	1.61 [0.205]	0.67 [0.412]	26.36 [0.000]
Wu-Hausman endogeneity test	2.459 [0.1169]	10.245 [0.001]	7.522 [0.006]	2.092 [0.148]	1.069 [0.3012]	12.674 [0.000]
Panel (b): Male Sample						
Second stage						
Education	0.589 (0.435)	0.451** (0.220)	1.220 (3.615)	-0.473 (0.390)	-0.168 (0.145)	1.737 (2.319)
First stage						
Compulsory schooling	0.056 (0.042)	0.098** (0.047)	0.016 (0.049)	-0.066 (0.050)	-0.085* (0.049)	0.015 (0.020)
F test	1.77 [0.183]	4.29 [0.039]	0.11 [0.739]	1.74 [0.187]	2.95 [0.0858]	0.55 [0.4584]
Wu-Hausman endogeneity test	24.130 [0.000]	31.239 [0.000]	6.359 [0.012]	17.229 [0.000]	4.548 [0.033]	65.521 [0.000]

Notes: Reported coefficients represent marginal effects from linear probability models. All equations include the following control variables: age, age square, number of children in the household that are under 6 years old, quarter of birth, birth cohort, metropolitan status, and marital status. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

between population groups. For Black Africans, males have a higher probability of being employed compared to females, although the estimate is statistically insignificant for some years. Similarly, for Whites, males have a higher probability of employment than females (2014), although the comparability is difficult due to the weak instrument problem for most of the years.

Table 2.5: 2SLS estimates of the effect of education employment probability of Whites

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): Female Sample						
Second stage						
Education	0.162*** (0.052)	0.960 (3.400)	0.354** (0.151)	0.545 (0.570)	0.265 (0.200)	0.223*** (0.054)
First stage						
Compulsory schooling	0.215*** (0.048)	0.016 (0.057)	0.121** (0.052)	0.054 (0.059)	0.080 (0.057)	0.112*** (0.023)
F test	20.38 [0.000]	0.07 [0.785]	5.47 [0.020]	0.84 [0.358]	1.96 [0.162]	22.68 [0.000]
Wu-Hausman endogeneity test	7.482 [0.006]	1.849 [0.1740]	12.324 [0.000]	6.516 [0.011]	2.877 [0.090]	19.615 [0.000]
Panel (b): Male Sample						
Second stage						
Education	0.188*** (0.057)	0.290** (0.114)	0.230*** (0.080)	1.411 (2.722)	0.651 (1.066)	0.307*** (0.067)
First stage						
Compulsory schooling	0.194*** (0.049)	0.143*** (0.052)	0.171*** (0.055)	0.033 (0.065)	0.035 (0.058)	0.113*** (0.024)
F test	15.41 [0.000]	7.53 [0.006]	9.56 [0.002]	0.26 [0.610]	0.36 [0.548]	22.58 [0.000]
Wu-Hausman endogeneity test	18.482 [0.000]	21.726 [0.000]	15.408 [0.000]	32.505 [0.000]	6.654 [0.010]	85.514 [0.000]

Notes: Reported coefficients represent marginal effects from linear probability models. All equations include the following control variables: age, age square, number of children in the household that are under 6 years old, quarter of birth, birth cohort, metropolitan status, and marital status. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

The observed gender and racial differences in the returns to education even after controlling for endogeneity likely reflect systemic inequalities, labour market segmentation, and persistent social and economic barriers. The instrument may not fully capture these complex dynamics, so some unobserved variables remain in the error term, explaining why education appears beneficial depending on one's group identity. More research is needed to better understand these underlying mechanisms and explore these relationships in greater detail.

2.4.3 Sensitivity analysis

The preceding analysis yielded positive education effects on employment for the overall sample, on average. In this section, I look deeper into the education mechanism, exploring the effect of completing a particular level of education on employment. I continue with the previous instrument. I present the overall results for the period 2014-2018 and pooled estimates in Table 2.6.

As before, there is strong evidence of endogeneity, justifying the use of an instrumental variable approach. At the same time, the first-stage results show that the instrument is strongly correlated with educational attainment (except for Grade 9 in 2017 and 2018), alleviating concerns about weak instruments. Furthermore, I continue to find that, on average, the CSL increased schooling at all levels, not just Grade 9, as the law primarily intended. In terms of effect sizes, the CSL had a larger impact on grade 10, 11 and 12 completion, than it did on Grade 9 or tertiary completion; Grade 9 was between 0.3 and 4.2 percent more likely to be completed, while Grade 10 was 0.6 to 4.2 percentage points more likely to be completed, as a result of the CSL. As expected, single grade completion effects are lower than previously found for completed years of schooling (see Table 2.2, which is an aggregation across each of these levels.

The effect of grade completion on employment probability, on the other hand, is larger for each grade completed than was estimated for years of schooling. Yet, I still find that increased schooling raises employment, overall, amongst those impacted by the law. On average, completing Grade 9, increases employment by approximately between 65 and 172 percentage points, overall, in 2015 and 2016, respectively. The 2017 and 2018 estimates are implausible due to the weak instrument

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table 2.6: 2SLS estimates of the effect of grade completion on employment probability

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): 2SLS Estimates: Grade 9						
Second stage Grade 9	0.061 (0.055)	0.653*** (0.118)	1.724*** (0.471)	4.751* (2.534)	3.991* (2.219)	0.585*** (0.059)
First stage Compulsory schooling	0.026*** (0.001)	0.015*** (0.001)	0.006*** (0.001)	0.003* (0.001)	0.003* (0.001)	0.012*** (0.001)
F test	407.12 [0.000]	112.16 [0.000]	17.73 [0.000]	3.58 [0.058]	3.36 [0.067]	432.90 [0.000]
Wu-Hausman endogeneity test	0.681 [0.409]	27.343 [0.000]	35.904 [0.000]	58.306 [0.058]	37.453 [0.067]	80.312 [0.000]
Panel (b): 2SLS Estimates: Grade 10						
Second stage Grade 10	0.037 (0.034)	0.336*** (0.056)	0.694*** (0.125)	1.651*** (0.386)	1.577*** (0.440)	0.348*** (0.033)
First stage Compulsory schooling	0.042*** (0.001)	0.028*** (0.002)	0.014*** (0.002)	0.008*** (0.002)	0.006*** (0.002)	0.020*** (0.001)
F test	907.59 [0.000]	350.53 [0.000]	89.83 [0.000]	23.28 [0.000]	16.92 [0.000]	988.70 [0.000]
Wu-Hausman endogeneity test	4.745 [0.029]	18.235 [0.000]	28.999 [0.000]	54.269 [0.000]	34.819 [0.000]	59.077 [0.000]
Panel (c): 2SLS Estimates: Grade 11						
Second stage Grade 11	0.033 (0.030)	0.251*** (0.041)	0.402*** (0.066)	0.850*** (0.137)	1.000*** (0.215)	0.264*** (0.025)
First stage Compulsory schooling	0.047*** (0.001)	0.038*** (0.002)	0.025*** (0.002)	0.015*** (0.002)	0.010*** (0.002)	0.027*** (0.001)
F test	1009.60 [0.000]	556.52 [0.000]	230.67 [0.000]	77.51 [0.000]	36.44 [0.000]	1513.28 [0.000]
Wu-Hausman endogeneity test	10.960 [0.001]	9.885 [0.002]	19.484 [0.000]	45.256 [0.000]	30.477 [0.000]	32.995 [0.000]
Panel (d): 2SLS Estimates: Grade 12						
Second stage Grade 12	0.038 (0.034)	0.284*** (0.046)	0.331*** (0.053)	0.564*** (0.078)	0.781*** (0.149)	0.264*** (0.025)
First stage Compulsory schooling	0.042*** (0.001)	0.034*** (0.002)	0.030*** (0.002)	0.022*** (0.002)	0.013*** (0.002)	0.027*** (0.001)
F test	818.32 [0.000]	448.18 [0.000]	342.88 [0.000]	174.81 [0.000]	58.55 [0.000]	1535.07 [0.000]
Wu-Hausman endogeneity test	11.544 [0.001]	8.883 [0.003]	12.307 [0.001]	33.105 [0.000]	26.177 [0.000]	21.877 [0.000]
Panel (e): 2SLS Estimates: Tertiary						
Second stage Tertiary	0.301 (0.271)	1.514*** (0.331)	1.546*** (0.333)	1.431*** (0.244)	1.418*** (0.302)	1.062*** (0.115)
First stage Compulsory schooling	0.005*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.009*** (0.001)	0.007*** (0.001)	0.007*** (0.000)
F test	28.62 [0.000]	33.91 [0.000]	33.97 [0.000]	56.09 [0.000]	36.20 [0.000]	201.81 [0.000]
Wu-Hausman endogeneity test	0.150 [0.699]	28.726 [0.000]	30.628 [0.000]	44.928 [0.000]	29.284 [0.000]	75.870 [0.000]

Notes: Reported coefficients represent marginal effects from linear probability models. All equations include the following control variables: age, age square, number of children in the household that are under 6 years old, sex where applicable, race (Black African, Coloured, Asian/Indian, and White with Black African serving as the base category), quarter of birth, birth cohort, metropolitan status, and marital status. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

bias.

For the completion of Grade 10, the first stage estimate are larger than those reported for Grade 9, while the second stage estimates are smaller. Similarly, for grades 11 and 12, the pattern mostly

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

continues, with larger first stage effects and smaller employment effects. The pattern changes at the tertiary level, although it should be remembered that completing tertiary education is equivalent (in South Africa) to completing three additional school years; thus, dividing the estimates by 3 would yield smaller effects of education on employment per year.

2.4.4 Discussion

One potential concern is that the estimated effects of compulsory schooling may be confounded by contemporaneous changes in labour legislation and broader socio-political reforms. The early 1990s, particularly around the transition to democracy in 1994, marked a period of significant legal and institutional reform in South Africa, including amendments to labour laws aimed at promoting equality, improving worker protection, and addressing historical labour market imbalances. These changes, alongside affirmative action policies and the introduction of the Labour Relations Act (1995) and Basic Conditions of Employment Act (1997), could plausibly influence employment outcomes independently of educational attainment.

Furthermore, it is conceivable that the impact of the compulsory schooling law on labour market outcomes is intertwined with the socio-economic restructuring of the post-apartheid period. However, as noted in Chapter 1, despite these substantial reforms, racial and gender disparities in both educational attainment and labour market participation have persisted well into the post-1994 period. This persistence suggests that the broader legislative and political changes may have had a limited direct impact on the structural inequalities that underlie educational and employment outcomes.

While it is difficult to fully disentangle the effects of CSL from other concurrent policy changes, the enduring disparities observed across cohorts implies that the estimated effects in this study are unlikely to be driven solely by confounding policy shifts. Nevertheless, I acknowledge this as an important limitation and therefore, the results must be interpreted with caution in light of these shortcomings.

Another concern is that the compulsory schooling law may have induced general equilibrium effects in the labour market, potentially confounding the estimated impact of education. Specifi-

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

cally, a significant increase in the supply of educated workers, without a matching rise in labour demand, could lower employment probabilities. Given the relatively stagnant labour demand in recent decades, it is plausible that such general equilibrium effects have dampened the observed returns to education in this context.

The signaling and screening theories suggest that education plays a signaling and screening role rather than a productivity enhancer as believed by the human capital theorists. The fact that both the human capital and the signalling theories assume that education is associated with labour market success makes it rather difficult to determine empirically whether the estimated employment returns to education are more influenced by signaling or skill development mechanisms (Clark and Martorell, 2014; Huntington-Klein, 2021). However, empirical evidence tends to support the human capital theory more than it does signaling and screening theories (Oreopoulos and Salvanes, 2011). Kaymak (2025) who investigates the human capital and signalling roles of education in the labour market finds that the role of signalling is small at 23% compared to the human capital role.

Although the human capital and signaling and screening theories predict a positive association between education and employment probability, previous empirical analyses find mixed results as noted in the literature review. This analysis is consistent with the theory. Abekah-Nkrumah et al. (2019), Boahen et al. (2021) and Drucker et al. (2022) find similar results. The former find that an extra year in senior high school increases the likelihood of employment by between 5.3 and 5.6 percentage points. Similarly, Boahen et al. (2021) estimate that education increases the probability of employment by between 7.8 and 8.2 percentage points while Drucker et al. (2022) find an estimate of 3 percentage points.

I find that the effects of education on employment is heterogeneous across race and gender. Torun (2018) and Abekah-Nkrumah et al. (2019) report similar results. Torun (2018) finds that education increases the probability of employment by 0.6 and 1.4 percentage points for men and women, respectively, although the estimates are statistically insignificant. Abekah-Nkrumah et al. (2019) find small marginal effects ranging from 1.5 to 2.5 percentage points for males and -0.3 to 1.1 for females, also statistically insignificant. Abekah-Nkrumah et al. (2019) attribute the

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

gender gaps in employment probability in Ghana to the socially imposed roles of men and women (Abekah-Nkrumah et al., 2019; Kercheval et al., 2013).

There is limited previous research attempting to address the endogeneity of education when estimating the importance of education on employment in South Africa. Only Festus et al. (2016) attempt to do so. They used several exclusion restrictions, including the number of children, the number of adult males, adult females and elderly members in the household. For these to be valid exclusion restrictions, such individuals in the household need to affect employment – indirectly and only through schooling. Given that household formation is likely endogenous³, this exclusion restriction is not likely to be valid.

Even if those exclusion restrictions are valid, they are likely to affect behaviour for a rather different set of compliers than the CSL. Therefore, the exclusion restriction is more likely to hold in my analysis, than in the previously mentioned study. Further, since the exclusion restriction likely affects a different subset of the population, my analysis offers a slightly different focus than previous studies.

2.5 Conclusion

The human capital and signalling theories suggest that people with higher educational attainment are more likely to be employed than those with lower educational achievement. However, in South Africa the unemployment rate has remained stubbornly high despite the expansion in education over the years. The government implemented the compulsory schooling policy in 1997 to ensure that children remain in school up until the 9th grade or their 15th birthday. However, the high dropout rates in grades 9 to 12 and government policies have limited educational attainment, leaving many young people with incomplete secondary schooling.

This study investigates the extent to which education is useful in increasing employment opportunities for both males and females. Although this question has been investigated previously,

³For example, Wilcox and Sahni (2023) finds that living with an older family member who requires assistance significantly reduces younger women's working hours and likelihood of full-time employment, while simultaneously increasing these labour market outcomes among younger men.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

both in South Africa and beyond, the literature is far from unified. Whilst some studies find that education increases the probability of employment, others find the opposite and yet others find no effect at all. Notably, the majority of them fail to control for the endogeneity of education, which biases their estimates. Whilst some control for endogeneity, they use different methodologies and identification strategies, which have yielded contrasting results.

Therefore, this study contributes to the literature by examining the impact of education on employment probability using the compulsory schooling policy for exogenous variation in education to control for endogeneity. The analysis applied two-stage least squares to assess the impact of education on employment outcomes for male and female South Africans, and further separated the results to different population groups and different levels of education. The data was obtained from the South African General Household Survey for the years 2014-2018.

The findings indicate that compulsory schooling increased education by between 0.04 and 0.18 years, on average. However, the effect of compulsory schooling on education is heterogeneous across gender. An extra year of compulsory schooling induces between 0.04 and 0.20 years of education for males and between 0.09 and 0.17 years for females. In addition, the results indicate that education is endogenous, suggesting the need for an instrument when analysing the relationship between education and employment.

The analysis finds that an additional year of education increases the probability of employment by between 7.3 and 28.4 percentage points overall, on average. Moreover, whilst education increases the probability of employment for women by between 3 and 18 percentage points, on average, it raises the employment probability of men by between 3.0 and 13.9 percentage points. The current study suggests that failing to control for unobserved individual heterogeneity may produce biased estimates of education returns to employment.

Given the numerous institutional and legislative changes in both the education system and the labour market that coincided with the implementation of the compulsory schooling policy, it is difficult to fully disentangle the effects of the CSL from other concurrent policy shifts. However, the persistent disparities observed across cohorts suggest that the estimated effects in this study

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

are unlikely to be driven solely by such confounding factors. Nevertheless, I acknowledge this as an important limitation, and the results should therefore be interpreted with caution.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Appendix A

Table A1: Years of education derived from highest education level

Highest education level	Code	Years of education
No schooling	98	0
Grade R/0	0	0
Grade 1/Sub A/Class 1	1	1
Grade 2/Sub B/Class 2	2	2
Grade 3/Standard 1/ABET/AET 1	3	3
Grade 4/Standard 2	4	4
Grade 5/Standard 3/ABET/AET 2	5	5
Grade 6/Standard 4	6	6
Grade 7/Standard 5/ABET/AET 3	7	7
Grade 8/Standard 6/Form 1	8	8
Grade 9/Standard 7/Form 2/NCV Level 1/Occupational Certificate-NQF Level 1	9	9
Grade 10/Standard 8/Form 3/NCV Level 2/Occupational Certificate-NQF Level 2	10	10
Grade 11/Standard 9/Form 4/NCV Level 3/Occupational Certificate-NQF Level 3	11	11
Grade 12/Standard 10/Form 5/ NCV Level 4/Occupational Certificate-NQF Level 4	12	12
NTC I/N1	13	9
NTC II/N2	14	10
NTC III/N3	15	11
N4/NTC 4/Occupational Certificate-NQF Level 5	16	13
N5/NTC 5/Occupational Certificate-NQF Level 5	17	13
N6/NTC 6/Occupational Certificate-NQF Level 5	18	13
Certificate with less than Grade 12/Standard 10	19	12
Diploma with less than Grade 12/Standard 10	20	12
Higher/National/Advance certificate with Grade 12/Occup. Certif.-NQF Level 5	21	13
Diploma with Grade 12/Standard 10/Occupational Certificate-NQF Level 6	22	14
Higher Diploma/Occupational Certificate (B-Tech Diploma)-NQF Level 7	23	15
Post Higher Diploma (Masters Diploma and Masters Degree)-NQF Level 9	24	18
Bachelors Degree/Occupational Certificate-NQF Level 7	25	15
Honours Degree/Postgraduate Diploma/Occupational Certificate-NQF Level 8	26	16
Doctoral Degrees (Doctoral Diploma and PhD)-NQF Level 10	27	21

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A2: Summary of the literature on the effect of education on employment probability

No.	Author	Country	Methods	Education Slope/Findings
1	Alam and Mamun (2016)	Australia	Simultaneous system of two-equations. Instruments: Whether attended a government school or not; and whether attended a university/technical and further college institute or not.	Education reduces the probability of employment by 7.6, 7.5 and 0.42 percentage points for all individuals, males and females, respectively.
2	Abekah-Nkrumah et al. (2019)	Ghana	Probit regression. IV: whether attended a 4-year or 3-year senior high school (SHS) system.	An extra year in SHS increases the likelihood of formal employment by between 5.3 and 5.6 percentage points, but has no effect on the likelihood of being in paid employment.
3	Baah-Boateng (2013)	Ghana	Probit regressions.	Basic and secondary education increases unemployment by 0.02 and 0.11 percentage points, respectively, whilst tertiary education is negative and statistically insignificant.
4	Baah-Boateng (2015)	Ghana	Logistic regressions.	An addition year of education increases unemployment probability by between 0.04 and 0.85 percentage points.
5	Berhe (2021)	Ethiopia	Logistic regression	The odds ratio of being unemployed increases by a factor of 1.44 and 0.73 for urban youth with high school and vocation training/diploma, respectively compared to youth with primary education.
6	Boahen et al. (2021)	Ghana	Fuzzy regression discontinuity design (RDD). IV: Pre-university education duration	Education increases the probability of employment by between 7.8 and 8.2 percentage points.
7	Bhorat and Kimani (2017)	South Africa	Heckprobit. Exclusion variable: presence of a social grant holder in a household.	Graduating from a university increases the employment probability by 7–10 percentage points compared to graduating from a college.
8	Branson and Leibbrandt (2013b)	South Africa	Linear Probability Model (LPM).	Males and females with tertiary education are 10 and 20 percentage points likely to be employed, respectively, compared to those with primary education.
9	Branson et al. (2013)	South Africa	Linear Probability Model (LPM).	Younger cohorts more likely to find employment than older cohorts.
10	Castel et al. (2010)	Malawi	Probit regressions.	Being literate without a degree is likely to reduce unemployment by 0.26 and having a high school certificate is likely to increase unemployment by 0.05 percentage points.
11	Diawara et al. (2011)	Senegal	Probit regressions.	University graduates are 11 percentage points more likely to be employed compared to primary.
12	Drucker et al. (2022)	Poland	DiD using quasi-panel data. IV: expansion of secondary education	Education increases the probability of employment by 3 percentage points.
13	Festus et al. (2016)	South Africa	Heckman 2-stage model with number of children, adult male, adult female and elderly members in the household as exclusion restriction variables.	A post-secondary certificate/diploma increases the probability of employment by 4 to 8 compared to 8 to 10 percentage points for a university degree.
14	Hall (2016)	Sweden	Fixed effect (Variation in municipalities and over time). IV: Extended vocational education	Education has a statistically insignificant effect on employment outcomes.
16	Kuepié and Nordman (2016)	DRC	Logistic regression.	Education increases the unemployment incidence by 7.4 to 13.8 percentage points in Brazzaville, and 5.7 to 10.3 percentage points in Pointe Norte.
17	Liwiński (2020)	Poland	Fuzzy RDD. IV: compulsory schooling	Education has a statistically insignificant effect on employment probability.
18	Torun (2018)	Turkey	2SLS. IV: compulsory schooling	Education has a statistically insignificant effect on the probability of employment of females and males.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A3: OLS estimates of the effect of education on employment probability for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Education	0.023*** (0.001)	0.023*** (0.001)	0.023*** (0.001)	0.022*** (0.001)	0.022*** (0.001)	0.022*** (0.001)
Age in years	0.069*** (0.002)	0.075*** (0.002)	0.084*** (0.002)	0.086*** (0.002)	0.083*** (0.003)	0.083*** (0.003)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Children Under 6	-0.044*** (0.002)	-0.056*** (0.003)	-0.052*** (0.003)	-0.062*** (0.003)	-0.060*** (0.003)	-0.060*** (0.003)
Coloured	0.049*** (0.007)	0.051*** (0.008)	0.040*** (0.008)	0.057*** (0.008)	0.039*** (0.008)	0.039*** (0.008)
Indian/Asian	0.055*** (0.015)	0.015 (0.015)	0.021 (0.016)	0.040** (0.017)	-0.018 (0.017)	-0.018 (0.017)
White	0.113*** (0.009)	0.110*** (0.009)	0.134*** (0.009)	0.121*** (0.009)	0.111*** (0.010)	0.111*** (0.010)
Quarter 1	0.004 (0.007)	0.004 (0.006)	0.006 (0.006)	0.004 (0.006)	-0.002 (0.006)	-0.002 (0.006)
Quarter 3	0.003 (0.007)	-0.002 (0.006)	-0.000 (0.006)	0.008 (0.006)	-0.009 (0.006)	-0.009 (0.006)
Quarter 4	0.004 (0.007)	0.004 (0.006)	0.004 (0.006)	0.003 (0.006)	-0.005 (0.007)	-0.005 (0.007)
Legally married	0.078*** (0.006)	0.066*** (0.006)	0.064*** (0.006)	0.083*** (0.006)	0.095*** (0.006)	0.095*** (0.006)
Metropolitan	0.044*** (0.005)	0.046*** (0.005)	0.050*** (0.005)	0.041*** (0.005)	0.044*** (0.005)	0.044*** (0.005)
Cohort 1950–59	0.146*** (0.036)	-0.060 (0.038)	-0.117*** (0.039)	-0.086** (0.040)	-0.154*** (0.041)	-0.154*** (0.041)
Cohort 1960–69	0.115*** (0.027)	-0.072** (0.028)	-0.115*** (0.029)	-0.100*** (0.030)	-0.098*** (0.032)	-0.098*** (0.032)
Cohort 1970–79	0.088*** (0.020)	-0.096*** (0.022)	-0.162*** (0.023)	-0.142*** (0.024)	-0.119*** (0.025)	-0.119*** (0.025)
Cohort 1980–89	0.095*** (0.013)	-0.055*** (0.017)	-0.103*** (0.017)	-0.089*** (0.018)	-0.087*** (0.019)	-0.087*** (0.019)
Cohort 1990–99	0.000 (.)	-0.115*** (0.008)	-0.119*** (0.007)	-0.099*** (0.008)	-0.071*** (0.009)	-0.071*** (0.009)
Constant	-1.139*** (0.032)	-1.096*** (0.027)	-1.278*** (0.030)	-1.319*** (0.032)	-1.313*** (0.036)	-1.313*** (0.036)
Observations	54910	45217	43937	44206	43360	43360

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A4: OLS estimates of the effect of education on employment probability for the female sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Education	0.029*** (0.001)	0.030*** (0.001)	0.031*** (0.001)	0.030*** (0.001)	0.031*** (0.001)	0.030*** (0.000)
Age in years	0.061*** (0.003)	0.064*** (0.003)	0.073*** (0.003)	0.073*** (0.003)	0.074*** (0.004)	0.065*** (0.001)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Children Under 6	-0.036*** (0.003)	-0.049*** (0.004)	-0.045*** (0.003)	-0.054*** (0.003)	-0.054*** (0.004)	-0.047*** (0.001)
Coloured	0.068*** (0.010)	0.048*** (0.011)	0.050*** (0.011)	0.076*** (0.011)	0.050*** (0.012)	0.058*** (0.005)
Indian/Asian	0.027 (0.022)	-0.045** (0.023)	-0.023 (0.024)	-0.001 (0.026)	-0.079*** (0.026)	-0.025** (0.011)
White	0.125*** (0.014)	0.105*** (0.014)	0.140*** (0.014)	0.119*** (0.014)	0.106*** (0.014)	0.119*** (0.006)
Quarter 1	-0.003 (0.009)	-0.001 (0.008)	0.003 (0.009)	0.009 (0.009)	0.006 (0.009)	0.002 (0.004)
Quarter 3	0.006 (0.009)	-0.006 (0.008)	0.003 (0.009)	0.009 (0.009)	-0.010 (0.009)	-0.001 (0.004)
Quarter 4	0.000 (0.009)	-0.006 (0.009)	-0.004 (0.009)	0.008 (0.009)	-0.004 (0.009)	-0.002 (0.004)
Legally married	-0.017* (0.009)	-0.022*** (0.008)	-0.033*** (0.008)	-0.016* (0.008)	-0.009 (0.008)	-0.019*** (0.004)
Metropolitan	0.040*** (0.008)	0.047*** (0.007)	0.054*** (0.007)	0.042*** (0.007)	0.044*** (0.007)	0.046*** (0.003)
Cohort 1950–59	0.233*** (0.050)	-0.110** (0.052)	-0.161*** (0.054)	-0.110** (0.055)	-0.132** (0.057)	-0.015 (0.022)
Cohort 1960–69	0.209*** (0.037)	-0.071* (0.039)	-0.121*** (0.040)	-0.080* (0.041)	-0.071 (0.044)	0.019 (0.016)
Cohort 1970–79	0.120*** (0.027)	-0.094*** (0.030)	-0.160*** (0.032)	-0.109*** (0.032)	-0.101*** (0.035)	-0.023* (0.013)
Cohort 1980–89	0.094*** (0.017)	-0.083*** (0.022)	-0.133*** (0.023)	-0.098*** (0.024)	-0.107*** (0.026)	-0.028*** (0.009)
Cohort 1990–99	0.000 (.)	-0.120*** (0.009)	-0.136*** (0.010)	-0.117*** (0.011)	-0.090*** (0.013)	-0.070*** (0.004)
Constant	-1.096*** (0.044)	-1.032*** (0.035)	-1.209*** (0.040)	-1.229*** (0.043)	-1.279*** (0.049)	-1.124*** (0.017)
Observations	29431	23941	23344	23479	22991	123186

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A5: OLS estimates of the effect of education on employment for the male sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Education	0.017*** (0.001)	0.017*** (0.001)	0.016*** (0.001)	0.016*** (0.001)	0.015*** (0.001)	0.016*** (0.001)
Age in years	0.080*** (0.003)	0.088*** (0.003)	0.097*** (0.003)	0.101*** (0.004)	0.093*** (0.004)	0.087*** (0.001)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Children Under 6	-0.038*** (0.003)	-0.042*** (0.004)	-0.038*** (0.004)	-0.047*** (0.004)	-0.046*** (0.004)	-0.042*** (0.002)
Coloured	0.030*** (0.010)	0.059*** (0.011)	0.031*** (0.011)	0.039*** (0.012)	0.028** (0.011)	0.038*** (0.005)
Indian/Asian	0.079*** (0.020)	0.066*** (0.018)	0.062*** (0.020)	0.071*** (0.020)	0.033 (0.020)	0.061*** (0.009)
White	0.105*** (0.012)	0.123*** (0.011)	0.135*** (0.012)	0.128*** (0.012)	0.122*** (0.012)	0.123*** (0.005)
Quarter 1	0.013 (0.009)	0.003 (0.009)	0.005 (0.009)	-0.003 (0.009)	-0.008 (0.009)	0.001 (0.004)
Quarter 3	0.002 (0.009)	-0.000 (0.009)	-0.006 (0.009)	0.009 (0.009)	-0.010 (0.009)	-0.002 (0.004)
Quarter 4	0.012 (0.010)	0.013 (0.009)	0.010 (0.009)	-0.001 (0.009)	-0.005 (0.009)	0.005 (0.004)
Legally married	0.200*** (0.009)	0.170*** (0.008)	0.182*** (0.008)	0.201*** (0.008)	0.216*** (0.008)	0.193*** (0.004)
Metropolitan	0.040*** (0.007)	0.040*** (0.006)	0.041*** (0.007)	0.035*** (0.007)	0.038*** (0.007)	0.039*** (0.003)
Cohort 1950–59	0.093* (0.051)	0.029 (0.053)	-0.046 (0.054)	-0.049 (0.055)	-0.145** (0.057)	0.010 (0.022)
Cohort 1960–69	0.040 (0.037)	-0.056 (0.040)	-0.097* (0.040)	-0.112*** (0.042)	-0.097** (0.044)	-0.017 (0.017)
Cohort 1970–79	0.055** (0.028)	-0.098*** (0.031)	-0.167*** (0.032)	-0.181*** (0.033)	-0.130*** (0.035)	-0.051*** (0.013)
Cohort 1980–89	0.097*** (0.019)	-0.027 (0.024)	-0.070*** (0.024)	-0.081*** (0.026)	-0.056** (0.027)	0.020** (0.010)
Cohort 1990–99	0.000 (.)	-0.109*** (0.011)	-0.101*** (0.010)	-0.080*** (0.012)	-0.048*** (0.014)	-0.036*** (0.004)
Constant	-1.226*** (0.046)	-1.213*** (0.039)	-1.387*** (0.043)	-1.455*** (0.047)	-1.378*** (0.052)	-1.281*** (0.018)
Observations	25479	21276	20593	20727	20369	108444

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A6: 2SLS estimates of the effect of education on employment for the overall sample: second stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Education	0.009 (0.008)	0.073*** (0.012)	0.113*** (0.020)	0.209*** (0.043)	0.284*** (0.089)	0.064*** (0.006)
Age in years	0.076*** (0.004)	0.054*** (0.006)	0.052*** (0.008)	0.028** (0.014)	0.012 (0.025)	0.060*** (0.002)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.001*** (0.000)
Children Under 6	-0.039*** (0.003)	-0.031*** (0.004)	-0.020*** (0.005)	-0.001 (0.012)	0.024 (0.025)	-0.033*** (0.002)
Coloured	0.053*** (0.007)	0.051*** (0.007)	0.041*** (0.008)	0.073*** (0.012)	0.056*** (0.015)	0.049*** (0.003)
Indian/Asian	0.074*** (0.016)	-0.060*** (0.023)	-0.108*** (0.033)	-0.219*** (0.062)	-0.378*** (0.125)	-0.037*** (0.010)
White	0.155*** (0.022)	-0.010 (0.032)	-0.085* (0.051)	-0.370*** (0.113)	-0.557** (0.230)	0.016 (0.016)
Quarter 1	0.006 (0.005)	-0.001 (0.006)	0.004 (0.007)	-0.002 (0.009)	0.007 (0.012)	0.001 (0.003)
Quarter 3	0.007 (0.005)	-0.013** (0.006)	-0.015* (0.007)	-0.035*** (0.013)	-0.057*** (0.020)	-0.009*** (0.003)
Quarter 4	0.009 (0.005)	-0.004 (0.006)	-0.005 (0.007)	-0.025** (0.012)	-0.011 (0.012)	-0.004 (0.003)
Male	0.128*** (0.004)	0.154*** (0.005)	0.162*** (0.007)	0.192*** (0.014)	0.206*** (0.027)	0.147*** (0.002)
Legally married	0.089*** (0.008)	0.018 (0.012)	-0.021 (0.020)	-0.099** (0.042)	-0.161* (0.087)	0.038*** (0.006)
Metropolitan	0.058*** (0.009)	-0.009 (0.014)	-0.043** (0.022)	-0.153*** (0.045)	-0.216** (0.089)	0.002 (0.007)
Cohort 1950–59	0.141*** (0.029)	-0.070* (0.040)	-0.227*** (0.050)	-0.367*** (0.087)	-0.451*** (0.128)	0.000 (0.015)
Cohort 1960–69	0.098*** (0.023)	-0.049 (0.033)	-0.151*** (0.035)	-0.241*** (0.057)	-0.268*** (0.084)	0.025** (0.012)
Cohort 1970–79	0.068*** (0.017)	-0.100*** (0.028)	-0.214*** (0.030)	-0.300*** (0.052)	-0.316*** (0.083)	-0.026*** (0.010)
Cohort 1980–89	0.087*** (0.010)	-0.093*** (0.026)	-0.190*** (0.030)	-0.285*** (0.054)	-0.336*** (0.093)	-0.021*** (0.008)
Cohort 1990–99	0.000 (.)	-0.187*** (0.027)	-0.254*** (0.035)	-0.358*** (0.062)	-0.415*** (0.120)	-0.094*** (0.008)
Constant	-1.170*** (0.039)	-1.286*** (0.042)	-1.662*** (0.079)	-2.230*** (0.198)	-2.763*** (0.475)	-1.405*** (0.026)
Observations	54910	45217	43937	44206	43360	231630

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A7: 2SLS estimates of the effect of education on employment probability for the the overall sample: first stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Age in years	0.555*** (0.014)	0.497*** (0.015)	0.405*** (0.016)	0.339*** (0.017)	0.291*** (0.018)	0.415*** (0.006)
Age Square	-0.007*** (0.000)	-0.006*** (0.000)	-0.006*** (0.000)	-0.005*** (0.000)	-0.004*** (0.000)	-0.005*** (0.000)
Children Under 6	-0.228*** (0.012)	-0.268*** (0.015)	-0.232*** (0.015)	-0.265*** (0.016)	-0.278*** (0.016)	-0.255*** (0.007)
Coloured	0.126*** (0.043)	0.049 (0.047)	0.001 (0.047)	-0.077 (0.048)	-0.061 (0.048)	0.002 (0.021)
Indian/Asian	1.368*** (0.077)	1.494*** (0.084)	1.435*** (0.084)	1.379*** (0.085)	1.373*** (0.083)	1.409*** (0.037)
White	2.747*** (0.047)	2.469*** (0.052)	2.469*** (0.052)	2.643*** (0.052)	2.565*** (0.053)	2.578*** (0.023)
Quarter 1	0.090*** (0.034)	0.068* (0.037)	0.008 (0.037)	0.031 (0.038)	-0.030 (0.038)	0.039** (0.016)
Quarter 3	0.175*** (0.034)	0.223*** (0.037)	0.173*** (0.037)	0.242*** (0.038)	0.190*** (0.038)	0.216*** (0.016)
Quarter 4	0.198*** (0.035)	0.191*** (0.039)	0.125*** (0.038)	0.173*** (0.039)	0.039 (0.039)	0.165*** (0.017)
Male	-0.267*** (0.025)	-0.233*** (0.027)	-0.253*** (0.027)	-0.293*** (0.027)	-0.292*** (0.027)	-0.268*** (0.012)
Legally married	0.871*** (0.032)	0.927*** (0.034)	0.941*** (0.034)	0.967*** (0.035)	0.973*** (0.034)	0.930*** (0.015)
Metropolitan	1.042*** (0.025)	1.080*** (0.028)	1.038*** (0.027)	1.036*** (0.028)	0.988*** (0.028)	1.042*** (0.012)
Cohort 1950–59	-0.983*** (0.187)	-0.046 (0.251)	1.002*** (0.239)	1.307*** (0.246)	1.023*** (0.248)	-0.365*** (0.095)
Cohort 1960–69	-1.526*** (0.133)	-0.584*** (0.207)	0.295 (0.188)	0.650*** (0.193)	0.590*** (0.194)	-0.643*** (0.075)
Cohort 1970–79	-0.873*** (0.099)	0.125 (0.178)	0.627*** (0.156)	0.854*** (0.158)	0.758*** (0.157)	-0.021 (0.063)
Cohort 1980–89	-0.324*** (0.061)	0.555*** (0.152)	0.862*** (0.125)	0.953*** (0.126)	0.894*** (0.126)	0.362*** (0.050)
Cohort 1990–99	0.000 (.)	1.172*** (0.125)	1.307*** (0.089)	1.241*** (0.085)	1.215*** (0.083)	0.808*** (0.036)
Compulsory schooling	0.184*** (0.009)	0.131*** (0.010)	0.088*** (0.010)	0.059*** (0.011)	0.035*** (0.010)	0.111*** (0.004)
Constant	-0.332 (0.272)	-0.104 (0.271)	1.926*** (0.269)	3.458*** (0.286)	4.663*** (0.292)	1.831*** (0.110)
Observations	54910	45217	43937	44206	43360	231630

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A8: 2SLS estimates of the effect of education on employment probability for the female sample: second stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Education	-0.012 (0.012)	0.030* (0.016)	0.101*** (0.025)	0.076*** (0.027)	0.180*** (0.059)	0.035*** (0.008)
Age in years	0.079*** (0.006)	0.064*** (0.008)	0.048*** (0.010)	0.058*** (0.009)	0.031* (0.018)	0.063*** (0.003)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)	-0.001*** (0.000)	-0.000 (0.000)	-0.001*** (0.000)
Children Under 6	-0.046*** (0.004)	-0.049*** (0.006)	-0.027*** (0.007)	-0.039*** (0.009)	-0.008 (0.018)	-0.045*** (0.003)
Coloured	0.076*** (0.009)	0.047*** (0.010)	0.048*** (0.011)	0.080*** (0.010)	0.056*** (0.014)	0.058*** (0.004)
Indian/Asian	0.078*** (0.022)	-0.046 (0.031)	-0.119*** (0.039)	-0.062 (0.041)	-0.272*** (0.080)	-0.032** (0.013)
White	0.236*** (0.034)	0.103** (0.042)	-0.032 (0.062)	0.002 (0.070)	-0.272* (0.151)	0.106*** (0.020)
Quarter 1	0.002 (0.007)	-0.001 (0.008)	0.003 (0.009)	0.007 (0.008)	0.008 (0.011)	0.002 (0.003)
Quarter 3	0.013* (0.008)	-0.007 (0.009)	-0.011 (0.010)	-0.000 (0.010)	-0.038** (0.016)	-0.002 (0.004)
Quarter 4	0.008 (0.008)	-0.006 (0.008)	-0.014 (0.010)	0.002 (0.009)	-0.008 (0.011)	-0.003 (0.004)
Legally married	0.007 (0.009)	-0.022* (0.013)	-0.077*** (0.018)	-0.048** (0.020)	-0.115*** (0.043)	-0.022*** (0.006)
Metropolitan	0.085*** (0.014)	0.046** (0.019)	-0.019 (0.027)	-0.005 (0.028)	-0.103* (0.059)	0.040*** (0.009)
Cohort 1950–59	0.168*** (0.044)	-0.110** (0.052)	-0.245*** (0.062)	-0.193*** (0.072)	-0.320*** (0.103)	-0.013 (0.020)
Cohort 1960–69	0.122*** (0.038)	-0.070 (0.045)	-0.142*** (0.044)	-0.120** (0.048)	-0.154** (0.065)	0.023 (0.017)
Cohort 1970–79	0.056** (0.028)	-0.094** (0.038)	-0.197*** (0.039)	-0.149*** (0.041)	-0.212*** (0.064)	-0.022 (0.013)
Cohort 1980–89	0.070*** (0.015)	-0.083** (0.034)	-0.198*** (0.037)	-0.145*** (0.039)	-0.239*** (0.064)	-0.030*** (0.011)
Cohort 1990–99	0.000 (.)	-0.121*** (0.035)	-0.239*** (0.042)	-0.179*** (0.041)	-0.282*** (0.079)	-0.075*** (0.011)
Constant	-0.955*** (0.058)	-1.034*** (0.057)	-1.455*** (0.099)	-1.436*** (0.131)	-2.047*** (0.311)	-1.143*** (0.034)
Observations	29431	23941	23344	23479	22991	123186

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A9: 2SLS estimates of the effect of education on employment probability for the female sample: first stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Age in years	0.550*** (0.019)	0.508*** (0.021)	0.417*** (0.022)	0.369*** (0.023)	0.321*** (0.024)	0.426*** (0.009)
Age Square	-0.007*** (0.000)	-0.007*** (0.000)	-0.006*** (0.000)	-0.006*** (0.000)	-0.005*** (0.000)	-0.006*** (0.000)
Children Under 6	-0.268*** (0.016)	-0.298*** (0.019)	-0.266*** (0.020)	-0.322*** (0.021)	-0.308*** (0.020)	-0.292*** (0.008)
Coloured	0.198*** (0.058)	0.046 (0.064)	0.021 (0.064)	-0.070 (0.066)	-0.036 (0.065)	0.028 (0.028)
Indian/Asian	1.250*** (0.107)	1.519*** (0.119)	1.367*** (0.118)	1.339*** (0.120)	1.291*** (0.118)	1.352*** (0.052)
White	2.711*** (0.064)	2.442*** (0.072)	2.466*** (0.072)	2.568*** (0.072)	2.546*** (0.072)	2.547*** (0.031)
Quarter 1	0.134*** (0.046)	0.031 (0.051)	0.015 (0.051)	0.037 (0.052)	-0.009 (0.052)	0.047** (0.023)
Quarter 3	0.206*** (0.046)	0.256*** (0.051)	0.224*** (0.051)	0.212*** (0.052)	0.198*** (0.051)	0.233*** (0.023)
Quarter 4	0.221*** (0.047)	0.157*** (0.052)	0.167*** (0.052)	0.168*** (0.054)	0.043 (0.053)	0.168*** (0.023)
Legally married	0.594*** (0.041)	0.713*** (0.046)	0.654*** (0.045)	0.716*** (0.046)	0.715*** (0.046)	0.673*** (0.020)
Metropolitan	1.084*** (0.035)	1.082*** (0.038)	1.045*** (0.038)	1.024*** (0.039)	0.990*** (0.038)	1.050*** (0.017)
Cohort 1950–59	-1.781*** (0.248)	-0.324 (0.344)	0.860*** (0.326)	1.534*** (0.334)	1.047*** (0.335)	-0.564*** (0.128)
Cohort 1960–69	-2.122*** (0.180)	-0.880*** (0.286)	0.089 (0.260)	0.694*** (0.264)	0.419 (0.264)	-0.873*** (0.103)
Cohort 1970–79	-1.204*** (0.136)	-0.116 (0.248)	0.542** (0.216)	0.875*** (0.217)	0.729*** (0.216)	-0.125 (0.086)
Cohort 1980–89	-0.473*** (0.084)	0.400* (0.212)	0.764*** (0.173)	0.888*** (0.173)	0.788*** (0.172)	0.281*** (0.068)
Cohort 1990–99	0.000 (.)	1.127*** (0.176)	1.241*** (0.124)	1.146*** (0.117)	1.148*** (0.114)	0.777*** (0.050)
Compulsory schooling	0.173*** (0.013)	0.131*** (0.014)	0.097*** (0.014)	0.085*** (0.015)	0.052*** (0.014)	0.119*** (0.006)
Constant	-0.038 (0.370)	-0.019 (0.373)	1.865*** (0.370)	3.078*** (0.390)	4.261*** (0.400)	1.807*** (0.150)
Observations	29431	23941	23344	23479	22991	123186

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A10: 2SLS estimates of the effect of education on employment probability for the male sample: second stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Education	0.030*** (0.010)	0.121*** (0.020)	0.139*** (0.035)	0.536** (0.228)	0.598 (0.443)	0.104*** (0.010)
Age in years	0.075*** (0.005)	0.045*** (0.009)	0.054*** (0.013)	-0.052 (0.068)	-0.057 (0.115)	0.055*** (0.004)
Age Square	-0.001*** (0.000)	-0.000*** (0.000)	-0.000** (0.000)	0.001 (0.001)	0.001 (0.002)	-0.001*** (0.000)
Children Under 6	-0.035*** (0.004)	-0.015** (0.007)	-0.013 (0.009)	0.067 (0.052)	0.108 (0.118)	-0.021*** (0.003)
Coloured	0.030*** (0.009)	0.053*** (0.012)	0.033** (0.013)	0.081* (0.043)	0.075 (0.056)	0.040*** (0.005)
Indian/Asian	0.061*** (0.022)	-0.087** (0.036)	-0.124** (0.057)	-0.664** (0.329)	-0.814 (0.647)	-0.068*** (0.017)
White	0.071** (0.030)	-0.137*** (0.052)	-0.171** (0.087)	-1.290** (0.623)	-1.390 (1.149)	-0.107*** (0.027)
Quarter 1	0.012* (0.007)	-0.007 (0.010)	0.005 (0.010)	-0.019 (0.031)	0.021 (0.040)	-0.001 (0.004)
Quarter 3	0.001 (0.007)	-0.019* (0.010)	-0.018* (0.011)	-0.131* (0.068)	-0.111 (0.084)	-0.018*** (0.005)
Quarter 4	0.010 (0.008)	-0.008 (0.011)	0.003 (0.011)	-0.087* (0.049)	-0.017 (0.036)	-0.008* (0.005)
Legally married	0.186*** (0.013)	0.056** (0.024)	0.032 (0.043)	-0.416 (0.272)	-0.484 (0.533)	0.092*** (0.012)
Metropolitan	0.028** (0.011)	-0.072*** (0.023)	-0.085** (0.036)	-0.503** (0.237)	-0.534 (0.434)	-0.051*** (0.011)
Cohort 1950–59	0.090** (0.041)	-0.050 (0.068)	-0.226*** (0.084)	-0.680** (0.337)	-0.803 (0.545)	0.008 (0.024)
Cohort 1960–69	0.050* (0.030)	-0.059 (0.054)	-0.180*** (0.058)	-0.472** (0.218)	-0.591 (0.411)	0.016 (0.019)
Cohort 1970–79	0.066*** (0.023)	-0.136*** (0.047)	-0.251*** (0.049)	-0.609*** (0.225)	-0.585 (0.372)	-0.041*** (0.015)
Cohort 1980–89	0.101*** (0.013)	-0.125*** (0.043)	-0.204*** (0.051)	-0.651** (0.268)	-0.663 (0.473)	-0.023* (0.013)
Cohort 1990–99	0.000 (.)	-0.267*** (0.044)	-0.294*** (0.059)	-0.828** (0.334)	-0.831 (0.599)	-0.127*** (0.014)
Constant	-1.261*** (0.051)	-1.398*** (0.065)	-1.763*** (0.120)	-3.594*** (0.953)	-4.318* (2.241)	-1.561*** (0.039)
Observations	25479	21276	20593	20727	20369	108444

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A11: 2SLS estimates of the effect of education on employment probability for the male sample: first stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Age in years	0.571*** (0.021)	0.492*** (0.022)	0.402*** (0.023)	0.313*** (0.025)	0.267*** (0.026)	0.412*** (0.009)
Age Square	-0.008*** (0.000)	-0.006*** (0.000)	-0.006*** (0.000)	-0.005*** (0.000)	-0.004*** (0.000)	-0.005*** (0.000)
Children Under 6	-0.187*** (0.020)	-0.249*** (0.024)	-0.209*** (0.024)	-0.218*** (0.026)	-0.265*** (0.025)	-0.229*** (0.011)
Coloured	0.044 (0.064)	0.058 (0.069)	-0.017 (0.068)	-0.082 (0.071)	-0.082 (0.070)	-0.023 (0.031)
Indian/Asian	1.478*** (0.111)	1.481*** (0.119)	1.507*** (0.119)	1.413*** (0.120)	1.453*** (0.118)	1.466*** (0.052)
White	2.785*** (0.069)	2.503*** (0.075)	2.481*** (0.075)	2.724*** (0.076)	2.592*** (0.076)	2.615*** (0.033)
Quarter 1	0.048 (0.050)	0.108** (0.054)	0.001 (0.054)	0.032 (0.056)	-0.050 (0.055)	0.033 (0.024)
Quarter 3	0.138*** (0.050)	0.194*** (0.054)	0.116** (0.054)	0.274*** (0.056)	0.176*** (0.055)	0.197*** (0.024)
Quarter 4	0.174*** (0.052)	0.230*** (0.057)	0.074 (0.056)	0.174*** (0.058)	0.024 (0.057)	0.156*** (0.025)
Legally married	1.135*** (0.050)	1.116*** (0.053)	1.232*** (0.052)	1.189*** (0.054)	1.205*** (0.053)	1.170*** (0.023)
Metropolitan	0.987*** (0.037)	1.070*** (0.040)	1.017*** (0.040)	1.034*** (0.041)	0.979*** (0.041)	1.025*** (0.018)
Cohort 1950–59	-0.017 (0.283)	0.250 (0.368)	1.130*** (0.351)	1.067*** (0.363)	1.041*** (0.367)	-0.110 (0.140)
Cohort 1960–69	-0.858*** (0.198)	-0.301 (0.298)	0.471* (0.273)	0.599** (0.283)	0.793*** (0.284)	-0.388*** (0.109)
Cohort 1970–79	-0.539*** (0.144)	0.339 (0.255)	0.668*** (0.224)	0.819*** (0.229)	0.779*** (0.230)	0.074 (0.091)
Cohort 1980–89	-0.168* (0.090)	0.700*** (0.218)	0.944*** (0.181)	1.031*** (0.183)	1.004*** (0.184)	0.449*** (0.072)
Cohort 1990–99	0.000 (.)	1.204*** (0.179)	1.363*** (0.128)	1.341*** (0.123)	1.286*** (0.122)	0.838*** (0.052)
Compulsory schooling	0.196*** (0.014)	0.132*** (0.015)	0.083*** (0.015)	0.037** (0.015)	0.021 (0.015)	0.106*** (0.006)
Constant	-1.049*** (0.400)	-0.500 (0.394)	1.629*** (0.389)	3.477*** (0.420)	4.690*** (0.426)	1.499*** (0.160)
Observations	25479	21276	20593	20727	20369	108444

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A12: 2SLS estimates of the effect of education on the likelihood of employment for black women: second stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Education	-0.037*** (0.014)	0.010 (0.019)	0.085*** (0.029)	0.050* (0.029)	0.210** (0.082)	0.020** (0.009)
Age in years	0.084*** (0.007)	0.064*** (0.009)	0.043*** (0.012)	0.060*** (0.011)	0.011 (0.028)	0.062*** (0.004)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.000* (0.000)	-0.001*** (0.000)	0.000 (0.000)	-0.001*** (0.000)
Children Under 6	-0.053*** (0.004)	-0.055*** (0.006)	-0.032*** (0.008)	-0.049*** (0.010)	-0.002 (0.024)	-0.050*** (0.003)
Quarter 1	0.009 (0.009)	-0.000 (0.009)	0.004 (0.009)	0.018** (0.009)	0.013 (0.013)	0.007* (0.004)
Quarter 3	0.023*** (0.009)	-0.003 (0.010)	-0.014 (0.011)	0.010 (0.011)	-0.043** (0.022)	0.003 (0.004)
Quarter 4	0.020** (0.009)	-0.002 (0.010)	-0.009 (0.011)	0.008 (0.010)	-0.012 (0.014)	0.003 (0.004)
Legally married	0.024** (0.011)	-0.004 (0.015)	-0.057*** (0.019)	-0.027 (0.021)	-0.125** (0.058)	-0.008 (0.006)
Metropolitan	0.127*** (0.017)	0.076*** (0.022)	0.005 (0.032)	0.033 (0.031)	-0.124 (0.084)	0.067*** (0.010)
Cohort 1950–59	0.208*** (0.050)	-0.071 (0.058)	-0.227*** (0.070)	-0.146* (0.081)	-0.447*** (0.156)	-0.009 (0.022)
Cohort 1960–69	0.141*** (0.043)	-0.038 (0.048)	-0.117** (0.047)	-0.085* (0.051)	-0.233** (0.092)	0.027 (0.019)
Cohort 1970–79	0.071** (0.030)	-0.049 (0.041)	-0.163*** (0.042)	-0.114*** (0.044)	-0.273*** (0.093)	-0.007 (0.014)
Cohort 1980–89	0.072*** (0.017)	-0.048 (0.036)	-0.165*** (0.039)	-0.120*** (0.038)	-0.272*** (0.082)	-0.023** (0.011)
Cohort 1990–99	0.000 (.)	-0.094** (0.038)	-0.215*** (0.045)	-0.158*** (0.040)	-0.331*** (0.104)	-0.072*** (0.011)
Constant	-0.774*** (0.066)	-0.867*** (0.062)	-1.273*** (0.109)	-1.228*** (0.133)	-2.092*** (0.410)	-0.994*** (0.036)
Observations	23543	19609	19259	19423	19081	100915

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

Table A13: 2SLS estimates of the effect of education on the likelihood of employment for black women: first stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Age in years	0.570*** (0.021)	0.530*** (0.023)	0.445*** (0.024)	0.403*** (0.026)	0.353*** (0.026)	0.450*** (0.010)
Age Square	-0.008*** (0.000)	-0.007*** (0.000)	-0.007*** (0.000)	-0.006*** (0.000)	-0.006*** (0.000)	-0.006*** (0.000)
Children Under 6	-0.249*** (0.017)	-0.280*** (0.021)	-0.246*** (0.021)	-0.318*** (0.022)	-0.290*** (0.021)	-0.276*** (0.009)
Quarter 1	0.182*** (0.052)	0.048 (0.057)	0.065 (0.058)	0.043 (0.058)	-0.001 (0.057)	0.073*** (0.025)
Quarter 3	0.242*** (0.053)	0.277*** (0.057)	0.253*** (0.058)	0.230*** (0.058)	0.213*** (0.057)	0.259*** (0.025)
Quarter 4	0.280*** (0.054)	0.218*** (0.059)	0.209*** (0.059)	0.217*** (0.060)	0.064 (0.058)	0.217*** (0.026)
Legally married	0.586*** (0.048)	0.710*** (0.052)	0.595*** (0.052)	0.696*** (0.053)	0.693*** (0.052)	0.651*** (0.023)
Metropolitan	1.131*** (0.040)	1.112*** (0.043)	1.071*** (0.043)	1.061*** (0.043)	1.019*** (0.042)	1.084*** (0.019)
Cohort 1950–59	-1.563*** (0.289)	-0.231 (0.388)	0.942** (0.373)	1.731*** (0.379)	1.361*** (0.380)	-0.531*** (0.146)
Cohort 1960–69	-1.940*** (0.205)	-0.841*** (0.315)	0.115 (0.290)	0.755* (0.293)	0.626** (0.292)	-0.857*** (0.114)
Cohort 1970–79	-0.992*** (0.151)	-0.036 (0.270)	0.579** (0.238)	0.928*** (0.237)	0.898*** (0.235)	-0.049 (0.095)
Cohort 1980–89	-0.433*** (0.092)	0.329 (0.230)	0.689*** (0.190)	0.766*** (0.188)	0.761*** (0.186)	0.217*** (0.075)
Cohort 1990–99	0.000 (.)	1.055*** (0.190)	1.146*** (0.136)	1.020*** (0.127)	1.101*** (0.124)	0.701*** (0.055)
Compulsory schooling	0.167*** (0.014)	0.122*** (0.015)	0.086*** (0.016)	0.083*** (0.016)	0.045*** (0.016)	0.114*** (0.007)
Constant	-0.078 (0.412)	0.032 (0.409)	1.840*** (0.410)	2.919*** (0.425)	4.130*** (0.435)	1.757*** (0.165)
Observations	23543	19609	19259	19423	19081	100915

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A14: 2SLS estimates of the effect of education on the likelihood of employment for black men: second stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Education	0.005 (0.011)	0.087*** (0.021)	0.110*** (0.038)	0.380*** (0.159)	0.528 (0.416)	0.071*** (0.010)
Age in years	0.078*** (0.006)	0.052*** (0.010)	0.057*** (0.016)	-0.015 (0.051)	-0.060 (0.128)	0.062*** (0.004)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.001** (0.000)	0.001 (0.001)	0.001 (0.002)	-0.001*** (0.000)
Children Under 6	-0.042*** (0.004)	-0.030*** (0.007)	-0.027*** (0.008)	0.022 (0.035)	0.074 (0.101)	-0.034*** (0.003)
Quarter 1	0.021** (0.008)	-0.014 (0.010)	-0.003 (0.011)	-0.039 (0.028)	-0.002 (0.033)	-0.004 (0.004)
Quarter 3	0.010 (0.009)	-0.009 (0.011)	-0.014 (0.012)	-0.093* (0.051)	-0.107 (0.088)	-0.010** (0.005)
Quarter 4	0.015* (0.009)	0.000 (0.012)	0.001 (0.012)	-0.057 (0.036)	-0.001 (0.034)	-0.003 (0.005)
Legally married	0.233*** (0.016)	0.097*** (0.026)	0.080 (0.049)	-0.250 (0.206)	-0.410 (0.521)	0.141*** (0.013)
Metropolitan	0.067*** (0.013)	-0.020 (0.024)	-0.039 (0.042)	-0.338** (0.168)	-0.471 (0.426)	-0.004 (0.011)
Cohort 1950–59	0.111** (0.049)	-0.055 (0.071)	-0.197** (0.096)	-0.555** (0.251)	-0.904 (0.614)	-0.021 (0.025)
Cohort 1960–69	0.073** (0.033)	-0.064 (0.055)	-0.164*** (0.063)	-0.438** (0.178)	-0.770 (0.530)	-0.009 (0.019)
Cohort 1970–79	0.095*** (0.024)	-0.120** (0.047)	-0.227*** (0.052)	-0.544*** (0.178)	-0.654 (0.407)	-0.047*** (0.016)
Cohort 1980–89	0.122*** (0.015)	-0.086** (0.042)	-0.170*** (0.051)	-0.498*** (0.186)	-0.634 (0.435)	-0.012 (0.013)
Cohort 1990–99	0.000 (.)	-0.220*** (0.042)	-0.262*** (0.060)	-0.615*** (0.224)	-0.754 (0.536)	-0.111*** (0.013)
Constant	-1.081*** (0.054)	-1.249*** (0.064)	-1.574*** (0.116)	-2.880*** (0.651)	-3.757** (1.904)	-1.380*** (0.036)
Observations	20234	17346	16856	17077	16854	88367

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

Table A15: 2SLS estimates of the effect of education on the likelihood of employment for black men: first stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Age in years	0.623*** (0.024)	0.534*** (0.025)	0.445*** (0.026)	0.338*** (0.028)	0.315*** (0.029)	0.449*** (0.010)
Age Square	-0.009*** (0.000)	-0.007*** (0.000)	-0.007*** (0.000)	-0.005*** (0.000)	-0.005*** (0.000)	-0.006*** (0.000)
Children Under 6	-0.155*** (0.022)	-0.226*** (0.026)	-0.167*** (0.026)	-0.209*** (0.027)	-0.240*** (0.027)	-0.202*** (0.011)
Quarter 1	0.104* (0.057)	0.160*** (0.061)	0.065 (0.060)	0.092 (0.061)	-0.009 (0.061)	0.089*** (0.027)
Quarter 3	0.193*** (0.057)	0.172*** (0.061)	0.141** (0.061)	0.284*** (0.062)	0.198*** (0.061)	0.214*** (0.027)
Quarter 4	0.217*** (0.059)	0.279*** (0.064)	0.083 (0.063)	0.171*** (0.064)	-0.003 (0.064)	0.170*** (0.028)
Legally married	1.234*** (0.058)	1.192*** (0.061)	1.268*** (0.060)	1.290*** (0.061)	1.254*** (0.060)	1.241*** (0.027)
Metropolitan	1.033*** (0.043)	1.067*** (0.045)	1.057*** (0.045)	1.049*** (0.045)	1.022*** (0.045)	1.053*** (0.020)
Cohort 1950–59	0.321 (0.332)	0.413 (0.422)	1.333*** (0.404)	1.029** (0.413)	1.276*** (0.420)	-0.051 (0.161)
Cohort 1960–69	-0.563** (0.226)	-0.225 (0.335)	0.615** (0.306)	0.695** (0.313)	1.151*** (0.316)	-0.255** (0.122)
Cohort 1970–79	-0.347** (0.160)	0.340 (0.283)	0.689*** (0.247)	0.929*** (0.249)	0.926*** (0.250)	0.153 (0.100)
Cohort 1980–89	-0.152 (0.100)	0.523** (0.242)	0.821*** (0.200)	0.992*** (0.199)	0.980*** (0.200)	0.380*** (0.079)
Cohort 1990–99	0.000 (.)	1.020*** (0.200)	1.226*** (0.143)	1.265*** (0.134)	1.222*** (0.132)	0.747*** (0.058)
Compulsory schooling	0.203*** (0.015)	0.127*** (0.016)	0.075*** (0.016)	0.041** (0.017)	0.022 (0.017)	0.107*** (0.007)
Constant	-1.699*** (0.446)	-0.689 (0.437)	1.337*** (0.429)	3.285*** (0.459)	4.199*** (0.466)	1.165*** (0.177)
Observations	20234	17346	16856	17077	16854	88367

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A16: 2SLS estimates of the effect of education on the likelihood of employment for coloured women: second stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Education	0.106** (0.043)	0.204*** (0.069)	0.678 (0.844)	0.230 (0.203)	0.264 (0.345)	0.162*** (0.042)
Age in years	0.065*** (0.021)	0.027 (0.028)	-0.099 (0.252)	0.051 (0.038)	0.041 (0.067)	0.046*** (0.014)
Age Square	-0.001** (0.000)	-0.000 (0.000)	0.002 (0.004)	-0.000 (0.001)	-0.000 (0.001)	-0.000** (0.000)
Children Under 6	-0.008 (0.018)	0.035 (0.034)	0.240 (0.366)	0.028 (0.075)	0.049 (0.156)	0.011 (0.018)
Quarter 1	0.000 (0.023)	0.079* (0.043)	0.060 (0.135)	0.020 (0.060)	0.049 (0.099)	0.019 (0.016)
Quarter 3	-0.002 (0.023)	-0.001 (0.035)	-0.092 (0.207)	0.020 (0.040)	-0.010 (0.045)	0.007 (0.014)
Quarter 4	-0.014 (0.024)	-0.003 (0.036)	-0.092 (0.162)	0.010 (0.041)	0.062 (0.071)	0.003 (0.014)
Legally married	-0.036 (0.026)	-0.095** (0.048)	-0.520 (0.628)	-0.168 (0.153)	-0.184 (0.242)	-0.099*** (0.028)
Metropolitan	-0.105* (0.054)	-0.243*** (0.091)	-0.844 (1.094)	-0.266 (0.242)	-0.317 (0.387)	-0.187*** (0.052)
Cohort 1950–59	-0.028 (0.123)	-0.440* (0.261)	-1.314 (1.419)	-0.291 (0.413)	-0.676 (0.576)	-0.176** (0.077)
Cohort 1960–69	-0.021 (0.114)	-0.298 (0.216)	-0.710 (0.745)	-0.272 (0.376)	-0.504 (0.488)	-0.096 (0.062)
Cohort 1970–79	-0.035 (0.103)	-0.336* (0.195)	-0.873 (0.851)	-0.322 (0.333)	-0.494 (0.417)	-0.150*** (0.054)
Cohort 1980–89	-0.015 (0.056)	-0.364* (0.194)	-1.110 (1.270)	-0.373 (0.392)	-0.478 (0.510)	-0.169*** (0.055)
Cohort 1990–99	0.000 (.)	-0.421** (0.199)	-1.438 (1.770)	-0.395 (0.415)	-0.410 (0.616)	-0.201*** (0.067)
Constant	-1.811*** (0.200)	-1.825*** (0.287)	-3.820 (3.163)	-2.564** (1.280)	-2.844 (2.231)	-1.912*** (0.192)
Observations	3491	2396	2358	2342	2257	12844

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

Table A17: 2SLS estimates of the effect of education on the likelihood of employment for coloured women: first stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Age in years	0.546*** (0.050)	0.451*** (0.059)	0.315*** (0.061)	0.198*** (0.068)	0.207*** (0.070)	0.347*** (0.024)
Age Square	-0.007*** (0.001)	-0.006*** (0.001)	-0.004*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.005*** (0.000)
Children Under 6	-0.370*** (0.045)	-0.452*** (0.061)	-0.434*** (0.058)	-0.361*** (0.064)	-0.450*** (0.069)	-0.406*** (0.026)
Quarter 1	-0.015 (0.124)	-0.370** (0.149)	-0.112 (0.143)	-0.210 (0.155)	-0.258* (0.151)	-0.189*** (0.064)
Quarter 3	-0.058 (0.120)	0.069 (0.147)	0.221 (0.142)	0.047 (0.152)	0.051 (0.148)	0.077 (0.063)
Quarter 4	0.146 (0.125)	-0.023 (0.151)	0.163 (0.144)	-0.007 (0.158)	-0.158 (0.156)	0.032 (0.065)
Legally married	0.449*** (0.099)	0.549*** (0.122)	0.746*** (0.120)	0.746*** (0.128)	0.694*** (0.126)	0.634*** (0.053)
Metropolitan	1.191*** (0.086)	1.259*** (0.106)	1.294*** (0.102)	1.188*** (0.111)	1.118*** (0.109)	1.224*** (0.045)
Cohort 1950–59	-0.786 (0.634)	0.995 (1.024)	1.439 (0.882)	1.523 (0.950)	1.385 (0.937)	0.422 (0.352)
Cohort 1960–69	-1.548*** (0.476)	0.311 (0.894)	0.608 (0.727)	1.503* (0.782)	1.201 (0.774)	0.092 (0.292)
Cohort 1970–79	-1.426*** (0.372)	0.556 (0.807)	0.871 (0.624)	1.442** (0.668)	1.083* (0.657)	0.263 (0.253)
Cohort 1980–89	-0.701*** (0.233)	1.200* (0.712)	1.395*** (0.501)	1.743*** (0.528)	1.356*** (0.521)	0.810*** (0.200)
Cohort 1990–99	0.000 (.)	1.628*** (0.615)	1.999*** (0.359)	1.858*** (0.352)	1.671*** (0.335)	1.284*** (0.146)
Compulsory schooling	0.154*** (0.035)	0.148*** (0.042)	0.031 (0.040)	0.057 (0.045)	0.036 (0.043)	0.090*** (0.017)
Constant	0.515 (0.986)	0.264 (1.145)	3.141*** (1.059)	5.227*** (1.171)	5.791*** (1.201)	2.842*** (0.434)
Observations	3491	2384	2358	2342	2257	12844

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A18: 2SLS estimates of the effect of education on the likelihood of employment for coloured men: second stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Education	0.589 (0.435)	0.451** (0.220)	1.220 (3.615)	-0.473 (0.390)	-0.168 (0.145)	1.737 (2.319)
Age in years	-0.080 (0.149)	0.017 (0.052)	-0.159 (0.776)	0.187* (0.096)	0.134*** (0.043)	-0.380 (0.639)
Age Square	0.001 (0.002)	-0.000 (0.001)	0.003 (0.012)	-0.003* (0.001)	-0.002*** (0.001)	0.005 (0.009)
Children Under 6	0.234 (0.190)	0.163 (0.100)	0.516 (1.608)	-0.191 (0.158)	-0.104 (0.070)	0.740 (1.017)
Quarter 1	0.042 (0.106)	0.159 (0.108)	0.916 (2.710)	-0.167 (0.200)	-0.107 (0.095)	0.792 (1.063)
Quarter 3	-0.068 (0.087)	-0.028 (0.080)	0.184 (0.685)	0.103 (0.104)	-0.031 (0.045)	-0.033 (0.132)
Quarter 4	-0.051 (0.096)	0.042 (0.084)	0.222 (0.657)	0.140 (0.149)	0.046 (0.051)	-0.133 (0.227)
Legally married	-0.474 (0.513)	-0.350 (0.262)	-1.825 (6.028)	0.834* (0.505)	0.467** (0.199)	-2.078 (3.058)
Metropolitan	-0.586 (0.414)	-0.613** (0.297)	-1.490 (4.262)	0.515 (0.428)	0.109 (0.126)	-1.897 (2.500)
Cohort 1950–59	0.967 (1.016)	0.403 (0.507)	-0.706 (2.380)	0.765 (0.662)	-0.017 (0.283)	1.419 (1.899)
Cohort 1960–69	0.929 (0.961)	0.059 (0.415)	-0.458 (1.619)	0.333 (0.458)	-0.004 (0.251)	1.673 (2.232)
Cohort 1970–79	0.627 (0.718)	-0.383 (0.388)	-0.655 (1.748)	0.339 (0.418)	-0.089 (0.203)	0.875 (1.312)
Cohort 1980–89	0.157 (0.247)	-0.677 (0.438)	-1.076 (3.298)	0.626 (0.451)	0.104 (0.147)	-0.322 (0.647)
Cohort 1990–99	0.000 (.)	-0.850* (0.476)	-1.883 (5.688)	0.769 (0.515)	0.282* (0.152)	-1.467 (2.098)
Constant	-4.034** (1.930)	-3.357*** (1.115)	-7.900 (19.684)	1.466 (2.145)	-0.243 (0.834)	-9.513 (11.047)
Observations	2966	2034	2015	1948	1914	10877

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

Table A19: 2SLS estimates of the effect of education on the likelihood of employment for coloured men: first stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Age in years	0.372*** (0.061)	0.259*** (0.070)	0.223*** (0.073)	0.191** (0.078)	0.215*** (0.081)	0.282*** (0.028)
Age Square	-0.005*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.004*** (0.000)
Children Under 6	-0.425*** (0.059)	-0.421*** (0.076)	-0.444*** (0.074)	-0.384*** (0.083)	-0.452*** (0.085)	-0.438*** (0.033)
Quarter 1	-0.140 (0.147)	-0.324* (0.174)	-0.748*** (0.171)	-0.451** (0.180)	-0.575*** (0.176)	-0.456*** (0.075)
Quarter 3	0.073 (0.142)	0.096 (0.171)	-0.179 (0.170)	0.128 (0.176)	-0.077 (0.174)	0.017 (0.074)
Quarter 4	0.103 (0.150)	-0.062 (0.181)	-0.169 (0.174)	0.283 (0.182)	0.157 (0.179)	0.082 (0.076)
Legally married	1.180*** (0.126)	1.190*** (0.152)	1.673*** (0.152)	1.259*** (0.161)	1.335*** (0.155)	1.322*** (0.066)
Metropolitan	0.936*** (0.103)	1.321*** (0.122)	1.177*** (0.122)	1.082*** (0.127)	0.845*** (0.124)	1.077*** (0.053)
Cohort 1950–59	-2.110*** (0.773)	-0.796 (1.095)	0.524 (1.035)	1.082 (1.096)	-0.232 (1.106)	-0.767* (0.407)
Cohort 1960–69	-2.046*** (0.569)	-0.290 (0.914)	0.327 (0.842)	0.263 (0.885)	-0.614 (0.895)	-0.924*** (0.330)
Cohort 1970–79	-1.419*** (0.438)	0.627 (0.790)	0.427 (0.712)	0.409 (0.743)	-0.523 (0.749)	-0.497* (0.282)
Cohort 1980–89	-0.402 (0.272)	1.273* (0.659)	0.869 (0.561)	0.950 (0.586)	0.186 (0.593)	0.222 (0.221)
Cohort 1990–99	0.000 (.)	1.646*** (0.522)	1.529*** (0.380)	1.398*** (0.388)	1.088*** (0.389)	0.870*** (0.157)
Compulsory schooling	0.056 (0.042)	0.098** (0.047)	0.016 (0.049)	-0.066 (0.050)	-0.085* (0.049)	0.015 (0.020)
Constant	3.186*** (1.185)	2.975** (1.237)	5.156*** (1.248)	6.468*** (1.322)	6.946*** (1.364)	4.517*** (0.494)
Observations	2966	2034	2015	1948	1914	10877

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A20: 2SLS estimates of the effect of education on the likelihood of employment for white women: second stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Education	0.162*** (0.052)	0.960 (3.400)	0.354** (0.151)	0.545 (0.570)	0.265 (0.200)	0.223*** (0.054)
Age in years	0.008 (0.033)	-0.424 (1.970)	0.010 (0.054)	-0.059 (0.188)	0.012 (0.079)	0.012 (0.025)
Age Square	-0.000 (0.000)	0.004 (0.020)	-0.000 (0.001)	0.001 (0.002)	-0.000 (0.001)	-0.000 (0.000)
Children Under 6	0.072*** (0.027)	0.130 (0.632)	0.019 (0.053)	-0.008 (0.088)	-0.030 (0.042)	0.003 (0.017)
Quarter 1	-0.007 (0.033)	-0.190 (0.597)	0.054 (0.065)	-0.158 (0.172)	-0.067 (0.059)	-0.039** (0.020)
Quarter 3	0.033 (0.035)	-0.057 (0.210)	0.006 (0.057)	-0.055 (0.103)	-0.129 (0.099)	-0.032 (0.020)
Quarter 4	0.012 (0.035)	0.121 (0.632)	0.038 (0.066)	-0.012 (0.102)	-0.113 (0.070)	-0.024 (0.020)
Legally married	-0.092*** (0.030)	-0.120 (0.268)	-0.150** (0.066)	-0.124 (0.148)	-0.092 (0.083)	-0.082*** (0.021)
Metropolitan	-0.045 (0.036)	-0.444 (1.942)	-0.087 (0.084)	-0.294 (0.339)	-0.109 (0.139)	-0.064* (0.034)
Cohort 1950–59	0.486* (0.289)	1.351 (6.459)	-0.546 (0.444)	-1.192 (1.053)	-0.021 (0.320)	0.144 (0.130)
Cohort 1960–69	0.518* (0.271)	1.403 (6.700)	-0.514 (0.379)	-0.922 (0.792)	0.117 (0.302)	0.168 (0.123)
Cohort 1970–79	0.415* (0.221)	1.071 (5.552)	-0.599* (0.346)	-0.938 (0.800)	-0.022 (0.257)	0.056 (0.102)
Cohort 1980–89	0.325*** (0.107)	0.046 (1.125)	-0.569* (0.342)	-1.031 (1.114)	-0.096 (0.224)	0.009 (0.073)
Cohort 1990–99	0.000 (.)	-1.248 (4.349)	-0.719* (0.380)	-1.035 (1.168)	-0.215 (0.287)	-0.183** (0.086)
Constant	-1.625*** (0.237)	-2.028 (1.450)	-3.159*** (0.724)	-3.932 (2.526)	-2.579*** (0.866)	-2.131*** (0.187)
Observations	1819	1413	1271	1287	1249	7039

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

Table A21: 2SLS estimates of the effect of education on the likelihood of employment for white women: first stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Age in years	0.726*** (0.065)	0.588*** (0.077)	0.404*** (0.077)	0.357*** (0.090)	0.431*** (0.093)	0.502*** (0.032)
Age Square	-0.007*** (0.001)	-0.006*** (0.001)	-0.004*** (0.001)	-0.004*** (0.001)	-0.004*** (0.001)	-0.005*** (0.000)
Children Under 6	-0.304*** (0.089)	-0.183 (0.121)	-0.248** (0.107)	-0.099 (0.130)	-0.125 (0.129)	-0.198*** (0.051)
Quarter 1	-0.073 (0.139)	0.170 (0.174)	-0.174 (0.166)	0.256 (0.182)	0.130 (0.187)	0.080 (0.075)
Quarter 3	-0.164 (0.140)	0.041 (0.176)	0.022 (0.162)	0.071 (0.183)	0.426** (0.182)	0.099 (0.075)
Quarter 4	-0.090 (0.145)	-0.176 (0.177)	-0.171 (0.169)	0.054 (0.191)	0.251 (0.187)	0.008 (0.077)
Legally married	0.249** (0.119)	0.071 (0.147)	0.308** (0.140)	0.218 (0.153)	0.353** (0.154)	0.247*** (0.063)
Metropolitan	0.499*** (0.102)	0.570*** (0.130)	0.487*** (0.124)	0.583*** (0.138)	0.665*** (0.138)	0.564*** (0.056)
Cohort 1950–59	-4.333*** (0.713)	-1.882 (1.211)	1.389 (1.056)	1.442 (1.136)	-0.132 (1.149)	-1.224*** (0.430)
Cohort 1960–69	-4.192*** (0.596)	-1.952* (1.109)	1.063 (0.938)	0.968 (1.014)	-0.665 (1.013)	-1.290*** (0.386)
Cohort 1970–79	-3.070*** (0.497)	-1.593 (1.010)	1.184 (0.837)	1.147 (0.895)	-0.328 (0.888)	-0.743** (0.347)
Cohort 1980–89	-1.290*** (0.315)	-0.247 (0.866)	1.421** (0.687)	1.751** (0.721)	0.371 (0.717)	0.229 (0.280)
Cohort 1990–99	0.000 (.)	1.231* (0.720)	1.967*** (0.499)	1.868*** (0.479)	1.070** (0.467)	1.075*** (0.206)
Compulsory schooling	0.215*** (0.048)	0.016 (0.057)	0.121** (0.052)	0.054 (0.059)	0.080 (0.057)	0.112*** (0.023)
Constant	-2.343* (1.316)	0.036 (1.474)	1.925 (1.399)	3.299** (1.597)	2.452 (1.632)	0.903 (0.592)
Observations	1819	1413	1271	1287	1249	7039

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A22: 2SLS estimates of the effect of education on the likelihood of employment for white men: second stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Education	0.188*** (0.057)	0.290** (0.114)	0.230*** (0.080)	1.411 (2.722)	0.651 (1.066)	0.307*** (0.067)
Age in years	0.028 (0.031)	-0.057 (0.066)	0.045 (0.034)	-0.393 (1.011)	-0.096 (0.261)	-0.018 (0.029)
Age Square	-0.000 (0.000)	0.001 (0.001)	-0.001 (0.000)	0.004 (0.010)	0.001 (0.002)	0.000 (0.000)
Children Under 6	0.049 (0.030)	0.075* (0.043)	0.073* (0.041)	-0.127 (0.348)	0.136 (0.242)	0.066*** (0.022)
Quarter 1	0.130*** (0.047)	0.028 (0.052)	0.019 (0.043)	0.097 (0.366)	0.001 (0.120)	0.049* (0.026)
Quarter 3	0.099** (0.042)	0.012 (0.051)	-0.030 (0.045)	-0.222 (0.571)	-0.028 (0.153)	0.016 (0.025)
Quarter 4	0.133*** (0.046)	-0.006 (0.052)	-0.016 (0.045)	-0.070 (0.316)	-0.087 (0.193)	0.015 (0.026)
Legally married	-0.000 (0.044)	0.006 (0.057)	-0.025 (0.052)	-0.042 (0.346)	-0.232 (0.575)	-0.026 (0.034)
Metropolitan	-0.090** (0.046)	-0.232** (0.105)	-0.140** (0.057)	-1.341 (2.621)	-0.528 (0.870)	-0.225*** (0.055)
Cohort 1950–59	0.186 (0.185)	-0.108 (0.338)	-0.219 (0.264)	-0.948 (2.591)	-0.108 (0.951)	0.126 (0.144)
Cohort 1960–69	0.218 (0.181)	0.000 (0.309)	-0.288 (0.232)	-0.730 (1.944)	0.221 (0.675)	0.141 (0.133)
Cohort 1970–79	0.204 (0.155)	-0.032 (0.273)	-0.409* (0.210)	-1.153 (2.277)	-0.223 (1.102)	-0.011 (0.111)
Cohort 1980–89	0.229** (0.093)	-0.212 (0.258)	-0.487** (0.219)	-2.182 (4.323)	-0.706 (1.918)	-0.185* (0.110)
Cohort 1990–99	0.000 (.)	-0.656** (0.331)	-0.487** (0.231)	-2.835 (5.742)	-0.806 (1.904)	-0.410*** (0.135)
Constant	-2.286*** (0.308)	-1.225*** (0.307)	-2.342*** (0.333)	-5.984 (8.440)	-4.293 (5.765)	-2.256*** (0.221)
Observations	1699	1333	1250	1238	1149	6669

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

Table A23: 2SLS estimates of the effect of education on the likelihood of employment for white men: first stage

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Age in years	0.624*** (0.071)	0.639*** (0.072)	0.474*** (0.081)	0.391*** (0.102)	0.263*** (0.094)	0.475*** (0.033)
Age Square	-0.007*** (0.001)	-0.007*** (0.001)	-0.005*** (0.001)	-0.004*** (0.001)	-0.002** (0.001)	-0.005*** (0.000)
Children Under 6	-0.298*** (0.104)	-0.205* (0.114)	-0.399*** (0.112)	0.105 (0.147)	-0.213 (0.135)	-0.204*** (0.054)
Quarter 1	-0.551*** (0.155)	0.046 (0.165)	0.005 (0.172)	-0.078 (0.212)	0.001 (0.186)	-0.112 (0.079)
Quarter 3	-0.401** (0.156)	0.066 (0.163)	0.222 (0.173)	0.187 (0.209)	0.094 (0.187)	0.042 (0.079)
Quarter 4	-0.517*** (0.155)	-0.024 (0.166)	0.200 (0.172)	0.049 (0.215)	0.151 (0.195)	-0.012 (0.081)
Legally married	0.539*** (0.143)	0.337** (0.148)	0.484*** (0.157)	0.081 (0.194)	0.534*** (0.164)	0.388*** (0.072)
Metropolitan	0.663*** (0.112)	0.868*** (0.121)	0.612*** (0.128)	0.967*** (0.158)	0.814*** (0.138)	0.787*** (0.058)
Cohort 1950–59	-0.982 (0.795)	-0.211 (1.087)	-0.085 (1.068)	0.619 (1.294)	0.477 (1.152)	-0.385 (0.443)
Cohort 1960–69	-1.902*** (0.642)	-0.661 (0.975)	-0.173 (0.941)	0.361 (1.142)	0.102 (1.018)	-0.626 (0.392)
Cohort 1970–79	-1.474*** (0.522)	-0.153 (0.872)	0.671 (0.826)	0.657 (0.999)	0.862 (0.891)	0.019 (0.349)
Cohort 1980–89	-0.849** (0.337)	0.887 (0.732)	1.624** (0.660)	1.471* (0.802)	1.672** (0.723)	0.926*** (0.280)
Cohort 1990–99	0.000 (.)	2.138*** (0.580)	2.158*** (0.447)	2.000*** (0.540)	1.655*** (0.482)	1.569*** (0.201)
Compulsory schooling	0.194*** (0.049)	0.143*** (0.052)	0.171*** (0.055)	0.033 (0.065)	0.035 (0.058)	0.113*** (0.024)
Constant	-0.245 (1.429)	-2.936** (1.327)	-0.353 (1.422)	2.397 (1.798)	4.682*** (1.658)	0.631 (0.599)
Observations	1699	1333	1250	1238	1149	6669

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A24: Impact of grade 9 completion on employment outcomes: Second stage estimates for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Grade 9	0.061 (0.055)	0.653*** (0.118)	1.724*** (0.471)	4.751* (2.534)	3.991* (2.219)	0.585*** (0.059)
Age in years	0.076*** (0.004)	0.053*** (0.006)	0.044*** (0.014)	-0.023 (0.062)	-0.020 (0.061)	0.060*** (0.003)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.000 (0.000)	0.001 (0.001)	0.001 (0.001)	-0.001*** (0.000)
Children Under 6	-0.039*** (0.002)	-0.033*** (0.004)	-0.005 (0.012)	0.068 (0.067)	0.059 (0.064)	-0.033*** (0.002)
Coloured	0.054*** (0.007)	0.062*** (0.008)	0.054*** (0.013)	0.035 (0.032)	0.058** (0.028)	0.051*** (0.003)
Indian/Asian	0.078*** (0.014)	-0.044** (0.022)	-0.162** (0.063)	-0.395 (0.254)	-0.443* (0.257)	-0.018* (0.009)
White	0.163*** (0.015)	0.020 (0.029)	-0.199* (0.109)	-0.828 (0.540)	-0.651 (0.459)	0.048*** (0.014)
Quarter 1	0.006 (0.005)	-0.001 (0.006)	0.002 (0.010)	0.006 (0.024)	0.021 (0.024)	0.001 (0.003)
Quarter 3	0.007 (0.005)	-0.011 (0.007)	-0.035** (0.014)	-0.083 (0.057)	-0.097* (0.055)	-0.009*** (0.003)
Quarter 4	0.008 (0.005)	0.001 (0.007)	-0.015 (0.012)	-0.064 (0.046)	-0.037 (0.028)	-0.006* (0.003)
Male	0.128*** (0.004)	0.159*** (0.006)	0.190*** (0.017)	0.336*** (0.111)	0.271*** (0.083)	0.153*** (0.003)
Legally married	0.092*** (0.006)	0.034*** (0.011)	-0.062 (0.041)	-0.270 (0.199)	-0.184 (0.167)	0.052*** (0.005)
Metropolitan	0.059*** (0.008)	-0.004 (0.014)	-0.123** (0.054)	-0.430 (0.264)	-0.340 (0.226)	0.003 (0.007)
Cohort 1950–59	0.143*** (0.029)	-0.236*** (0.055)	-0.680*** (0.177)	-1.425* (0.750)	-1.043** (0.531)	-0.051*** (0.016)
Cohort 1960–69	0.100*** (0.024)	-0.207*** (0.043)	-0.575*** (0.141)	-1.143** (0.581)	-0.698* (0.362)	-0.020 (0.012)
Cohort 1970–79	0.070*** (0.019)	-0.255*** (0.042)	-0.655*** (0.146)	-1.228** (0.598)	-0.720** (0.356)	-0.069*** (0.011)
Cohort 1980–89	0.089*** (0.011)	-0.247*** (0.046)	-0.632*** (0.156)	-1.213** (0.618)	-0.705* (0.363)	-0.063*** (0.011)
Cohort 1990–99	0.000 (.)	-0.353*** (0.054)	-0.709*** (0.173)	-1.361** (0.693)	-0.823* (0.437)	-0.146*** (0.013)
Constant	-1.135*** (0.030)	-0.926*** (0.051)	-1.435*** (0.071)	-2.582*** (0.701)	-2.482*** (0.687)	-1.196*** (0.014)
Observations	54910	45217	43937	44206	43360	231630

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A25: Impact of grade 9 completion on employment outcomes: First stage estimates for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Compulsory schooling	0.026*** (0.001)	0.015*** (0.001)	0.006*** (0.001)	0.003* (0.001)	0.003* (0.001)	0.012*** (0.001)
Age in years	0.076*** (0.002)	0.056*** (0.002)	0.031*** (0.002)	0.026*** (0.002)	0.029*** (0.002)	0.046*** (0.001)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.001*** (0.000)
Children Under 6	-0.026*** (0.002)	-0.027*** (0.002)	-0.024*** (0.002)	-0.026*** (0.002)	-0.029*** (0.002)	-0.027*** (0.001)
Coloured	0.003 (0.006)	-0.011* (0.006)	-0.007 (0.006)	0.004 (0.006)	-0.005 (0.006)	-0.004 (0.003)
Indian/Asian	0.128*** (0.011)	0.143*** (0.011)	0.126*** (0.011)	0.098*** (0.011)	0.114*** (0.011)	0.122*** (0.005)
White	0.250*** (0.006)	0.230*** (0.007)	0.229*** (0.007)	0.213*** (0.007)	0.206*** (0.007)	0.225*** (0.003)
Quarter 1	0.010** (0.005)	0.007 (0.005)	0.002 (0.005)	-0.000 (0.005)	-0.006 (0.005)	0.004 (0.002)
Quarter 3	0.018*** (0.005)	0.021*** (0.005)	0.023*** (0.005)	0.021*** (0.005)	0.023*** (0.005)	0.024*** (0.002)
Quarter 4	0.032*** (0.005)	0.014*** (0.005)	0.014*** (0.005)	0.016*** (0.005)	0.009* (0.005)	0.021*** (0.002)
Male	-0.043*** (0.003)	-0.034*** (0.004)	-0.033*** (0.004)	-0.043*** (0.004)	-0.037*** (0.004)	-0.038*** (0.002)
Legally married	0.068*** (0.004)	0.080*** (0.005)	0.085*** (0.005)	0.078*** (0.005)	0.075*** (0.005)	0.077*** (0.002)
Metropolitan	0.118*** (0.003)	0.112*** (0.004)	0.114*** (0.004)	0.104*** (0.004)	0.101*** (0.004)	0.111*** (0.002)
Cohort 1950–59	-0.173*** (0.026)	0.248*** (0.034)	0.329*** (0.032)	0.280*** (0.032)	0.221*** (0.033)	0.048*** (0.013)
Cohort 1960–69	-0.249*** (0.018)	0.176*** (0.028)	0.265*** (0.026)	0.218*** (0.025)	0.150*** (0.025)	0.007 (0.010)
Cohort 1970–79	-0.165*** (0.014)	0.251*** (0.024)	0.297*** (0.021)	0.233*** (0.021)	0.155*** (0.021)	0.070*** (0.008)
Cohort 1980–89	-0.083*** (0.008)	0.299*** (0.021)	0.313*** (0.017)	0.237*** (0.016)	0.156*** (0.016)	0.111*** (0.007)
Cohort 1990–99	0.000 (.)	0.386*** (0.017)	0.350*** (0.012)	0.266*** (0.011)	0.189*** (0.011)	0.176*** (0.005)
Constant	-0.617*** (0.037)	-0.563*** (0.037)	-0.005 (0.036)	0.226*** (0.037)	0.262*** (0.038)	-0.158*** (0.015)
Observations	54910	45217	43937	44206	43360	231630

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A26: Impact of grade 10 completion on employment outcomes: Second stage estimates for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Grade 10	0.037 (0.034)	0.336*** (0.056)	0.694*** (0.125)	1.651*** (0.386)	1.577*** (0.440)	0.348*** (0.033)
Age in years	0.076*** (0.004)	0.057*** (0.005)	0.055*** (0.007)	0.030** (0.015)	0.030* (0.017)	0.061*** (0.002)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.001*** (0.000)
Children Under 6	-0.040*** (0.002)	-0.040*** (0.003)	-0.029*** (0.004)	-0.010 (0.012)	-0.006 (0.014)	-0.039*** (0.001)
Coloured	0.055*** (0.007)	0.066*** (0.008)	0.064*** (0.009)	0.117*** (0.019)	0.103*** (0.022)	0.061*** (0.003)
Indian/Asian	0.080*** (0.013)	-0.014 (0.017)	-0.056* (0.025)	-0.138** (0.054)	-0.201*** (0.064)	-0.000 (0.008)
White	0.167*** (0.012)	0.076*** (0.018)	0.007 (0.035)	-0.223** (0.096)	-0.194* (0.103)	0.087*** (0.010)
Quarter 1	0.007 (0.005)	-0.001 (0.006)	0.003 (0.007)	-0.005 (0.011)	0.005 (0.010)	0.001 (0.003)
Quarter 3	0.007 (0.005)	-0.009 (0.006)	-0.011 (0.007)	-0.024* (0.014)	-0.028** (0.012)	-0.005* (0.003)
Quarter 4	0.009* (0.005)	0.000 (0.006)	0.002 (0.007)	-0.015 (0.012)	-0.006 (0.010)	-0.001 (0.003)
Male	0.127*** (0.004)	0.150*** (0.005)	0.163*** (0.007)	0.212*** (0.020)	0.196*** (0.022)	0.146*** (0.002)
Legally married	0.094*** (0.005)	0.058*** (0.007)	0.021 (0.013)	-0.039 (0.034)	-0.017 (0.038)	0.068*** (0.004)
Metropolitan	0.061*** (0.006)	0.024*** (0.009)	-0.022 (0.018)	-0.127*** (0.045)	-0.122** (0.053)	0.023*** (0.005)
Cohort 1950–59	0.144*** (0.029)	-0.098** (0.040)	-0.363*** (0.067)	-0.736*** (0.175)	-0.631*** (0.158)	-0.054*** (0.015)
Cohort 1960–69	0.100*** (0.024)	-0.075** (0.032)	-0.285*** (0.048)	-0.599*** (0.133)	-0.450*** (0.117)	-0.014 (0.012)
Cohort 1970–79	0.070*** (0.019)	-0.118*** (0.028)	-0.348*** (0.046)	-0.672*** (0.136)	-0.480*** (0.114)	-0.054*** (0.010)
Cohort 1980–89	0.088*** (0.010)	-0.116*** (0.027)	-0.337*** (0.052)	-0.665*** (0.145)	-0.500*** (0.126)	-0.050*** (0.009)
Cohort 1990–99	0.000 (.)	-0.212*** (0.029)	-0.405*** (0.060)	-0.743*** (0.160)	-0.573*** (0.150)	-0.127*** (0.011)
Constant	-1.111*** (0.040)	-0.791*** (0.065)	-0.871*** (0.081)	-0.929*** (0.105)	-1.069*** (0.083)	-0.964*** (0.024)
Observations	54910	45217	43937	44206	43360	231630

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A27: Impact of grade 10 completion on employment outcomes: First stage estimates for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Compulsory schooling	0.042*** (0.001)	0.028*** (0.002)	0.014*** (0.002)	0.008*** (0.002)	0.006*** (0.002)	0.020*** (0.001)
Age in years	0.120*** (0.002)	0.098*** (0.002)	0.062*** (0.002)	0.042*** (0.003)	0.041*** (0.003)	0.075*** (0.001)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Children Under 6	-0.030*** (0.002)	-0.033*** (0.002)	-0.025*** (0.002)	-0.028*** (0.002)	-0.031*** (0.002)	-0.030*** (0.001)
Coloured	-0.023*** (0.006)	-0.034*** (0.007)	-0.033*** (0.007)	-0.037*** (0.007)	-0.041*** (0.007)	-0.034*** (0.003)
Indian/Asian	0.157*** (0.011)	0.190*** (0.013)	0.160*** (0.013)	0.126*** (0.012)	0.135*** (0.012)	0.153*** (0.006)
White	0.304*** (0.007)	0.282*** (0.008)	0.271*** (0.008)	0.246*** (0.008)	0.232*** (0.008)	0.266*** (0.003)
Quarter 1	0.008 (0.005)	0.013** (0.006)	0.003 (0.006)	0.006 (0.006)	-0.004 (0.006)	0.006** (0.002)
Quarter 3	0.027*** (0.005)	0.037*** (0.006)	0.022*** (0.006)	0.024*** (0.006)	0.015*** (0.006)	0.027*** (0.002)
Quarter 4	0.039*** (0.005)	0.029*** (0.006)	0.010* (0.006)	0.016*** (0.006)	0.004 (0.006)	0.022*** (0.003)
Male	-0.048*** (0.004)	-0.037*** (0.004)	-0.043*** (0.004)	-0.049*** (0.004)	-0.046*** (0.004)	-0.045*** (0.002)
Legally married	0.073*** (0.005)	0.084*** (0.005)	0.093*** (0.005)	0.086*** (0.005)	0.084*** (0.005)	0.082*** (0.002)
Metropolitan	0.139*** (0.004)	0.137*** (0.004)	0.139*** (0.004)	0.115*** (0.004)	0.118*** (0.004)	0.130*** (0.002)
Cohort 1950–59	-0.296*** (0.028)	0.071* (0.038)	0.361*** (0.036)	0.389*** (0.036)	0.299*** (0.037)	0.088*** (0.014)
Cohort 1960–69	-0.418*** (0.020)	-0.050 (0.031)	0.241*** (0.028)	0.299*** (0.028)	0.222*** (0.029)	-0.006 (0.011)
Cohort 1970–79	-0.275*** (0.015)	0.082*** (0.027)	0.297*** (0.023)	0.334*** (0.023)	0.241*** (0.023)	0.075*** (0.009)
Cohort 1980–89	-0.120*** (0.009)	0.189*** (0.023)	0.353*** (0.019)	0.351*** (0.019)	0.265*** (0.019)	0.149*** (0.007)
Cohort 1990–99	0.000 (.)	0.330*** (0.019)	0.431*** (0.013)	0.391*** (0.012)	0.319*** (0.012)	0.242*** (0.005)
Constant	-1.661*** (0.040)	-1.496*** (0.041)	-0.825*** (0.040)	-0.350*** (0.042)	-0.234*** (0.043)	-0.931*** (0.016)
Observations	54910	45217	43937	44206	43360	231630

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A28: Impact of grade 11 completion on employment outcomes: Second stage estimates for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Grade 11	0.033 (0.030)	0.251*** (0.041)	0.402*** (0.066)	0.850*** (0.137)	1.000*** (0.215)	0.264*** (0.025)
Age in years	0.076*** (0.003)	0.060*** (0.005)	0.062*** (0.006)	0.050*** (0.007)	0.046*** (0.010)	0.063*** (0.002)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.001*** (0.000)
Children Under 6	-0.040*** (0.002)	-0.042*** (0.003)	-0.035*** (0.003)	-0.030*** (0.005)	-0.022*** (0.008)	-0.041*** (0.001)
Coloured	0.055*** (0.007)	0.069*** (0.008)	0.067*** (0.009)	0.122*** (0.014)	0.108*** (0.018)	0.065*** (0.004)
Indian/Asian	0.080*** (0.013)	-0.002 (0.015)	-0.024 (0.019)	-0.067** (0.027)	-0.162*** (0.041)	0.004 (0.007)
White	0.168*** (0.012)	0.097*** (0.014)	0.076*** (0.021)	-0.047 (0.038)	-0.091 (0.058)	0.104*** (0.008)
Quarter 1	0.007 (0.005)	-0.000 (0.006)	0.005 (0.006)	-0.000 (0.007)	0.005 (0.008)	0.002 (0.003)
Quarter 3	0.007 (0.005)	-0.008 (0.006)	-0.004 (0.006)	-0.007 (0.008)	-0.020** (0.008)	-0.003 (0.003)
Quarter 4	0.009* (0.005)	0.002 (0.006)	0.004 (0.006)	-0.003 (0.008)	0.005 (0.008)	0.002 (0.003)
Male	0.127*** (0.004)	0.148*** (0.004)	0.153*** (0.005)	0.175*** (0.009)	0.177*** (0.013)	0.143*** (0.002)
Legally married	0.093*** (0.005)	0.063*** (0.006)	0.046*** (0.008)	0.017 (0.015)	0.016 (0.022)	0.072*** (0.003)
Metropolitan	0.062*** (0.006)	0.033*** (0.007)	0.018* (0.010)	-0.051*** (0.019)	-0.065** (0.029)	0.032*** (0.004)
Cohort 1950–59	0.142*** (0.029)	-0.029 (0.038)	-0.160*** (0.040)	-0.393*** (0.072)	-0.511*** (0.099)	-0.036** (0.015)
Cohort 1960–69	0.099*** (0.024)	-0.007 (0.033)	-0.107*** (0.030)	-0.313*** (0.052)	-0.357*** (0.073)	0.003 (0.012)
Cohort 1970–79	0.068*** (0.018)	-0.045 (0.028)	-0.171*** (0.025)	-0.372*** (0.050)	-0.387*** (0.070)	-0.035*** (0.010)
Cohort 1980–89	0.087*** (0.010)	-0.035 (0.023)	-0.147*** (0.023)	-0.355*** (0.052)	-0.380*** (0.073)	-0.025*** (0.008)
Cohort 1990–99	0.000 (.)	-0.126*** (0.020)	-0.210*** (0.025)	-0.414*** (0.058)	-0.435*** (0.086)	-0.097*** (0.008)
Constant	-1.106*** (0.043)	-0.819*** (0.060)	-0.882*** (0.071)	-0.804*** (0.089)	-0.863*** (0.099)	-0.926*** (0.027)
Observations	54910	45217	43937	44206	43360	231630

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A29: Impact of grade 11 completion on employment outcomes: First stage estimates for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Compulsory schooling	0.047*** (0.001)	0.038*** (0.002)	0.025*** (0.002)	0.015*** (0.002)	0.010*** (0.002)	0.027*** (0.001)
Age in years	0.128*** (0.002)	0.120*** (0.002)	0.090*** (0.003)	0.058*** (0.003)	0.048*** (0.003)	0.090*** (0.001)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Children Under 6	-0.029*** (0.002)	-0.034*** (0.002)	-0.028*** (0.002)	-0.030*** (0.003)	-0.034*** (0.003)	-0.031*** (0.001)
Coloured	-0.046*** (0.007)	-0.057*** (0.007)	-0.064*** (0.008)	-0.077*** (0.008)	-0.070*** (0.008)	-0.063*** (0.003)
Indian/Asian	0.191*** (0.012)	0.205*** (0.013)	0.196*** (0.013)	0.161*** (0.013)	0.174*** (0.013)	0.185*** (0.006)
White	0.313*** (0.007)	0.293*** (0.008)	0.295*** (0.008)	0.271*** (0.008)	0.263*** (0.008)	0.287*** (0.004)
Quarter 1	0.005 (0.005)	0.016*** (0.006)	-0.001 (0.006)	0.006 (0.006)	-0.007 (0.006)	0.004 (0.003)
Quarter 3	0.026*** (0.005)	0.043*** (0.006)	0.022*** (0.006)	0.027*** (0.006)	0.016*** (0.006)	0.028*** (0.003)
Quarter 4	0.027*** (0.005)	0.030*** (0.006)	0.013** (0.006)	0.016*** (0.006)	-0.005 (0.006)	0.018*** (0.003)
Male	-0.051*** (0.004)	-0.043*** (0.004)	-0.048*** (0.004)	-0.052*** (0.004)	-0.054*** (0.004)	-0.050*** (0.002)
Legally married	0.093*** (0.005)	0.092*** (0.005)	0.097*** (0.005)	0.102*** (0.005)	0.099*** (0.005)	0.095*** (0.002)
Metropolitan	0.139*** (0.004)	0.145*** (0.004)	0.140*** (0.004)	0.134*** (0.004)	0.130*** (0.004)	0.138*** (0.002)
Cohort 1950–59	-0.279*** (0.029)	-0.180*** (0.040)	0.117*** (0.039)	0.352*** (0.039)	0.351*** (0.039)	0.050*** (0.015)
Cohort 1960–69	-0.414*** (0.021)	-0.339*** (0.033)	-0.028 (0.030)	0.244*** (0.030)	0.257*** (0.031)	-0.071*** (0.012)
Cohort 1970–79	-0.244*** (0.016)	-0.183*** (0.028)	0.071*** (0.025)	0.295*** (0.025)	0.286*** (0.025)	0.029*** (0.010)
Cohort 1980–89	-0.095*** (0.010)	-0.068*** (0.024)	0.136*** (0.020)	0.317*** (0.020)	0.298*** (0.020)	0.100*** (0.008)
Cohort 1990–99	0.000 (.)	0.096*** (0.020)	0.258*** (0.014)	0.372*** (0.013)	0.365*** (0.013)	0.203*** (0.006)
Constant	-1.995*** (0.043)	-1.891*** (0.043)	-1.397*** (0.043)	-0.828*** (0.045)	-0.575*** (0.047)	-1.370*** (0.017)
Observations	54910	45217	43937	44206	43360	231630

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A30: Impact of grade 12 completion on employment outcomes: Second stage estimates for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Grade 12	0.038 (0.034)	0.284*** (0.046)	0.331*** (0.053)	0.564*** (0.078)	0.781*** (0.149)	0.264*** (0.025)
Age in years	0.077*** (0.003)	0.060*** (0.004)	0.067*** (0.005)	0.060*** (0.005)	0.062*** (0.006)	0.065*** (0.002)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Children Under 6	-0.040*** (0.002)	-0.041*** (0.003)	-0.037*** (0.003)	-0.037*** (0.004)	-0.028*** (0.006)	-0.041*** (0.001)
Coloured	0.054*** (0.006)	0.056*** (0.007)	0.044*** (0.007)	0.074*** (0.008)	0.055*** (0.009)	0.052*** (0.003)
Indian/Asian	0.077*** (0.014)	-0.030* (0.018)	-0.035* (0.019)	-0.061*** (0.023)	-0.177*** (0.039)	-0.014 (0.008)
White	0.164*** (0.015)	0.065*** (0.019)	0.067*** (0.022)	-0.020 (0.030)	-0.102* (0.053)	0.081*** (0.010)
Quarter 1	0.007 (0.005)	-0.000 (0.006)	0.004 (0.006)	0.002 (0.006)	-0.002 (0.007)	0.002 (0.003)
Quarter 3	0.007 (0.005)	-0.006 (0.006)	0.000 (0.006)	0.005 (0.006)	-0.015** (0.007)	-0.001 (0.003)
Quarter 4	0.009* (0.005)	0.002 (0.006)	0.005 (0.006)	-0.001 (0.007)	0.001 (0.007)	0.002 (0.003)
Male	0.127*** (0.004)	0.144*** (0.004)	0.145*** (0.005)	0.152*** (0.005)	0.158*** (0.008)	0.140*** (0.002)
Legally married	0.092*** (0.006)	0.055*** (0.007)	0.049*** (0.008)	0.036*** (0.011)	0.025 (0.018)	0.068*** (0.003)
Metropolitan	0.062*** (0.006)	0.035*** (0.007)	0.035*** (0.008)	-0.003 (0.010)	-0.030 (0.019)	0.037*** (0.003)
Cohort 1950–59	0.142*** (0.029)	0.004 (0.039)	-0.075** (0.037)	-0.142*** (0.042)	-0.407*** (0.071)	-0.017 (0.014)
Cohort 1960–69	0.098*** (0.023)	0.024 (0.035)	-0.037 (0.031)	-0.097*** (0.031)	-0.292*** (0.054)	0.016 (0.012)
Cohort 1970–79	0.068*** (0.017)	-0.013 (0.030)	-0.093*** (0.026)	-0.145*** (0.026)	-0.308*** (0.049)	-0.019* (0.010)
Cohort 1980–89	0.087*** (0.010)	-0.003 (0.024)	-0.068*** (0.019)	-0.130*** (0.022)	-0.291*** (0.049)	-0.006 (0.008)
Cohort 1990–99	0.000 (.)	-0.094*** (0.019)	-0.126*** (0.015)	-0.197*** (0.022)	-0.320*** (0.055)	-0.075*** (0.007)
Constant	-1.106*** (0.043)	-0.817*** (0.060)	-0.941*** (0.061)	-0.856*** (0.070)	-0.925*** (0.077)	-0.933*** (0.026)
Observations	54910	45217	43937	44206	43360	231630

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A31: Impact of grade 12 completion on employment outcomes: First stage estimates for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Compulsory schooling	0.042*** (0.001)	0.034*** (0.002)	0.030*** (0.002)	0.022*** (0.002)	0.013*** (0.002)	0.027*** (0.001)
Age in years	0.107*** (0.002)	0.105*** (0.002)	0.093*** (0.003)	0.070*** (0.003)	0.042*** (0.003)	0.081*** (0.001)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Children Under 6	-0.029*** (0.002)	-0.035*** (0.002)	-0.030*** (0.002)	-0.035*** (0.003)	-0.036*** (0.003)	-0.033*** (0.001)
Coloured	-0.000 (0.007)	-0.005 (0.007)	-0.010 (0.008)	-0.031*** (0.008)	-0.021*** (0.008)	-0.014*** (0.003)
Indian/Asian	0.248*** (0.012)	0.281*** (0.013)	0.272*** (0.013)	0.231*** (0.013)	0.242*** (0.013)	0.254*** (0.006)
White	0.398*** (0.007)	0.372*** (0.008)	0.387*** (0.008)	0.361*** (0.008)	0.351*** (0.008)	0.374*** (0.004)
Quarter 1	0.010* (0.005)	0.014** (0.006)	0.002 (0.006)	0.004 (0.006)	-0.000 (0.006)	0.006** (0.003)
Quarter 3	0.024*** (0.005)	0.031*** (0.006)	0.014** (0.006)	0.019*** (0.006)	0.015** (0.006)	0.020*** (0.003)
Quarter 4	0.023*** (0.005)	0.027*** (0.006)	0.011* (0.006)	0.021*** (0.006)	-0.002 (0.006)	0.016*** (0.003)
Male	-0.035*** (0.004)	-0.026*** (0.004)	-0.035*** (0.004)	-0.038*** (0.004)	-0.044*** (0.004)	-0.036*** (0.002)
Legally married	0.108*** (0.005)	0.109*** (0.005)	0.109*** (0.005)	0.118*** (0.005)	0.115*** (0.006)	0.110*** (0.002)
Metropolitan	0.114*** (0.004)	0.121*** (0.004)	0.121*** (0.004)	0.118*** (0.004)	0.122*** (0.004)	0.120*** (0.002)
Cohort 1950–59	-0.237*** (0.029)	-0.274*** (0.039)	-0.115*** (0.038)	0.087** (0.039)	0.316*** (0.040)	-0.022 (0.015)
Cohort 1960–69	-0.347*** (0.021)	-0.407*** (0.032)	-0.243*** (0.030)	-0.015 (0.030)	0.245*** (0.031)	-0.122*** (0.012)
Cohort 1970–79	-0.199*** (0.015)	-0.275*** (0.028)	-0.150*** (0.025)	0.043* (0.025)	0.265*** (0.025)	-0.035*** (0.010)
Cohort 1980–89	-0.077*** (0.010)	-0.173*** (0.024)	-0.073*** (0.020)	0.079*** (0.020)	0.267*** (0.020)	0.028*** (0.008)
Cohort 1990–99	0.000 (.)	-0.027 (0.020)	0.062*** (0.014)	0.175*** (0.013)	0.321*** (0.013)	0.122*** (0.006)
Constant	-1.785*** (0.042)	-1.680*** (0.043)	-1.517*** (0.043)	-1.155*** (0.045)	-0.656*** (0.047)	-1.347*** (0.017)
Observations	54910	45217	43937	44206	43360	231630

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A32: Impact of tertiary education on employment outcomes: Second stage estimates for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Tertiary	0.301 (0.271)	1.514*** (0.331)	1.546*** (0.333)	1.431*** (0.244)	1.418*** (0.302)	1.062*** (0.115)
Age in years	0.075*** (0.004)	0.054*** (0.007)	0.056*** (0.008)	0.061*** (0.006)	0.055*** (0.008)	0.063*** (0.002)
Age Square	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Children Under 6	-0.038*** (0.003)	-0.027*** (0.006)	-0.025*** (0.006)	-0.031*** (0.005)	-0.025*** (0.007)	-0.033*** (0.002)
Coloured	0.058*** (0.007)	0.091*** (0.013)	0.085*** (0.014)	0.102*** (0.012)	0.087*** (0.014)	0.077*** (0.005)
Indian/Asian	0.064*** (0.023)	-0.046* (0.027)	-0.064* (0.031)	-0.090*** (0.032)	-0.105*** (0.030)	-0.034*** (0.011)
White	0.111* (0.061)	-0.157** (0.072)	-0.145* (0.074)	-0.200*** (0.066)	-0.175** (0.075)	-0.070** (0.027)
Quarter 1	0.005 (0.005)	0.008 (0.008)	0.009 (0.008)	0.006 (0.008)	-0.003 (0.008)	0.003 (0.003)
Quarter 3	0.006 (0.005)	0.001 (0.008)	0.000 (0.008)	0.005 (0.008)	-0.021** (0.008)	-0.002 (0.003)
Quarter 4	0.008 (0.005)	0.006 (0.008)	0.008 (0.008)	0.002 (0.008)	-0.000 (0.008)	0.003 (0.003)
Male	0.130*** (0.005)	0.158*** (0.007)	0.158*** (0.008)	0.154*** (0.007)	0.149*** (0.008)	0.147*** (0.003)
Legally married	0.073*** (0.022)	-0.039 (0.028)	-0.045 (0.029)	-0.025 (0.023)	-0.026 (0.031)	0.005 (0.010)
Metropolitan	0.054*** (0.012)	-0.013 (0.019)	0.001 (0.017)	-0.015 (0.015)	-0.015 (0.018)	0.013** (0.006)
Cohort 1950–59	0.146*** (0.030)	0.008 (0.052)	0.000 (0.053)	0.043 (0.051)	-0.050 (0.052)	0.063*** (0.019)
Cohort 1960–69	0.099*** (0.024)	0.030 (0.048)	0.043 (0.049)	0.052 (0.044)	0.056 (0.048)	0.078*** (0.017)
Cohort 1970–79	0.060*** (0.015)	-0.019 (0.040)	-0.033 (0.041)	-0.017 (0.036)	0.017 (0.041)	0.019 (0.013)
Cohort 1980–89	0.073*** (0.013)	-0.013 (0.032)	-0.020 (0.029)	-0.025 (0.026)	-0.007 (0.028)	0.016* (0.009)
Cohort 1990–99	0.000 (.)	-0.057** (0.026)	-0.051*** (0.019)	-0.053*** (0.016)	-0.037** (0.016)	-0.028*** (0.006)
Constant	-1.078*** (0.064)	-0.734*** (0.097)	-0.794*** (0.111)	-0.853*** (0.087)	-0.787*** (0.112)	-0.884*** (0.035)
Observations	54910	45217	43937	44206	43360	231630

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

CHAPTER 2. RACE, GENDER, EDUCATION AND EMPLOYMENT OUTCOMES

Table A33: Impact of tertiary education on employment outcomes: First stage estimates for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Compulsory schooling	0.005*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.009*** (0.001)	0.007*** (0.001)	0.007*** (0.000)
Age in years	0.018*** (0.001)	0.024*** (0.002)	0.027*** (0.002)	0.027*** (0.002)	0.028*** (0.002)	0.023*** (0.001)
Age Square	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
Children Under 6	-0.010*** (0.001)	-0.016*** (0.002)	-0.014*** (0.002)	-0.018*** (0.002)	-0.022*** (0.002)	-0.015*** (0.001)
Coloured	-0.014*** (0.004)	-0.024*** (0.005)	-0.028*** (0.005)	-0.032*** (0.005)	-0.034*** (0.005)	-0.027*** (0.002)
Indian/Asian	0.075*** (0.008)	0.063*** (0.009)	0.076*** (0.009)	0.112*** (0.009)	0.083*** (0.009)	0.082*** (0.004)
White	0.224*** (0.005)	0.216*** (0.006)	0.220*** (0.006)	0.267*** (0.006)	0.244*** (0.006)	0.235*** (0.002)
Quarter 1	0.006 (0.004)	-0.003 (0.004)	-0.003 (0.004)	-0.001 (0.004)	0.001 (0.004)	0.000 (0.002)
Quarter 3	0.006* (0.004)	0.001 (0.004)	0.003 (0.004)	0.008* (0.004)	0.013*** (0.004)	0.006*** (0.002)
Quarter 4	0.007* (0.004)	0.003 (0.004)	0.001 (0.004)	0.007 (0.004)	-0.000 (0.004)	0.003* (0.002)
Male	-0.014*** (0.003)	-0.014*** (0.003)	-0.016*** (0.003)	-0.017*** (0.003)	-0.018*** (0.003)	-0.016*** (0.001)
Legally married	0.079*** (0.003)	0.082*** (0.004)	0.085*** (0.004)	0.089*** (0.004)	0.100*** (0.004)	0.087*** (0.002)
Metropolitan	0.041*** (0.003)	0.055*** (0.003)	0.048*** (0.003)	0.055*** (0.003)	0.056*** (0.003)	0.051*** (0.001)
Cohort 1950–59	-0.044** (0.019)	-0.054** (0.027)	-0.073*** (0.026)	-0.095*** (0.027)	-0.077*** (0.028)	-0.081*** (0.010)
Cohort 1960–69	-0.048*** (0.014)	-0.080*** (0.022)	-0.104*** (0.021)	-0.110*** (0.021)	-0.110*** (0.022)	-0.088*** (0.008)
Cohort 1970–79	0.001 (0.010)	-0.047** (0.019)	-0.071*** (0.017)	-0.073*** (0.017)	-0.083*** (0.018)	-0.044*** (0.007)
Cohort 1980–89	0.035*** (0.006)	-0.026 (0.016)	-0.047*** (0.014)	-0.042*** (0.014)	-0.053*** (0.014)	-0.014** (0.005)
Cohort 1990–99	0.000 (.)	-0.030** (0.013)	-0.035*** (0.010)	-0.031*** (0.009)	-0.023** (0.009)	-0.014*** (0.004)
Constant	-0.316*** (0.028)	-0.370*** (0.029)	-0.420*** (0.029)	-0.457*** (0.031)	-0.459*** (0.033)	-0.381*** (0.012)
Observations	54910	45217	43937	44206	43360	231630

Notes: Reported coefficients represent marginal effects from linear probability models. Standard errors are in parentheses whilst p-values are in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical level.

IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

Compulsory schooling laws are designed to keep children in school, under the expectation that doing so increases employment opportunities and wages. However, in South Africa, the education system does not perform as well as it does in other countries, which may, in turn, mitigate the effectiveness of the compulsory schooling reform on labour market outcomes. Gender wage gap research in the country finds significant differences that need further explanation. This research investigates whether the 1997 compulsory schooling law (CSL) partially explains gender differences in the distribution of wages, and whether selection into employment moderates that relationship. The analysis is underpinned by conditional quantile regression, along with a selection-corrected version that incorporates the CSL. The data is taken from the 2018 South African General Household Survey. The results suggest that: returns are higher for women than men, the schooling reform increased returns, and are not constant across either the wage distribution or the level of schooling, suggesting that education increases wage inequality. Furthermore, I find also that correcting for selection into employment reduces the returns for women, but does not alter the ranking of returns along the wage distribution before and after the CSL implementation. Policies aimed at expanding access to higher education for formerly disadvantaged groups and improving the quality of skills training could help reduce disparities in labour market earnings.

3.1 Introduction

Education is widely recognised as a key driver of individual and societal progress. It contributes to the development of human capital (Becker, 1962), a key driver of sustainable economic growth (Draissi et al., 2023; Solarin and Yen, 2016). Governments around the world have long used policy tools to promote access to education, with compulsory schooling laws (CSLs) among the most prominent interventions (Harmon, 2017). By legally requiring children to attend school up to a certain age or grade, CSLs seek to raise educational attainment, especially for disadvantaged groups who might otherwise exit the schooling system early. The underlying assumption is that prolonged schooling improves human capital and leads to better labour market outcomes, such as higher wages and improved employment prospects (de New et al., 2021; Harmon, 2017). However, the actual impact of CSLs on education and earnings remains a subject of empirical investigation, particularly in contexts marked by inequality and limited schooling access.

The South African government introduced a compulsory schooling policy in 1997, which requires young people to attend school from the beginning of the year they reach the age of 7, until the last day of the school year in which they turn 15 or reach Grade 9, whichever comes first (Republic of South Africa, 1996b). The Basic Education Amendment Laws Act (BELA) of 2024 extended mandatory schooling to Grade R (reception year) for 6 year old children.

However, there are concerns about the policy's effectiveness. There is evidence of extensive grade repetition and high dropout rates in the South African education system. van der Berg et al. (2019) report that average repetition rates in grades 9 to 11 are highest, relative to all other grades, exceeding the national repetition average of 10%; furthermore, repetition in Grade 10 is highest, averaging 21% during the period 2014-2018. Moreover, between 58% and 60% of young people dropped out of the schooling system after the compulsory schooling phase, i.e., after the year they turned 15 or reached Grade 9 (Branson and Leibbrandt, 2013b). Thus, completed education levels, which were low before the policy was instituted, remain low (Case and Deaton, 1999; Case and Yogo, 1999; Branson and Leibbrandt, 2013a), raising questions about the policy's impact on schooling success, and, by extension, labour market outcomes.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

This research attempts to offer insight into the policy's effectiveness, as measured by wage outcomes corrected for selection into employment. Specifically, I examine whether the distribution of the returns to education for individuals subjected to the CSL are different to those who were not affected by the law. The analysis employs a quantile selection model, recently introduced by [Arel-lano and Bonhomme \(2017\)](#), to estimate the returns to education along the earnings distribution for Black male and female South Africans. This approach, spanning the entire wage distribution, offers advantages over mean regressions, enabling a nuanced exploration of the variation in returns, as well as its interaction with unobserved factors that influence employment outcomes. Thus, the analysis offer insights crucial for addressing labour market related inequality. The investigation draws on data from the 2018 South African General Household Survey ([Statistics South Africa, 2019](#)).

The study contributes to the literature in a number of ways. Firstly, it presents a comparison of the returns to education received by individuals exposed to the compulsory education reform against those who were not. Given that the reform itself is exogenously imposed, the comparison allows for some inference into the effectiveness of the policy. Secondly, it applies a recently introduced approach for correcting sample selection bias in quantile regressions; few researchers have considered the effect of sample selection when it comes to the distribution of returns to education ([Buchinsky, 2001](#); [Staneva et al., 2022](#)).

However, both analyses apply the sample selection correction technique developed by [Buchinsky \(1998\)](#), which has been criticized in the literature. The major drawback of this method is its implicit assumption that the error term and the regressors are independent, conditional on the selection probability. This assumption implies that the quantile regression curves are parallel across the distribution, indicating that the coefficients do not vary with the quantile level and are effectively equal to those from a mean regression ([Huber and Melly, 2015](#)).

In contrast, I use the sample selection correction procedure introduced by [Arellano and Bonhomme \(2017\)](#), which overcomes the challenge associated with conventional approaches. Given the extent of unemployment in the country, 46% in 2023 among Black Africans (see Figure 1.1), selection into employment could be an important component of the total returns estimate. While other

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

researchers such as [Maasoumi and Wang \(2019\)](#) and [Elass \(2024\)](#), have used a similar approach to estimate gender wage gaps across different quantiles, they do not examine returns to education as I do.

Thirdly, it adds to the literature by estimating returns to education for both males and females across the wage distribution taking into account selection into employment. Fourthly, given the estimates, it is possible to infer whether selection into employment worsens inequality, especially amongst the younger employment cohorts for whom the end of apartheid might be expected to benefit. Fifthly, the analysis also offers the first update of estimated returns to education along the wage distribution in South Africa in nearly 27 years ([Mwabu and Schultz, 1996](#)).

South Africa, marked by a Gini coefficient of 0.69, stands as one of the most unequal nations globally ([Mosomi et al., 2020](#)). The labour market is a significant contributor to income inequality, evidenced by a wage Gini of 0.634 ([Wittenberg, 2017b](#); [Mosomi et al., 2020](#)). A notable 83% of households without employed members experience poverty, a figure that contrasts with 50% for those with at least one wage employee ([Finn, 2015](#)). The study focuses on Black South Africans for two reasons. First, they are the population majority, representing 81.4% of the total populace ([Statistics South Africa, 2023](#)). Second, they are believed to have been impacted the most by the compulsory schooling policy. According to [Patrinos et al. \(2021\)](#), compulsory schooling reform is expected to impact those at the lowest end of the distribution of educational attainment. Black Africans lag in educational attainment compared to others, because apartheid limited their access to quality education ([Thobejane, 2013](#)).

In South Africa, the relationship between education and the earnings distribution has included both race and gender; however, most analyses focus on mean returns¹. Furthermore, Black African women are under-represented in high-skilled and management positions ([Espinoza et al., 2019](#)), suggesting that there are likely to be differences in the earnings distributions for men and women. With the exception of [Mwabu and Schultz \(1996\)](#), most distributional analyses examine the gender wage gap along the wage distribution. For instance, [Ntuli \(2009\)](#) investigate this

¹See, for example, [Moll \(1992, 1998\)](#); [Mwabu and Schultz \(2000\)](#); [Bhorat \(2000\)](#); [Keswell and Poswell \(2002, 2004\)](#); [Fryer and Vencatachellum \(2005\)](#); [Bhorat and Goga \(2012\)](#); [Branson and Leibbrandt \(2013b\)](#); [Branson et al. \(2013\)](#); [Salisbury \(2016\)](#); [Hertz \(2003\)](#); [Nikolov et al. \(2020\)](#); [Biyase and Zwane \(2015\)](#); [Depken et al. \(2019\)](#); [Zwane \(2020\)](#)

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

gap between Black African men and women, while [Mosomi \(2019a\)](#) explore its evolution from 1993 to 2015. However, neither consider sample selection, the inclusion of which could alter their conclusions.

Although assessing schooling returns across the wage distribution dates back at least to [Buchinsky \(1994\)](#), few have considered CSLs. [Buchinsky \(1994\)](#) assesses returns to education across the earnings spectrum using U.S. data, finding that returns to education are higher at upper quantiles of the earnings distribution. The results suggest that individuals with greater abilities receive more substantial rewards in the labour market, compared to those with lower abilities – an indication of complementarity between ability and education ([Patrinos and Sakellariou, 2006](#)). Such results also imply that education increases earnings inequality. Similar results are available for Portugal ([Hartog et al., 2001](#)), a group of 16 developed countries ([Martins and Pereira, 2004](#)), a different group of European countries ([Prieto-Rodriguez et al., 2008](#)), and India ([Mitra, 2019](#)). In Ethiopia, [Girma and Kedir \(2005\)](#) find that returns decline along the distribution of wages in their two-stage quantile regression model.

The literature also points to differences by gender and level of economic development. [Patrinos and Sakellariou \(2006\)](#) find higher returns for women than for men in Venezuela, a phenomenon that is common ([Dougherty, 2005](#)). On the development front, [Patrinos et al. \(2006\)](#) observed decreasing returns at higher levels of the wage distribution in low-income countries (Mongolia, Cambodia, Vietnam, Indonesia, Thailand, and the Philippines) and increasing patterns in high-income countries (Singapore). In South Africa, [Mwabu and Schultz \(1996\)](#) find differences across race groups – declining returns at higher levels of education for Black South Africans and increasing returns for White South Africans.

One concern that arises in the analysis of education returns is the potential for sample selection. The impact of sample selection on the mean wage function goes back to [Gronau \(1974\)](#) and [Heckman \(1974\)](#); selection has received limited attention in the quantile wage function. [Buchinsky \(1998\)](#) introduces a quantile selection model that [Buchinsky \(2001\)](#) follows to estimate returns to education for women in the U.S finding significant selection bias over the wage distribution across all ages. [Staneva et al. \(2022\)](#), on the other hand, using the same method for Bulgaria,

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

Russia, Kazakhstan and Serbia, finds selection bias in Russia and Kazakhstan for both males and females; correcting for the bias leads to lower returns for women in both countries.

However, the [Buchinsky \(1998\)](#) correction makes an implicit independence assumption between the error term and the regressors, conditional on the selection probability, and, therefore, may not be appropriate. The assumption implies that all quantile regression curves are parallel to each other, such that the quantile regression coefficients are identical to each other and equal to the mean regression coefficients ([Huber and Melly, 2015](#)). [Arellano and Bonhomme \(2017\)](#) overcome this challenge through the non-random selection process that may be present in the data across the distribution. This allows for more accurate estimation of the conditional quantile functions and, hence, heterogeneous returns to education across the wage distribution. Therefore, I follow [Arellano and Bonhomme \(2017\)](#).

Although focused on the gender wage gap, one recent analysis to follow [Arellano and Bonhomme \(2017\)](#) is [Elass \(2024\)](#), who investigates the evolution of the gender wage gap across the distribution for three European countries - the UK, France and Finland - finding different selection patterns across the countries. They observe positive selection into employment in France, but negative selection in both the UK and Finland, finding negative selection for both males and females in the UK, but only for females in Finland. Positive (negative) selection implies that the wages of those that are employed are higher (lower) compared to those that are unemployed. In a related analysis, [Maasoumi and Wang \(2019\)](#) find negative selection for U.S women when “employment rates are extremely low and positive selection with relatively higher employment.”

The findings of the analysis reveal that returns to education are heterogeneous by compulsory schooling exposure, educational level, gender, and along the wage distribution. I find, similar to others, that returns for women exceed those for men. This could be due to education’s “double effects” on women’s earnings ([Dougherty, 2005](#)): education increases women’s productivity, as it does for men, but can mitigate gender norms that contribute to the gender wage gap (such as labour market discrimination). I also find that women exposed to the schooling law tend to earn more than women not exposed. For men, a similar conclusion arises, but only at higher quantiles of the wage distribution; it is the opposite for men at the bottom of the wage distribution.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

In terms of selection, female returns suffer; correcting for selection lowers their returns between 0.9 and 4.4 percentage points, depending on the location on the wage distribution. Despite that, correcting for selection does not reverse the returns ranking for women, at least when the comparison is before and after the compulsory schooling law was implemented. Furthermore, the results suggest that compulsory schooling has worsened skills-bias and wage inequality. This arises from an analysis focused specifically on levels of education, rather than years of education. I find that returns increase by level of education, rising from the bottom end to the top end of the wage spectrum for both females and males. The results – increasing returns across the wage distribution, except at the extreme top, suggest a skills-biased structure to the labour market that increases inequality.

3.2 Estimation Strategy

The main objectives of this research are to examine returns along the wage distribution, investigate whether there is a statistically significant difference based on exposure to the CSL, and to determine whether selection increases or decreases returns. To address potential returns heterogeneity, I apply the conditional quantile regression (QR) model [Koenker and Bassett Jr \(1978\)](#). I incorporate a CSL dummy and interaction with schooling to capture the CSL differences. Finally, I extend each of these to account for selection into the workplace via [Arellano and Bonhomme \(2017\)](#). The details follow.

3.2.1 Quantile regression

Let y be a continuous outcome variable and x be a vector of covariates. The quantile regression is specified as

$$Q_{\tau}(y|x) = x' \beta_{\tau} \quad (3.1)$$

where $\tau \in (0, 1)$, and $x' \beta_{\tau}$ is strictly increasing in τ . The estimation of β_{τ} involves minimizing the weighted absolute deviations from the regression line, and is solved via linear programming:

$$\min_{\beta} \left\{ \sum_{i: y \geq k} \tau |y - x' \beta_{\tau}| + \sum_{i: y \leq k} (1 - \tau) |y - x' \beta_{\tau}| \right\} \quad (3.2)$$

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

where $k \equiv x' \beta_\tau$. Furthermore, $\tau|k$ is the penalty for under-prediction, when the actual value of y is higher than its predicted value. Conversely, $(1 - \tau)|k$ is the penalty for over-prediction, if y is lower than its prediction. Because the penalties are linear, rather than squared, quantile regression is more robust to outliers than is ordinary regression.

To estimate returns, I include schooling in the regression. To assess the difference across CSL exposure, I include a CSL exposure indicator, as well as its interaction with schooling (as well as additional covariates):

$$Q_\tau(y|X, S) = X' \alpha_\tau + CSL \theta_\tau + S \gamma_\tau + \delta S_\tau \times CSL \quad (3.3)$$

where y denotes the logarithm of monthly wage or earnings; S , is years of education; X , is a vector of covariates consisting of age, number of children in the household, marital status, and province of residence. CSL is a dummy variable to capture compulsory schooling (exposed = 1, and 0 = otherwise). α_τ is the return to X at the τ^{th} quantile, and γ_τ is the return to education at the τ^{th} quantile for those who were not exposed to the CSL. With respect to the CSL, θ captures the average difference in (log) monthly wages for those exposed, relative to those not, while δ is the marginal difference in returns to education between individuals who were affected by the CSL and those who were not. Equation (3.3) is estimated separately for females and males.

3.2.2 Quantile selection model

Two concerns often arise when considering the returns to schooling. One is the potential endogeneity of schooling; the other is the potential non-random selection into employment. Addressing endogeneity requires an instrument that yields variation in schooling outcomes that is exogenous to wages – certainly the compulsory schooling law could be used ([Balestra and Backes-Gellner, 2017](#); [Brunello et al., 2009](#)). Although both are likely to impact returns estimates, I focus my attention on employment selection in this analysis. The reason for doing so arises from the depth of unemployment in the country – approximately 46% for the population that underpins the sample – which means that a large share of possible workers are not working and, thus, are not earning a wage.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

As a result, estimates derived from the standard quantile wage regression are prone to bias, leading to inaccurate statistical inferences about the broader population (Blau et al., 2024). In addition, there are complications arising from attempting to separate schooling endogeneity and sample selection across the quantiles. Thus, implicitly, I am assuming that the underlying endogeneity is constant over time. I also assume that the CSL is exogenously given. Thus, the CSL interaction can be interpreted as the difference in the returns to education arising from exposure to the CSL.

Moreover, Arellano and Bonhomme (2017) offer more flexibility, since their model does not require rank invariance, and, therefore, does not disregard the plethora of potential patterns associated with sample selection. While the rank invariance assumption postulates that the effect of the instrument (in this study, number of children below the age of 6) on the conditional quantiles of wages remains consistent across all quantiles, that assumption is empirically unlikely (Elass, 2024). Thus, I accommodate quantile-specific heterogeneity in the treatment effects of the outcome variable, unlike conventional models.

To set the stage, I consider a general outcome specification, where the quantile function is a linear index, as in Arellano and Bonhomme (2017).

$$Y^* = Q(U, X) = X' \beta_\tau \quad (3.4)$$

where Y^* is the latent outcome variable, i.e wage offers, the function Q is the τ th conditional quantile of Y^* , given the covariates X , and U is the error term of the wage equation. X is a vector of covariates as defined before, consisting of years of education (or grade completion), age, number of children under the age of eighteen in the household, marital status, metropolitan status, and province of residence. The selection equation (employment equation) is defined as

$$D = 1\{V \leq p(Z)\} \quad (3.5)$$

In this model, $D = 1$ indicates employment, while $D = 0$ denotes non-employment. The vector Z includes all variables in X along with number of children in the household that are under the age of six, which serves as an instrument for employment because it is likely to impact employment status but should not directly affect wages. Previous studies such as Staneva et al. (2022) show

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

that the number of young children in the household is correlated with employment outcomes. The term $p(Z)$ represents the propensity score, which captures the probability of employment for individuals characterized by Z . V is an error term in the selection equation. The model's key assumption is the exclusion restriction, which asserts that the number of young children in the household is independent of both U and V , conditional on X (Elass, 2024).

The connection between the quantiles of the distribution of observable outcomes Y and those of the distribution of potential outcomes Y^* is created through the following expression (Arellano and Bonhomme, 2017):

$$\begin{aligned} P(Y^* \leq X' \beta_\tau \mid D = 1, Z = z) &= P(U \leq \tau \mid V \leq p(Z), Z = z) \\ &= C_{U,V|X=x}\{\tau, p(z)\} \\ &= G_x\{\tau, p(z)\} \end{aligned} \tag{3.6}$$

where $G_x \equiv C(\tau, p)/p$ is the conditional copula function, which measures the dependence between U and V . Estimation is carried out via three steps. First, equation (3.5) is estimated via probit to obtain the propensity scores. Second, equation (3.6) is estimated by minimising the method-of-moments objective function to recover the copula parameter, which measures the correlation between the error terms in the wage and employment equations. Third, the propensity scores and G_x are used to consistently estimate β at each quantile of the wage distribution (Muñoz and Siravegna, 2021). Maasoumi and Wang (2019) and Elass (2024) apply this approach to estimate selection-corrected gender wage gaps across different quantiles.

The copula parameter reflects the nature of selection into employment, with its sign indicating the direction of selection (Maasoumi and Wang, 2019). A negative parameter suggests positive selection, meaning that employed individuals tend to have characteristics associated with higher wages relative to the general population — for example, individuals with higher ability may be both more likely to work and to earn higher wages (Maasoumi and Wang, 2019). In contrast, a positive parameter indicates negative selection, where those who are employed tend to have lower potential wages compared to the broader population. This may occur due to labour market constraints or policies that encourage the employment of individuals with lower earning potential (Granados et al., 2020; Dolado et al., 2020). It is important to note that this interpretation differs

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

from the conventional Heckman selection model, where a positive or negative selection coefficient directly indicates positive or negative selection into employment, respectively. In contrast, with the copula approach, the sign of the copula parameter has the opposite interpretation.

For comparison purposes, I also report selection estimates at the mean arising from [Heckman \(1979\)](#). The main difference between [Heckman \(1979\)](#) and [Arellano and Bonhomme \(2017\)](#) arises in the additivity assumption associated with the unobserved factors. [Heckman \(1979\)](#) allows the propensity score to enter linearly, as is common in the control function literature, while [Arellano and Bonhomme \(2017\)](#) show this assumption is unlikely to be appropriate in the non-linear quantile setting. The effect of the selection correction can vary across quantiles. For lower quantiles, the selection effect might be small while for upper quantiles, the selection effect might be large. They account for this by allowing a flexible, non-linear correction.

3.2.3 Instrument

As noted earlier, the quantile selection model attempts to address potential selection bias, which requires an instrument for the identification of the correlation between the unobserved factors affecting both wages and employment. In this analysis, I use the number of children in the household below age of 6 as the instrument for employment. [Staneva et al. \(2022\)](#) use a similar variable for identification. The number of children, especially younger children, is believed to negatively impact female labour supply ([Liu and Marois, 2023](#)), because it takes more time and energy to care for more children leaving less time and energy for employment. In Vietnam, [Vu et al. \(2021\)](#) find that the number of children has a negative impact on maternal participation in the labour market and on working hours but has no effect on earnings.

[Markussen and Strøm \(2022\)](#) find that whilst women's labour supply is negatively affected by the number of children, men's labour supply increases. Further, they find the impact of children on labour supply to be heterogeneous along the distribution. Given the persistent gender wage gap in South Africa ([Oyenubi and Mosomi, 2024](#)), and the fact that women are the primary caregivers to children ([Posel and Bruce-Brand, 2021](#)), the number of young children is likely to impact the labour supply of women ([Posel, 2025](#)). As shown in the previous chapter, the number of children

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

under the age of 6 has a negative effect on the employment probability of both men and women (see Table A2). Moreover, women move into and out of full-time homemaking over time (Posel and Bruce-Brand 2021).

However, the number of children might not be purely exogenous. In South Africa, household formation and membership are often more fluid and complex than in developed countries. Individuals may belong to multiple households or be regarded as household members even if they are not resident for much of the year, reflecting flexible kinship networks and economic coping strategies (Posel, 2010). In a study of factors influencing rural remittances, Posel (2001) finds that the presence of young children (aged six and under) is a significant determinant of remittance flows. This suggests that the number of young children may be correlated with income or wages, potentially confounding wage equation estimates.

Nevertheless, the prevalence of labour migration and remittance receipt has declined substantially in recent years. In 1993, around 61% of households receiving income transfers also reported having migrant workers — that is, household members recorded as absent due to migration. By 2008, this figure had fallen to just 28%, reflecting a marked decline in the prevalence of labour migration among transfer-receiving households (Posel, 2010).

Moreover, Branson and Leibbrandt (2013b) use marital status, number of children, and the number of employed adults in the household as instruments for employment when estimating earnings. However, they find that these variables are statistically significant in both the employment and wage equations, potentially violating the exclusion restriction. Nevertheless, the estimated coefficients on education remain stable in both magnitude and significance, whether or not these instruments are included, suggesting robustness of the education effect despite concerns about instrument validity.

3.2.4 Data

The data used for this analysis come from the 2018 South African General Household Survey (GHS). I excluded non-Black Africans (11,596), those with incomplete birth date information (166), those younger than 15 or older than 65 (22,674), those who are economically inactive

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

(15,660), those with zero or missing wage data (141), and those with missing educational levels (360). I limit the analysis to labour force participants between the ages of 21 and 65; thus, I focus on those who have completed their compulsory phase, as well as those who have not yet reached retirement age. Consequently, the analysis is based on 20,540 individuals, including 9,901 women and 10,639 men.

The primary variables in the analysis are gross earnings, years of education (or the categorical completed grades measure), exposure to the CSL, and the number of children in the household that are under the age of six. Importantly, I was able to access birth dates for each individual from Statistics South Africa (StatsSA), which are used to determine whether or not an individual was subjected to the CSL.²

The dependent variable is gross earnings, which includes overtime, allowances, bonuses and any deductions before tax. The years of education for each individual are derived from the highest level of education completed.³ Years of education is commonly used in the literature; see [Andini \(2010\)](#), [Bhorat and Kimani \(2017\)](#) and [Liwiński \(2020\)](#) for South Africa and Poland, respectively, for just a few examples.

In addition to the primary variables, I also control for a number of potential confounding variables. I include age to proxy for labour market experience, similar to [\(Torun, 2018\)](#). Age is interacted with CSL to control for differences in age between the treated and untreated groups. In the South Africa context, the number of children in the household is correlated with both employment and wages, hence I include the number of children in the household that are under the age of 18 [\(Branson and Leibbrandt, 2013b\)](#). Metropolitan status and province of residence are included to capture local differences in the labour market; geographical variables are also relatively common in the literature, such as [Depken et al. \(2019\)](#) and [Biyase and Zwane \(2015\)](#) for South Africa, and [Stefani et al. \(2006\)](#) for Brazil. I also include marital status, which is correlated with earnings in

²The CSL went into effect on January 1, 1997. Schooling then became compulsory for all individuals aged 7 to either age 15 or Grade 9. Thus, everyone with a 15th birthday in 1997 was affected, as was anyone younger than that. In other words, everyone born on or after January 1, 1982 was affected by the law.

³The South African education system offers a number of different routes and qualifications; this includes a number of different options related to the completion of secondary education, for example, Table [A1](#) in Appendix A describes the derivation of education from the highest level completed variable.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

a number of studies (Biyase and Zwane, 2015; Bardasi and Taylor, 2008; Depken et al., 2019).⁴

As we will see, the quantile selection model requires the use of instrument to mitigate non-random selection into employment. In this case I use the number of children in the household that are under the age of six to instrument for employment. Finally, I separate the analysis by gender. Depken et al. (2019) find that, in South Africa, being female is associated with 32.5% lower wages. Such large differences in average wages, as well as gender wage gap decomposition research suggesting differences in returns to various endowments (Mosomi, 2019a), suggests that it is beneficial to allow for gender differences in the returns to education.

3.3 Results and Discussion

Below, I describe the main features of the data and present the model estimates.

3.3.1 Descriptive statistics

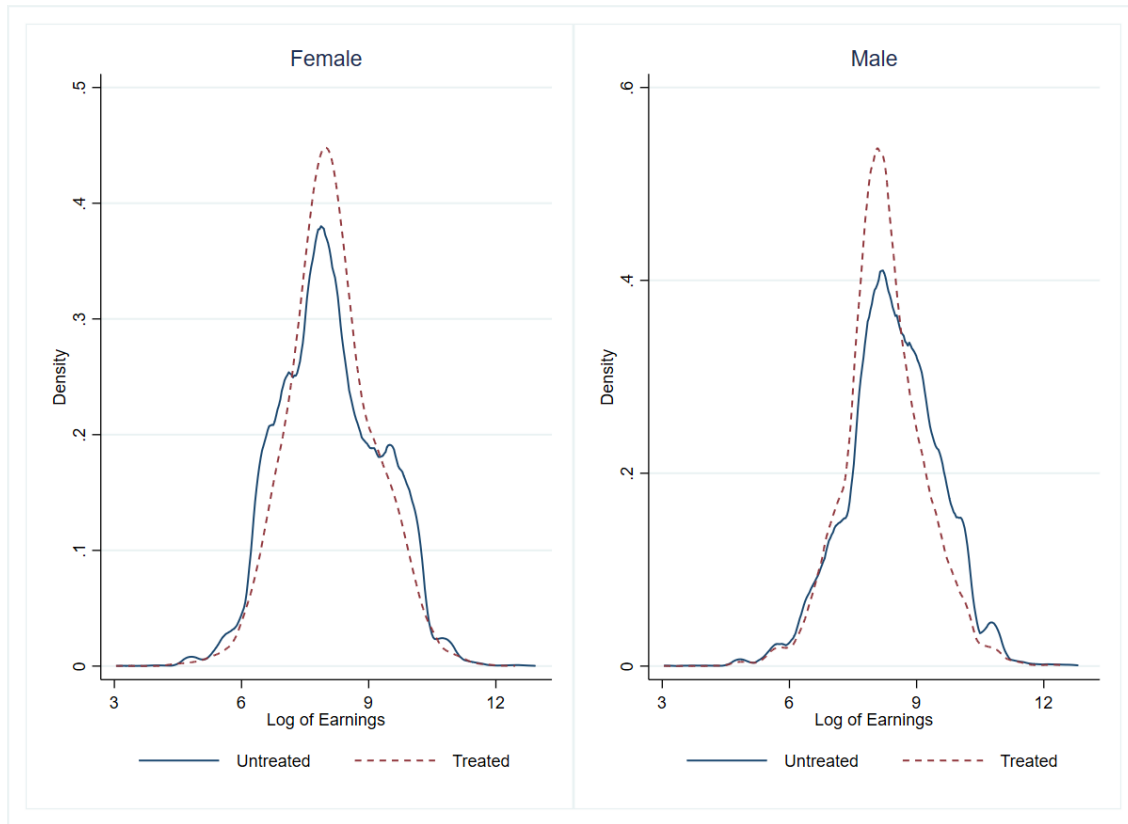
Given my focus on the distribution of returns to wages, I present wage densities, conditional on employment, for both men and women, before and after the CSL. See Figure 3.1, where "treated" refers to those who were subjected to the CSL. For both males and females, exposure to the CSL is associated with a smoother density, which may reflect differences in labour market experience and educational attainment between the groups.

As suggested – see the summary statistics, split by CSL exposure and gender in Table 3.1 – there are differences in both education and age, the latter of which is expected to relate to work experience, across the treatment groups. Exposed females have more education, 11.4 years, on average, compared to 10.1 for their unexposed counterparts. Similarly, the proportion of treated females completing grades 11 (19%) and 12 (45%) exceeded that of the untreated women for Grade 11 (14%) and Grade 12 (27%). However, the untreated group has a higher proportion of individuals with tertiary education (16%) compared to the treated group's 14%.

⁴There are a multiple legal marriage options in the country, including traditional/customary or civil/religious marriages; therefore, all of these are used to define legally married; in terms of interpretation, the reference category is 'not legally married'. Although marital status might also be endogenous, I am not directly interested in the effect of marriage on wages; thus, I do not worry about the possibility here.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

Figure 3.1: Distribution of log monthly earnings by gender and treatment status



Source: Author's calculations based on the sample data.

Despite more years of education, the treated group earns only 82.1% (7,028.49/8,563.62) of the untreated group's monthly wages. Similarly, the employment rate is lower amongst the treated (56%) compared to the untreated (82%), and there is an approximately double marital rate. Since the CSL intervention targeted younger cohorts, the observed wage gap likely reflects timing and life-cycle effects. The lower wages among CSL-exposed individuals can be attributed to their younger age and relative inexperience in the labour market compared to the untreated group. As shown in Table 3.1, the treated cohort is, on average, 18 years younger and earns R1,535.13 less than their untreated counterparts. These differences are consistent with human capital theory (Mincer, 1974), which posits that wages typically increase with experience, even when educational attainment is held constant.

As was the case with the women, CSL-exposed males are 17.5 years younger. The difference in average years of education between the two groups is 1.1 years in favour of the CSL group, with

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

Table 3.1: Summary statistics for women and men by compulsory schooling treatment exposure

	Females			Males		
	Control Group Mean	Treated Group Mean	Difference b	Control Group Mean	Treated Group Mean	Difference b
Age	47.12	29.49	17.63***	46.95	29.44	17.51***
Education	10.09	11.42	-1.33***	9.77	10.89	-1.12***
Primary	0.20	0.04	0.16***	0.23	0.08	0.15***
Grade 8	0.05	0.03	0.02***	0.06	0.04	0.02***
Grade 9	0.06	0.05	0.01***	0.06	0.08	-0.02***
Grade 10	0.11	0.10	0.01	0.12	0.13	-0.01
Grade 11	0.14	0.19	-0.04***	0.14	0.18	-0.05***
Grade 12	0.27	0.45	-0.18***	0.27	0.39	-0.12***
Tertiary	0.16	0.14	0.02**	0.12	0.10	0.02***
Employed	0.82	0.56	0.26***	0.85	0.66	0.20***
Monthly Earnings	8563.62	7028.49	1535.13***	11082.30	8066.22	3016.08***
Legally married	0.35	0.15	0.20***	0.48	0.12	0.36***
Western Cape	0.05	0.06	-0.01***	0.06	0.06	-0.00
Eastern Cape	0.26	0.19	0.06***	0.21	0.19	0.02*
Northern Cape	0.03	0.03	0.00	0.03	0.03	0.00
Free State	0.07	0.06	0.00	0.06	0.06	0.00
KwaZulu-Natal	0.17	0.21	-0.04***	0.15	0.20	-0.04***
North West	0.06	0.06	0.00	0.08	0.07	0.01***
Gauteng	0.29	0.29	0.00	0.32	0.31	0.01
Mpumalanga	0.10	0.11	-0.01**	0.10	0.10	-0.00
Limpopo	0.11	0.08	0.03***	0.10	0.09	0.01*
Observations	5025	4876	9901	5122	5517	10639

Notes: p-values come from the t test comparing individuals exposed and not exposed to the compulsory schooling law.

a higher Grade 12 (39%), Grade 11 (18%), and Grade 10 (13%) completion rate, compared to those not exposed at 27%, 14%, and 12%, respectively. However, the untreated have more tertiary education (12%) compared to the treated's 10%. CSL-males earn only 72.8% (8,066.22/11,082.30) of the wages of non-CSL men, a lower share than observed for women. This wage difference is statistically significant. Similar to what was observed for women, some of the wage differences are likely due to life cycle patterns, since the untreated are 1.6 times older and have a marriage rate that is 4 times higher. Turning to employment, 85% of untreated men are employed, compared to 66% of the treated.

As expected, labour market outcomes vary by gender. Men have higher employment rates compared to women (85% vs 82% and 66% vs 56%). Women earn less than men. Whilst untreated women earn R8,563.62 per month, this is only 77.3% of untreated men's monthly wage of R11,082.30. Similarly, treated women earn only 87.1% (7,028.49/8,066.22) of their male counterparts. The treated-untreated male-female, approximately a difference-in-difference (DD), wage gap is in favour of men, despite the DD education gap (at least for grades 11 and 12) favouring women. Relatively low female wages suggests their over-representation in low-paying jobs, as

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

suggested by [Espí et al. \(2019\)](#).

3.3.2 Quantile regression results

The primary objectives of this chapter are to analyse the returns to years of education and levels of education across the wage distribution, assess whether exposure to the CSL leads to a statistically significant difference, and evaluate whether selection increases or decreases returns. To answer these questions, education and Grade 9 completion are interacted with a dummy variable for compulsory schooling exposure to assess whether there is a statistically significant difference between the returns to education and levels of education of individuals who were exposed to the law and those who were not. Estimates are available for both males and females, and, as a benchmark to the main results, the mean returns are estimated via OLS. The estimates are reported at 5 distinct locations of the wage distribution, that is, at the 10th, 25th, 50th, 75th and 90th quantiles (Q10, Q25, Q50, Q75 and Q90, respectively). The selection question is considered under the the quantile selection model below.

Returns to education

Those exposed to the CSL are earning only approximately 77%⁵ of those who were not exposed. As explain in the descriptive statistics section, the difference in wages between these two groups is driven by life-cycle patterns as predicted by the human capital theory. The densities, see Figure 3.1, suggest fewer high wage earners, either male or female, amongst those exposed; this is also observed in the regressions – see the "CSL" parameter estimates in Table 3.2⁶. For females at the 75th and 90th quantiles of wages, exposure to the CSL was associated with at least a 30% reduction in earnings. For males at the same location of the distribution, exposure was associated with a 17% (or greater) reduction in earnings. Given the large age differences between the groups, and, presumably, large differences in experience, such differences are not unexpected.

Even though wages are, on average, lower, returns to education are generally found to increase

⁵Table 3.1: Wages for treated individuals = R8,066.22 + R7,028.49 = R15,090.71. Wages for untreated individuals = R11,082.30 + R8,563.62 = R19,645.92. Wages of the treated group as a percent of wages of the untreated group = R15,090.71/R19,645.92 = 77%.

⁶The rest of the results are shown in tables B1 and B2 in Appendix B.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

along the distribution of earnings, and higher, although not always statistically significantly so, for the treated.

Table 3.2: Returns to education for women and men estimated via OLS and quantile regression

	(1) Mean	(2) Q10	(3) Q25	(4) Q50	(5) Q75	(6) Q90
Females						
Education	0.139*** (0.004)	0.093*** (0.007)	0.145*** (0.007)	0.150*** (0.005)	0.143*** (0.005)	0.130*** (0.006)
CSL	-0.565** (0.237)	-1.427*** (0.321)	-0.701** (0.293)	-0.480* (0.253)	-1.064*** (0.282)	-0.293 (0.363)
CSL × Education	0.023** (0.011)	0.067*** (0.019)	0.025* (0.014)	0.026** (0.012)	0.034*** (0.013)	0.021 (0.013)
Males						
Education	0.101*** (0.005)	0.092*** (0.007)	0.103*** (0.007)	0.109*** (0.004)	0.108*** (0.004)	0.097*** (0.008)
CSL	-0.542*** (0.195)	-1.012*** (0.282)	-0.604*** (0.211)	-0.338* (0.205)	-0.584*** (0.210)	-0.174 (0.321)
CSL × Education	0.001 (0.008)	-0.006 (0.012)	-0.003 (0.010)	0.001 (0.009)	0.003 (0.008)	-0.001 (0.010)

Notes: (1) All equations include the following additional control variables: age, marital status, and province. (2) In parentheses are standard errors. (3) ***, **, and * indicate significance at the 1, 5, and 10% critical level. (4) CSL = 1 if individual i was subjected to the compulsory schooling law (CSL), CSL = 0 otherwise. (5) F-test for females: Education = 2.40 with p-value = 0.048. F-test for females: CSL × Education = 32.86 with p-value = 0.000. (6) F-test for males: Education = 0.19 with p-value = 0.946. F-test for males: CSL × Education = 2.51 with p-value = 0.040.

Among females, the return to an additional year of education increases along the wage distribution, from 9.3% and Q10 to 15.0% at Q50 before slightly dropping to 14.3% at the Q75 and further to 13.0% at Q90. These estimates represent the returns for untreated women. The F-test focusing on education across the distribution suggests that the returns are heterogeneous along the wage distribution (see note (5) under Table 3.2). The treatment effect on returns, i.e., the change in returns arising from exposure to the CSL is captured by the coefficient on the interaction, CSL × Education. For women, this is positive and statistically significant at all selected quantiles, except the 90th percentile, suggesting that women's returns to education are 2.5-6.7% higher, because of CSL exposure, at any (selected) quantile.

The additional return appears to follow a declining trend, with relatively higher improvements at the lower end of the distribution. When combined, the returns to education for CSL-exposed women are increasing along most of the wage distribution. The return to education is calculated as the sum of the coefficient on Education and the product of the coefficient on the interaction term (CSL × Education) and the mean proportion of individuals exposed to CSL. At the mean,

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

the return to education is calculated as follows: $\text{return to education} = 0.139 + (0.023 \times 0.4925) = 0.1503$ or 15.03%. Accordingly, at Q10, Q25, Q50, Q75, and Q90, the return is: 12.6%, 15.7%, 16.3%, 16.0%, and 14.0% , respectively. Moreover, the F-test for the interaction term along the distribution is statistically significant (see note (5) in Table 3.2), suggesting that the treatment effect is heterogeneous along the earnings distribution.

For men, estimated returns are increasing along the earnings distribution, rising from 9.2% at Q10 to 10.9% at Q50, before dropping to 10.8% at Q75 and further to 9.7% at Q90. The F test is not statistically significant (see note (6) in Table 3.2), and, therefore, one can conclude that wages are homogeneous along the distribution. Moreover, the coefficient of the interaction term, $\text{CSL} \times \text{Education}$, is not statistically significant at any of the selected quantiles of the wage distribution. This suggests that exposure to the CSL did not affect the conditional distribution of earnings.

Returns to education levels

In addition to estimating the returns to an additional year of education, the returns are also estimated for each grade completed. The levels of education considered are primary, grades 8, 9, 10, 11 and 12 as well as tertiary education. Primary is the reference category. To asses the difference between returns of treated and untreated individuals, the dummy variable for CSL is interacted with the dummy variable variable for grade completion. For example, $\text{CSL} \times \text{Grade 9}$ is the interaction term between CSL and Grade 9 completion. I focus on that, because the CSL is meant to affect Grade 9 completion.

Table 3.3 shows the rate of return to grade completion along the wage distribution for females and males, respectively. Only results for key variables are shown in these tables. The rest of the results are contained in tables B3 and B4 in Appendix B. The top panel of Table 3.3 presents the results for females. I find two important differences between tables 3.2 and 3.3. The first is that the return from one grade to the next is not constant (see Table B3 in Appendix B), regardless of whether we focus on the mean or any specific quantile. Second, although there is evidence of heterogeneity in returns to any particular grade, returns generally follow an inverse U -shape, at least for higher levels of completed education. Thus, an assumption of constant returns, as is

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

Table 3.3: Returns to grade completion for women and men estimated via OLS and quantile regression

	(1) Mean	(2) Q10	(3) Q25	(4) Q50	(5) Q75	(6) Q90
Females						
Grade 9	0.206*** (0.070)	0.206*** (0.074)	0.243** (0.095)	0.199*** (0.073)	0.084 (0.070)	0.249** (0.123)
CSL	-0.467** (0.184)	-0.776*** (0.257)	-0.426** (0.201)	-0.310* (0.185)	-0.473** (0.226)	-0.262 (0.371)
CSL × Grade 9	0.021 (0.035)	-0.030 (0.043)	-0.009 (0.046)	0.054* (0.031)	0.075* (0.045)	-0.050 (0.085)
Males						
Grade 9	0.178*** (0.069)	0.219** (0.087)	0.129 (0.101)	0.135* (0.075)	0.205* (0.124)	0.293*** (0.096)
CSL	-0.639*** (0.163)	-1.186*** (0.283)	-0.928*** (0.206)	-0.585*** (0.147)	-0.391** (0.188)	0.002 (0.265)
CSL × Grade 9	0.024 (0.029)	0.094** (0.039)	0.082** (0.038)	0.050* (0.028)	-0.014 (0.047)	-0.056 (0.038)

Notes: All equations include the following additional control variables: age, marital status, and province. Bootstrapped standard errors are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10% critical level. CSL = 1 if individual i was subjected to the compulsory schooling law, CSL = 0 otherwise.

assumed in the model underscoring the previous results, is not appropriate. Instead, much of the difference in returns across the treated and untreated groups appears to relate to the completion of Grade 12 and university education, not exposure to the CSL on its own.

The coefficient for each completed grade is statistically significant and positive throughout most grades/levels of education and across most percentiles of the wage distribution being considered (except for Grade 8 at Q10 and Q25, and Grade 9 at Q75). Although I do not undertake a formal test, it is also clear that the confidence intervals for coefficients for most of the education levels do not overlap (within a column); thus, the returns differ by grade level.

The return to Grade 9 for untreated female workers is estimated to be approximately 20.6%, on average. It increases from 20.6% at Q10 to 24.3% at Q25, falls at the median to 19.9% and increases again at Q90. The coefficient for the interaction term between Grade 9 and the compulsory schooling dummy is positive and statistically significant only at Q50 and Q75. Thus, I find that the return to Grade 9 is approximately 5.4% to 7.5% higher for those around the second to third quartiles of earnings. As before, the CSL leads to improved earnings for women.

The results for male workers are depicted in the bottom panel of Table 3.3. We also see a number of differences between the results reported in Table 3.2 and those reported in Table 3.3. Specifically, we see evidence that the CSL had led to lower wages at the bottom of the distribution, but higher

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

wages at the top of the distribution. I also find, as I did with women, that the constant returns assumption is not entirely plausible. Furthermore, we do not see any obvious returns patterns across the earnings distribution, even though there is heterogeneity.

Recalling that the CSL is expected to impact returns related to Grade 9, more than for any other grade, I focus my attention on Grade 9 and its interaction with the CSL. Male returns to Grade 9 are slightly lower than those for women: 17.8% for men (20.6% for women). However, returns do not follow either an increasing or decreasing pattern across the distribution, although they are lowest at Q25 (12.9%) and highest at Q90 (29.3%). For men, the CSL had the greatest benefit for the bottom half of wage earners, ranging from an additional 5.0% to 9.4%. Thus, for men, the CSL offered some respite by reducing earnings inequality amongst Grade 9 completers.

3.3.3 Quantile selection results

It is possible that the estimated returns reported to this point are biased, due to the fact that they do not account for selection into employment. Thus, I also estimated quantile selection models for females and males, which I report in Table 3.4. For comparison, I report the results estimated from the Heckman selection model, in addition to the quantile selection results. The probit estimates are estimated with employment as the dependent variable. Only key results are reported here; the rest can be found in Appendix B tables B5 and B6.

Table 3.4: Returns to education for women estimated via the quantile selection model

	Mean		Quantile Selection Model					
	Probit	Heckman	Probit	Q10	Q25	Q50	Q75	Q90
Education	0.048*** (0.006)	0.126*** (0.005)	0.048*** (0.005)	0.094*** (0.007)	0.133*** (0.009)	0.141*** (0.007)	0.138*** (0.007)	0.115*** (0.011)
CSL	-0.848*** (0.253)	0.149 (0.300)	-0.848*** (0.248)	-1.003*** (0.376)	-0.240 (0.364)	-0.097 (0.384)	0.172 (0.681)	0.683 (0.616)
CSL × Education	0.022** (0.011)	0.012 (0.010)	0.022** (0.010)	0.057*** (0.020)	0.017 (0.014)	0.019 (0.012)	0.018 (0.018)	-0.008 (0.021)
Children under 6	-0.161*** (0.017)		-0.161*** (0.017)					
$G_x(\rho)$	-0.634*** (0.187)		0.494** (0.236)					

Notes: All equations include the following additional control variables: age, marital status, and province. Standard errors are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10% critical level. CSL = 1 if individual i was subjected to the compulsory schooling law, CSL = 0 otherwise. In the quantile selection model a positive sign for G_x (or ρ) indicates negative selection into employment (Maasoumi and Wang, 2019; Arellano and Bonhomme, 2017). Negative selection bias implies that potential wages of the working individuals are somewhat lower than those of the entire population (Granados et al., 2020).

The first takeaway from Table 3.4 is that the coefficient on the selection variable is statistically

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

significant only for women, even though the coefficient on the instrument (Children under 6) is statistically significant (and similar in magnitude) for both men (see Table B6) and women. In other words, for women, there is a correlation between the unobserved components of employment and wages; therefore, there is evidence of non-random selection into employment, which biases the estimates of the previously reported quantile regressions. [Staneva et al. \(2022\)](#) find similar evidence for Russia and Kazakhstan. Given that I find this only for women, I only discuss results for them below.

There are two relevant comparisons, one across the quantiles within the selection model, and the other comparing the selection model with the non-selection model. Beginning with the selection model – see Table 3.4 – we continue to see that returns to earnings are increasing from Q10 to Q50, falling slightly at Q75 and Q90. I also find that compulsory schooling has reduced earnings at Q10, but there is no statistically significant CSL effect at other locations across the distribution. At the margin, the CSL offers a small increase in returns to education at Q10, but not at other locations.

Comparatively speaking – see the top panel in Table 3.2 and Table 3.4 – returns to education are slightly lower, after correcting for selection. Furthermore, the treatment effect of the CSL, is estimated to be smaller, as well, and statistically significant only at Q10 following selection correction. Table 3.4 indicates that G_x is positive, which corresponds to the presence of negative selection (refer to the section on the estimation strategy for further details). Similar results are available for the US, UK and Finland ([Elass, 2024](#); [Maasoumi and Wang, 2019](#)). [Maasoumi and Wang \(2019\)](#) find that negative selection among US women happens during periods of low employment rates. This result implies that women who opt out of employment tend to have higher potential wages compared to those who remain employed.

This finding is somewhat unexpected, particularly in the context of South Africa, where structural and societal factors would typically suggest the opposite trend. One possible explanation for this phenomenon could be that higher-earning women may have greater flexibility or alternative sources of income, enabling them to leave the labour market voluntarily. For instance, these women might prioritize other pursuits such as education, caregiving, or entrepreneurial activities.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

On the other hand, lower-earning women may feel compelled to remain employed due to financial necessity, even if the returns to their labour are relatively low. This observation challenges conventional assumptions about labour market participation in South Africa, where economic pressures and limited opportunities for women often drive different patterns of workforce engagement. Further investigation into the underlying causes of this negative selection is necessary to fully understand its implications for gender dynamics in the labour market.

In addition to correcting the returns to education for sample selection bias, I attempted to correct the returns to grade completion for selection into employment. However, the results for both women and men show the absence of selection bias (see Tables B7 and B8). Hence, no further analysis was conducted.

3.3.4 Discussion

A potential threat to identification in this study is the use of the number of children in the household, which may not be strictly exogenous. Branson and Leibbrandt (2013b) find that the number of children is statistically significant in both the employment and wage equations, raising concerns about the validity of the exclusion restriction. Nevertheless, their results indicate that the estimated coefficients on education remain stable in both magnitude and significance, regardless of whether this instrument is included. Accordingly, while the estimates of returns to education appear robust, they should nonetheless be interpreted with caution given the potential endogeneity of the exclusion variable.

Although I generally find that average returns to education amongst the younger (CSL-exposed) cohort are lower, the returns to an additional year of education (or for completing Grade 9, which was directly targeted through the CSL) are estimated to be larger for the same exposed cohort. Depending on whether we look at average returns or returns at particular parts of the earnings distribution, or whether we account for selection into employment, the overall picture is that CSL treatment did not worsen returns for anyone considered in the analysis, and raised returns for at least some; therefore, the CSL offered benefits, at least on the margin, to the population it was meant to help.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

It is difficult to determine whether returns increases are driven by signalling/screening or by productive education. The human capital theory assumes that additional education increases marginal productivity, and, therefore, leads to higher returns to schooling (Becker, 1962; Schultz, 1960). The signalling/screening assumes that additional education comes with a cost that is more easily borne by those that are more productive at school activities, which provides information to the labour market about the individual's potential (Arrow et al., 1973; Spence, 1973; Stiglitz, 1975). Again, this model implies higher returns to education for those with more education. In South Africa, there are certainly concerns over the quality of education (Mlachila and Moeletsi, 2019; Spaull, 2015b).⁷ On the other hand, not completing Grade 9 within the context of a compulsory schooling law that pushes for Grade 9 completion, provides information – that is very likely negative – to the labour market. Thus, I interpret the CSL-exposure $\times$ education interaction results, especially for Grade 9 completion, as more likely driven by signalling.

As described in Table 3.1, those exposed to compulsory education have more education; the one exception to that is tertiary education, which is likely the result of the younger ages of the exposed cohort. Thus, in keeping with both theories, the CSL-exposed have more education, and conditionally, ought to have greater returns. The findings are similar to several other studies that find a positive relationship between compulsory schooling and completed years of education (Korwatanasakul, 2023; Clay et al., 2021; Oyelere, 2010; Angrist and Keueger, 1991; Brunello et al., 2009).

The non-corrected quantile regression overestimates the returns to education for females. The bias ranges from -0.9 and -4.4 percentage points; others have similarly found that the bias is small (Dougherty, 2005; Harmon et al., 2000). Controlling for selection into employment, however, does not alter the basic conclusion: (marginal or average) returns are constant neither across years of schooling nor across the distribution of earnings. The non-linear pattern of returns by

⁷In 2016, South African learners ranked last in reading scores in a sample of Grade 4 learners from 50 countries participating in PIRLS. The Progress in International Reading Literacy Study (PIRLS) is an international assessment and research project designed to measure reading achievement at the fourth-grade level. Even though PIRLS consists of mostly advanced countries, South Africa still performs poorly among nations with comparable economic status. For example, South African learners are outperformed by their counterparts in Botswana, Swaziland (now Eswatini), Seychelles, Kenya and Mauritius in Grade 6 reading and mathematics achievement, despite the fact that these countries spend less on education as a proportion of GDP and in per-student terms (Mlachila and Moeletsi, 2019).

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

level of education is consistent with the literature ([Mwabu and Schultz, 2000](#); [Branson et al., 2012](#)). The pattern can be attributed to increased demand for skilled workers ([Bhorat, Goga and Stanwix, 2014](#)), who are rewarded with higher wages than their lesser qualified counterparts. The results reinforce the idea of complementarity between education and ability in the sense that more able individuals tend to have higher returns to education than less able individuals ([Stefani and Biderman, 2009](#); [Fasih et al., 2012](#)), suggesting that education is more beneficial to individuals with higher ability than those with low ability.

Not counting the CSL-exposure $\times$ education (or Grade 9) effect, returns to education, for the most part, increase from the bottom to the top of the distribution of earnings. Additionally, returns rise by level of education with tertiary education experiencing much higher returns than for lower levels of education. However, that pattern also implies that education increases wage inequality, since those at the top benefit more from education than those at the bottom of the distribution ([Tansel and Bodur, 2012](#); [Branson et al., 2012](#)).

Furthermore, I estimate returns to education for women to be higher than for men, which is common in the literature ([Dougherty, 2005](#); [Rahman and Al-Hasan, 2018](#)). In purely mechanical terms, due to the fact that base wages for women are lower than for men, the lower base is more amenable to larger percentage changes. In order for the amenable to occur, however, it is necessary for higher levels of education to allow women to overcome discrimination and be employed in high-paying traditionally male occupations ([Dougherty, 2005](#)).

Although returns are often increasing, at least through the middle of the distribution, they tend to decline at the 90th percentile, which could be due to earnings under-reporting for those at the top of the wage distribution. Both [Seekings \(2016\)](#) and [Wittenberg \(2017a\)](#) find that the South African earnings data are significantly lower in surveys than the reported earnings in tax assessments, especially at the top end of the earnings distribution. However, tax assessment data do not contain human capital variables and, as such, cannot be used for returns to education research.

3.4 Conclusion

South Africa introduced the compulsory schooling law in 1997, which compels young people of school-going age to be in school up until the last day of the school year in which they turn 15 or reach Grade 9. However, the effectiveness of the reform is limited, due to the high repetition and dropout rates that characterise the education system in the country. Those dropout rates are particularly high for those who have completed the compulsory schooling phase.

In this research, I investigate whether individuals subjected to mandatory schooling benefited, in terms of returns to schooling, compared to those not affected. I estimated returns for Black African men and women, the majority of the population in South Africa, as well as the group that was disadvantaged under the apartheid regime and still experiences the highest rates of poverty and unemployment in the country.

Using conditional quantile regression, as well as an appropriate selection correction model, I estimated the corrected and uncorrected returns along the wage distribution, as well as at the mean. The focus on quantiles, rather than the mean, allows me to uncover heterogeneities in returns across the wage distribution. The analysis is predicated on the 2018 South African General Household Survey.

Although several global studies have examined similar questions, this analysis is unique in several ways. First, it draws comparisons and contrasts between the education returns for individuals subjected to the mandatory schooling reform and those who were not. Second, it corrects for sample selection bias and estimates returns that are consistent. In a country like South Africa, where unemployment is rife, selection into employment might have had a large impact on the resulting estimates. Third, the most recent South African analysis was completed 27 years ago; thus, I offer a timely update; I also extend that analysis by examining returns for both men and women.

I find returns to each year of education, as well as the completion of Grade 9 (which was specifically targeted by the CSL), to be higher, on average, as well as across multiple points of the wage distribution for individuals subjected to the CSL. Furthermore, I find returns from the

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

10th to the 50th percentiles of the wage distribution to generally be increasing, before dropping at the 90th percentile. The results suggest that returns are non-linear both across the wage distribution and the years of schooling completed. Finally, not controlling for selection does yield upward biased estimates of returns; the effect of correcting that bias is highly dependent on the modelling assumptions.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

Appendix B

Table B1: Returns to education for women estimated via the quantile regression

	(1) OLS	(2) Q10	(3) Q25	(4) Q50	(5) Q75	(6) Q90
Education	0.139*** (0.004)	0.093*** (0.007)	0.145*** (0.007)	0.150*** (0.005)	0.143*** (0.005)	0.130*** (0.006)
CSL	-0.565** (0.237)	-1.427*** (0.321)	-0.701** (0.293)	-0.480* (0.253)	-1.064*** (0.282)	-0.293 (0.363)
CSL × Education	0.023** (0.011)	0.067*** (0.019)	0.025* (0.014)	0.026** (0.012)	0.034*** (0.013)	0.021 (0.013)
Age	0.010*** (0.003)	-0.005 (0.003)	0.007** (0.003)	0.016*** (0.003)	0.014*** (0.003)	0.015*** (0.003)
CSL × Age	0.010* (0.005)	0.023*** (0.007)	0.014** (0.006)	0.007 (0.006)	0.020*** (0.007)	0.003 (0.010)
Legally married	0.179*** (0.031)	0.152*** (0.045)	0.147*** (0.032)	0.198*** (0.036)	0.218*** (0.039)	0.093*** (0.034)
Eastern Cape	-0.543*** (0.063)	-0.808*** (0.082)	-0.677*** (0.058)	-0.506*** (0.061)	-0.373*** (0.099)	-0.309*** (0.073)
Northern Cape	-0.201** (0.093)	-0.626*** (0.132)	-0.275*** (0.080)	-0.141 (0.119)	0.011 (0.143)	0.008 (0.107)
Free State	-0.432*** (0.072)	-0.897*** (0.093)	-0.580*** (0.085)	-0.343*** (0.073)	-0.343*** (0.089)	-0.316*** (0.114)
KwaZulu-Natal	-0.366*** (0.060)	-0.635*** (0.084)	-0.413*** (0.055)	-0.279*** (0.058)	-0.230*** (0.079)	-0.206*** (0.069)
North West	-0.276*** (0.075)	-0.531*** (0.125)	-0.363*** (0.074)	-0.242*** (0.073)	-0.142 (0.102)	-0.220** (0.086)
Gauteng	0.041 (0.057)	-0.194** (0.077)	-0.059 (0.047)	0.049 (0.055)	0.077 (0.072)	0.045 (0.071)
Mpumalanga	-0.337*** (0.067)	-0.701*** (0.098)	-0.419*** (0.061)	-0.292*** (0.062)	-0.176* (0.092)	-0.190** (0.079)
Limpopo	-0.501*** (0.065)	-0.898*** (0.095)	-0.668*** (0.066)	-0.440*** (0.063)	-0.313*** (0.084)	-0.257*** (0.077)
Constant	6.623*** (0.149)	6.734*** (0.218)	6.142*** (0.199)	6.133*** (0.167)	6.960*** (0.173)	7.595*** (0.176)
Observations	6875	6875	6875	6875	6875	6875

Notes: In parentheses are bootstrapped standard errors. ***, **, and * indicate significance at the 1, 5, and 10% critical level. CSL = 1 if individual i was subjected to the compulsory schooling law (CSL), CSL = 0 otherwise.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

Table B2: Returns to education for men estimated via the quantile regression

	(1) OLS	(2) Q10	(3) Q25	(4) Q50	(5) Q75	(6) Q90
Education	0.101*** (0.005)	0.092*** (0.007)	0.103*** (0.007)	0.109*** (0.004)	0.108*** (0.004)	0.097*** (0.008)
CSL	-0.542*** (0.195)	-1.012*** (0.282)	-0.604*** (0.211)	-0.338* (0.205)	-0.584*** (0.210)	-0.174 (0.321)
CSL × Education	0.001 (0.008)	-0.006 (0.012)	-0.003 (0.010)	0.001 (0.009)	0.003 (0.008)	-0.001 (0.010)
Age	0.008*** (0.003)	-0.005 (0.003)	0.002 (0.003)	0.015*** (0.003)	0.012*** (0.003)	0.011** (0.005)
CSL × Age	0.016*** (0.005)	0.032*** (0.006)	0.018*** (0.005)	0.011** (0.005)	0.017*** (0.005)	0.004 (0.007)
Legally married	0.311*** (0.028)	0.341*** (0.040)	0.282*** (0.028)	0.323*** (0.030)	0.317*** (0.034)	0.206*** (0.050)
Eastern Cape	-0.630*** (0.057)	-0.917*** (0.076)	-0.690*** (0.064)	-0.500*** (0.055)	-0.582*** (0.069)	-0.552*** (0.099)
Northern Cape	-0.245*** (0.079)	-0.603*** (0.122)	-0.225** (0.093)	-0.086 (0.095)	-0.047 (0.086)	-0.319*** (0.083)
Free State	-0.445*** (0.066)	-0.625*** (0.089)	-0.436*** (0.066)	-0.263*** (0.066)	-0.467*** (0.076)	-0.424*** (0.094)
KwaZulu-Natal	-0.381*** (0.051)	-0.511*** (0.072)	-0.310*** (0.046)	-0.244*** (0.047)	-0.400*** (0.063)	-0.408*** (0.070)
North West	-0.209*** (0.062)	-0.305*** (0.093)	-0.177*** (0.056)	-0.037 (0.063)	-0.198*** (0.068)	-0.259*** (0.076)
Gauteng	-0.027 (0.048)	-0.086 (0.057)	0.011 (0.039)	0.030 (0.042)	-0.122* (0.057)	-0.107* (0.065)
Mpumalanga	-0.230*** (0.058)	-0.323*** (0.077)	-0.197*** (0.051)	-0.130** (0.052)	-0.293*** (0.061)	-0.342*** (0.088)
Limpopo	-0.507*** (0.055)	-0.743*** (0.090)	-0.497*** (0.060)	-0.372*** (0.051)	-0.495*** (0.075)	-0.471*** (0.075)
Constant	7.370*** (0.144)	6.978*** (0.196)	6.975*** (0.162)	6.851*** (0.150)	7.721*** (0.147)	8.503*** (0.270)
Observations	8006	8006	8006	8006	8006	8006

Notes: In parentheses are bootstrapped standard errors. ***, **, and * indicate significance at the 1, 5, and 10% critical level. CSL = 1 if individual i was subjected to the compulsory schooling law, CSL = 0 otherwise.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

Table B3: Returns to grade completion for women estimated via the quantile regression

	(1) OLS	(2) Q10	(3) Q25	(4) Q50	(5) Q75	(6) Q90
Grade 8	0.197** (0.077)	0.074 (0.089)	0.138 (0.093)	0.206** (0.084)	0.265*** (0.099)	0.340* (0.180)
Grade 9	0.206*** (0.070)	0.206*** (0.074)	0.243** (0.095)	0.199*** (0.073)	0.084 (0.070)	0.249** (0.123)
Grade 10	0.480*** (0.055)	0.258*** (0.068)	0.585*** (0.067)	0.468*** (0.060)	0.549*** (0.064)	0.552*** (0.106)
Grade 11	0.520*** (0.049)	0.322*** (0.066)	0.630*** (0.056)	0.532*** (0.056)	0.510*** (0.061)	0.527*** (0.092)
Grade 12	0.875*** (0.045)	0.694*** (0.058)	0.921*** (0.048)	0.843*** (0.055)	0.975*** (0.049)	0.881*** (0.083)
Tertiary	1.698*** (0.049)	1.325*** (0.079)	1.793*** (0.062)	1.961*** (0.056)	1.871*** (0.044)	1.492*** (0.082)
CSL	-0.467** (0.184)	-0.776*** (0.257)	-0.426** (0.201)	-0.310* (0.185)	-0.473** (0.226)	-0.262 (0.371)
CSL × Grade 9	0.021 (0.035)	-0.030 (0.043)	-0.009 (0.046)	0.054* (0.031)	0.075* (0.045)	-0.050 (0.085)
Age	0.005** (0.003)	-0.004 (0.004)	0.003 (0.003)	0.009*** (0.003)	0.010*** (0.003)	0.010*** (0.003)
CSL × Age	0.013** (0.005)	0.026*** (0.007)	0.013** (0.006)	0.009* (0.005)	0.013* (0.007)	0.009 (0.011)
Legally married	0.125*** (0.030)	0.097** (0.044)	0.143*** (0.033)	0.118*** (0.031)	0.118*** (0.033)	0.094** (0.039)
Eastern Cape	-0.565*** (0.064)	-0.716*** (0.086)	-0.702*** (0.060)	-0.542*** (0.062)	-0.385*** (0.072)	-0.414*** (0.092)
Northern Cape	-0.260** (0.092)	-0.533*** (0.130)	-0.336*** (0.113)	-0.178* (0.105)	-0.102 (0.093)	-0.043 (0.112)
Free State	-0.436*** (0.071)	-0.777*** (0.103)	-0.537*** (0.094)	-0.362*** (0.061)	-0.363*** (0.083)	-0.365*** (0.100)
KwaZulu-Natal	-0.416*** (0.061)	-0.661*** (0.089)	-0.472*** (0.056)	-0.335*** (0.052)	-0.258*** (0.066)	-0.289*** (0.086)
North West	-0.314*** (0.074)	-0.547*** (0.144)	-0.395*** (0.070)	-0.257*** (0.061)	-0.199*** (0.077)	-0.244** (0.102)
Gauteng	-0.025 (0.057)	-0.191** (0.080)	-0.090* (0.050)	-0.026 (0.049)	0.065 (0.062)	0.027 (0.089)
Mpumalanga	-0.409*** (0.067)	-0.719*** (0.126)	-0.480*** (0.062)	-0.297*** (0.059)	-0.301*** (0.070)	-0.271*** (0.101)
Limpopo	-0.553*** (0.065)	-0.844*** (0.095)	-0.711*** (0.066)	-0.430*** (0.059)	-0.384*** (0.070)	-0.412*** (0.097)
Constant	7.640*** (0.143)	7.177*** (0.214)	7.117*** (0.157)	7.359*** (0.161)	7.822*** (0.143)	8.516*** (0.200)
Observations	6875	6875	6875	6875	6875	6875

Notes: In parentheses are bootstrapped standard errors. ***, **, and * indicate significance at the 1, 5, and 10% critical level. CSL = 1 if individual i was subjected to the compulsory schooling law, CSL = 0 otherwise.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

Table B4: Returns to grade completion for men estimated via the quantile regression

	(1) OLS	(2) Q10	(3) Q25	(4) Q50	(5) Q75	(6) Q90
Grade 8	0.138** (0.056)	0.153 (0.109)	0.113 (0.077)	0.142*** (0.054)	0.140** (0.068)	0.120 (0.090)
Grade 9	0.178*** (0.069)	0.219** (0.087)	0.129 (0.101)	0.135* (0.075)	0.205* (0.124)	0.293*** (0.096)
Grade 10	0.315*** (0.048)	0.352*** (0.093)	0.354*** (0.057)	0.336*** (0.044)	0.303*** (0.055)	0.285*** (0.068)
Grade 11	0.352*** (0.041)	0.520*** (0.072)	0.450*** (0.054)	0.389*** (0.041)	0.297*** (0.047)	0.206*** (0.065)
Grade 12	0.604*** (0.039)	0.682*** (0.060)	0.660*** (0.049)	0.659*** (0.040)	0.621*** (0.041)	0.532*** (0.054)
Tertiary	1.265*** (0.048)	1.135*** (0.076)	1.236*** (0.067)	1.446*** (0.045)	1.376*** (0.047)	1.240*** (0.063)
CSL	-0.639*** (0.163)	-1.186*** (0.283)	-0.928*** (0.206)	-0.585*** (0.147)	-0.391** (0.188)	0.002 (0.265)
CSL <i>times</i> Grade 9	0.024 (0.029)	0.094** (0.039)	0.082** (0.038)	0.050* (0.028)	-0.014 (0.047)	-0.056 (0.038)
Age	0.005* (0.002)	-0.007 (0.005)	-0.002 (0.003)	0.007*** (0.002)	0.011*** (0.002)	0.011*** (0.004)
CSL <i>times</i> Age	0.019*** (0.005)	0.034*** (0.008)	0.025*** (0.006)	0.017*** (0.004)	0.013** (0.005)	0.001 (0.007)
Legally married	0.283*** (0.027)	0.345*** (0.045)	0.281*** (0.029)	0.318*** (0.028)	0.280*** (0.034)	0.231*** (0.042)
Eastern Cape	-0.630*** (0.057)	-0.776*** (0.087)	-0.672*** (0.073)	-0.565*** (0.058)	-0.615*** (0.065)	-0.638*** (0.081)
Northern Cape	-0.286** (0.077)	-0.396*** (0.135)	-0.251*** (0.090)	-0.225*** (0.068)	-0.084 (0.101)	-0.305*** (0.080)
Free State	-0.436*** (0.066)	-0.536*** (0.149)	-0.385*** (0.082)	-0.368*** (0.064)	-0.479*** (0.075)	-0.422*** (0.101)
KwaZulu-Natal	-0.404*** (0.051)	-0.434*** (0.083)	-0.325*** (0.046)	-0.337*** (0.049)	-0.465*** (0.065)	-0.505*** (0.072)
North West	-0.223*** (0.062)	-0.266** (0.123)	-0.198*** (0.076)	-0.114 (0.072)	-0.193*** (0.072)	-0.285*** (0.078)
Gauteng	-0.057 (0.048)	-0.009 (0.076)	0.013 (0.042)	-0.047 (0.046)	-0.182*** (0.059)	-0.166* (0.070)
Mpumalanga	-0.254*** (0.058)	-0.293*** (0.099)	-0.173*** (0.053)	-0.236*** (0.055)	-0.304*** (0.068)	-0.420*** (0.076)
Limpopo	-0.519*** (0.055)	-0.619*** (0.085)	-0.484*** (0.071)	-0.444*** (0.052)	-0.587*** (0.065)	-0.542*** (0.085)
Constant	8.108*** (0.129)	7.413*** (0.238)	7.764*** (0.156)	7.876*** (0.122)	8.401*** (0.131)	9.056*** (0.203)
Observations	8006	8006	8006	8006	8006	8006

Notes: All equations include survey year as an additional control variable. Standard errors are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10% critical level. CSL = 1 if individual *i* was subjected to the compulsory schooling law, CSL = 0 otherwise.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

Table B5: Returns to education for women estimated via the quantile selection model

	Mean		Quantile Selection Model					
	Probit	Heckman	Probit	Q10	Q25	Q50	Q75	Q90
Education	0.048*** (0.006)	0.126*** (0.005)	0.048*** (0.005)	0.094*** (0.007)	0.133*** (0.009)	0.141*** (0.007)	0.138*** (0.007)	0.115*** (0.011)
CSL	-0.848*** (0.253)	0.149 (0.300)	-0.848*** (0.248)	-1.003*** (0.376)	-0.240 (0.364)	-0.097 (0.384)	0.172 (0.681)	0.683 (0.616)
CSL × Education	0.022** (0.011)	0.012 (0.010)	0.022** (0.010)	0.057*** (0.020)	0.017 (0.014)	0.019 (0.012)	0.018 (0.018)	-0.008 (0.021)
Age	0.046*** (0.003)	0.002 (0.004)	0.046*** (0.003)	-0.008* (0.004)	0.002 (0.004)	0.007 (0.005)	0.006 (0.006)	0.006 (0.007)
CSL × Age	0.021*** (0.006)	-0.007 (0.007)	0.021*** (0.006)	0.014* (0.008)	0.003 (0.008)	-0.003 (0.009)	-0.012 (0.016)	-0.014 (0.013)
Legally married	0.098*** (0.035)	0.145*** (0.030)	0.098*** (0.036)	0.130*** (0.042)	0.088** (0.043)	0.210*** (0.045)	0.146*** (0.052)	0.088 (0.059)
Eastern Cape	-0.011 (0.073)	-0.518*** (0.064)	-0.011 (0.074)	-0.839*** (0.067)	-0.643*** (0.056)	-0.485*** (0.080)	-0.350*** (0.104)	-0.147 (0.140)
Northern Cape	-0.099 (0.101)	-0.128 (0.091)	-0.099 (0.102)	-0.572*** (0.141)	-0.269*** (0.076)	-0.006 (0.146)	0.091 (0.138)	0.036 (0.158)
Free State	-0.242*** (0.080)	-0.367*** (0.076)	-0.242*** (0.081)	-0.857*** (0.101)	-0.437*** (0.100)	-0.232** (0.106)	-0.287*** (0.107)	-0.140 (0.157)
KwaZulu-Natal	-0.145** (0.067)	-0.266*** (0.062)	-0.145** (0.068)	-0.586*** (0.082)	-0.330*** (0.067)	-0.167* (0.091)	-0.120 (0.093)	-0.061 (0.148)
North West	-0.122 (0.083)	-0.207*** (0.076)	-0.122 (0.084)	-0.544*** (0.099)	-0.247*** (0.090)	-0.147 (0.113)	-0.186* (0.103)	-0.076 (0.166)
Gauteng	-0.130** (0.064)	0.077 (0.058)	-0.130** (0.065)	-0.170*** (0.063)	-0.009 (0.043)	0.171** (0.087)	0.083 (0.081)	0.261 (0.194)
Mpumalanga	-0.238*** (0.072)	-0.219*** (0.070)	-0.238*** (0.074)	-0.630*** (0.095)	-0.311*** (0.072)	-0.166* (0.100)	-0.071 (0.095)	-0.007 (0.189)
Limpopo	0.159** (0.076)	-0.504*** (0.066)	0.159** (0.077)	-0.872*** (0.082)	-0.624*** (0.059)	-0.388*** (0.082)	-0.334*** (0.099)	-0.187 (0.121)
Children under 6	-0.161*** (0.017)		-0.161*** (0.017)					
Constant	-1.534*** (0.197)	7.278*** (0.270)	-1.534*** (0.190)	6.977*** (0.251)	6.557*** (0.266)	6.719*** (0.355)	7.570*** (0.400)	8.225*** (0.485)
λ/ρ	-0.634*** (0.187)		0.494** (0.236)					
Observations	9901	9901	9901	9901	9901	9901	9901	9901

Notes: All equations include survey year as an additional control variable. Standard errors are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10% critical level. CSL = 1 if individual i was subjected to the compulsory schooling law, CSL = 0 otherwise.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

Table B6: Returns to education for men estimated via the quantile selection model

	Mean		Quantile Selection Model					
	Probit	Heckman	Probit	Q10	Q25	Q50	Q75	Q90
Education	0.019*** (0.006)	0.100*** (0.004)	0.019*** (0.006)	0.091*** (0.009)	0.101*** (0.008)	0.107*** (0.005)	0.104*** (0.005)	0.091*** (0.008)
CSL	-2.209*** (0.237)	-0.322 (0.333)	-2.209*** (0.228)	-0.827 (0.527)	-0.375 (0.458)	-0.137 (0.447)	-0.367 (0.701)	0.061 (0.633)
CSL × Education	-0.000 (0.010)	0.001 (0.007)	-0.000 (0.009)	-0.006 (0.012)	-0.001 (0.010)	0.004 (0.008)	0.007 (0.012)	-0.003 (0.011)
Age	0.005 (0.003)	0.007*** (0.002)	0.005 (0.003)	-0.006 (0.004)	0.003 (0.003)	0.015*** (0.003)	0.010*** (0.003)	0.008 (0.005)
CSL × Age	0.063*** (0.005)	0.010 (0.009)	0.063*** (0.005)	0.025* (0.015)	0.010 (0.013)	0.004 (0.012)	0.009 (0.018)	-0.003 (0.017)
Legally married	0.674*** (0.041)	0.284*** (0.057)	0.674*** (0.040)	0.297*** (0.069)	0.235*** (0.065)	0.275*** (0.071)	0.247** (0.102)	0.157 (0.114)
Eastern Cape	-0.324*** (0.077)	-0.608*** (0.063)	-0.324*** (0.078)	-0.904*** (0.095)	-0.624*** (0.096)	-0.447*** (0.079)	-0.519*** (0.096)	-0.425*** (0.155)
Northern Cape	-0.284*** (0.103)	-0.180** (0.082)	-0.284*** (0.104)	-0.538*** (0.176)	-0.200* (0.117)	0.047 (0.139)	0.001 (0.102)	-0.251 (0.170)
Free State	-0.381*** (0.085)	-0.401*** (0.072)	-0.381*** (0.086)	-0.619*** (0.119)	-0.357*** (0.099)	-0.207** (0.088)	-0.387*** (0.100)	-0.363*** (0.136)
KwaZulu-Natal	-0.305*** (0.071)	-0.346*** (0.059)	-0.305*** (0.072)	-0.475*** (0.096)	-0.279*** (0.062)	-0.212*** (0.067)	-0.376*** (0.103)	-0.327*** (0.101)
North West	-0.228*** (0.083)	-0.171*** (0.063)	-0.228*** (0.084)	-0.299*** (0.112)	-0.111 (0.078)	-0.011 (0.096)	-0.147* (0.079)	-0.223* (0.120)
Gauteng	-0.209*** (0.069)	-0.027 (0.051)	-0.209*** (0.069)	-0.089 (0.073)	0.035 (0.048)	0.064 (0.058)	-0.099 (0.080)	-0.013 (0.114)
Mpumalanga	-0.358*** (0.077)	-0.211*** (0.065)	-0.358*** (0.078)	-0.328*** (0.094)	-0.146** (0.074)	-0.112 (0.079)	-0.248*** (0.087)	-0.220 (0.152)
Limpopo	0.031 (0.081)	-0.500*** (0.055)	0.031 (0.083)	-0.728*** (0.100)	-0.461*** (0.065)	-0.367*** (0.056)	-0.475*** (0.083)	-0.443*** (0.095)
Children under 6	-0.135*** (0.019)		-0.135*** (0.019)					
Constant	0.680*** (0.196)	7.465*** (0.156)	0.680*** (0.186)	7.110*** (0.263)	7.024*** (0.169)	6.899*** (0.149)	7.905*** (0.240)	8.736*** (0.345)
λ/ρ	-0.201 (0.230)		0.262 (0.319)					
Observations	10639	10639	10639	10639	10639	10639	10639	10639

Notes: All equations include survey year as an additional control variable. Standard errors are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10% critical level. CSL = 1 if individual i was subjected to the compulsory schooling law, CSL = 0 otherwise.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

Table B7: Returns to grade completion for women estimated via the quantile selection models

	Mean		Quantile Selection Model					
	Probit	Heckman	Probit	Q10	Q25	Q50	Q75	Q90
Grade 8	-0.043 (0.082)	0.188*** (0.066)	-0.043 (0.084)	0.011 (0.110)	0.137 (0.089)	0.174** (0.085)	0.246* (0.138)	0.277* (0.151)
Grade 9	-0.175** (0.089)	0.208*** (0.072)	-0.175* (0.090)	0.240*** (0.087)	0.231** (0.116)	0.223*** (0.073)	0.106 (0.134)	0.247* (0.147)
Grade 10	-0.025 (0.061)	0.492*** (0.050)	-0.025 (0.061)	0.288*** (0.071)	0.592*** (0.095)	0.494*** (0.065)	0.512*** (0.078)	0.563*** (0.124)
Grade 11	0.015 (0.057)	0.518*** (0.047)	0.015 (0.057)	0.334*** (0.099)	0.628*** (0.066)	0.553*** (0.066)	0.483*** (0.081)	0.508*** (0.114)
Grade 12	0.208*** (0.052)	0.861*** (0.043)	0.208*** (0.053)	0.697*** (0.068)	0.941*** (0.054)	0.869*** (0.067)	0.958*** (0.066)	0.904*** (0.124)
Tertiary	0.736*** (0.064)	1.699*** (0.063)	0.736*** (0.063)	1.358*** (0.114)	1.799*** (0.076)	2.009*** (0.097)	1.885*** (0.096)	1.543*** (0.177)
CSL	-0.777*** (0.205)	-0.328 (0.225)	-0.777*** (0.201)	-0.782** (0.379)	-0.541* (0.316)	-0.302 (0.224)	-0.529 (0.470)	-0.415 (0.568)
CSL × Grade 9	0.064* (0.039)	0.024 (0.037)	0.064 (0.039)	-0.020 (0.054)	-0.007 (0.052)	0.060 (0.053)	0.053 (0.062)	-0.049 (0.100)
Age	0.042*** (0.003)	0.005 (0.003)	0.042*** (0.003)	-0.002 (0.005)	0.005 (0.005)	0.012** (0.005)	0.012*** (0.004)	0.013** (0.006)
CSL × Age	0.025*** (0.006)	0.010 (0.007)	0.025*** (0.006)	0.025** (0.011)	0.017* (0.009)	0.009 (0.007)	0.014 (0.013)	0.014 (0.017)
Legally married	0.070* (0.036)	0.117*** (0.027)	0.070* (0.036)	0.099** (0.049)	0.147*** (0.046)	0.116*** (0.036)	0.133*** (0.034)	0.114** (0.055)
Eastern Cape	-0.041 (0.074)	-0.552*** (0.059)	-0.041 (0.074)	-0.729*** (0.137)	-0.712*** (0.065)	-0.580*** (0.077)	-0.422*** (0.094)	-0.407*** (0.120)
Northern Cape	-0.148 (0.101)	-0.212** (0.085)	-0.148 (0.102)	-0.516*** (0.156)	-0.384*** (0.121)	-0.244** (0.123)	-0.149 (0.129)	-0.053 (0.155)
Free State	-0.266*** (0.080)	-0.427*** (0.071)	-0.266*** (0.081)	-0.759*** (0.132)	-0.548*** (0.117)	-0.383*** (0.079)	-0.411*** (0.115)	-0.367** (0.168)
KwaZulu-Natal	-0.192*** (0.068)	-0.370*** (0.058)	-0.192*** (0.068)	-0.659*** (0.131)	-0.505*** (0.092)	-0.347*** (0.063)	-0.308*** (0.092)	-0.267** (0.124)
North West	-0.162* (0.083)	-0.284*** (0.071)	-0.162* (0.084)	-0.625*** (0.179)	-0.402*** (0.085)	-0.264*** (0.065)	-0.229** (0.098)	-0.243* (0.129)
Gauteng	-0.175*** (0.065)	-0.012 (0.054)	-0.175*** (0.065)	-0.225* (0.119)	-0.111* (0.059)	-0.050 (0.074)	0.026 (0.082)	0.008 (0.120)
Mpumalanga	-0.287*** (0.073)	-0.360*** (0.066)	-0.287*** (0.074)	-0.732*** (0.172)	-0.496*** (0.096)	-0.335*** (0.081)	-0.364*** (0.099)	-0.294** (0.148)
Limpopo	0.115 (0.077)	-0.540*** (0.060)	0.115 (0.078)	-0.837*** (0.148)	-0.697*** (0.067)	-0.459*** (0.068)	-0.420*** (0.098)	-0.416*** (0.124)
Children under 6	-0.153*** (0.017)		-0.153*** (0.017)					
Constant	-0.942*** (0.184)	7.623*** (0.208)	-0.942*** (0.179)	7.055*** (0.352)	6.985*** (0.312)	7.206*** (0.331)	7.755*** (0.228)	8.325*** (0.445)
λ/ρ	-0.091 (0.166)		-0.128 (0.294)					
Observations	9901	9901	9901	9901	9901	9901	9901	9901

Notes: All equations include survey year as an additional control variable. Standard errors are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10% critical level. CSL = 1 if individual i was subjected to the compulsory schooling law, CSL = 0 otherwise.

CHAPTER 3. IMPACT OF COMPULSORY SCHOOLING ON RETURNS TO EDUCATION

Table B8: Returns to grade completion for men estimated via the quantile selection models

	Mean		Quantile Selection Model					
	Probit	Heckman	Probit	Q10	Q25	Q50	Q75	Q90
Grade 8	-0.055 (0.071)	0.137** (0.055)	-0.055 (0.073)	0.117 (0.112)	0.111 (0.082)	0.158** (0.067)	0.177** (0.083)	0.092 (0.120)
Grade 9	-0.059 (0.093)	0.161** (0.066)	-0.059 (0.095)	0.260** (0.103)	0.163* (0.098)	0.158** (0.071)	0.225** (0.114)	0.294** (0.124)
Grade 10	-0.148*** (0.054)	0.326*** (0.045)	-0.148*** (0.055)	0.278** (0.127)	0.363*** (0.054)	0.329*** (0.049)	0.280*** (0.073)	0.290** (0.125)
Grade 11	0.003 (0.053)	0.376*** (0.040)	0.003 (0.053)	0.485*** (0.081)	0.438*** (0.048)	0.416*** (0.066)	0.307*** (0.063)	0.211** (0.097)
Grade 12	0.097** (0.047)	0.634*** (0.036)	0.097** (0.047)	0.640*** (0.077)	0.684*** (0.054)	0.666*** (0.047)	0.637*** (0.064)	0.544*** (0.073)
Tertiary	0.348*** (0.064)	1.315*** (0.049)	0.348*** (0.062)	1.106*** (0.095)	1.254*** (0.064)	1.466*** (0.064)	1.411*** (0.093)	1.281*** (0.094)
CSL	-2.290*** (0.199)	-0.728** (0.299)	-2.290*** (0.195)	-1.501** (0.703)	-1.230** (0.605)	-0.692 (0.447)	-0.535 (0.523)	-0.330 (0.593)
CSL × Grade 9	0.027 (0.037)	0.039 (0.029)	0.027 (0.037)	0.086 (0.060)	0.070* (0.038)	0.054 (0.041)	-0.013 (0.058)	-0.049 (0.059)
Age	0.003 (0.003)	0.003 (0.002)	0.003 (0.003)	-0.007 (0.005)	-0.003 (0.003)	0.008*** (0.002)	0.012*** (0.003)	0.011** (0.005)
CSL × Age	0.065*** (0.005)	0.021** (0.009)	0.065*** (0.005)	0.044** (0.021)	0.033* (0.017)	0.020 (0.013)	0.018 (0.015)	0.011 (0.017)
Legally married	0.661*** (0.041)	0.318*** (0.053)	0.661*** (0.040)	0.397*** (0.104)	0.322*** (0.088)	0.344*** (0.070)	0.320*** (0.096)	0.264** (0.131)
Eastern Cape	-0.328*** (0.077)	-0.646*** (0.061)	-0.328*** (0.078)	-0.812*** (0.100)	-0.725*** (0.118)	-0.578*** (0.091)	-0.646*** (0.100)	-0.661*** (0.173)
Northern Cape	-0.298*** (0.103)	-0.252*** (0.080)	-0.298*** (0.104)	-0.440*** (0.143)	-0.256** (0.119)	-0.201* (0.118)	-0.148 (0.139)	-0.314** (0.127)
Free State	-0.379*** (0.085)	-0.439*** (0.070)	-0.379*** (0.086)	-0.585*** (0.166)	-0.481*** (0.146)	-0.384*** (0.106)	-0.511*** (0.123)	-0.460*** (0.162)
KwaZulu-Natal	-0.317*** (0.072)	-0.407*** (0.058)	-0.317*** (0.072)	-0.465*** (0.102)	-0.340*** (0.066)	-0.347*** (0.071)	-0.510*** (0.113)	-0.510*** (0.130)
North West	-0.229*** (0.083)	-0.208*** (0.061)	-0.229*** (0.084)	-0.359** (0.151)	-0.241** (0.099)	-0.111 (0.095)	-0.234** (0.099)	-0.270** (0.109)
Gauteng	-0.220*** (0.069)	-0.071 (0.050)	-0.220*** (0.069)	-0.038 (0.093)	-0.006 (0.058)	-0.044 (0.061)	-0.205* (0.095)	-0.194 (0.126)
Mpumalanga	-0.361*** (0.077)	-0.281*** (0.064)	-0.361*** (0.078)	-0.347*** (0.119)	-0.215*** (0.082)	-0.248*** (0.074)	-0.333*** (0.096)	-0.443*** (0.154)
Limpopo	0.024 (0.081)	-0.520*** (0.054)	0.024 (0.083)	-0.646*** (0.097)	-0.490*** (0.084)	-0.430*** (0.063)	-0.589*** (0.088)	-0.540*** (0.102)
Children under 6	-0.130*** (0.020)		-0.130*** (0.019)					
Constant	0.912*** (0.178)	8.105*** (0.136)	0.912*** (0.171)	7.399*** (0.247)	7.762*** (0.149)	7.786*** (0.156)	8.311*** (0.221)	9.019*** (0.253)
λ/ρ	0.057 (0.215)		-0.194 (0.399)					
Observations	10639	10639	10639	10639	10639	10639	10639	10639

Notes: All equations include survey year as an additional control variable. Standard errors are in parentheses. ***, **, and * indicate significance at the 1, 5, and 10% critical level. CSL = 1 if individual i was subjected to the compulsory schooling law, CSL = 0 otherwise.

TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

This study investigates the effect of education on labour market earnings among Black South Africans. The compulsory schooling law and school entry rules permit children to begin school (Grade 1) in the year they turn seven years old as long as they reach that age by 30 June of the admission year. I exploit the exogenous variation in education induced by the school entry rule and estimate the returns to educations using a regression discontinuity design with data from the South African General Household Survey covering the period 2014–2018. This work contributes to the limited body of literature on the effects of school entry policies in Sub-Saharan Africa. The findings suggest evidence of manipulation of birth dates around January 1st, likely to take advantage of school entry rules. To address this issue, I exclude individuals born exactly on January 1st from the estimation sample. This adjustment helps mitigate potential bias in the treatment effect. The findings indicate that starting school early is associated with reduced years of education and labour market earnings. Moreover, among early school entrants, education significantly increased earnings, on average. These findings raise important policy questions regarding the appropriate age for school entry and underscore the need for targeted academic support for younger children who may be disadvantaged.

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

4.1 Introduction

The concept of returns to education is not new. As early as 1776, Adam Smith recognized the economic value of education and skills. In his pioneering book, *An Inquiry into the Nature and Causes of the Wealth of Nations*, he described human capital as the "extraordinary dexterity and skill" acquired through education and learning, noting that individuals in skilled employments earn wages "over and above the usual wages of common labour" to compensate for the cost of acquiring such education and training (Smith, 2023; Chiswick, 2024).

This early insight laid the foundation for the modern theory of human capital, which was later developed by Becker (1962) and Schultz (1960, 1961). The theory provides a foundational framework for analyzing education as an economic investment (Mincer, 1974). According to this theory, education entails the allocation of current resources such as time and money in anticipation of future returns, primarily in the form of higher productivity and increased earnings. The estimation of returns to education is a widely studied topic in the economics of education literature (Ashenfelter and Krueger, 1994; Card, 2001; Psacharopoulos and Patrinos, 2018; Duflo, 2001).

This chapter examines the returns to education of Black South Africans using data from the General Household Survey (GHS) for the years 2014-2018 (Statistics South Africa, 2017a, 2016, 2017b, 2018, 2019). Estimating the effects of education on earnings is a challenging task because of many unobserved variables that are likely to be correlated with both education and earnings, such as ability, motivation, and family connections, which could lead to biased and inconsistent estimates (Card, 2001; Patrinos et al., 2021). To overcome the challenge this analysis follows studies such as Solli (2017), Oosterbeek et al. (2021) Guo et al. (2023) and Balestra et al. (2020) that use school entry rules to estimate the causal effect of education on labour market earnings using regression discontinuity (RD) designs.

In South Africa, the age at which children start primary school is governed by the compulsory schooling law (Republic of South Africa, 1996b, 2024). School entry rules allow children to begin Grade 1 in the year they turn seven years old, provided they reach the required age before 30 June

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

of that year.¹ If they turn the required age after 30 June, they typically start school the following year, in which they turn eight. The RD approach leverages exogenous variations introduced by the school entry law, to identify causal effects. By comparing individuals situated just above and just below the cut-off point, the design provides a more credible framework compared to other approaches such as instrumental variables (IV) and difference-in-differences (DiD) (Guo et al., 2023). The study employs a nonparametric local polynomial regression, which has become the standard method for estimating treatment effects in RD designs (Gelman and Imbens, 2019; Clark, 2023; Liwiński, 2020; Dolton and Sandi, 2017).

Angrist and Krueger (1992) pioneered the use of school entry laws as a source of exogenous variation in education for causal inference. Despite its influence, there remains a limited body of research estimating the long-run effects of school starting age (SSA) on final educational attainment and labour market outcomes, particularly wages (Dobkin and Ferreira, 2010; Kawaguchi, 2011; Fredriksson and Öckert, 2014; Larsen and Solli, 2017; Solli, 2017; Görlitz et al., 2025).

The literature on this topic finds mixed results. Most such studies use reduced form estimates with some finding that older children earn more, on average, than their younger counterparts in countries like China (Guo et al., 2023), Brazil (Matta et al., 2016; Peña, 2017) and Mexico (Peña, 2017) while others such as Oosterbeek et al. (2021) in the Netherlands find that younger cohorts earn more, on average. Yet, others find that starting school early has no statistically significant effect on earnings in the United States (Dobkin and Ferreira, 2010), Finland (Pehkonen et al., 2015) and Norway (Larsen and Solli, 2017).

However, to the best of my knowledge, no study has estimated the causal effect of age at school entry on adult earnings using a regression discontinuity design within the Sub-Saharan African context, a region characterized by distinct social, economic, and cultural environments. Similar analyses include Oyelere (2010) for Nigeria and Sakaue et al. (2024) for Uganda, both of which used universal primary education (UPE) to identify the effect of education on earnings using IV methods. This chapter addresses this gap by leveraging the exogenous variation introduced by

¹Grade R (the reception year) was made compulsory only in 2024, following the enactment of the Basic Education Laws Amendment Act (Republic of South Africa, 2024).

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

school entry laws to identify the causal impact of education on labour market earnings in South Africa, a country marked by exceptionally high income inequality (Wittenberg, 2017b; Mosomi et al., 2020).

A key threat to identification in this context is that parents may manipulate the timing of births to benefit from school entry rules, as documented in studies from China (Huang et al., 2020), Argentina (González and Dip, 2024), and Norway (Beck, 2025). Such endogenous sorting around the cut-off violates the continuity assumption of the RD design and can bias estimates if not properly addressed (Barreca et al., 2011). To mitigate this concern, this study implements a donut RD approach, excluding observations that fall directly on the cut-off date and are therefore most likely to be affected by manipulation (Barreca et al., 2011; Fan et al., 2018) as we will see later.

South Africa presents a unique context for this analysis. The racial discriminatory policies and practices of the apartheid regime have lasting negative impacts on socioeconomic conditions of the majority non-white population (Thobejane, 2013; Gradín, 2019). Despite the various labour laws that have been enacted since the end of apartheid in 1994 to eliminate discriminatory gender and racial practices, significant socioeconomic challenges persist, particularly among the Black population (see Chapter 1), which constitutes 81.4% of the total population (Statistics South Africa, 2023). The country is one of the most unequal countries in the world, with a wage Gini coefficient of 0.634 (Wittenberg, 2017b; Mosomi et al., 2020). As such Black South Africans experience the highest unemployment rate of 46% (see Figure 1.1) and the poverty rate at 71% (Kollamparambil and Razak, 2016; Finn, 2015; Francis and Webster, 2019).

In South Africa, several empirical analyses have estimated the returns to education using either ordinary least squares (OLS)², fixed effects (Hertz, 2003), or two-stage least squares applying various instruments for education including parental education (Depken et al., 2019), lagged value of the household head's education (Biyase and Zwane, 2015; Zwane, 2020), and quarter of birth (Nikolov et al., 2020). However, Hertz (2003) and Nikolov et al. (2020) are not

²Moll (1992, 1998), Mwabu and Schultz (1996, 2000), Bhorat (2000), Keswell and Poswell (2002, 2004), Fryer and Vencatachellum (2005), Bhorat and Goga (2012), Branson and Leibbrandt (2013b), Branson et al. (2013) and Salisbury (2016).

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

representative³ while [Depken et al. \(2019\)](#), [Biyase and Zwane \(2015\)](#) and [Zwane \(2020\)](#) use weak identification strategies.⁴ Moreover, the gender wage gap literature ([Mosomi, 2019b,a](#); [Kollamparambil and Razak, 2016](#); [Hinks, 2002](#)) often use methodologies that are not designed to control for endogeneity bias.

The results reveal a spike in births on 1 January, suggesting strategic behaviour by parents aiming to take advantage of the school entry cut-off. To address potential manipulation, a donut RD design is employed, excluding observations born on 1 January. The analysis shows that early school entry is associated with reduced educational attainment, with early starters completing between 9 and 12 months less schooling, on average. This educational disadvantage translates into significantly lower monthly earnings, between 16% and 25% on average. However, for early entrants, each additional year of education is associated with a 20% to 30% increase in monthly earnings, highlighting the substantial returns to education even within this group.

The effect of education on earnings varies by year and gender. Across years, the highest returns are observed in 2014, followed by 2016 and 2017, while no statistically significant effects are found in the other years. Regarding gender, statistically significant returns are observed only for women in 2017, at 23%. For men, the estimated returns to education are not statistically significant in any year.

The remainder of the chapter is outlined as follows: section 4.2 discusses the literature, followed by section 4.3, which explains the estimation strategy, whilst section 4.4 presents and discusses the results, concluding with section 4.5.

4.2 Literature review

[Angrist and Krueger \(1992\)](#) were among the earliest to estimate the effect of school entry age on educational attainment. Following this pioneering work several studies have used school

³[Hertz \(2003\)](#) focuses only on a small sample of Blacks in the KwaZulu-Natal province whilst [Nikolov et al. \(2020\)](#) focus on the Cape area of the country.

⁴Although parental education is correlated with child education, it can also be directly correlated with wages, due to increased parental social capital, for example. The lagged value of the household head's education (derived from a relatively short panel) will only differ from household head education at the beginning of the panel, if the household head has chosen to complete additional schooling, and, therefore, may not be appropriately exogenous.

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

entry laws for identification of the causal effect of education on labour market earnings in different countries using either IV, DiD or RD (Guo et al., 2023; Li et al., 2022; Matta et al., 2016; Kawaguchi, 2011). The current study follows this long standing tradition to identify the causal effect of education on labour market earnings in South Africa using the RD design.

Relative to other quasi-experimental designs like DID and IVs, RD demonstrates greater robustness, necessitating only the less stringent identifying assumptions of randomly distributed observations and a continuous running variable (Dai et al., 2022). This method exploits exogenous variations created by policy cut-offs, such as school entry laws, to identify causal effects. This design offers a more credible approach by comparing individuals who are just above and just below the cut-off, assuming that these individuals are similar in all respects except for the treatment induced by the policy. Although the RD design demonstrates strong internal validity, its external validity is generally considered relatively weak. This implies that RD estimates should not be generalized beyond the local region around the cut-off (Lewis and Nguyen, 2020).

Empirical analyses employing RD designs to examine the impact of school entry age on educational and labour market outcomes report mixed findings. In Japan, Kawaguchi (2011) finds that older children tend to obtain 0.17 (males) and 0.08 (females) more years of education than their younger peers, which translates into higher annual earnings among males (3.9%). Similar patterns are observed in Norway by Solli (2017), who reports that relatively older children achieve higher grade point averages (about 20% of a standard deviation) and earn 4% more, on average. In Brazil, Matta et al. (2016) show that delayed school entry improves aptitude test scores and raises earnings by between 15% and 22%. In Mexico, younger cohorts of women and men are 9.2% and 6.4% less likely to have attained college education and their earnings are 3.8% and 1.9% lower than the older cohort (Peña, 2017). In China, Guo et al. (2023) find that individuals born in or after the school entry cut-off (older cohort) experience increases of 0.53 in years of schooling and 59% in annual earnings, particularly in non-agricultural sectors.

In contrast, several studies report null or negative results. Li et al. (2022), also using data from China although applying within-twin fixed effects, find that although late enrollment reduces years of schooling by 0.51 years, it does not have a statistically significant effect on earnings.

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Similarly, [Dobkin and Ferreira \(2010\)](#), [Pehkonen et al. \(2015\)](#), [Balestra et al. \(2020\)](#), [Larsen and Solli \(2017\)](#) and [Solli \(2017\)](#) for the United States, Finland, Switzerland, and Norway all find no significant relationship between school entry age or month of birth and labour market outcomes, respectively. For Sweden, [Fredriksson and Öckert \(2014\)](#) find that although school entry age raises educational attainment, it does not affect prime-age earnings on average. However, it increases earnings for individuals with low-educated parents. [Oosterbeek et al. \(2021\)](#), studying the Netherlands, show that while school entry age has no effect on educational attainment, relatively younger cohorts have higher earnings before the age of 40, primarily because they enter the labour market earlier and accumulate more work experience.

There is an ongoing debate in the RD literature regarding the appropriateness of using global polynomial approximations of the running variable, as opposed to local linear or quadratic methods. This issue has drawn particular attention in the context of compulsory schooling reforms. A prominent example is the United Kingdom's 1947 reform, which raised the minimum school-leaving age from 14 to 15. [Oreopoulos \(2006b\)](#) analyses this reform using a global polynomial approximation of year of birth (the running variable) with data from the General Household Survey (GHS) spanning 1984–1998. He reports a sizable return to education of approximately 15%. However, this estimate has been subject to significant scrutiny.

[Devereux and Hart \(2010\)](#) re-analyse the same reform using the same dataset, but for a broader period (1979–1998) and employing a local linear RD approach. They find substantially smaller returns, an average of 3% for the overall sample, 4–7% for men, and no evidence of positive returns for women. More recently, [Clark \(2023\)](#), using the same 1947 reform but a different dataset, the Lifetime Labour Market (L2) administrative data from 1978–2009, and applying a local linear regression, find no returns attributable to the reform (0%).

However, [Gelman and Imbens \(2019\)](#) strongly caution against the use of high-order global polynomials in RD analysis, arguing that they can lead to overfitting, poor extrapolation near the boundaries, and highly sensitive and unstable estimates. Instead, they advocate for the use of local linear or quadratic estimators with appropriate bandwidth selection and robust bias correction methods. This recommendation has since become widely influential in the empirical literature,

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

prompting a shift away from global polynomials toward more localized, flexible approaches in identification strategies.

A key identifying assumption in RD designs is that the running variable, in this case, date of birth, is not manipulable and thus exogenously assigned around the cut-off. However, evidence from other settings suggests otherwise. In China, [Huang et al. \(2020\)](#) find that advantaged mothers may manipulate the timing of birth to ensure early school entry. This endogenous sorting violates the RD design's continuity assumption and may bias causal estimates if not accounted for. Similar findings have been reported in Argentina where [González and Dip \(2024\)](#) find that parents defer their children to take advantage of the schooling system and Norway where [Beck \(2025\)](#) finds that 10.1% of births are strategically shifted to fall in the week after the school entry cut-off, rather than the week before, a behaviour which is significantly higher among parents with higher educational attainment.

[Barreca et al. \(2011\)](#) demonstrate that when manipulation occurs near the threshold, the estimated treatment effect becomes sensitive to the exclusion of observations close to the cut-off, thereby weakening the reliability of RD estimates. In response, researchers have increasingly employed the so-called "Donut RD" design as a robustness check, which involves excluding observations in a narrow window around the threshold to test for sensitivity to local manipulation. For instance, [Fan et al. \(2018\)](#) apply this method in the context of estimating returns to education and find that excluding observations within 0.01 and 0.001 of the cut-off does not materially affect the treatment effect, thus reinforcing the credibility of their estimates.

However, in Sub-Saharan Africa, evidence on the effectiveness of school entry laws remains limited. Related studies include [Oyelere \(2010\)](#) and [Sakaue et al. \(2024\)](#), which focus on Nigeria and Uganda, respectively. These analyses are different from the current study as they use universal primary education (UPE) to identify the causal effect of education. [Oyelere \(2010\)](#) takes advantage of Nigeria's UPE reform of 1976 to generate an IV that is applied to estimate the causal effect of education on earnings, finding a return to education of 2.8%, on average. Similarly, [Sakaue et al. \(2024\)](#) apply an IV approach by using Uganda's UPE policy of 1997 to identify the effect of education on real consumption expenditure, finding returns of 4.8% and 5.2% for males

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

and females, respectively.

In summary, the literature presents mixed evidence on the impact of school starting age on labour market outcomes. While some studies find positive and significant effects on earnings, others report statistically insignificant results. These discrepancies can be attributed to differences in context, institutional features of education systems, empirical specifications, the functional form of the running variable, and the quality and scope of datasets used. This study contributes to the literature by providing new evidence from a Sub-Saharan African country, offering insights within a unique institutional and socioeconomic context.

4.3 Estimation strategy

This section presents the empirical strategy used to estimate the causal impact of school entry age on labour market earnings. The data and variables used in the analysis are also discussed.

4.3.1 Fuzzy RD

Children are eligible to enter Grade 1 if they turn seven years old on or before 30 June of the admission year. This school-entry cut-off generates a discontinuity in educational attainment, which I exploit using a fuzzy regression discontinuity design. Although individuals born before 30 June and who turn seven are more likely to be influenced by the admission policy compared to those born later in the year, not everybody mandated to attend school actually did, particularly in rural areas. Enrollment in those localities tends to be delayed due to poor socioeconomic conditions ([van der Berg et al., 2025](#)).

The Imperfect compliance with the admission policy aligns with the Fuzzy Regression Discontinuity (FRD) Design approach: there's a discernible increase in treatment probability for individuals born on or before 30 June, yet it's not absolute. Hence, this study applies the FRD to evaluate the causal effect of the admission rule on labour market earnings in South Africa. The main approach of this analysis is to combine the date of birth and the policy to construct an indicator variable, which is used as an instrument for the probability of being treated.

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Let Y_i represent the labour market earnings of individual i , x the assignment or running variable (birthdays), $\tilde{x}$ birthdays centered on 1 January and x_0 the cut-off date of birth, i.e. 1 January. Following [Balestra et al. \(2020\)](#), I estimate a nonparametric local polynomial regression with the data generating process expressed as

$$S_i = \alpha_0 + \alpha_1 \tilde{x} + 1(x > x_0) \cdot \alpha_2 \tilde{x} + \pi U_i + \epsilon_i \quad (4.1)$$

$$Y_i = \beta_0 + \beta_1 \tilde{x} + S_i \cdot \beta_2 \tilde{x} + \tau V_i + \varepsilon_i \quad (4.2)$$

where S_i , is years of schooling of individual i , U_i and V_i are pre-determined covariates, which include province of residence at the time of the survey, marital status, age group, sex, language of interview and relation to head of household, ϵ_i and ε_i are error terms for the treatment and earnings equations, respectively. $1(x > x_0) = 1$ for individuals born between 1 January and 30 June and 0, for individuals born between 1 July and 31 December.

Equations 4.1 and 4.2 are estimated by two-stage least squares using the `rdrobust` command in Stata suggested by [Calonico et al. \(2017\)](#). Given that $\tilde{x}$ is exogenous and statistically significant, the procedure ensures that the return to education (β_2) is unbiased; however, the interpretation of β_2 changes. The interpretation is that β_2 represents the average marginal benefit from exposure to early school entry for those who were affected by the admission policy ([Harmon et al., 2000](#)). This change in interpretation is due to the fact that IV estimates local average treatment effects (LATE), instead of the average treatment effects (ATE).

To estimate β_2 , a local polynomial RD is preferred to the global RD ([Clark, 2023](#)). The problems associated with the global approach include noisy estimates, sensitivity to the degree of the polynomial, and poor coverage of confidence intervals ([Gelman and Imbens, 2019](#)). This implies that the observations far away from the cut-off might influence the estimate of β_2 .

One of the most crucial decisions to make is selecting the bandwidth, which, in this context, refers to the number of birthdays on either side of the cut-off date that should be included when estimating β_2 . Although a larger bandwidth can enhance the precision of the treatment estimate, it may also introduce more bias. This is because accurately modeling the regression slopes on either side of the cut-off becomes more challenging with a larger bandwidth. If these

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

slopes are estimated incorrectly, the treatment effect will be biased. Hence, I rely on recent recommendations and advancements in RD methodology and use a data-driven, non-parametric approach suggested by [Skovron and Titiunik \(2015\)](#). Specifically, I apply the method developed by [Calonico et al. \(2014a,b\)](#) to determine the optimal bandwidth.

Another important decision involves selecting the kernel function, which determines the weight given to observations near the cut-off. I use a triangular kernel, the standard choice in the literature ([Skovron and Titiunik, 2015](#)). With a triangular kernel, observations further from the cut-off receive less weight whilst those close to the cut-off date on either side receive more weight.

Estimating the LATE involves using date of birth as an instrument for treatment received, relying on three key assumptions. Firstly, there should be a discontinuity in the probability of exposure to treatment at the reform's cut-off point. Secondly, the assignment variable was not manipulated by the units. Although, it is reasonable to assume that birthday was not manipulated in anticipation of the policy, there is evidence that parents can act strategically if they believe that such a move will benefit their child ([González and Dip, 2024](#); [Huang et al., 2020](#)). Hence, the [McCrary \(2008\)](#) density test is used to detect manipulation.

4.3.2 Data

To estimate the impact of education on labour market earnings, this study relies on data from the GHS surveys for the years 2014-2018. Although the GHS is a cross-sectional survey, since 2015 some respondents have been re-interviewed in subsequent years, which gives it an unbalanced panel structure. Hence, I estimate equation [4.1](#) and [4.2](#) for each survey year, but also show the pooled estimates for comparison purposes.

The choice of the sample years (2014-2018) is due to the availability of full birth date information for all individuals of interest. Attempts to get the date of birth information for 2019 from Statistics South Africa were not successful as the data were not readily available. I did not use data from 2020 onward due to Covid-19, which resulted in massive layoffs and in the process altering employment patterns and wages ([Ranchhod and Daniels, 2021](#)). The analysis is limited to Black Africans aged 15 to 65 years who are employed in either the formal or informal sector and receive

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

a wage, commission, or salary. This specific age group is chosen because they constitute the labour force and are likely to have not yet reached retirement age.

Earnings, which is the outcome variable, is measured as monthly gross earnings. It includes overtime, allowances, bonuses and any deductions before tax. Previous studies, which used gross monthly earnings include [Keswell and Poswell \(2004\)](#), [Aydemir and Kirdar \(2017\)](#), and [Mishra and Smyth \(2015\)](#). The main variable of interest is the number of years of completed schooling. Other studies that have used this variable include [Depken et al. \(2019\)](#), [Balestra and Backes-Gellner \(2017\)](#) and [Mishra and Smyth \(2015\)](#).

The years of education were derived from the highest level of education. Each completed grade was assigned one year of education. Thus, a respondent who completed primary school was allocated seven years of education whilst somebody who completed Grade 12 was assigned twelve years of completed education. However, the South African National Qualification Framework (NQF) was used to allocate years of education to vocational and tertiary qualifications. Table A1 in Appendix A depicts how the variable was constructed.

Birthdays is used to divide the sample into control and treated groups based on the school entry cut-off date of 1 January. Several studies have used school starting age as an instrumental variable or as a running variable to mitigate for the endogeneity of education ([Guo et al., 2023](#); [Peña, 2017](#); [Matta et al., 2016](#)).

I include province of residence to account for geographical variation in earnings. It is coded from 1 to 9 representing the 9 provinces of South Africa. Similar studies that have used a geographical variable include [Depken et al. \(2019\)](#), and [Biyase and Zwane \(2015\)](#) for South Africa, and [Stefani et al. \(2006\)](#) for Brazil. I also account for marital status, as it has been shown to correlate with earnings in various studies ([Biyase and Zwane, 2015](#); [Bardasi and Taylor, 2008](#); [Depken et al., 2019](#))⁵. Moreover, I also include age group to account for the fact that individuals who share the same birthday come from different birth cohorts.

⁵The country recognizes multiple legal marriage types, including traditional/customary and civil/religious marriages. Consequently, all of these are categorized as "legally married," with the reference group being those who are "not legally married." While marital status could be endogenous, my primary focus is not on the impact of marriage on wages; therefore, I do not address this potential concern here.

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Initially, I examined whether the running variable was correlated with predetermined covariates, specifically, the language of interview and the respondent's relationship to the household head. The language variable includes categories such as Zulu, Xhosa, Ndebele, Swati, Sesotho, Setswana, Tsonga, Venda, English, Afrikaans, and Sepedi. The relationship to household head variable includes head, spouse, child, sibling, parent, grandchild, grandparent, other relative, and non-relative. Upon analysis, I found that some language and relationship categories were significantly associated with the running variable, suggesting potential imbalance. As a result, I included these covariates in the relevant models to improve the validity of the RD estimates.

4.4 Results

This section presents the main findings on the effects of education on labour market earnings using school entry age as an instrument for education in a regression discontinuity design. Results are shown for the overall sample, and separately by gender, over the period 2014–2018. The section proceeds with descriptive statistics, manipulation checks, RD estimates for educational attainment and log earnings, followed by robustness and sensitivity tests.

The GHS 2014-2018 has repeated observations throughout the period, however, not all individuals were interviewed in each of the years, making it an unbalanced panel. To avoid the bias, which may arise due to the unbalanced panel nature of the data, I provide RD estimates for each survey year. For comparison purposes, I also show results for the pooled sample. Two estimation strategies are undertaken, the standard fuzzy RD, which may be affected by manipulation around the cut-off, and the donut RD, which excludes observations near the cut-off to correct for potential manipulation.

4.4.1 Descriptive statistics

In this section I present and discuss the descriptive statistics for both men and women (see Table 4.1). The total sample consists of 69,673 individuals of which 37,352 (53.6%) are men and 32,321 (46.4%) are women. There are more men than women in the sample because there is a higher proportion of men who are employed compared to women. However, it should be noted

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

that the actual number of observations used in the econometric analysis will be limited by the bandwidths. Overall, the treated group consist of 35,614 (51.1%) individuals compared to 34,059 (48.9%) workers in the control group. A comparison is made between the control and treated groups.

Table 4.1: Summary statistics of key variables for all individuals, women and men

	All individuals			Women			Men		
	Control	Treated	Difference	Control	Treated	Difference	Control	Treated	Difference
	Mean	Mean	b	Mean	Mean	b	Mean	Mean	b
Age (years)	39.02	39.46	-0.44***	39.88	40.34	-0.46***	38.25	38.71	-0.46***
Education (years)	10.38	10.08	0.30***	10.59	10.29	0.30***	10.20	9.91	0.29***
Real Wage (Rands)	6079.32	5833.01	246.31***	5649.81	5467.62	182.19**	6461.40	6140.92	320.47***
Female	0.47	0.46	0.01***	1.00	1.00	0.00	0.00	0.00	0.00
Legally married	0.31	0.32	-0.01**	0.27	0.28	-0.01*	0.34	0.34	-0.00
Metropolitan	0.39	0.38	0.01***	0.39	0.37	0.01**	0.40	0.38	0.02***
Household size	3.76	3.80	-0.03*	4.18	4.26	-0.08***	3.39	3.41	-0.02
Western Cape	0.06	0.06	0.00	0.06	0.05	0.01***	0.06	0.06	-0.00
Eastern Cape	0.23	0.23	0.00	0.26	0.25	0.01	0.21	0.21	-0.00
Northern Cape	0.03	0.03	-0.00	0.03	0.03	0.00	0.03	0.04	-0.00*
Free State	0.08	0.07	0.00**	0.07	0.07	0.00	0.08	0.07	0.01***
KwaZulu-Natal	0.16	0.16	0.01**	0.17	0.17	0.00	0.16	0.15	0.01**
North West	0.07	0.08	-0.00*	0.07	0.07	0.00	0.08	0.08	-0.01**
Gauteng	0.28	0.27	0.01**	0.26	0.26	0.01	0.29	0.28	0.01**
Mpumalanga	0.11	0.11	-0.00	0.10	0.11	-0.01**	0.11	0.11	0.00
Limpopo	0.10	0.11	-0.02***	0.10	0.12	-0.02***	0.09	0.11	-0.02***
Observations	34059	35614	69673	16034	16287	32321	18025	19327	37352

Note: Difference represents the difference in means between the control and treated groups. b is the coefficient of the difference in means between the control and treated groups. ***, **, and * indicate statistical significance at the 1, 5, and 10% critical levels.

The treated group is slightly older compared to the control group with a statistically significant difference of 0.44 years between them. However, the latter is more educated with an additional 0.3 years of schooling compared to the former. Hence, the control group has higher earnings with an additional R246.31, on average per month compared to the treated group. Additionally, the control group has a higher proportion of workers living in metropolitan areas. The earnings of women and men mirrors that of the overall sample with the control group earning higher on average compared to their treated counterparts. As with the overall sample, this difference is statistically significant seems to be driven by differences in years of education as one would expect.

Table 4.1 indicates also that there are gender differences in wages with women earning less than men, on average, despite having more years of education, on average. The monthly average earnings for untreated women is only 87.4% of untreated men's average monthly earnings whilst

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

the monthly average earnings of treated women is just 89.0% of their treated male counterparts. One possible reason for this difference in wages could be the over-representation of women in low-paying jobs due to a historical patriarchal society (Espi et al., 2019; Mosomi, 2019b). It is observable also that there are differences in the proportion of treated and untreated women and men across the 9 provinces.

4.4.2 Evidence of manipulation

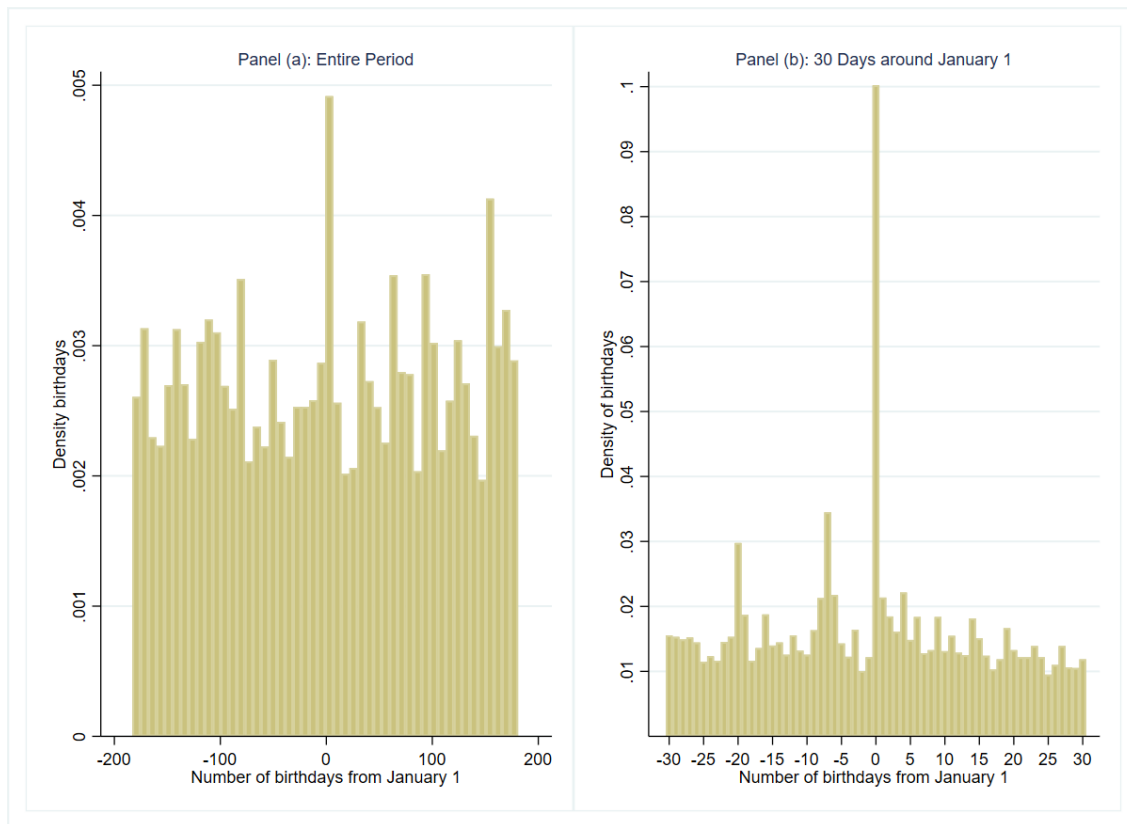
As stated earlier, the RD design is based on the assumption that birthdays, which is the running variable, cannot be manipulated by participants. However, this analysis shows evidence of manipulation. To illustrate this strategic behaviour, I present two histograms displaying the distribution of birth dates. One histogram shows the distribution of birth dates for the overall (or full) sample, while the other focuses on the period between 1 December and 30 January, where manipulation is suspected to have occurred. Figure 4.1 shows the two histograms side-by-side. Panel (a) of the figure depicts the histogram of birth dates for the entire sample while panel (b) focuses on the suspected period covering 30 days before and after the actual cut-off date (1 January).

Panel (a) shows heaping of birth date at the cut-off. However, to show a more clear picture of the problem I present a histogram depicting birthdays within 30 days around 1 January on either side of the threshold. Panel (b) indicates a clear evidence of heaping on January 1, which implies that some births were possibly misreported or registered to fall just at the cut-off, likely to meet the school entry age requirements. This behaviour is non-random and violates the RD assumption of no precise control over the running variable near the cut-off. I conduct the McCrary manipulation test to formally investigate the issue.

This test involves examining the distribution of the running variable (birthdays) around the threshold to ensure that there is no manipulation or sorting just above or below the cut-off. The McCrary density test involves testing for a significant jump in the density of the running variable at the threshold (McCrary, 2008). A smooth density function indicates that individuals did not manipulate their position relative to the threshold. The test is implemented using the local

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Figure 4.1: Birth date heaping around January 1



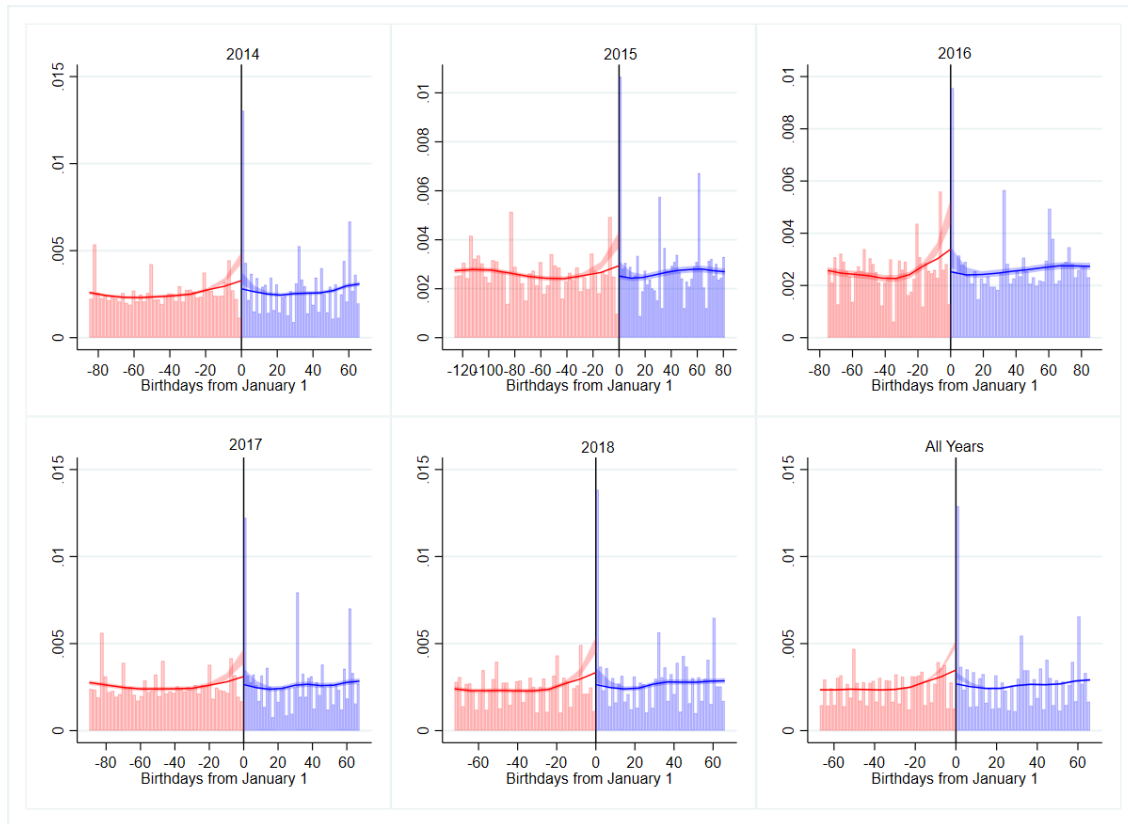
Source: Author's calculations based on the 2014-2018 GHS data.

polynomial density estimators proposed in [Cattaneo, Jansson and Ma \(2020\)](#). It is designed to test the continuity assumption around the cut-off point. It offers a clear and easy way to determine empirically if units near the cut-off point can change their assignment to treatment.

As noted before, units are assigned to the control group if born on or between 1 July and 31 December (or $x < x_0$) and to the treatment group if born between 1 January and 30 June (or $x \geq x_0$). The main aim is to test whether the density of birthdays is continuous at the cut-off point (1 January) using the control and treatment groups. Thus, the null hypothesis is that the density of birthdays is continuous around the cut-off point (no manipulation) against the alternative hypothesis that the density is discontinuous (manipulation) ([Cattaneo, Jansson and Ma, 2020](#)). The null hypothesis is rejected if the p-value related to the test is less than a chosen level of significance. The results are depicted in Figure 4.2. Figure 4.2 shows a big jump just at January 1 for each of the survey years, suggesting that birthdays is not continuous around the

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Figure 4.2: Manipulation test for birthdays: 2014-2018



Note: Source: Author's calculations based on data from the GHS 2014-2018.

Table 4.2: McCrary manipulation test

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Full sample						
Birthdays	1.476 [0.140]	1.502 [0.133]	-0.686 [0.493]	0.701 [0.484]	1.549 [0.121]	1.106 [0.269]
Female sample						
Birthdays	-1.384 [0.166]	-1.823* [0.068]	-1.516 [0.130]	-1.581 [0.114]	0.121 [0.903]	-1.984** [0.047]
Male sample						
Birthdays	1.488 [0.137]	-2.307** [0.021]	-1.062 [0.288]	1.975** [0.048]	0.971 [0.332]	3.6955 [0.000]

Note: The coefficients for birthdays are robust test statistic for the McCrary density test. Robust p-values in square brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical levels. The optimal bandwidths were estimated via the `rdwselect` command in Stata (Calónico et al., 2014a).

cut-off. However, a formal density test suggests otherwise. The test results are depicted in Table 4.2 for each year under consideration as well as for all years combined. For the overall sample, the results indicate that the null hypothesis of no manipulation is not rejected for each of the years as well as for the pooled sample, suggesting that manipulation is not a problem. Similarly, for the male and female sub-samples, the p-value of the test in each year is greater than the conventional significance levels except for 2015. The fact that the p-value is greater than the

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

level of significance suggests that the null hypothesis of continuity at the cut-off point cannot be rejected, indicating that birthdays is not manipulated.

Although the formal density test is not statistically significant for the overall sample in each year and for the sub-samples in most years, the visible heaping in birth dates at the cut-off is consistent with strategic behaviour by parents. Although the McCrary test does not reject the null hypothesis of continuity in density, its power is limited with discrete running variables (Barreca et al., 2011). Given the evidence of potential manipulation around the cut-off, I adopt a conservative approach by implementing a donut RD design. Following best practices in the literature (Barreca et al., 2011; Fan et al., 2018), I exclude observations with a date of birth exactly on the cut-off (January 1). To assess the robustness of the results, I extend the exclusion window incrementally to ± 1 and ± 2 days around the cut-off. For transparency, I report both the baseline RD estimates including all observations and the donut RD estimates that account for potential manipulation.

4.4.3 Impact on Education

As outlined in the estimation strategy, the fuzzy RD design employs school entry age as an instrumental variable for years of education. School entry age is proxied by the running variable, the number of birthdays from the school entry cut-off (January 1). For the fuzzy RD approach to be valid, the first stage, the effect of school entry age on years of education, must be statistically significant to justify estimating the second stage, which captures the impact of years of education on earnings. Accordingly, I estimate the relationship between school entry age (SEA) and years of education to assess the strength of the instrument. Since province, marital status, age group, and sex exhibit discontinuities at the cut-off (see Appendix C), I include these variables as controls in the analysis. Additionally, where appropriate, I control for the language of interview and the respondent's relationship to the household head. I first present the overall sample estimates, followed by a gender-disaggregated discussion.

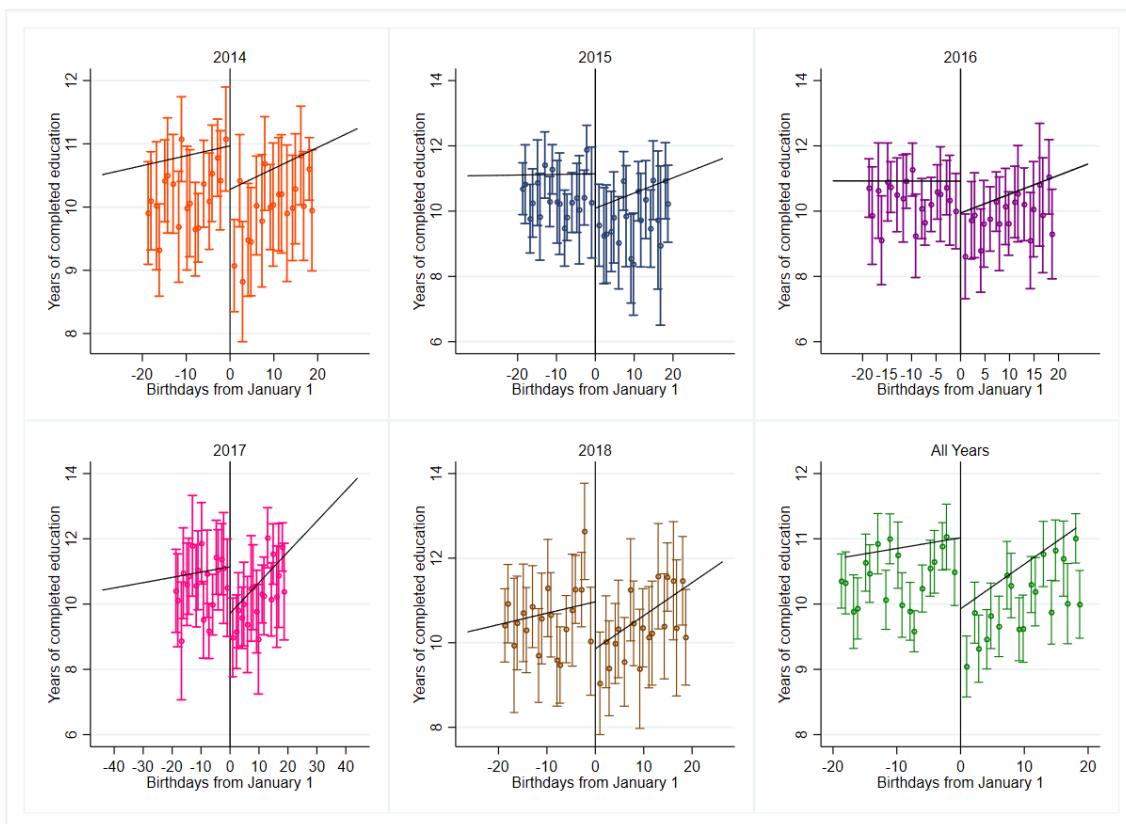
CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Overall Sample

The results for the overall sample is presented in Figure 4.3 and Table 4.3. The vertical axis of the graph measures completed years of education, while the horizontal axis shows the number of birthdays away from January 1, the school entry cut-off. The vertical line at zero divides the sample into two groups: the control group (to the left) and the treatment group (to the right, including zero).

The solid curves on either side of the cutoff represent regression lines for the control and treatment groups, estimated using Stata's `rdplot` command. The bars depict 95% confidence intervals. The graph shows that early school entrants (younger cohort) tend to have fewer years of completed education than later entrants (older cohort), indicating a negative relationship between SEA and educational attainment. Notably, the pooled effect across all years is larger than the individual yearly estimates. The corresponding numerical estimates are provided in Table 4.3. Panel (a)

Figure 4.3: Impact of school entry age on years of completed education for the overall sample



Note: The graphs are based on the optimal bandwidths estimated in Tables 4.3. Source: Author's calculations based on data from the GHS 2014 and 2016.

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

presents fuzzy RD estimates from the model potentially affected by birth date manipulation, while Panel (b) reports estimates after correcting for this issue by excluding individuals born on January 1. Both models control for sex, marital status, province, age group, language of interview (Afrikaans, Sesotho, Siswati, Xhosa), and relationship to the household head, covariates selected based on prior literature and data availability.

The corrected estimates in Panel (b) consistently yield smaller effects of SEA on years of education compared to the uncorrected results, both for individual years and in the pooled sample. For example, in the corrected model, the coefficient on SEA ranges from -0.771 in 2017 to -1.066 in 2018, whereas the uncorrected estimates range from -1.121 in 2016 to -1.438 in 2018. This discrepancy highlights the bias introduced by birth date manipulation around the cutoff.

Across all years and model specifications, the SEA coefficient is negative and statistically significant at the 1% level. The corrected estimates suggest that starting school one year earlier reduces completed years of education by approximately 9.3 to 12.8 months, depending on the year. This result aligns with international evidence ([Kawaguchi, 2011](#); [Li et al., 2022](#); [Valdés and Requena, 2024](#); [Marcenaro-Gutierrez and Lopez-Agudo, 2021](#); [Solli, 2017](#)). For instance, [Kawaguchi \(2011\)](#) find that in Japan, males who started school later accumulated 13.20 years of education, compared to 13.03 years among younger entrants. Similarly, [Valdés and Requena \(2024\)](#) report that in Spain, later school starters gained, on average, 0.2 additional years of education.

Older entrants likely benefit from greater maturity and skill development before starting school ([Valdés and Requena, 2024](#)). Conversely, although younger children start earlier, they tend to face higher repetition rates, leading to lower overall attainment ([Matta et al., 2016](#)). In South Africa, [van der Berg et al. \(2025\)](#) show that younger school entrants experience higher repetition rates, for example, in the Eastern Cape, repetition rates among younger boys reach 31%, compared to 19% for girls in the same age group. These high repetition rates are often linked to poor academic performance, increasing the risk of dropout and reducing educational achievement. The fuzzy RD estimates corrected for manipulation (Panel (b)) consistently show smaller effects of school entry age (SEA) on years of completed education compared to the uncorrected estimates, across all

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Table 4.3: Impact on years education and logarithm of monthly earnings for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): Fuzzy RD with potential manipulation						
First stage						
School entry age	-1.418*** (0.266)	-1.318*** (0.302)	-1.121*** (0.260)	-1.333*** (0.257)	-1.438*** (0.304)	-1.530*** (0.153)
Reduced Form						
School entry age	-0.296*** (0.107)	-0.226** (0.097)	-0.193** (0.094)	-0.270*** (0.078)	-0.066 (0.089)	-0.242*** (0.053)
Second stage						
Education	0.209*** (0.070)	0.171** (0.067)	0.172** (0.079)	0.203*** (0.056)	0.046 (0.060)	0.158*** (0.031)
Covariates	✓	✓	✓	✓	✓	✓
Polynomial Order	1	1	1	1	1	1
Optimal Bandwidth (h)	22.229	29.524	33.947	45.208	26.867	18.586
Bandwidth Bias (b)	62.838	72.980	89.571	112.887	66.239	64.321
Panel (b): Fuzzy RD corrected for manipulation (donut RD)						
First stage						
School entry age	-0.817*** (0.267)	-0.984*** (0.345)	-1.039*** (0.343)	-0.771*** (0.285)	-1.066*** (0.351)	-1.221*** (0.186)
Reduced Form						
School entry age	-0.246** (0.101)	-0.105 (0.109)	-0.217* (0.126)	-0.156* (0.086)	-0.070 (0.106)	-0.225*** (0.064)
Second stage						
Education	0.302** (0.133)	0.106 (0.102)	0.209* (0.114)	0.203* (0.104)	0.065 (0.098)	0.184*** (0.049)
Covariates	✓	✓	✓	✓	✓	✓
Polynomial Order	1	1	1	1	1	1
Optimal Bandwidth (h)	33.679	32.055	27.804	45.264	26.352	18.575
Bandwidth Bias (b)	84.208	85.749	80.704	112.318	66.225	63.253

Note: The dependent variable in the first-stage equation is years of schooling while the dependent variable in the reduced form and second stage equations is the logarithm of earnings. ***, **, and * indicate significance at the 1, 5, and 10% critical levels. The bandwidth size is measured in birthdays away from 1 January, on either side of the cut-off point. The optimal bandwidths were estimated via the `rdwselect` command in Stata (Calónico et al., 2014a). Covariates include province of residency, sex and marital status. The symbols $\times$ denotes no, and $\checkmark$ stands for yes.

years and in the pooled results. In the first-stage equation, the coefficient of SEA in the corrected model ranges from -0.771 in 2017 to -1.066 in 2018, while the uncorrected estimates range from -1.121 in 2016 to -1.438 in 2018. This notable difference underscores the impact of birth date manipulation around the cut-off on estimated treatment effects.

The coefficient of SEA is negative and statistically significant at the 1% level of significance in the two specifications and throughout the survey years. Therefore, the estimates suggest that the school entry age reduced the years of completed education by between 9.3 and 12.8 months between 2014 and 2018. This finding aligns with the literature (Kawaguchi, 2011; Li et al., 2022; Valdés and Requena, 2024; Marcenaro-Gutierrez and Lopez-Agudo, 2021; Solli, 2017). For example, the findings of Kawaguchi (2011) in Japan indicate that males who entered primary school later due to school entry laws accumulated 13.20 years of education while their younger counterparts only had 13.03 years of education. Moreover, Valdés and Requena (2024) find that in Spain men who started school later gained 0.2 more years of schooling, on average, than their

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

younger peers.

The reason for this advantage among older school entrants is that they benefit from an extra year of maturation and skill development (Valdés and Requena, 2024). Although younger cohorts start school earlier the higher repetition rates among them results in reduced years of schooling (Matta et al., 2016). In South Africa, van der Berg et al. (2025) find that children who start school at a younger age within the typical entry range tend to experience higher repetition rates. For example, in the Eastern Cape, the repetition rate among boys who begin school at a younger age is as high as 31%, compared to 19% for girls in the same age group. These high repetition rates are often associated with poor academic performance, which may ultimately result in school dropout or reduced educational attainment.

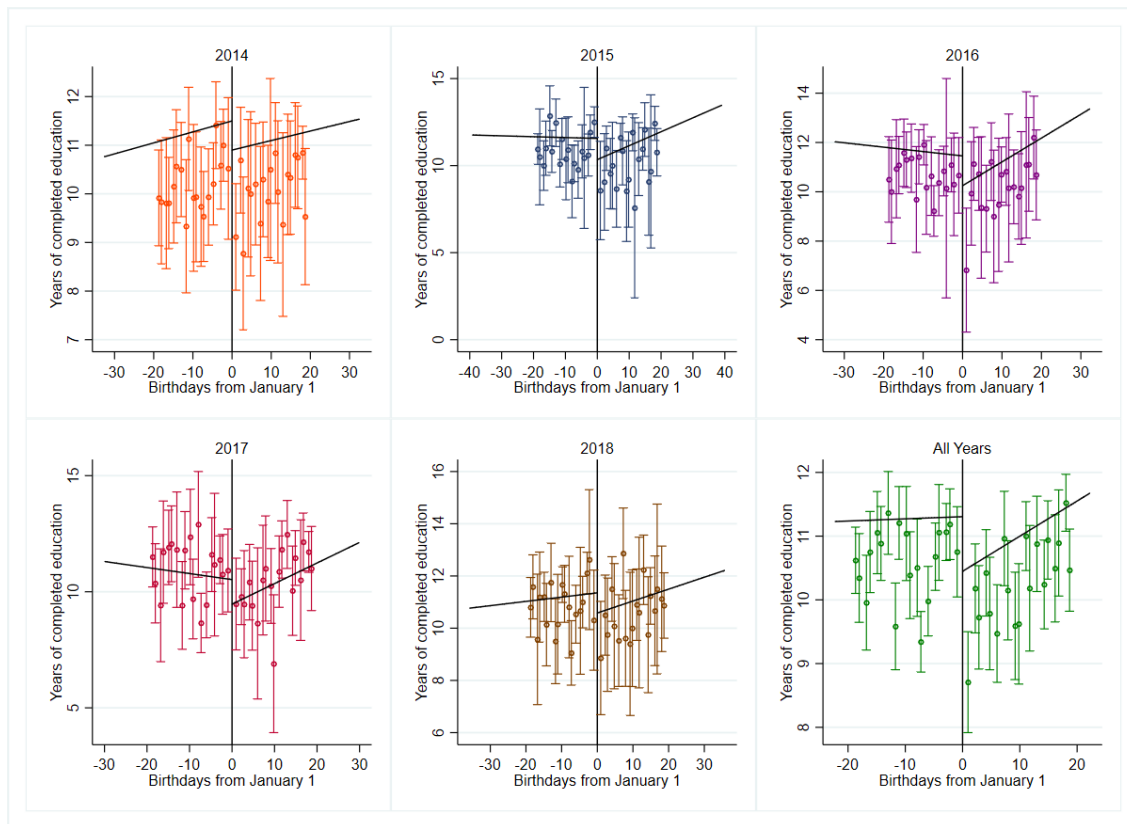
Females

The impact of the school entry law on women's years of education is depicted in Figure 4.6 and Table 4.4. As with the overall sample, the said figure shows that females who started school early experience reduced years of education throughout the study period. Table 4.4 confirms this negative impact. Panel (b), which presents estimates based on the donut RD (excluding observations whose birth date falls on 1 January) indicates that the SEA significantly reduced the years of completed education by between 9 and 15 months. The estimates are considerably smaller compared to the uncorrected ones (Panel (a)) except for 2016, again emphasizing the significance of the manipulation. These results are consistent with the literature as discussed earlier.

The school entry age strongly and significantly reduces the years of education throughout the five-year period. The estimates ranges from -0.55 in 2018 to -1.2 years in 2016. These results indicate that a one-year increase in the age at which an individual started school is associated with a reduction of between 0.55 and 1.2 years of education. However, the pooled estimate is -1.1 years, suggesting some biases in years of education especially when compared to 2014 and 2018. These results suggest that school entry age is a strong instrument for education. Moreover they align with some of the studies, which find that starting school early reduces educational

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Figure 4.4: Impact of school entry age on years of completed education for women



Note: The graphs are based on the optimal bandwidths estimated in Tables 4.3. Source: Author's calculations based on data from the GHS 2014 and 2016.

attainment among women (Kawaguchi, 2011).

Males

The effect of school entry age on men's years of education is shown in Figure 4.4 and Table 4.5. The figure shows that men who started school early have lower years of education compared to those who started late. Table 4.5 confirms the negative effect. Two sets of estimates are shown: Panel (a) shows the fuzzy RD estimates that are based on all available data, including the cut-off (0), which may be subject to manipulation as seen before. Panel (b) depicts estimates that exclude the cut-off, mitigating manipulation concerns.

Panel (b) of the table under the first stage shows a negative coefficient for school entry age, which is statistically significant throughout the survey years indicating that SEA significantly affects the probability of receiving more education. The negative sign is an indication that individuals

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Table 4.4: Impact on years of education and the logarithm of monthly earnings for women

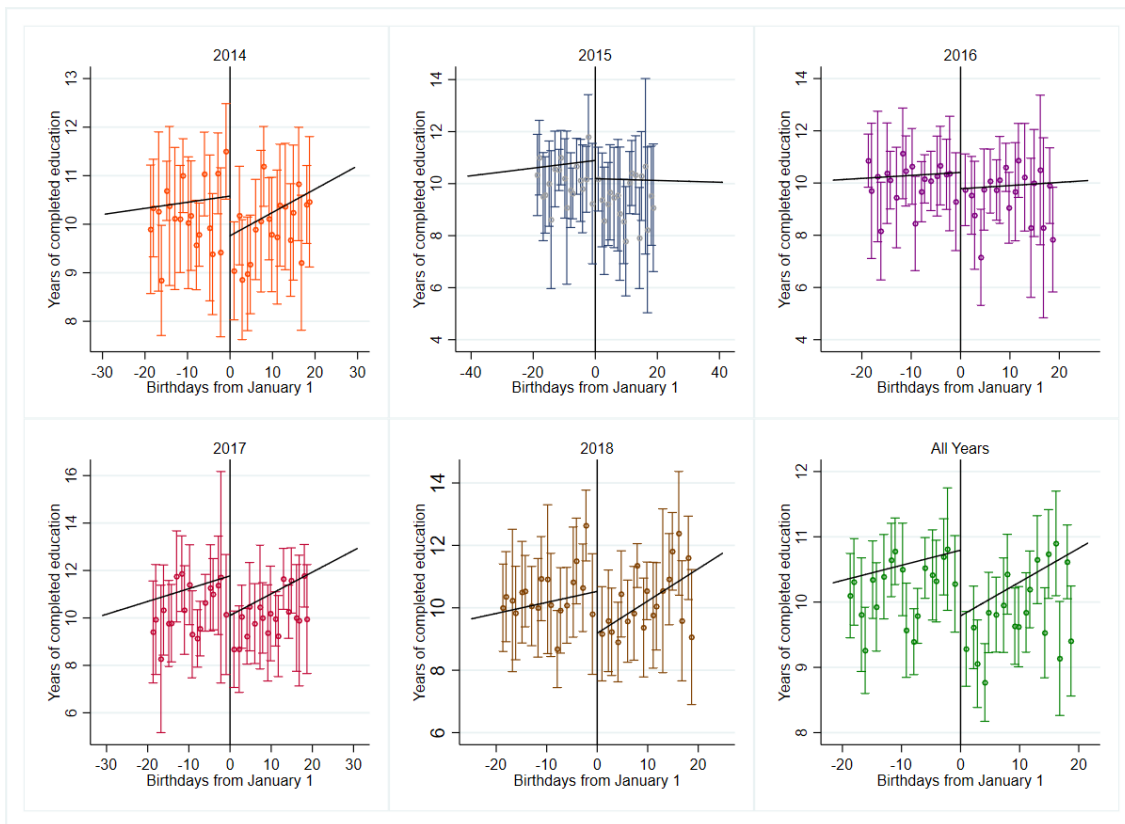
	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): Fuzzy RD with potential manipulation						
First stage						
School entry age	-0.946*** (0.327)	-1.233*** (0.370)	-1.092*** (0.406)	-1.317*** (0.425)	-1.021** (0.410)	-1.486*** (0.230)
Reduced Form						
School entry age	-0.192* (0.114)	-0.350*** (0.120)	-0.222 (0.155)	-0.287** (0.131)	-0.140 (0.117)	-0.337*** (0.083)
Second stage						
Education	0.203* (0.113)	0.284*** (0.099)	0.203 (0.133)	0.218** (0.091)	0.137 (0.113)	0.227*** (0.049)
Covariates	✓	✓	✓	✓	✓	✓
Polynomial Order	1	1	1	1	1	1
Optimal Bandwidth (h)	34.792	45.770	32.743	30.074	34.895	15.706
Bandwidth Bias (b)	83.917	105.976	77.248	73.470	80.283	52.529
Panel (b): Fuzzy RD corrected for manipulation (donut RD)						
First stage						
School entry age	-0.772** (0.393)	-0.984** (0.439)	-1.230** (0.487)	-1.117** (0.494)	-0.475 (0.466)	-1.072*** (0.246)
Reduced Form						
School entry age	-0.276* (0.152)	-0.184 (0.143)	-0.273 (0.183)	-0.259 (0.162)	-0.131 (0.147)	0.299*** (0.082)
Second stage						
Education	0.357 (0.222)	0.188 (0.133)	0.222 (0.141)	0.232* (0.129)	0.275 (0.374)	0.299*** (0.082)
Covariates	✓	✓	✓	✓	✓	✓
Polynomial Order	1	1	1	1	1	1
Optimal Bandwidth (h)	31.989	42.519	32.339	28.945	34.006	22.230
Bandwidth Bias (b)	74.071	105.926	84.851	72.958	81.391	62.191

Note: The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equation is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. Bandwidths below 4 were not feasible due to few observations. All specifications are estimated with the following covariates: metropolitan status, marital status, sex and household size. The optimal bandwidths were estimated via the `rdwselect` command in Stata (Calónico et al., 2014a). Bandwidths below 4 were not feasible due to few observations.

who started school earlier (younger cohort) have lower educational attainment compared to those who enrolled later. The corrected coefficients are smaller in magnitude compared to the uncorrected ones, which indicates the extent of the manipulation effect. However, the coefficients remain statistically significant throughout the survey years. It is evident that school entry age had the smallest impact on years of education in 2015 with a coefficient of -0.692. This estimate shows that starting school one day earlier significantly reduces education years by about 8 months. The impact is larger in the other years, with the largest effect in 2017 in which a one-day earlier entry into school is associated with a reduction of 17 months of education. These results are consistent with the literature, which finds that individuals who start school later end up benefiting more in terms of educational attainment compared to those who starts school earlier. In Spain, for example, Valdés and Requena (2024) find that males who started school later (older

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Figure 4.5: Impact of school entry age on years of completed education for men



Note: The graphs are based on the optimal bandwidths estimated in Tables 4.4. Source: Author's calculations based on data from the GHS 2014 and 2016.

cohort) accumulate 0.2 more schooling years than males who started school earlier (younger cohort). In South Africa, as discussed before, [van der Berg et al. \(2025\)](#) find that children who start school at a younger age tend to experience higher repetition rates. For instance, in the Eastern Cape, the repetition rate among boys who begin school at a younger age is as high as 31%. These elevated repetition rates are often linked to poor academic performance, which may ultimately lead to school dropout or reduced educational attainment.

4.4.4 Impact on earnings

Having established that school entry age (SEA) is a strong instrument for education, I examine the effect of years of education on labour market earnings. Before I do that I discuss the reduced-form effects. As stated earlier, most of the literature assesses the impact of school entry age on earnings using reduced-form estimates. The reduced form captures how being just eligible for

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Table 4.5: Impact on years of education and logarithm of monthly earnings for men

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): Fuzzy RD with potential manipulation						
First stage						
School entry as	-1.379*** (0.337)	-1.148 *** (0.373)	-1.214 *** (0.381)	-2.211 *** (0.481)	-1.737 *** (0.422)	-1.538*** (0.182)
Reduced form						
School entry age	-0.318** (0.145)	-0.105 (0.117)	-0.125 (0.123)	-0.343** (0.137)	0.060 (0.123)	-0.167*** (0.060)
Second stage						
Education	0.231** (0.101)	0.092 (0.092)	0.103 (0.098)	0.155*** (0.057)	-0.035 (0.076)	0.109*** (0.037)
Covariates	✓	✓	✓	✓	✓	✓
Polynomial Order	1	1	1	1	1	1
Optimal Bandwidth (h)	29.528	33.837	27.839	26.434	26.228	26.933
Bandwidth Bias (b)	71.478	86.607	72.559	73.731	61.135	80.935
Panel (b): Fuzzy RD corrected for manipulation (donut RD)						
First stage						
School entry age	-1.078*** (0.394)	-0.692* (0.407)	-0.785* (0.469)	-1.427*** (0.529)	-1.376*** (0.495)	-1.196*** (0.238)
Reduced form						
School entry age	-0.255* (0.154)	0.008 (0.127)	-0.082 (0.166)	-0.179 (0.140)	0.140 (0.141)	-0.077 (0.077)
Second stage						
Education	0.238 (0.148)	-0.013 (0.188)	0.104 (0.207)	0.125 (0.092)	-0.102 (0.124)	0.064 (0.063)
Covariates	✓	✓	✓	✓	✓	✓
Polynomial Order	1	1	1	1	1	1
Optimal Bandwidth (h)	30.774	38.473	25.270	28.770	24.783	21.542
Bandwidth Bias (b)	74.554	100.141	67.219	70.575	60.768	73.707

Note: The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equation is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. All specifications are estimated with the following covariates: metropolitan status, marital status, sex and household size. The optimal bandwidths were estimated via the `rdbwselect` command in Stata (Calonico et al., 2014a).

early school entry affects labour market earnings regardless of whether individuals actually obtained more education. As with education, I include control variables such as marital status, age group, province and sex. These covariates show evidence of discontinuities at the cut-off (see Appendix C), and are therefore included as control variables in the analysis. As before, I discuss the overall sample estimates before I split the discussion by gender.

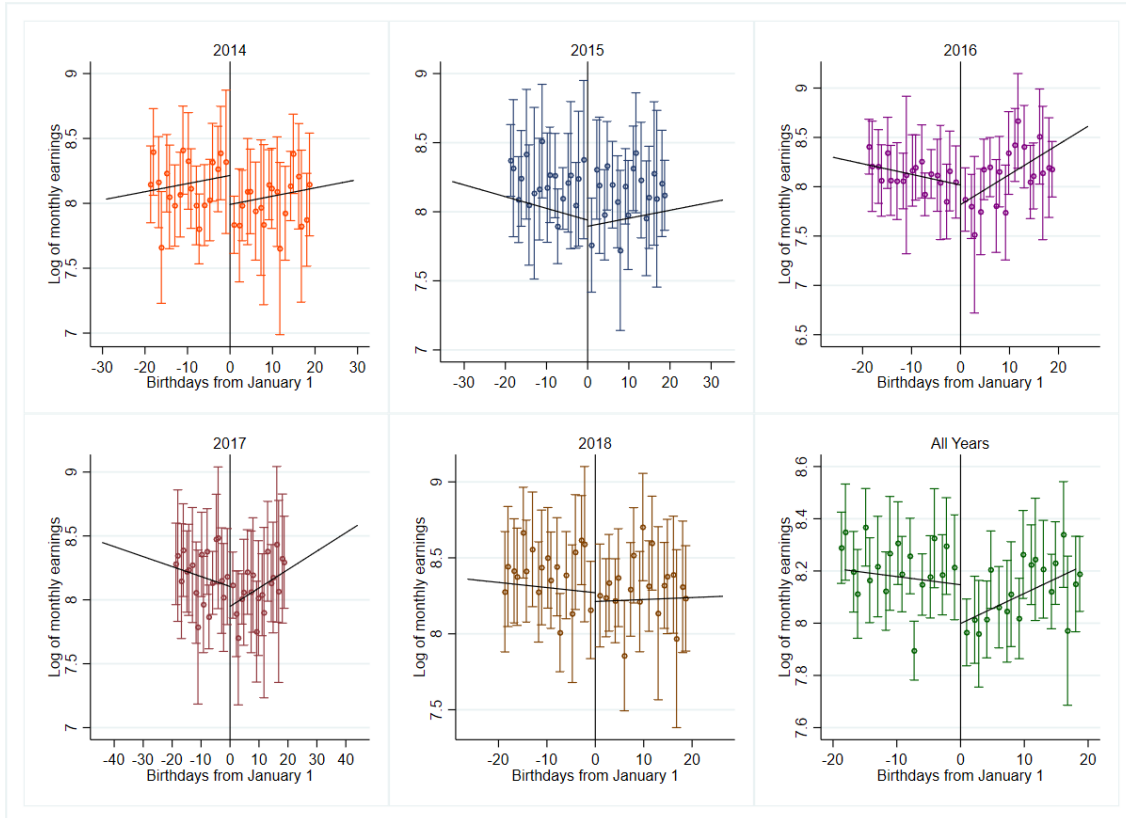
Overall Sample

The reduced form impact of SEA on earnings for the overall sample is presented in Figure 4.5 and Table 4.3. The vertical axis in the graph represents the logarithm of real monthly earnings, while the horizontal axis corresponds to the number of birthdays, which is the running variable centered around January 1. The vertical line at zero divides the sample into two groups: the control group (to the left of zero) and the treated group (to the right of zero, including zero). The

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

solid curves on either side of zero show the regression functions estimated using the ‘rdplot’ command in Stata. Figure 4.5 shows a clear downward jump in log monthly earnings at the

Figure 4.6: Impact of school entry age on the logarithm of monthly earnings for the overall sample



Note: The graphs are based on the optimal bandwidths estimated in Table 4.3. Source: Author's calculations based on data from the GHS 2014 and 2016.

cut-off in all survey years, indicating that individuals who started school earlier (younger cohort) earn significantly less as adults compared to those who started school later (older cohort). This discontinuity suggests a causal effect of age at school entry on adult earnings. The downward jump in log monthly earnings is confirmed by the negative sign of the coefficient on school entry age in the reduced-form equation in Panel (b) of Table 4.3. However, SEA is statistically significant only in 2014, 2016 and 2017 while the pooled estimate is also significant. Accordingly, early school entry is associated with 24.6, 21.7 and 15.6% reduction in monthly earnings in the respective years. The pooled estimate is 22.5%, on average.

The result aligns with the findings in the literature. Peña (2017) finds that in Mexico, the younger cohort earn 1.9% (men) and 3.8% (women) less than their older counterparts. In Brazil, the

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

impact is even larger. [Matta et al. \(2016\)](#) find that the earnings of the older cohort is between 15% and 22% higher than those of the younger cohort. [Pehkonen et al. \(2015\)](#) and [Larsen and Solli \(2017\)](#) find similar results for Finland and Norway, respectively. [Kawaguchi \(2011\)](#) finds that older children take advantage of their maturation at school entry to obtain more education years than their younger counterparts, which eventually leads to higher earnings.

However, the reduced form effects are limited to the effect of the SEA on earnings, they reflect the intention to treat regardless of whether actual treatment was received or not. Hence, they do not tell us whether those who eventually started school early benefited in terms of earnings. To investigate the effect on education on earnings, I turn to the second-stage estimates and assess whether the earnings of those who started school early increased or not. We have seen already that SEA is a strong instrument for education. Panel (b) of Table 4.3 shows the estimates for the corrected version of the RD model. The coefficient on education is positive and statistically significant in 2014, 2016 and 2017, but larger in magnitude compared to the uncorrected estimates (see Panel (a)) of the table. The pooled estimate is also positive, statistically significant and larger.

The effect of education on earnings although positive and statistically significant is heterogeneous across the years, with 2014 experiencing higher returns of 30.2% compared to 20.9 and 20.3% in 2016 and 2017, respectively. Therefore, an additional year of education attained among early school entrants increased their monthly earnings by between 20.3% and 30.2%, on average. Next, I examine the impact on earnings by gender.

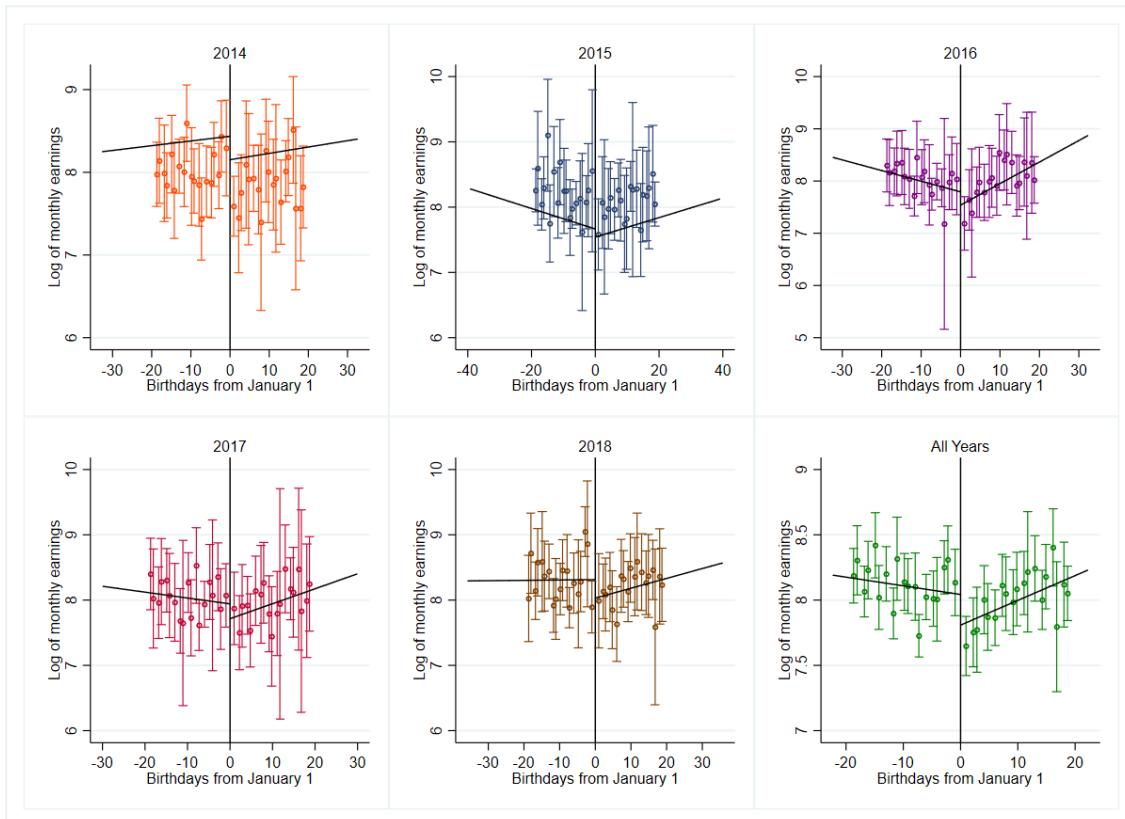
Females

Two effects are examined here. The effect of the instrument on earnings (reduced-form estimates) and the impact of education on earnings (second stage). Figure 4.7 indicates that the effect of school entry age on years of education is negative throughout the years. The size of the effect is heterogeneous across the years with 2015 having the smallest effect and 2014 the largest. Table 4.4 provides the estimates. Panel (a) shows the estimates with potential manipulation of birth dates while Panel (b) depicts results corrected for manipulation. As seen before, the difference between the two versions of estimates is evident. The RD version without correction indicates

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

that all coefficients are negative, with statistical significance in 2014, 2015, 2017, and the pooled sample. While the estimates corrected for manipulation are also negative they are statistically

Figure 4.7: Impact of school entry age on the logarithm of monthly earnings for women



Note: The graphs are based on the optimal bandwidths estimated in Tables 4.3. Source: Author's calculations based on data from the GHS 2014-2018.

significant only in 2014 and for the pooled sample, which underscores the importance of the correction. The coefficient of -0.276 for 2014 suggests that early school entry reduces earnings by about 28% among women, on average. This finding is in line with several studies. For example, [Liu and Marois \(2023\)](#) find that starting school later significantly increases annual earnings by 59% for non-agricultural jobs in China. In Mexico, however, the size of the effect is relatively smaller with [Peña \(2017\)](#) finding that the earnings of women who started school early are 3.8% lower than for late school entrants.

Turning to the impact of education on earnings, it is evident that the corrected estimates differ from the uncorrected ones. Focusing on the corrected version, one sees that the effect of education on earnings is positive and heterogeneous across the years. The coefficient on education is

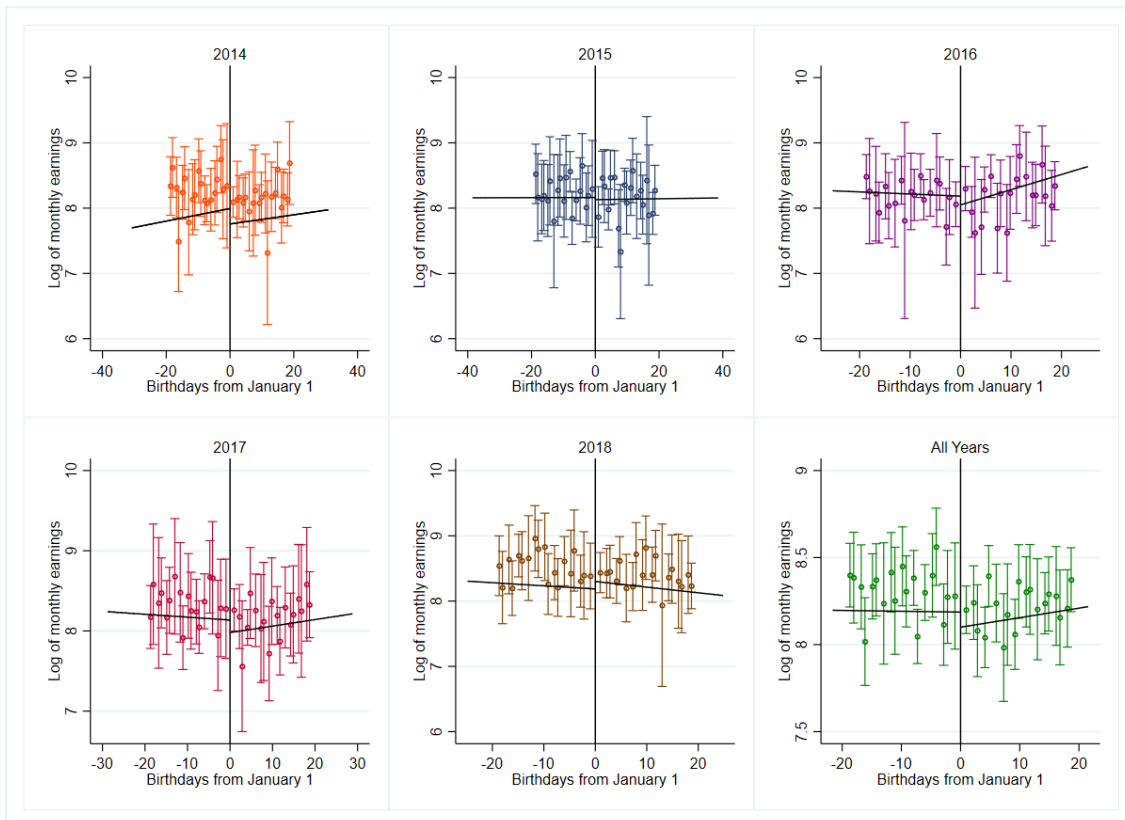
CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

statistically significant in 2017 with the treatment effect of 0.232, respectively. Therefore, an additional year of education increased the monthly earnings of early school entrants by 23.2% in 2017.

Males

As with the overall and female samples, I analyse the effect on earnings in two ways. The first is to look at the reduced-form estimates, which represents the intention to treat effect. This is the most common way the literature examines the effect of school entry age on labour market earnings. However, I go beyond that and estimate the causal effect of education on earnings using school entry age as an instrument for education. Thus, the second way I examine the impact on earnings is to estimate the returns to education among individuals who actually started school early due to the school entry rule. In the first analysis, I rely on both figure 4.8 and Table 4.5.

Figure 4.8: Impact of school entry age on the logarithm of monthly earnings for men



Note: The graphs are based on the optimal bandwidths estimated in Tables 4.4. Source: Author's calculations based on data from the GHS 2014-2018.

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

The figure appears to indicate that school entry age had a negative effect on years of education in 2014, 2016 and 2017. However, on a closer examination, Table 4.5 shows that the effect of SEA is statistically significant only in 2014. The uncorrected version shows statistical significance in 2017 as well. Thus, the coefficient of -0.255 in 2014 suggest that starting school earlier was associated with a reduction of 25.5% in monthly earnings in 2014 compared to starting school later. As noted before, this result is consistent with findings in other countries. For example, [Peña \(2017\)](#) find that younger cohorts of men earn 1.9% lower than the older cohort in Mexico.

Moving on to the second analysis, the causal effect of education on earnings is positive in 2014, 2016 and 2017 and negative in the other years, but statistically insignificant in all the years including the pooled sample. Studies such as [Dobkin and Ferreira \(2010\)](#) and [Pehkonen et al. \(2015\)](#) find similar results even though their estimates are based on reduced-form effects.

4.4.5 Robustness checks

I conduct a range of robustness checks to ensure that the findings are reliable and not the result of spurious relationships in the data. The two robustness checks that I consider are (1) estimating the treatment effect at placebo cut-off points of the running variable, and (2) checking for balance on pre-determined covariates.

Placebo cut-off points

The analysis of placebo cut-off points is a robustness check used to validate the credibility of the RD design. Specifically, it involves testing for discontinuities at arbitrary points along the running variable, points at which no treatment should occur. The absence of discontinuities at these placebo thresholds provides support for the claim that the observed discontinuity at the true cut-off is attributable to the treatment. Conversely, if significant discontinuities are detected at these artificial thresholds, it raises concerns about the internal validity of the RD design, suggesting that the observed effects may be driven by other unobserved factors correlated with the running variable. In this analysis, I conduct placebo tests by estimating the RD at eight alternative cut-off points, four to the left and four to the right of the actual threshold. The results

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

are presented in Table 4.6.

Table 4.6: Impact of education on the logarithm of monthly earnings at different placebo cut-offs for the overall sample

	(1) c=-40	(2) c=-30	(3) c=-20	(4) c=-10	(5) c=10	(6) c=20	(7) c=30	(8) c=40
Panel (a): Overall Sample								
2014	0.448 (0.343)	-1.079 (3.024)	-0.644 (14.290)	0.435 (0.877)	0.111 (0.337)	-1.410 (5.273)	0.152 (0.159)	-0.101 (0.391)
2015	0.108 (0.193)	0.157 (0.115)	0.310 (0.485)	0.860 (1.020)	0.210 (0.218)	-0.442 (2.468)	0.042 (0.243)	-0.113 (0.553)
2016	1.099 (1.189)	0.002 (0.147)	-0.349 (2.004)	0.282 (0.321)	1.764 (1.961)	-0.279 (0.320)	0.091 (0.137)	0.804 (1.135)
2017	-0.306 (0.969)	0.318 (0.707)	-0.201 (0.778)	0.042 (0.254)	0.138 (0.155)	-0.060 (0.350)	6.158 (53.454)	0.521 (0.407)
2018	0.379** (0.190)	0.249 (3.792)	-1.118 (7.105)	3.646 (31.107)	0.224 (0.173)	-0.100 (0.285)	0.351* (0.201)	-1.640 (26.104)
All Years	0.612 (0.636)	0.101 (0.191)	0.473 (2.783)	0.365 (0.322)	0.479* (0.269)	0.077 (0.293)	0.251** (0.109)	-0.021 (0.302)
Panel (b): Female Sample								
2014	0.396* (0.224)	-0.169 (0.743)	0.560 (2.714)	-1.452 (5.788)	0.072 (0.561)	-1.311 (100.961)	0.143 (0.136)	0.187 (0.365)
2015	4.636 (28.839)	0.242* (0.131)	0.247* (0.133)	0.406* (0.240)	0.127 (0.455)	0.131 (0.110)	0.388 (0.733)	0.078 (0.229)
2016	0.750 (0.561)	0.209 (0.162)	0.083 (0.443)	0.657 (0.535)	0.656** (0.319)	0.481 (0.506)	0.367 (0.274)	0.433* (0.229)
2017	-2.536 (9.843)	0.546 (0.549)	0.203 (0.337)	-0.233 (0.730)	0.170 (0.227)	-1.817 (20.595)	0.499 (0.655)	-0.629 (2.883)
2018	0.274 (0.208)	0.340 (0.878)	-0.168 (1.089)	-0.185 (1.083)	0.784 (1.496)	0.280 (0.584)	-0.015 (0.797)	-0.505 (1.302)
All Years	3.475 (41.812)	0.111 (0.184)	0.250 (0.378)	0.777 (0.574)	0.404** (0.169)	0.174 (0.214)	0.183 (0.128)	0.141 (0.488)
Panel (c): Male Sample								
2014	0.414 (0.833)	-1.449 (4.838)	-0.076 (1.372)	-0.215 (0.465)	0.162 (0.453)	-0.723 (1.647)	0.230 (0.242)	-0.394 (0.935)
2015	0.038 (0.152)	0.070 (0.256)	0.455 (0.974)	-0.593 (88.953)	0.206 (0.378)	0.091 (0.131)	0.057 (0.146)	0.072 (0.243)
2016	1.122 (2.062)	-0.124 (0.228)	0.550 (0.529)	-0.577 (1.852)	-1.634 (338.023)	-0.151 (0.156)	0.039 (0.250)	0.196 (0.201)
2017	0.315 (0.651)	0.723 (15.312)	-0.270 (0.752)	-0.118 (0.598)	0.189 (0.202)	0.023 (0.167)	2.979 (12.919)	1.021 (1.813)
2018	1.106 (5.801)	-0.081 (0.506)	-0.032 (0.874)	1.169 (3.412)	0.024 (0.161)	-0.356 (0.523)	0.302** (0.149)	-0.286 (1.047)
All Years	-0.073 (0.934)	-0.021 (0.324)	0.298 (0.438)	-0.043 (0.390)	0.330 (0.238)	0.352 (0.437)	0.238* (0.122)	0.819 (2.030)

Note: The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equation is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. All specifications are estimated with the following covariates: metropolitan status, marital status, sex and household size. The optimal bandwidths were estimated via the `rdbwselect` command in Stata (Calonic et al., 2014a).

The premise of the placebo test is that if treatment is truly assigned based on the January 1 cut-off, then no significant discontinuities should be observed at any other point along the running variable. The results confirm this expectation for most cases. For the overall sample, no statistically significant discontinuities are found in 2014 through 2017. However, two significant jumps are observed in 2018, one on either side of the true cut-off. In the pooled sample, two discontinuities appear on the right side of the threshold.

When disaggregated by gender, the female sample shows some evidence of discontinuity: one in 2014, three in 2015, two in 2016, and one in the pooled sample. By contrast, the male sample

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

exhibits no statistically significant discontinuities at any placebo cut-off point. Overall, while a few significant discontinuities emerge, primarily within the female subsample, most placebo estimates are statistically insignificant. This pattern provides reassurance that the RD estimates are not driven by arbitrary breaks in the running variable, thereby supporting the internal validity of the identification strategy.

Covariate balance

Balance on covariates involves examining whether the observable characteristics of units just below and just above the threshold are similar. If these covariates are balanced, this supports the idea that units near the threshold are comparable, reinforcing the validity of the RD design. Tables 4.7, 4.8 and 4.9 show the impact of school entry age on the pre-determined characteristics of units. Statistically significant differences in units around the threshold would call the RD design into question.

Table 4.7 depicts the impact of school entry age on language of interview for the overall and female samples. The respondents were interviewed in eleven languages as shown in the table. If the RD design is valid, school entry age should not systematically influence the language of interview, i.e., the coefficients should be small and statistically insignificant. Most of the coefficients are small in magnitude and statistically insignificant, indicating no systematic relationship between school entry age and language of interview. This supports the validity of the RD design.

Table 4.8 presents the results of covariate balance tests evaluating whether the school entry age cut-off affects the language of interview for the male sample across the survey years 2014 to 2018, as well as for the pooled sample. For the majority of years and language groups, the estimated coefficients are small in magnitude and statistically insignificant, suggesting that the running variable is generally balanced with respect to this covariate, a key requirement for the validity of the RD design.

Nonetheless, a few notable exceptions are observed. In 2014, Afrikaans shows a modest but statistically significant negative coefficient. In 2015, Sesotho and Siswati exhibit significant discontinuities. For 2016, Setswana and Zulu show positive and statistically significant differences.

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Table 4.7: Covariate balance: assessing impact of school entry age on language of interview for the overall and female samples

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): Overall Sample						
Afrikaans	-0.024** (0.012)	-0.007 (0.008)	0.008 (0.010)	0.005 (0.011)	0.004 (0.006)	-0.000 (0.004)
English	0.010 (0.012)	-0.004 (0.022)	-0.004 (0.022)	0.016 (0.024)	-0.009 (0.021)	0.002 (0.008)
Ndebele	-0.007 (0.007)	0.014 (0.009)	0.007 (0.009)	0.026*** (0.010)	0.006 (0.008)	0.007 (0.004)
Sepedi	0.009 (0.017)	-0.010 (0.023)	0.002 (0.020)	0.063*** (0.024)	0.004 (0.019)	0.014 (0.010)
Sesotho	-0.047** (0.020)	-0.026 (0.020)	-0.021 (0.019)	-0.032 (0.020)	0.017 (0.019)	-0.022** (0.009)
Siswati	0.020* (0.012)	-0.004 (0.011)	-0.008 (0.011)	-0.028 (0.019)	-0.017 (0.016)	-0.004 (0.007)
Tsonga	-0.007 (0.009)	0.019 (0.018)	0.007 (0.016)	0.014 (0.012)	0.014 (0.012)	0.007 (0.006)
Setswana	0.014 (0.015)	0.045** (0.020)	-0.026 (0.021)	-0.069*** (0.022)	-0.014 (0.021)	-0.005 (0.009)
Tshivenda	0.000 (0.006)	-0.016 (0.012)	-0.005 (0.010)	0.007 (0.016)	-0.001 (0.013)	-0.005 (0.005)
Xhosa	0.019* (0.011)	0.003 (0.017)	-0.000 (0.014)	-0.013 (0.017)	-0.006 (0.012)	0.004 (0.007)
Zulu	-0.007 (0.018)	0.004 (0.030)	0.048** (0.022)	-0.011 (0.028)	-0.015 (0.029)	0.005 (0.009)
Panel (b): Female Sample						
Afrikaans	-0.011 (0.014)	-0.027* (0.015)	0.009 (0.016)	0.016 (0.011)	0.009 (0.008)	0.001 (0.006)
English	0.024 (0.020)	-0.075** (0.030)	0.001 (0.029)	-0.022 (0.030)	-0.056** (0.028)	-0.016 (0.012)
Ndebele	-0.014 (0.009)	0.030* (0.016)	0.010 (0.013)	0.020 (0.017)	0.005 (0.013)	0.009 (0.007)
Sepedi	0.024 (0.023)	-0.006 (0.037)	0.018 (0.033)	0.025 (0.031)	-0.005 (0.027)	0.007 (0.012)
Sesotho	-0.037 (0.025)	0.008 (0.029)	-0.028 (0.031)	0.002 (0.027)	0.053* (0.028)	-0.005 (0.012)
Siswati	0.031** (0.015)	-0.049** (0.022)	-0.004 (0.022)	-0.049* (0.028)	-0.053*** (0.020)	-0.016 (0.010)
Tsonga	-0.003 (0.014)	0.001 (0.017)	0.019 (0.020)	0.010 (0.019)	0.051** (0.022)	0.010 (0.008)
Setswana	0.013 (0.021)	0.107*** (0.037)	0.005 (0.027)	-0.028 (0.027)	-0.018 (0.026)	0.013 (0.012)
Tshivenda	0.003 (0.010)	-0.021 (0.016)	-0.017 (0.014)	-0.015 (0.023)	-0.004 (0.018)	-0.010 (0.007)
Xhosa	0.002 (0.015)	0.007 (0.021)	-0.019 (0.022)	-0.023 (0.026)	0.003 (0.017)	-0.004 (0.009)
Zulu	-0.034 (0.026)	0.029 (0.037)	0.014 (0.041)	0.070 (0.042)	0.008 (0.035)	0.008 (0.014)

Note: The dependent variable is the the respective language while the independent variable is the running variable (birthdays). Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. All statistically significant languages were included in the estimation of the treatment effect as a covariate. All specifications are estimated with the following covariates: age group, marital status, sex and province of residence. The optimal bandwidths were estimated via the rdbwselect command in Stata (Calonico et al., 2014a).

In 2017, significant imbalances are detected for Ndebele, Sesotho, and Setswana, while Sepedi shows a significant increase. Finally, in the pooled sample, only Sesotho is associated with a statistically significant discontinuity.

Overall, although a handful of statistically significant differences are detected in specific years and languages, the general pattern supports covariate balance on the language of interview for the male sample, thereby reinforcing the validity of the RD design.

Table 4.9 tests whether school entry age assignment is balanced across different household

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Table 4.8: Covariate balance: assessing impact of school entry age on language of interview for the male sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Afrikaans	-0.027* (0.016)	0.010 (0.010)	0.008 (0.012)	-0.010 (0.015)	0.012 (0.011)	-0.008 (0.007)
English	0.002 (0.019)	0.038 (0.034)	-0.002 (0.031)	0.034 (0.032)	0.031 (0.027)	0.017 (0.011)
Ndebele	-0.001 (0.009)	0.002 (0.010)	-0.003 (0.010)	0.027*** (0.010)	0.002 (0.008)	0.005 (0.005)
Sepedi	-0.004 (0.022)	-0.019 (0.036)	-0.016 (0.030)	0.103*** (0.036)	0.010 (0.025)	0.013 (0.015)
Sesotho	-0.038 (0.024)	-0.069** (0.029)	-0.021 (0.026)	-0.061** (0.028)	-0.013 (0.024)	-0.044*** (0.014)
Siswati	0.007 (0.018)	0.026* (0.015)	-0.017 (0.013)	-0.011 (0.023)	0.013 (0.022)	0.005 (0.009)
Tsonga	-0.014 (0.014)	0.023 (0.027)	-0.005 (0.024)	0.023 (0.017)	0.003 (0.019)	0.003 (0.009)
Setswana	0.033 (0.022)	0.006 (0.033)	-0.052* (0.030)	-0.095*** (0.032)	-0.005 (0.029)	-0.017 (0.013)
Tshivenda	-0.000 (0.007)	-0.013 (0.016)	0.007 (0.014)	0.024 (0.021)	-0.002 (0.017)	0.003 (0.007)
Xhosa	0.015 (0.014)	-0.005 (0.024)	0.012 (0.022)	-0.004 (0.020)	-0.011 (0.021)	0.011 (0.010)
Zulu	0.021 (0.024)	0.013 (0.044)	0.081** (0.033)	-0.067* (0.040)	-0.031 (0.039)	0.003 (0.014)

Note: The dependent variable is the the respective language while the independent variable is the running variable (birthdays). Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. All statistically significant languages were included in the estimation of the treatment effect as a covariate. All specifications are estimated with the following covariates: age group, marital status, sex and province of residence. The optimal bandwidths were estimated via the rdbwselect command in Stata (Calonico et al., 2014a).

relationships across survey years. The goal is to ensure there are no systematic differences in household composition that could bias the RD design. Most coefficients are small and statistically insignificant, suggesting no strong relationship between school entry age and relationship to the head of household. This supports the validity of the RD design.

In general, I find that the pre-determined covariates are balanced. Where I find imbalance, I incorporate the the concerned covariates in the estimation of relevant RD models.

4.4.6 Sensitivity analysis

I conduct three sensitivity tests. These include varying the bandwidth size, using donut RD, and varying the polynomial order.

Bandwidth size

Selecting an optimal bandwidth involves a variance-bias trade-off. A narrower bandwidth yields a more precise estimate around the cut-off point, yet it risks amplifying noise. Conversely, a broader bandwidth diminishes noise but may obscure the treatment effect. Hence, careful consideration of data characteristics and the research question is vital in determining the optimal bandwidth

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Table 4.9: Covariate balance: assessing impact of school entry age on relation to head of household for the overall sample

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): Overall Sample						
Head of Household (HHD)	-0.060* (0.032)	0.013 (0.039)	-0.017 (0.035)	0.032 (0.035)	0.042 (0.045)	-0.010 (0.016)
Spouse	0.033 (0.021)	-0.026 (0.025)	0.011 (0.030)	-0.010 (0.029)	-0.004 (0.026)	-0.002 (0.011)
Child	-0.026 (0.022)	0.017 (0.023)	-0.038 (0.031)	-0.038 (0.029)	-0.041 (0.032)	-0.024* (0.014)
Sibling	-0.008 (0.013)	0.004 (0.016)	-0.033* (0.018)	-0.005 (0.013)	0.003 (0.018)	-0.010 (0.006)
Parent	0.001 (0.002)	0.003 (0.003)	0.001 (0.001)	0.002 (0.002)	0.000 (0.000)	0.001 (0.001)
Grandchild	0.015 (0.011)	-0.007 (0.011)	0.024** (0.011)	0.009 (0.008)	0.010 (0.009)	0.010* (0.005)
Other relative	0.039** (0.017)	0.012 (0.016)	0.020 (0.014)	0.005 (0.012)	0.021 (0.013)	0.021*** (0.007)
Not related	0.002 (0.006)	-0.005 (0.008)	-0.004 (0.008)	-0.006 (0.008)	-0.003 (0.008)	-0.000 (0.003)
Panel (b): Female Sample						
Head of Household	-0.042 (0.035)	0.017 (0.061)	-0.001 (0.054)	0.029 (0.057)	0.009 (0.060)	-0.016 (0.026)
Spouse	0.065* (0.034)	-0.061 (0.040)	-0.035 (0.053)	-0.054 (0.051)	-0.054 (0.044)	-0.031* (0.016)
Child	-0.028 (0.034)	0.005 (0.042)	0.015 (0.043)	0.020 (0.044)	0.026 (0.042)	-0.000 (0.020)
Sibling	-0.000 (0.014)	-0.011 (0.021)	-0.036 (0.025)	-0.006 (0.017)	-0.011 (0.019)	-0.016 (0.010)
Parent	0.003 (0.003)	0.007 (0.008)	0.002 (0.003)	0.003 (0.004)	0.000 (0.002)	0.002 (0.002)
Grandchild	0.029* (0.015)	-0.013 (0.017)	0.011 (0.015)	0.018 (0.015)	0.016 (0.015)	0.016** (0.007)
Other relative	0.052** (0.026)	0.026 (0.023)	0.016 (0.024)	0.012 (0.020)	0.019 (0.023)	0.026** (0.012)
Not related	-0.005 (0.007)	-0.007 (0.005)	-0.003 (0.011)	-0.008 (0.011)	0.002 (0.007)	-0.003 (0.005)
Panel (b): Male Sample						
Head of Household	-0.002 (0.040)	-0.007 (0.045)	0.021 (0.045)	0.075* (0.043)	0.089* (0.052)	0.016 (0.019)
Spouse	-0.019 (0.016)	-0.007 (0.017)	-0.008 (0.015)	-0.011 (0.016)	-0.008 (0.016)	-0.008 (0.006)
Child	-0.000 (0.032)	0.017 (0.031)	-0.034 (0.034)	-0.082** (0.040)	-0.109** (0.044)	-0.025 (0.017)
Sibling	-0.018 (0.021)	0.010 (0.021)	-0.030 (0.021)	0.008 (0.019)	0.014 (0.027)	-0.005 (0.010)
Parent	N/A	N/A	N/A	N/A	N/A	N/A
Grandchild	0.000 (0.015)	-0.002 (0.013)	0.037*** (0.012)	0.004 (0.009)	0.006 (0.011)	0.007 (0.007)
Other relative	0.023 (0.018)	-0.007 (0.018)	0.025 (0.020)	0.007 (0.014)	0.019 (0.017)	0.015* (0.008)
Not related	0.023 (0.018)	-0.007 (0.018)	0.025 (0.020)	0.007 (0.014)	0.019 (0.017)	0.015* (0.008)

Note: The dependent variable is the the respective relationship to the head of the household while the independent variable is the running variable (birthdays). Robust standard errors are shown in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. All statistically significant variables were included in the estimation of the treatment effect as a covariate. All specifications are estimated with the following covariates: age group, marital status, sex and province of residence. The optimal bandwidths were estimated via the `rdwselect` command in Stata (Calonico et al., 2014a).

method (Cattaneo, Idrobo and Titiunik, 2020). Therefore, the first sensitivity test that I applied involves examining how changes in bandwidth size affect the main results estimated in Tables 4.3, 4.4 and 4.5. I re-estimate the equations for all individuals, women and men for the period 2014-2018 (including the pooled sample) using bandwidth sizes of 10 - 90 birthdays. I use a data-driven bandwidth selection method as proposed by Calonico et al. (2017). The results are

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

presented in Table 4.10. The optimal bandwidth for 2014 is approximately 34 days, yielding a

Table 4.10: Impact of years of education on the logarithm of monthly earnings at different bandwidths

	(1) h=10	(2) h=20	(3) h=30	(4) h=40	(5) h=50	(6) h=60	(7) h=70	(8) h=80	(9) h=90
Panel (b): Overall Sample									
2014	0.319*** (0.112)	0.275** (0.114)	0.287** (0.126)	0.354** (0.166)	0.390* (0.225)	0.389 (0.253)	0.382 (0.281)	0.394 (0.306)	0.413 (0.316)
2015	0.398 (0.279)	0.121 (0.145)	0.102 (0.105)	0.100 (0.108)	0.125 (0.111)	0.127 (0.104)	0.116 (0.112)	0.115 (0.124)	0.115 (0.124)
2016	0.225** (0.106)	0.288** (0.123)	0.189* (0.112)	0.138 (0.141)	0.127 (0.169)	0.139 (0.168)	0.143 (0.205)	0.154 (0.248)	0.158 (0.269)
2017	0.100 (0.069)	0.121* (0.072)	0.157** (0.073)	0.184* (0.095)	0.213* (0.115)	0.231* (0.127)	0.244 (0.154)	0.235 (0.160)	0.223 (0.151)
2018	0.072 (0.086)	0.068 (0.083)	0.074 (0.106)	0.071 (0.174)	0.081 (0.268)	0.061 (0.260)	-0.029 (0.317)	-0.192 (0.448)	-0.377 (0.559)
All Years	0.204*** (0.048)	0.178*** (0.049)	0.173*** (0.051)	0.185*** (0.066)	0.197** (0.082)	0.191** (0.085)	0.177* (0.103)	0.164 (0.127)	0.148 (0.146)
Panel (b): Female Sample									
2014	0.533*** (0.195)	0.333** (0.152)	0.351 (0.215)	0.493 (0.404)	0.551 (0.635)	0.530 (0.783)	0.427 (1.040)	0.276 (1.186)	0.135 (1.300)
2015	0.293* (0.158)	0.109 (0.148)	0.156 (0.120)	0.175 (0.130)	0.214 (0.136)	0.219* (0.132)	0.209 (0.150)	0.212 (0.172)	0.205 (0.168)
2016	0.334*** (0.109)	0.352** (0.157)	0.234* (0.138)	0.178 (0.176)	0.183 (0.202)	0.201 (0.187)	0.214 (0.220)	0.234 (0.273)	0.262 (0.321)
2017	0.289** (0.138)	0.295** (0.144)	0.230* (0.132)	0.232 (0.163)	0.301 (0.218)	0.314 (0.241)	0.309 (0.303)	0.332 (0.343)	0.414 (0.359)
2018	0.138 (0.190)	0.248 (0.241)	0.270 (0.320)	0.290 (0.475)	0.329 (0.658)	0.302 (0.495)	0.268 (0.632)	0.113 (1.091)	-0.175 (1.501)
All Years	0.365*** (0.076)	0.309*** (0.082)	0.270*** (0.087)	0.280** (0.115)	0.300** (0.147)	0.286* (0.148)	0.257 (0.191)	0.208 (0.256)	0.161 (0.297)
Panel (c): Male Sample									
2014	0.155 (0.134)	0.212 (0.156)	0.238 (0.149)	0.275* (0.166)	0.308 (0.209)	0.333 (0.231)	0.400 (0.265)	0.528* (0.307)	0.671** (0.339)
2015	-0.566 (2.629)	0.162 (0.389)	0.022 (0.188)	-0.016 (0.190)	-0.006 (0.184)	0.003 (0.150)	-0.009 (0.147)	-0.022 (0.152)	-0.027 (0.152)
2016	-0.243 (0.679)	0.148 (0.217)	0.074 (0.200)	0.042 (0.248)	0.029 (0.315)	0.047 (0.334)	0.100 (0.439)	0.182 (0.527)	0.183 (0.521)
2017	0.005 (0.098)	0.062 (0.090)	0.136 (0.095)	0.148 (0.141)	0.096 (0.157)	0.075 (0.179)	0.047 (0.219)	0.008 (0.239)	-0.038 (0.258)
2018	-0.102 (0.133)	-0.091 (0.108)	-0.090 (0.141)	-0.175 (0.318)	-0.211 (0.635)	-0.264 (0.534)	-0.419 (0.552)	-0.793 (0.597)	-1.199* (0.690)
All Years	0.028 (0.076)	0.058 (0.065)	0.078 (0.063)	0.083 (0.082)	0.082 (0.101)	0.084 (0.105)	0.088 (0.120)	0.102 (0.132)	0.108 (0.139)

Note: The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equation is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. All specifications are estimated with the following covariates: metropolitan status, marital status, sex and household size. The optimal bandwidths were estimated via the rdbwselect command in Stata (Calonico et al., 2014a).

statistically significant treatment effect of 30.2%. At a 30-day bandwidth, the estimate is 28.7%, and it rises to 35.4% at 40 days. Although the estimates are similar across bandwidths, statistical significance weakens as the bandwidth increases, suggesting the true effect is likely captured near the optimal range.

For 2015, the optimal bandwidth is around 33 days, producing a treatment effect of 10.6%, which is statistically insignificant. Comparable results, both in magnitude and lack of significance, are observed across other bandwidth sizes. 2016 shows an optimal bandwidth of 28 days with a statistically significant effect of 20.9%. At 30 days, the estimate is 18.9%, and it increases to

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

28.8% at 20 days. While the estimates remain fairly stable, both the size and significance of the effect decline as the bandwidth widens.

In 2017, the optimal bandwidth is approximately 45 days, associated with a significant treatment effect of 20.3%. The estimates remain consistent at 40 days (18.4%) and 50 days (21.3%), though significance disappears once the bandwidth exceeds 70 days. The bandwidth for 2018 is optimally set at 26 days, producing an effect of 6.5%, which is not statistically significant. Estimates across other bandwidths remain similar in both size and significance. Lastly, the pooled sample (All Years) yields an optimal bandwidth of 19 days, with a statistically significant treatment effect of 18.4%. This estimate remains stable across a range of bandwidth choices, both in magnitude and significance.

For the female sample, in 2014, the optimal bandwidth is approximately 32 days with a treatment estimate of 35.10%. The estimate increases slightly to 35.70% at a bandwidth of 30 days but remains statistically insignificant. While the point estimates vary across bandwidths, they are statistically insignificant for most bandwidth sizes. For 2015 and 2016, the estimates remain stable and statistically insignificant across bandwidths, indicating robustness to bandwidth choice.

In contrast, the 2017 estimate is statistically significant and stable at both the 29-day optimal bandwidth (23.2%) and the 30-day alternative (23.0%). Notably, the effect also remains significant at 10-day and 20-day bandwidths, underscoring its robustness. In 2018, the estimates are consistent across bandwidths (27.5% at optimal and 27.0% at alternative), suggesting similar robustness. The pooled estimate is both stable and statistically significant across bandwidths (29.9% at 22-day optimal and 30.9% at 20-day alternative), highlighting a robust average treatment effect for the female sample.

For the male sample, the estimates at most alternative bandwidth sizes are statistically insignificant across the years, consistent with the main results reported in Table 4.5.

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Donut RD

Donut RD involves re-estimating the fuzzy RD model while excluding observations that are very close to the cut-off date, effectively creating a "donut hole" around the threshold. The main objective is to test whether the estimated treatment effects are sensitive to observations near the cut-off (Barreca et al., 2011; Fan et al., 2018). In this study, I have already implemented this approach by excluding individuals whose date of birth falls on 1 January. The corresponding results are presented in Tables 4.3, 4.4, and 4.5 for the overall, female, and male samples, respectively. In this sub-section, I extend the robustness check by excluding observations born within one and two days of 1 January. The results are reported for the first- and second-stage equations in Tables 4.11 and 4.12. For brevity, I do not present reduced-form estimates. However, for convenience, I replicate the ± 0 donut RD estimates.

Table 4.11: Impact of school entry age on years of education at different donut holes for overall, female and male samples

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): Overall Sample						
± 0 -day donut	-0.817*** (0.267)	-0.984*** (0.345)	-1.039*** (0.343)	-0.771*** (0.285)	-1.066*** (0.351)	-1.221*** (0.186)
± 1 -day donut	-0.357 (0.252)	-1.060*** (0.354)	-1.008** (0.394)	-0.640** (0.304)	-1.147*** (0.382)	-1.029*** (0.181)
± 2 -day donut	-0.716** (0.343)	-0.963** (0.438)	-0.959** (0.392)	-0.434 (0.316)	-1.002** (0.425)	-1.0399*** (0.218)
Panel (b): Female Sample						
± 0 -day donut	-0.772** (0.393)	-0.984** (0.439)	-1.230** (0.487)	-1.117** (0.494)	-0.475 (0.466)	-1.072*** (0.246)
± 1 -day donut	-0.642 (0.458)	-0.754* (0.452)	-0.490 (0.494)	-1.068** (0.532)	-0.412 (0.451)	-0.723*** (0.255)
± 2 -day donut	-0.871 (0.565)	-0.670 (0.528)	-0.517 (0.536)	-0.945 (0.618)	-0.282 (0.478)	-0.710 (0.282)
Panel (c): Male Sample						
± 0 -day donut	-1.078*** (0.394)	-0.692* (0.407)	-0.785* (0.469)	-1.427*** (0.529)	-1.376*** (0.495)	-1.196*** (0.238)
± 1 -day donut	-0.586 (0.468)	-1.132** (0.527)	-1.353*** (0.519)	-1.219** (0.527)	-1.503*** (0.529)	-1.336*** (0.269)
± 2 -day donut	-0.741 (0.536)	-1.219** (0.565)	-1.481** (0.580)	-0.915* (0.536)	-1.437** (0.589)	-1.253*** (0.278)

The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equations is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. Covariates include metropolitan, sex, marital status and household size. The symbol $\checkmark$ denotes yes.

The first-stage estimates are presented in Table 4.11. As previously noted, for the overall sample, the coefficient on school entry age is negative and statistically significant across all survey years, including the pooled sample. These results are reported as ± 0 -day donut estimates in the table. I further compare them with the ± 1 -day and ± 2 -day donut estimates. As shown in the table, both the ± 1 -day and ± 2 -day estimates suggest a negative and statistically significant effect of school

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

entry age on education in all survey years, except for 2014 and 2017, respectively.

For the female sample, the ± 0 -day donut estimates indicate that school entry age has a negative and statistically significant effect on education in all years except 2018, where the estimate is insignificant. The ± 1 -day and ± 2 -day donut estimates are negative throughout the years, including in the pooled sample, but are statistically significant only in 2015, 2017, and the pooled results. Despite these differences in statistical significance, the magnitude of the coefficients remains broadly consistent across specifications. The estimates for males closely mirror those of the overall sample.

I now turn to the second-stage estimates, presenting the effect of education on earnings using the ± 0 -day, ± 1 -day and ± 2 -day donut samples (see Table 4.12). As previously observed, for the overall sample, the ± 0 donut RD specification shows that education has a statistically significant positive effect on earnings in 2014, 2016, and 2017, with no significant effects in 2015 and 2018. The pooled estimate is also positive and statistically significant. Increasing the “donut hole” by one day yields similar results: the effect remains statistically significant in 2016 and 2017, though the significance in 2014 disappears. The pooled estimate continues to be positive and statistically significant, and its magnitude is similar to the ± 0 estimate.

The estimated effects are also comparable across most years, except in 2015, where the sign reverses and the estimate becomes notably smaller. When the “donut hole” is expanded to two days, the statistical significance disappears for all individual years, except for the pooled estimate, which remains significant but with a reduced effect size. Nonetheless, the magnitude of the estimates remains relatively stable across years, suggesting that the preferred model is robust to moderate changes in bandwidth around the cut-off.

For the female sample, the ± 0 donut RD estimates indicate that education has a statistically significant effect on monthly log earnings in 2017, with the pooled estimate also being positive and statistically significant. Increasing the “donut hole” by one day yields similar results: the effect remains significant in 2017, and the pooled estimate continues to be positive and significant. However, when the “donut hole” is expanded to two days, statistical significance disappears,

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Table 4.12: Impact of education on the logarithm of monthly earnings at different donut holes for overall, female and male samples

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): Overall Sample						
± 0 -day donut	0.302** (0.133)	0.106 (0.102)	0.209* (0.114)	0.203* (0.104)	0.065 (0.098)	0.184*** (0.049)
± 1 -day donut	0.497 (0.394)	-0.033 (0.115)	0.277* (0.165)	0.294* (0.156)	0.113 (0.117)	0.178*** (0.062)
± 2 -day donut	0.135 (0.155)	-0.041 (0.152)	0.144 (0.173)	0.390 (0.270)	0.063 (0.147)	0.115* (0.070)
Panel (b): Female Sample						
± 0 -day donut	0.357 (0.222)	0.188 (0.133)	0.222 (0.141)	0.232* (0.129)	0.275 (0.374)	0.299*** (0.082)
± 1 -day donut	0.270 (0.301)	0.127 (0.181)	0.171 (0.404)	0.285* (0.164)	0.880 (3.130)	0.317** (0.149)
± 2 -day donut	-0.133 (0.288)	0.109 (0.241)	-0.009 (0.495)	0.218 (0.197)	-0.306 (65.824)	0.129 (0.148)
Panel (c): Male Sample						
± 0 -day donut	0.238 (0.148)	-0.013 (0.188)	0.104 (0.207)	0.125 (0.092)	-0.102 (0.124)	0.064 (0.063)
± 1 -day donut	0.523 (0.415)	-0.114 (0.165)	0.167 (0.156)	0.220 (0.137)	-0.093 (0.131)	0.077 (0.067)
± 2 -day donut	0.469 (0.334)	-0.072 (0.158)	0.126 (0.171)	0.311 (0.215)	-0.101 (0.143)	0.086 (0.074)

The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equations is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. Covariates include metropolitan, sex, marital status and household size. The symbol ✓ denotes yes.

although the magnitude of the effect remains broadly similar. Overall, these results suggest that the estimated effect of education on female earnings is relatively stable across small changes in the exclusion window around the cut-off.

For the male sample, the ± 0 donut RD estimates indicate no statistically significant effect of education on monthly log earnings in any year or in the pooled sample. The ± 1 -day and ± 2 -day donut RD specifications yield similar findings. These results suggest that the estimated effect of education on male earnings is consistently insignificant and relatively insensitive to changes in the number of observations excluded around the threshold.

Polynomial order

Gelman and Imbens (2019) suggests that higher-order polynomials should not be used in regression discontinuity. The main results presented in Tables 4.3 – 4.5 were estimated using a first-order polynomial. In Table 4.13, I increase the polynomial order to two and compare the results from both specifications. For convenience, the first-order polynomial estimates are replicated in Table 4.13. As previously observed, for the overall sample, the first-order polynomial results show a statistically significant positive effect of education on log monthly earnings in

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

2014, 2016, and 2017, with no significant effects in 2015 and 2018. The pooled estimate is also positive and statistically significant.

Table 4.13: Impact of education on logarithm of monthly earnings at different polynomial orders for the overall, female and male samples

	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018	(6) All Years
Panel (a): Overall Sample (± 0 Donut RD)						
Polynomial Order 1	0.302** (0.133)	0.106 (0.102)	0.209* (0.114)	0.203* (0.104)	0.065 (0.098)	0.184*** (0.049)
Polynomial Order 2	0.272** (0.110)	0.111 (0.121)	0.231** (0.105)	0.087 (0.081)	0.056 (0.097)	0.193*** (0.051)
Panel (b): Female Sample (± 0 Donut RD)						
Polynomial Order 1	0.357 (0.222)	0.188 (0.133)	0.222 (0.141)	0.232* (0.129)	0.275 (0.374)	0.299*** (0.082)
Polynomial Order 2	0.428* (0.244)	0.131 (0.137)	0.230 (0.151)	0.303** (0.147)	0.266 (0.303)	0.364*** (0.088)
Panel (c): Male Sample (± 0 Donut RD)						
Polynomial Order 1	0.238 (0.148)	-0.013 (0.188)	0.104 (0.207)	0.125 (0.092)	-0.102 (0.124)	0.064 (0.063)
Polynomial Order 2	0.203 (0.148)	0.032 (0.218)	0.093 (0.227)	0.124 (0.099)	-0.061 (0.105)	0.024 (0.084)

The dependent variable in the first-stage equations is years of schooling while the dependent variable in the second stage equations is the logarithm of monthly earnings. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10% critical levels. Covariates include metropolitan, sex, marital status and household size. The symbol ✓ denotes yes.

Increasing the polynomial order to two yields broadly similar results: education has a statistically significant effect in 2014 and 2016, and the pooled estimate remains positive and significant. The magnitude of the estimated effects under both specifications is comparable across most years: 2014, 2015, 2016, 2018, and in the pooled sample. However, in 2017, the second-order polynomial estimate is smaller and statistically insignificant, contrasting with the significant first-order result. Overall, the findings suggest that the estimated treatment effect of education on earnings is largely robust to the choice of polynomial order across survey years.

The first-order polynomial results indicate that education has a statistically significant positive effect on monthly log earnings for females in 2017, as well as in the pooled sample. When the polynomial order is increased to two, the estimates remain statistically significant in 2017 and the pooled sample, and additionally become significant in 2014. The estimated effect sizes are broadly similar across the two specifications. These results suggest that the impact of education on female earnings is both positive and relatively stable across different model specifications. The first-order polynomial results for males indicate that education has no statistically significant effect on monthly log earnings in any year or in the pooled sample. The second-order polynomial yields

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

similar findings. These results suggest that education does not have a statistically significant impact on male earnings, regardless of the polynomial specification or survey year.

4.5 Conclusion

The admission policy for first time school entry into South African schools requires that children turn seven years of age by 30 June for enrollment in Grade 1, otherwise they should start school the following year in which they turn eight years of age. The study exploits exogenous variation in educational attainment arising from the school entry policy to address the endogeneity of education and estimate the causal returns to schooling. The analysis uses data from the General Household Survey for the years 2014–2018, which includes exact information on individuals' dates of birth.

The evidence presented in this chapter indicates that there is a relatively high concentration of births around 1 January. This fact puts the RD design into question. The uneven distribution of date of birth around 1 January suggests that parents act strategically to take advantage of school entry rules. In this case parents assume that there are advantages of having their children start school early. To address the problem of manipulation, I employ a donut fuzzy regression discontinuity. While numerous studies globally have investigated the impact of school starting age on labour market earnings, similar analyses remain scarce in Sub-Saharan Africa. As a result, the effects of early school entry in these contexts are largely unknown.

This study seeks to address this gap by focusing on Black South Africans, a population whose socioeconomic and institutional experiences were profoundly shaped by apartheid. The legacy of racial discrimination permeated all areas of life, including education and labour markets, relegating the Black majority to the economic margins. Although the democratic transition in 1994 abolished openly discriminatory laws, many Black South Africans continue to experience adverse socioeconomic outcomes, including poor-quality education and pervasive poverty. Most Black students still attend formerly disadvantaged schools, which often fail to deliver adequate cognitive skills. Furthermore, Black people continue to experience particularly poor labour market

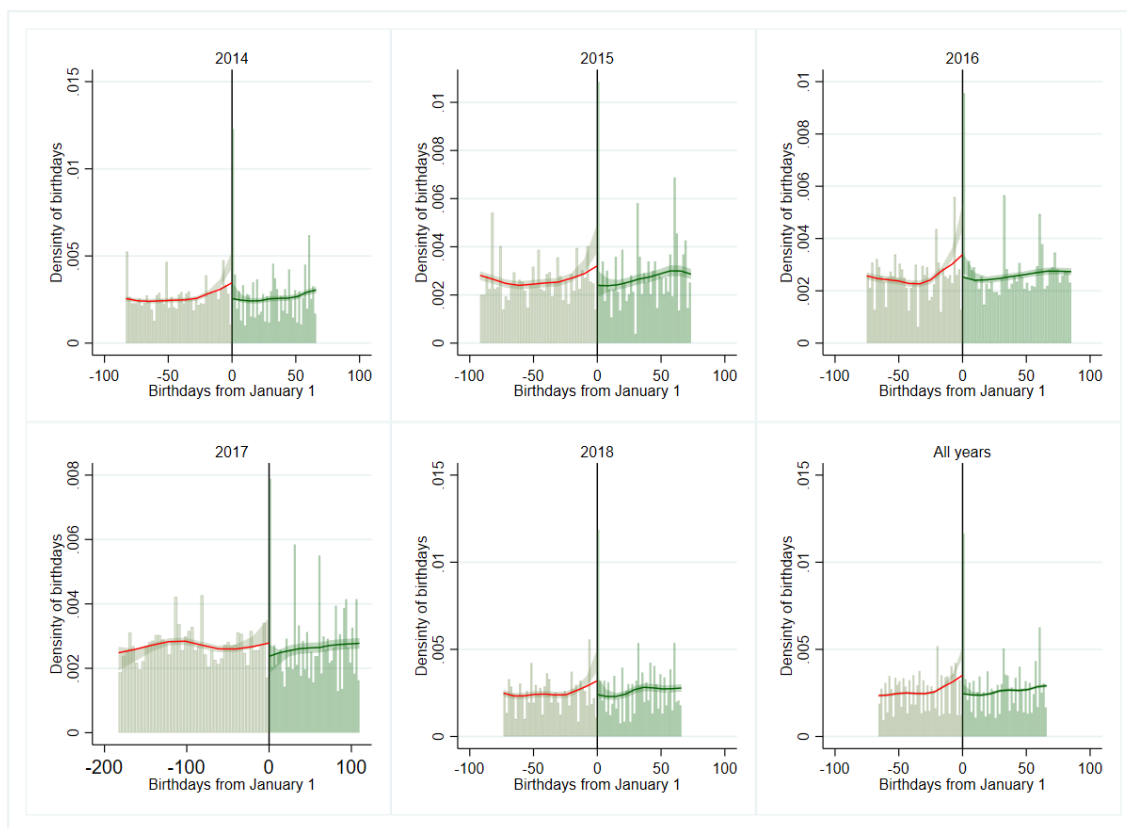
CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

outcomes, such as elevated unemployment rates and disproportionately low-paying jobs among women, reflecting both historical and ongoing gender and racial inequality.

The results reveal that early school entry reduced the total number of years of education and monthly earnings, on average. However, the findings also show that among individuals who started school earlier, education significantly increased monthly earnings. These findings raise important questions about the optimal age for school entry. Early school entry could lead to higher repetition rates and poor academic performance among children who are not developmentally ready for school. Conversely, delaying school entry may improve school readiness and long-term outcomes, but it may delay entry into the labour market. Therefore, finding the right balance is critical to ensuring that children begin formal education at an age that maximizes both educational success and long-term economic outcomes.

APPENDIX C

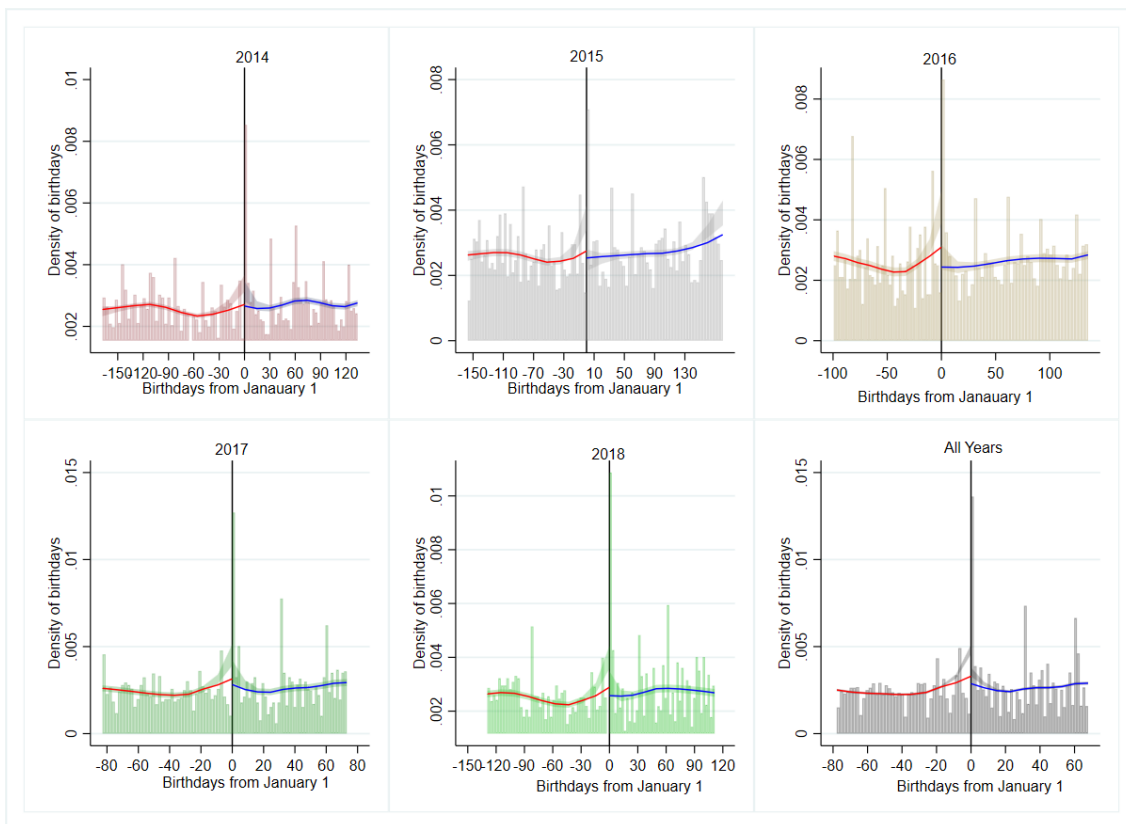
Figure C1: McCrarry manipulation test for women: 2014-2018



Source: Author's calculations based on data from the GHS 2014-2018.

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

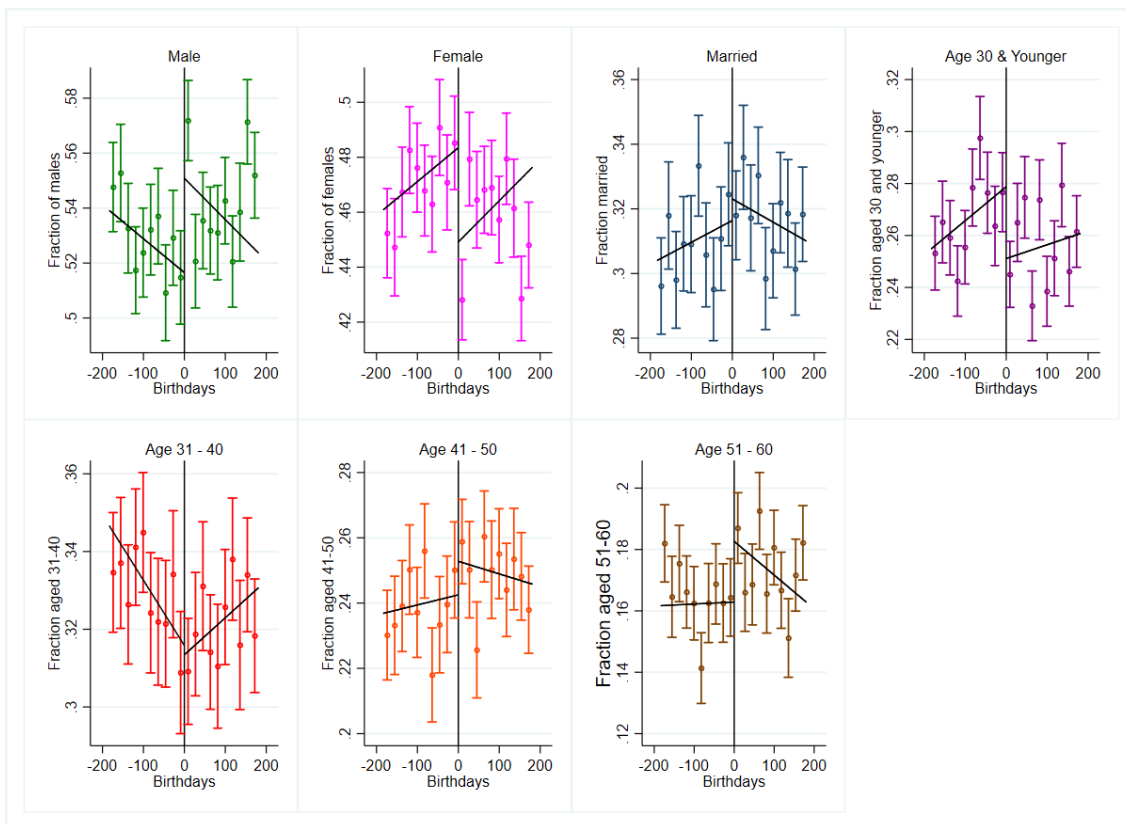
Figure C2: McCrary manipulation test for men: 2014-2018



Source: Author's calculations based on data from the GHS 2014-2018.

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

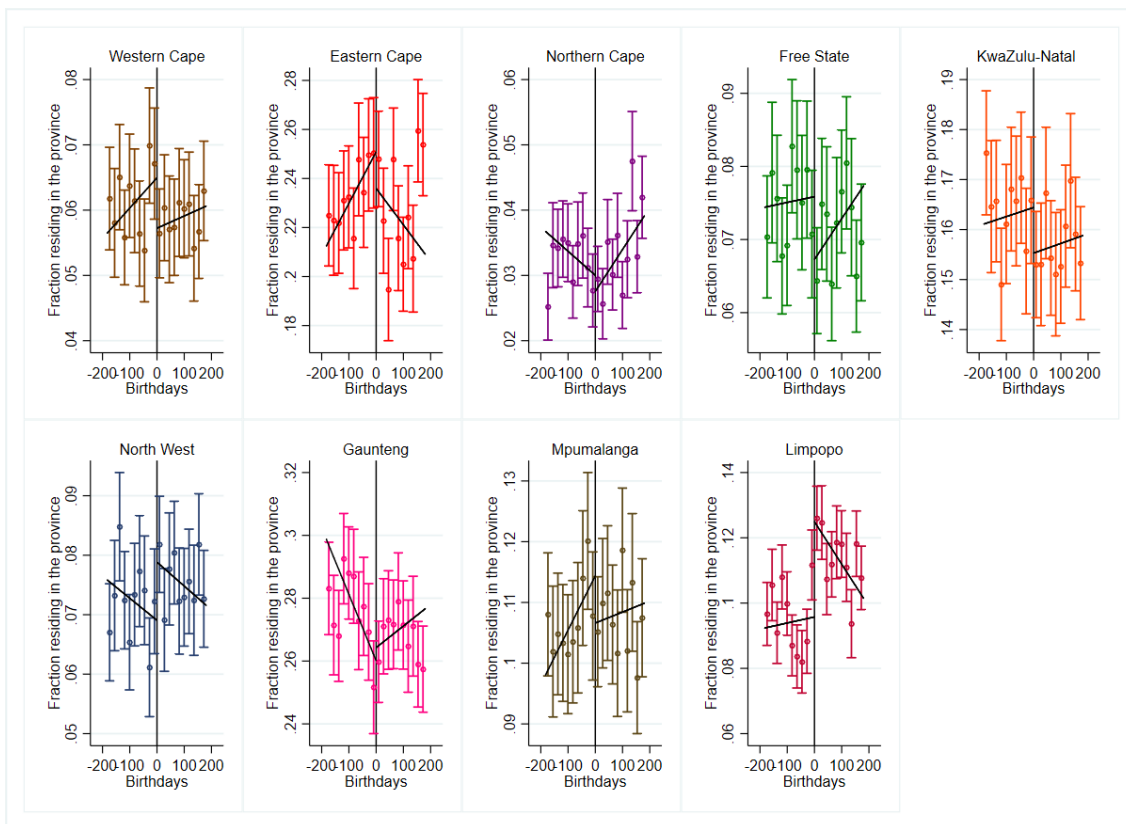
Figure C3: Covariate balance tests: gender, marital status and age group



Source: Author's calculations based on data from the GHS 2014-2018.

CHAPTER 4. TOO EARLY TO BENEFIT? SCHOOL ENTRY LAWS AND RETURNS TO EDUCATION

Figure C4: Covariate balance tests: province of residence at the time of the survey



Source: Author's calculations based on data from the GHS 2014-2018.

CHAPTER

CONCLUSION

The apartheid regime in South Africa affected all spheres of life including the education system and the labour market ([Mtapuri and Tinarwo, 2021](#)). Through the racial discriminatory policies and practices of the apartheid era Black people were subjected to a separate and poor-quality education system that was meant to impart only basic skills needed for the manual labour and menial jobs that the government deemed suitable for those of their race ([Thobejane, 2013](#); [Mtapuri and Tinarwo, 2021](#)). For the majority Black population, this meant limited career development and confinement to low-paying jobs. These policies and practices increased the income inequality between Blacks and Whites. Black women were worse off than their fellow Black men as unfavourable patriarchal beliefs and practices of the past limited their career development as they were confined to rural areas or worked in low-paying jobs such as domestic work in urban centres ([Espí et al., 2019](#); [Mosomi, 2019b](#)).

The democratic transition of 1994 marked a turning point in South Africa's history, as the new government sought to dismantle apartheid-era structures and create a more inclusive society. Central to these efforts were reforms in education and the labour market. The 1996 Constitution, which enshrined equality and human rights, provided the legal foundation for these changes ([Republic of South Africa, 1996a](#)). One of the first major reforms was the introduction of the

CHAPTER 5. CONCLUSION

Schools Act in 1996, which established a new national curriculum aimed at promoting equity in education. This act also introduced a compulsory schooling policy, ensuring that children of all races had access to basic education ([Republic of South Africa, 1996b](#)).

In tandem with the education reforms, the government also implemented significant legislative changes to transform the labour market. Key among these were the Labour Relations Act (LRA) of 1995 ([Republic of South Africa, 1995](#)), the Basic Conditions of Employment Act (BCEA) ([Republic of South Africa, 1997](#)), the Employment Equity Act (EEA) of 1998 ([Republic of South Africa, 1998](#)), and the Broad-Based Black Economic Empowerment (B-BBEE) Act of 2003 ([Republic of South Africa, 2004](#)). The LRA expanded workers' rights by institutionalizing collective bargaining, enabling union membership, and protecting the right to strike.

The BCEA aimed to create fairer employment conditions by mandating minimum standards, such as annual leave, sick leave, severance pay, and other basic entitlements. Importantly, the BCEA also introduced sector-specific minimum wages for vulnerable workers in industries with low union representation or low average earnings. These industries included domestic work, contract cleaning, private security, wholesale and retail, farming, forestry, taxis, learnerships, hospitality, civil engineering, and children's roles in advertising, creative arts, and cultural endeavors ([Millea et al., 2017](#); [Nattrass, 2016](#)). However, it was not until 2019 that a comprehensive national minimum wage law was enacted, signaling the slow yet progressive journey towards addressing economic inequalities.

The impact of these sector-specific minimum wage laws on employment and wages remains an area of extensive study and debate, with findings offering a mixed picture. Some studies report no statistically significant effects on employment rates following the implementation of minimum wage laws. For instance, research by [Bhorat et al. \(2013\)](#); [Conradie \(2003, 2005\)](#); [Murray and Van Walbeek \(2007\)](#) found limited evidence of employment changes across various sectors. Conversely, other studies, such as [Bhorat, Kanbur and Stanwix \(2014\)](#), suggest that minimum wage laws reduced the likelihood of employment, particularly in certain vulnerable industries.

The effect of minimum wage policies on wages is similarly inconclusive. [Bhorat et al. \(2013\)](#) found that minimum wage laws successfully raised real hourly wages in sectors like private security, retail, and domestic work. However, the same study noted wage decreases in the taxi sector and statistically insignificant impacts in forestry. In the agricultural sector, minimum wage regulations introduced notable wage increases for farm workers, as evidenced by [Bhorat, Kanbur and Stanwix \(2014\)](#). These findings highlight the sector-specific nature of minimum wage policies and their varying implications for workers and employers, underscoring the need for nuanced approaches to labour market interventions.

Unemployment levels in South Africa are high with over 40% of the labour force without jobs (see Figure 1.1). High levels of unemployment have significantly impacted the socioeconomic landscape of the country by contributing to elevated levels of income inequality and poverty ([Anand et al., 2016](#)). This issue affects not only individuals but also has a widespread effect on communities, leading to a cycle of deprivation that is hard to break ([The World Bank, 2018](#)).

Although access to educational opportunities for the formerly disadvantaged non-White population has improved with the introduction of a non-racial education system and compulsory schooling ([UNESCO, 2020](#)), educational outcomes remain poor, particularly within the Black African and Coloured communities. These groups often face numerous challenges, including under-resourced schools, inadequate facilities, and a lack of qualified teachers, which hinder academic progress and limit students' future job prospects ([Spaull, 2013](#)).

Therefore, South Africa provides a unique context to study the impact of education on the labour market. This thesis addresses three specific objectives. The first objective is to investigate whether education offers a channel for improving employment outcomes among men and women within the Black African, Coloured, and White communities (the Indian/Asian race was excluded due to small sample size). The second is to assess whether exposure to the compulsory schooling reform increases returns to education at different locations of the wage distribution for Black South Africans. The analysis is performed for both men and women. And the third objective is to estimate the impact of education on wages of Black men and women South Africans.

In Chapter 2 I took advantage of the compulsory schooling policy of 1997 and used it as a source of exogenous variation to identify the causal effect of education on employment outcomes. The instrumental variable is measured as the number of years of compulsory schooling an individual was exposed to. Using a two-stage least squares method to estimate the causal effect of education on employment with data from the General Household Surveys for the years 2014-2018, I showed that the instrument is plausible since it has a strong correlation with the variable years of education.

Also, the IV is exogenous since the compulsory education policy was merely intended to raise the educational levels of the underserved population groups, and not designed to improve employment outcomes directly. Thus, as an IV, it affects the years of education, but is unrelated to ability, which is omitted in the analyses due to lack of data. The analysis of employment outcomes provides evidence that education is endogenous, indicating that earlier South African research probably have endogeneity bias.

The findings indicate that an additional year of education raises the likelihood of employment by 7.3 to 28.4 percentage points, varying across survey years. The estimates point to notable differences by gender and race — both in the effect of compulsory schooling on educational attainment and in the impact of education on employment prospects. Disaggregating by gender, the employment gains linked to an additional year of education range from 3.0 to 18.0 percentage points for women and 3.0 to 13.9 percentage points for men. These results underscore the importance of education in enhancing employment prospects in South Africa. Strengthening access to quality education, especially at higher levels, can be an effective policy lever for improving labour market outcomes. By promoting educational attainment, particularly among disadvantaged groups, policymakers can help increase individuals' chances of employment and contribute to addressing South Africa's persistently high unemployment rates.

The analysis of returns to education along the wage distribution in Chapter 3 involved estimating the impact of compulsory schooling on returns to education along the distribution of earnings. I applied the quantile regression and the quantile selection model to control for selection into employment to find the rate of return to education at different locations of the earnings distribu-

tion for both Black men and women who were exposed and those who were not exposed to the compulsory schooling reform.

The findings indicate that returns to education vary based on exposure to compulsory schooling, educational level, gender, and across the wage distribution. Returns are higher for women than for men. In addition, women exposed to the compulsory schooling reform tend to earn more than those who were not. A similar pattern is observed among men, but only at certain quantiles of the wage distribution.

As expected, I find that female returns suffer from sample selection bias, and correcting for the selection lowers their returns, depending on the location on the wage distribution. Moreover, the results suggest a skills-biased labour market that increases wage inequality. This is supported by the analysis that is based on levels of education, rather than years of education. I find increasing returns along educational levels, rising from the lower to the higher end of the wage distribution both for females and males. This increasing pattern of returns to education suggests a skills-biased structure to the labour market, which increases wage inequality.

Chapter 4 explores the impact of education on labour market earnings among Black South Africans. Under South Africa's compulsory schooling law, children are required to begin Grade R or Grade 1 in the calendar year they turn six or seven, provided their birthday falls on or before June 30 of that admission year. Leveraging the natural variation in educational attainment induced by this school entry cutoff, I estimate the returns to education using a regression discontinuity design based on data from the General Household Survey (GHS) spanning 2014 to 2018. This study contributes to the limited body of research on the effects of school entry policies within Sub-Saharan Africa.

The analysis detects indications of birth date manipulation around January 1st, likely intended to circumvent school admission regulations. To mitigate bias arising from this manipulation, individuals born on January 1st are excluded from the estimation sample. This exclusion helps to reduce potential distortions in the estimated treatment effects. The findings suggest that starting school at a younger age is associated with fewer completed years of schooling and lower

CHAPTER 5. CONCLUSION

earnings in the labour market. Nevertheless, for those who do enter school early, education exerts a statistically significant positive influence on earnings. These results raise important policy questions regarding the optimal age for school entry and underscore the need for targeted support to younger entrants who may experience academic disadvantages.

BIBLIOGRAPHY

Aakvik, A., Salvanes, K. G. and Vaage, K. (2010), 'Measuring heterogeneity in the returns to education using an education reform', *European Economic Review* **54**(4), 483–500.

<https://doi.org/10.1016/j.euroecorev.2009.09.001>.

Abekah-Nkrumah, G., Asuming, P. O. and Yusif, H. (2019), 'Duration of high school education and youth labour market outcomes: Evidence from a policy experiment in Ghana', *Journal of International Development* **31**(7), 617–631.

<https://doi.org/10.1002/jid.3429>.

Acemoglu, D. and Angrist, J. (2000), 'How large are human-capital externalities? evidence from compulsory schooling laws', *NBER Macroeconomics Annual* **15**, 9–59.

<https://doi.org/10.1086/654403>.

Alam, K. and Mamun, S. A. K. (2016), 'The relationship between labour force status and educational attainment: Evidence from a system of simultaneous equations model', *Economic Analysis and Policy* **52**, 55–65.

<https://doi.org/10.1016/j.eap.2016.07.005>.

Anand, R., Kothari, S. and Kumar, N. (2016), *South Africa: Labor market dynamics and inequality*, International Monetary Fund.

Andini, C. (2010), 'Within-groups wage inequality and schooling: further evidence for Portugal', *Applied Economics* **42**(28), 3685–3691.

<https://doi.org/10.1080/00036840802314564>.

Angrist, J. D. and Keueger, A. B. (1991), 'Does compulsory school attendance affect schooling and earnings?', *The Quarterly Journal of Economics* **106**(4), 979–1014.

<https://doi.org/10.2307/2937954>.

Angrist, J. D. and 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.

<https://doi.org/10.2307/2290263>.

Arellano, M. and Bonhomme, S. (2017), 'Quantile selection models with an application to understanding changes in wage inequality', *Econometrica* **85**(1), 1–28.

<https://doi.org/10.3982/ECTA14030>.

Arrow, K. J. et al. (1973), 'Higher education as a filter', *Journal of Public Economics* **2**(3), 193–216.

[https://doi.org/10.1016/0047-2727\(73\)90013-3](https://doi.org/10.1016/0047-2727(73)90013-3).

Ashenfelter, O. and Krueger, A. (1994), 'Estimates of the economic return to schooling from a new sample of twins', *The American Economic Review* pp. 1157–1173.

<https://www.jstor.org/stable/2117766>.

Aydemir, A. and Kirdar, M. G. (2017), 'Low wage returns to schooling in a developing country: Evidence from a major policy reform in Turkey', *Oxford Bulletin of Economics and Statistics* **79**(6), 1046–1086.

<https://doi.org/10.1111/obes.12174>.

Baah-Boateng, W. (2013), 'Determinants of unemployment in Ghana', *African Development Review* **25**(4), 385–399.

<https://doi.org/10.1111/1467-8268.12037>.

Baah-Boateng, W. (2015), 'Unemployment in Ghana: a cross sectional analysis from demand and supply perspectives', *African Journal of Economic and Management Studies* .

<https://doi.org/10.1108/AJEMS-11-2014-0089>.

- Baigrie, N. and Eyal, K. (2014), 'An evaluation of the determinants and implications of panel attrition in the national income dynamics survey (2008-2010)', *South African Journal of Economics* **82**(1), 39–65.
<https://doi.org/10.1111/saje.12024>.
- Baker and McKenzie (2023), 'Guidebook to broad-based black economic empowerment'.
- Baldry, K. (2016), 'Graduate unemployment in South Africa: social inequality reproduced', *Journal of Education and Work* **29**(7), 788–812.
<https://doi.org/10.1080/13639080.2015.1066928>.
- Balestra, S. and Backes-Gellner, U. (2017), 'Heterogeneous returns to education over the wage distribution: Who profits the most?', *Labour Economics* **44**, 89–105.
<https://doi.org/10.1016/j.labeco.2017.01.001>.
- Balestra, S., Eugster, B. and Liebert, H. (2020), 'Summer-born struggle: The effect of school starting age on health, education, and work', *Health economics* **29**(5), 591–607.
<https://doi.org/10.1002/hec.4005>.
- Bardasi, E. and Taylor, M. (2008), 'Marriage and wages: A test of the specialization hypothesis', *Economica* **75**(299), 569–591.
<https://doi.org/10.1111/j.1468-0335.2007.00630.x>.
- Barreca, A. I., Guldi, M., Lindo, J. M. and Waddell, G. R. (2011), 'Saving babies? revisiting the effect of very low birth weight classification', *The Quarterly Journal of Economics* **126**(4), 2117–2123.
<https://doi.org/10.1093/qje/qjr042>.
- Beck, K. C. (2025), 'The later the better? a novel approach to estimating the effect of school starting age on adhd and academic skills', *European Sociological Review* p. jcaf004.
<https://doi.org/10.1093/esr/jcaf004>.
- Becker, G. S. (1962), 'Investment in human capital: A theoretical analysis', *Journal of Political Economy* **70**(5, Part 2), 9–49.

<https://doi.org/10.1086/258724>.

Berhe, M. W. (2021), 'Empirical analysis of urban youth unemployment in Ethiopia', *African Development Review* .

<https://doi.org/10.1111/1467-8268.12514>.

Bhorat, H. (2000), *Wage premia and wage differentials in the South African labour market*, University of Cape Town.

<https://open.uct.ac.za/items/56f7acb6-08a4-4a36-b29c-b6bcd1ea7f92>.

Bhorat, H. and Goga, S. (2012), 'The gender wage gap in the post-apartheid South African labour market'.

<http://dx.doi.org/10.2139/ssrn.2184258>.

Bhorat, H., Goga, S. and Stanwix, B. (2014), Skills-biased labour demand and the pursuit of inclusive growth in South Africa, Technical report, WIDER working paper.

<https://doi.org/10.35188/UNU-WIDER/2014/851-3>.

Bhorat, H., Kanbur, R. and Mayet, N. (2013), 'The impact of sectoral minimum wage laws on employment, wages, and hours of work in South Africa', *IZA Journal of Labor & Development* **2**(1), 1–27.

<http://www.izajold.com/content/2/1/1>.

Bhorat, H., Kanbur, R. and Stanwix, B. (2014), 'Estimating the impact of minimum wages on employment, wages, and non-wage benefits: the case of agriculture in South Africa', *American Journal of Agricultural Economics* **96**(5), 1402–1419.

<https://doi.org/10.1093/ajae/aau049>.

Bhorat, H. and Kimani, E. (2017), 'The role of post-school education and training institutions in predicting labour market outcomes', *HSRC LIMP Report* **23**.

Bhorat, H., Stanwix, B. et al. (2021), 'The impact of the national minimum wage in South Africa: Early quantitative evidence'.

<https://coilink.org/20.500.12592/1w7g30>.

- Biyase, M. and Zwane, T. (2015), 'Does education pay in South Africa? estimating returns to education using two stage least squares approach', *The International Business & Economics Research Journal (Online)* **14**(6), 807.
<https://core.ac.uk/download/pdf/268108007.pdf>.
- Blau, F. D., Kahn, L. M., Boboshko, N. and Comey, M. (2024), 'The impact of selection into the labor force on the gender wage gap', *Journal of Labor Economics* **42**(4), 1093–1133.
<https://doi.org/10.1086/725032>.
- Boahen, E. A., Opoku, K. and Schotte, S. (2021), 'Duration of pre-university education and labour market outcomes: Evidence from a quasi-experiment in Ghana', *Journal of International Development* **33**(1), 208–232.
<https://doi.org/10.1002/jid.3519>.
- Boockmann, B. and Steiner, V. (2006), 'Cohort effects and the returns to education in West Germany', *Applied Economics* **38**(10), 1135–1152.
<https://doi.org/10.1080/00036840500439168>.
- Bosch, A. and Koch, S. F. (2020), 'The South African financial cycle and its relation to household deleveraging', *South African Journal of Economics* **88**(2), 145–173.
<https://doi.org/10.1111/saje.12245>.
- Bound, J., Jaeger, D. A. and Baker, R. M. (1995), 'Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak', *Journal of the American statistical association* **90**(430), 443–450.
<https://doi.org/10.1080/01621459.1995.10476536>.
- Branson, N., Ardington, C., Lam, D. and Leibbrandt, M. (2013), 'Changes in education, employment and earnings in South Africa: A cohort analysis'.
<http://hdl.handle.net/11090/621>.
- Branson, N., Garlick, J., Lam, D. and Leibbrandt, M. (2012), 'Education and inequality: The South African case'.
<http://hdl.handle.net/11090/168>.

- Branson, N., Hofmeyr, C. and Lam, D. (2014), 'Progress through school and the determinants of school dropout in South Africa', *Development Southern Africa* **31**(1), 106–126.
<https://doi.org/10.1080/0376835X.2013.853610>.
- Branson, N. and Leibbrandt, M. (2013a), 'Education quality and labour market outcomes in South Africa'.
- <https://doi.org/10.1787/5k4c0vvbv0q-en>.
- Branson, N. and Leibbrandt, M. (2013b), 'Educational attainment and labour market outcomes in South Africa, 1994-2010'.
- <https://doi.org/10.1787/5k4c0vvbv0q-en>.
- Brunello, G., Fort, M. and Weber, G. (2009), 'Changes in compulsory schooling, education and the distribution of wages in Europe', *The Economic Journal* pp. 516–539.
<https://doi.org/10.1111/j.1468-0297.2008.02244.x>.
- Buchinsky, M. (1994), 'Changes in the US wage structure 1963-1987: Application of quantile regression', *Econometrica* pp. 405–458.
<https://doi.org/10.2307/2951618>.
- Buchinsky, M. (1998), 'The dynamics of changes in the female wage distribution in the USA: a quantile regression approach', *Journal of Applied Econometrics* **13**(1), 1–30.
[https://doi.org/10.1002/\(SICI\)1099-1255\(199801/02\)13:1<1::AID-JAE474>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1099-1255(199801/02)13:1<1::AID-JAE474>3.0.CO;2-A).
- Buchinsky, M. (2001), 'Quantile regression with sample selection: Estimating women's return to education in the US', *Empirical Economics* **26**(1), 87–113.
- Burger, R., Jafta, R. et al. (2010), 'Affirmative action in South Africa: an empirical assessment of the impact on labour market outcomes', *CRISE (Centre for Research on Inequality, Human Security and Ethnicity) Working Paper* **76**, 09–36.

Burger, R., Van der Berg, S. and Von Fintel, D. (2015), 'The unintended consequences of education policies on South African participation and unemployment', *South African Journal of Economics* **83**(1), 74–100.

<https://doi.org/10.1111/saje.12049>.

Burger, R., Von Fintel, D. et al. (2009), 'Determining the causes of the rising South African unemployment rate: An age, period and generational analysis', *Economic Research Southern Africa (ERSA) Working Papers* **158**.

Calonico, S., Cattaneo, M. D., Farrell, M. H. and Titiunik, R. (2017), 'rdrobust: Software for regression-discontinuity designs', *The Stata Journal* **17**(2), 372–404.

<https://doi.org/10.1177/1536867X1701700208>.

Calonico, S., Cattaneo, M. D. and Titiunik, R. (2014a), 'Robust data-driven inference in the regression-discontinuity design', *The Stata Journal* **14**(4), 909–946.

<https://doi.org/10.1177/1536867X1401400413>.

Calonico, S., Cattaneo, M. D. and Titiunik, R. (2014b), 'Robust nonparametric confidence intervals for regression-discontinuity designs', *Econometrica* **82**(6), 2295–2326.

<https://doi.org/10.3982/ECTA11757>.

Card, D. (2001), 'Estimating the return to schooling: Progress on some persistent econometric problems', *Econometrica* **69**(5), 1127–1160.

<https://doi.org/10.1111/1468-0262.00237>.

Casale, D. (2004), 'What has the feminisation of the labour market 'bought' women in South Africa? Trends in labour force participation, employment and earnings, 1995–2001', *Journal of Interdisciplinary Economics* **15**(3-4), 251–275.

<https://doi.org/10.1177/02601079X04001500302>.

Case, A. and Deaton, A. (1999), 'School inputs and educational outcomes in South Africa', *The Quarterly Journal of Economics* **114**(3), 1047–1084.

<https://doi.org/10.1162/003355399556124>.

- Case, A. and Yogo, M. (1999), 'Does school quality matter? returns to education and the characteristics of schools in South Africa'.
<https://doi.org/10.3386/w7399>.
- Castel, V., Phiri, M. and Stampini, M. (2010), 'Education and employment in Malawi', *African Development Bank Working Paper* .
- Cattaneo, M. D., Idrobo, N. and Titiunik, R. (2020), 'A practical introduction to regression discontinuity designs: Extensions', *Cambridge Elements: Quantitative and Computational Methods for Social Science* .
<https://doi.org/10.1017/9781009441896>.
- Cattaneo, M. D., Jansson, M. and Ma, X. (2020), 'Simple local polynomial density estimators', *Journal of the American Statistical Association* **115**(531), 1449–1455.
<https://doi.org/10.1080/01621459.2019.1635480>.
- Chiswick, B. R. (2024), 'Estimating returns to schooling and experience: a history of thought', *Education Economics* **32**(5), 566–580.
<https://doi.org/10.1080/09645292.2024.2326015>.
- Chuanchuan, Z. and Jingwen, W. (2021), 'Gender roles and women's labor market outcomes', *China Economic Quarterly International* **1**(2), 97–108.
<https://doi.org/10.1016/j.ceqi.2021.04.002>.
- Clark, D. (2023), 'School quality and the return to schooling in Britain: New evidence from a large-scale compulsory schooling reform', *Journal of Public Economics* **223**, 104902.
<https://doi.org/10.1093/ajae/aau049>.
- Clark, D. and Martorell, P. (2014), 'The signaling value of a high school diploma', *Journal of Political Economy* **122**(2), 282–318.
<https://doi.org/10.1086/675238>.

- Clay, K., Lingwall, J. and Stephens Jr, M. (2021), 'Laws, educational outcomes, and returns to schooling evidence from the first wave of US state compulsory attendance laws', *Labour Economics* **68**, 101935.
<https://doi.org/10.1016/j.labeco.2020.101935>.
- Conradie, B. (2003), 'Labour, wages and minimum wage compliance in the Breërivier valley six months after the introduction of minimum wages', *Centre for Social Science Research Working Paper No. 51* .
<https://hdl.handle.net/11427/19725>.
- Conradie, B. (2005), 'Wages and wage elasticities for wine and table grapes in South Africa', *Agrekon* **44**(1), 138–156.
<https://doi.org/10.1080/03031853.2005.9523706>.
- Dai, F., Cai, F. and Zhu, Y. (2022), 'Returns to higher education in China—evidence from the 1999 higher education expansion using a fuzzy regression discontinuity', *Applied Economics Letters* **29**(6), 489–494.
<https://doi.org/10.1080/13504851.2020.1871465>.
- de New, S. C., Schurer, S. and Sulzmaier, D. (2021), 'Gender differences in the lifecycle benefits of compulsory schooling policies', *European Economic Review* **140**, 103910.
<https://doi.org/10.1016/j.euroecorev.2021.103910>.
- Department of Employment and Labour (2023), 'Commission for Employment Equity: Annual Report 2022 - 2023'.
- Depken, C., Chiseni, C. and 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.
<https://doi.org/10.2478/zireb-2019-0009>.
- Devereux, P. J. and Hart, R. A. (2010), 'Forced to be rich? returns to compulsory schooling in Britain', *The Economic Journal* **120**(549), 1345–1364.
<https://doi.org/10.1111/j.1468-0297.2010.02365.x>.

- Diawara, B., Mughal, M. et al. (2011), 'Education and employment in Sub-Saharan Africa: Some evidence from Senegal', *Theoretical and Applied Economics* **5**(5), 59.
- Dobkin, C. and Ferreira, F. (2010), 'Do school entry laws affect educational attainment and labor market outcomes?', *Economics of education review* **29**(1), 40–54.
<https://doi.org/10.1016/j.econedurev.2009.04.003>.
- Dolado, J. J., García-Peñalosa, C. and Tarasonis, L. (2020), 'The changing nature of gender selection into employment over the great recession', *Economic Policy* **35**(104), 635–677.
<https://doi.org/10.1093/epolic/eiaa025>.
- Dolton, P. and Sandi, M. (2017), 'Returning to returns: Revisiting the British education evidence', *Labour Economics* **48**, 87–104.
<https://doi.org/10.1016/j.labeco.2017.07.001>.
- Dougherty, C. (2005), 'Why are the returns to schooling higher for women than for men?', *Journal of Human Resources* **40**(4), 969–988.
<https://doi.org/10.3368/jhr.XL.4.969>.
- Draissi, Z., Rong, Y. and Ebrahim Suliman, M. A. (2023), 'Employability and education mismatch in the Moroccan urban workplace', *SAGE Open* **13**(4), 21582440231217890.
<https://doi.org/10.1177/21582440231217890>.
- Drucker, L. F., Horn, D. and Jakubowski, M. (2022), 'The labour market effects of the Polish educational reform of 1999', *Journal for Labour Market Research* **56**(1), 1–21.
<https://doi.org/10.1186/s12651-022-00317-4>.
- Duflo, E. (2001), 'Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment', *American economic review* **91**(4), 795–813.
- Elass, K. (2024), 'Male and female selection effects on gender wage gaps in three countries', *Labour Economics* **87**, 102506.
<https://doi.org/10.1016/j.labeco.2024.102506>.

- Espi, G., Francis, D. and Valodia, I. (2019), 'Gender inequality in the South African labour market: Insights from the employment equity act data', *Agenda* **33**(4), 44–61.
<https://doi.org/10.1080/10130950.2019.1674675>.
- Fan, E., Meng, X., Wei, Z. and Zhao, G. (2018), 'Rates of return to four-year university education: An application of regression discontinuity design', *The Scandinavian Journal of Economics* **120**(4), 1011–1042.
<https://doi.org/10.1111/sjoe.12251>.
- Fasih, T., Kingdon, G. G., Patrinos, H. A., Sakellariou, C. and Söderbom, M. (2012), 'Heterogeneous returns to education in the labor market'.
<https://ssrn.com/abstract=2129612>.
- Festus, L., Kasongo, A., Moses, M. and Yu, D. (2016), 'The South African labour market, 1995–2015', *Development Southern Africa* **33**(5), 579–599.
<https://doi.org/10.1080/0376835X.2016.1203759>.
- Finn, A. (2015), 'A National Minimum Wage in the Context of the South African Labour Market', Technical report, Southern Africa Labour and Development Research Unit Working Paper Number 153, University of Cape Town.
- Finn, A. and Leibbrandt, M. (2018), The evolution and determination of earnings inequality in post-apartheid South Africa, Technical report, WIDER Working Paper.
<https://hdl.handle.net/10419/190132>.
- Francis, D. and Webster, E. (2019), 'Poverty and inequality in South Africa: critical reflections', *Development Southern Africa* **36**(6), 788–802.
<https://doi.org/10.1080/0376835X.2019.1666703>.
- Fredriksson, P. and Öckert, B. (2014), 'Life-cycle effects of age at school start', *The Economic Journal* **124**(579), 977–1004.
<https://doi.org/10.1111/ecoj.12047>.

- Fryer, D. and Vencatachellum, D. (2005), 'Returns to education in South Africa: Evidence from the Machibisa township', *African Development Review* **17**(3), 513–535.
<https://doi.org/10.1111/j.1017-6772.2006.00126.x>.
- Gelman, A. and Imbens, G. (2019), 'Why high-order polynomials should not be used in regression discontinuity designs', *Journal of Business & Economic Statistics* **37**(3), 447–456.
<https://doi.org/10.1080/07350015.2017.1366909>.
- Giotis, G., Gogas, T. and Gouda, K. (2025), 'Education and employment', *Encyclopedia* **5**(2), 85.
<https://doi.org/10.3390/>.
- Girma, S. and Kedir, A. (2005), 'Heterogeneity in returns to schooling: Econometric evidence from Ethiopia', *The Journal of Development Studies* **41**(8), 1405–1416.
<https://doi.org/10.1080/00220380500187026>.
- González, F. A. I. and Dip, J. A. (2024), 'School entry cutoff and the timing of births: evidence from Argentina', *Education Economics* **32**(1), 47–58.
<https://doi.org/10.1080/09645292.2023.2173148>.
- Görlitz, K., Heß, P. and Tamm, M. (2025), 'Should states allow early school enrollment? an analysis of individuals' long-term labor market effects', *Empirical Economics* pp. 1–29.
<https://doi.org/10.1007/s00181-024-02702-z>.
- Gradín, C. (2019), 'Occupational segregation by race in South Africa after apartheid', *Review of Development Economics* **23**(2), 553–576.
<https://doi.org/10.1111/rode.12551>.
- Granados, P. G., Wrohlich, K. et al. (2020), *Selection into Employment and the Gender Wage Gap across the Distribution and over Time*, Deutsches Institut für Wirtschaftsforschung (DIW).
<https://hdl.handle.net/10419/215255>.
- Grenet, J. (2013), 'Is extending compulsory schooling alone enough to raise earnings? evidence from French and British compulsory schooling laws', *The Scandinavian Journal of Economics* **115**(1), 176–210.

<https://doi.org/10.1111/j.1467-9442.2012.01739.x>.

Gronau, R. (1974), 'Wage comparisons—a selectivity bias', *Journal of Political Economy* **82**(6), 1119–1143.

<https://doi.org/10.1086/260267>.

Guo, C., Wang, X. and Meng, C. (2023), 'Does the early bird catch the worm? evidence and interpretation on the long-term impact of school entry age in China', *China Economic Review* **77**, 101900.

<https://doi.org/10.1016/j.chieco.2022.101900>.

Hall, C. (2016), 'Does more general education reduce the risk of future unemployment? evidence from an expansion of vocational upper secondary education', *Economics of Education Review* **52**, 251–271.

<https://doi.org/10.1016/j.econedurev.2016.03.005>.

Harmon, C., Oosterbeek, H. and Walker, I. (2000), *The returns to education: a review of evidence, issues and deficiencies in the literature*, Centre for the Economics of Education, London School of Economics and Political Science.

Harmon, C., Oosterbeek, H. and Walker, I. (2003), 'The returns to education: Microeconomics', *Journal of economic surveys* **17**(2), 115–156.

Harmon, C. P. (2017), 'How effective is compulsory schooling as a policy instrument?', *IZA World of Labor* .

<https://doi.org/10.15185/izawol.348>.

Harmon, C. and Walker, I. (1995), 'Estimates of the economic return to schooling for the United Kingdom', *The American Economic Review* **85**(5), 1278–1286.

<https://www.jstor.org/stable/2950988>.

Hartog, J., Pereira, P. T. and Vieira, J. A. (2001), 'Changing returns to education in Portugal during the 1980s and early 1990s: OLS and quantile regression estimators', *Applied Economics* **33**(8), 1021–1037.

<https://doi.org/10.1080/00036840122679>.

Heckman, J. (1974), 'Shadow prices, market wages, and labor supply', *Econometrica* pp. 679–694.

<https://doi.org/10.2307/1913937>.

Heckman, J. J. (1979), 'Sample selection bias as a specification error', *Econometrica* pp. 153–161.

<https://doi.org/10.2307/1912352>.

Hertz, T. (2003), 'Upward bias in the estimated returns to education: Evidence from South Africa', *American Economic Review* **93**(4), 1354–1368.

<https://doi.org/10.1257/000282803769206331>.

Hinks, T. (2002), 'Gender wage differentials and discrimination in the new South Africa', *Applied Economics* **34**(16), 2043–2052.

<https://doi.org/10.1080/00036840210124991>.

Horwitz, F. M. and Jain, H. (2011), 'An assessment of employment equity and broad based black economic empowerment developments in South Africa', *Equality, Diversity and Inclusion: An International Journal* **30**(4), 297–317.

<https://doi.org/10.1108/02610151111135750>.

Huang, C., Zhang, S. and Zhao, Q. (2020), 'The early bird catches the worm? school entry cutoff and the timing of births', *Journal of Development Economics* **143**, 102386.

<https://doi.org/10.1016/j.jdeveco.2019.102386>.

Huber, M. and Melly, B. (2015), 'A test of the conditional independence assumption in sample selection models', *Journal of Applied Econometrics* **30**(7), 1144–1168.

<https://doi.org/10.1002/jae.2431>.

Huntington-Klein, N. (2021), 'Human capital versus signaling is empirically unresolvable', *Empirical Economics* **60**(5), 2499–2531.

<https://doi.org/10.1007/s00181-020-01837-z>.

ILO (2024), *World employment and social outlook; Trends 2024*, ILO, Geneva.

<https://doi.org/10.54394/HQAE1085>.

- Kawaguchi, D. (2011), 'Actual age at school entry, educational outcomes, and earnings', *Journal of the Japanese and International Economies* **25**(2), 64–80.
<https://doi.org/10.1016/j.jjie.2009.02.002>.
- Kaymak, B. (2025), 'Quantifying the signaling role of education', *Federal Reserve Bank of Cleveland Working Paper Series Working Paper* .
<https://doi.org/10.26509/frbc-wp-202502>.
- Kercheval, J., Markowitz, D., Monson, K., Read, B. and Brown, D. (2013), 'Women in leadership—research on barriers to employment and decent work for women. literature review', *International Labour Organization (ILO)* .
- Keswell, M. and Poswell, L. (2002), 'How important is education for getting ahead in South Africa?'.
<http://hdl.handle.net/11427/20157>.
- Keswell, M. and Poswell, L. (2004), 'Returns to education in South Africa: A retrospective sensitivity analysis of the available evidence', *South African Journal of Economics* **72**(4), 834–860.
<https://doi.org/10.1111/j.1813-6982.2004.tb00136.x>.
- Kingdon, G. G. and Knight, J. (2004), 'Race and the incidence of unemployment in South Africa', *Review of Development Economics* **8**(2), 198–222.
<https://doi.org/10.1111/j.1467-9361.2004.00228.x>.
- Klasen, S. and Minasyan, A. (2021), 'Affirmative action and intersectionality at the top: Evidence from South Africa', *Industrial Relations* **60**(1), 3–35.
<https://doi.org/10.1111/irel.12272>.
- Koenker, R. and Bassett Jr, G. (1978), 'Regression quantiles', *Econometrica* pp. 33–50.
<https://doi.org/10.2307/1913643>.

- Kollamparambil, U. and Razak, A. (2016), 'Trends in gender wage gap and discrimination in South Africa: A comparative analysis across races', *Indian Journal of Human Development* **10**(1), 49–63.
<https://doi.org/10.1177/0973703016636446>.
- Korwatanasakul, U. (2023), 'Returns to schooling in Thailand: Evidence from the 1978 compulsory schooling law', *The Developing Economies* **61**(1), 3–35.
<https://doi.org/10.1111/deve.12341>.
- Kuepié, M. and Nordman, C. J. (2016), 'Where does education pay off in Sub-Saharan Africa? Evidence from two cities of the Republic of Congo', *Oxford Development Studies* **44**(1), 1–27.
<https://doi.org/10.1080/13600818.2015.1110568>.
- Lam, D. (2016), 'The impact of race on earnings and human capital in Brazil, South Africa and the United States', *Anais* pp. 1–26.
- Lam, D., Ardington, C. and Leibbrandt, M. (2011), 'Schooling as a lottery: Racial differences in school advancement in urban South Africa', *Journal of Development Economics* **95**(2), 121–136.
<https://doi.org/10.1016/j.jdeveco.2010.05.005>.
- Lam, D., Leibbrandt, M. and Mlatsheni, C. (2009), Education and youth unemployment in South Africa, in 'Labor Markets and Economic Development', Routledge, pp. 90–109.
- Landman, M. S., O'Clery, N. et al. (2020), The impact of the employment equity act on female inter-industry labour mobility and the gender wage gap in South Africa, Technical report, World Institute for Development Economic Research (UNU-WIDER).
<https://hdl.handle.net/10419/229276>.
- Larsen, E. R. and Solli, I. F. (2017), 'Born to run behind? persisting birth month effects on earnings', *Labour Economics* **46**, 200–210.
<http://dx.doi.org/10.1016/j.labeco.2016.10.005>.
- Lewis, B. D. and Nguyen, H. T. (2020), 'Assessing the causal impact of compulsory schooling policy in Indonesia', *International Journal of Educational Research* **104**, 101693.

<https://doi.org/10.1016/j.ijer.2020.101693>.

Li, X., Lou, X. and Zhang, J. (2022), 'Does the early bird catch the worm? the effect of school starting age on educational attainment and labor market outcomes: Evidence from Chinese urban twins data', *Journal of Comparative Economics* **50**(3), 832–848.

<https://doi.org/10.1016/j.jce.2022.04.004>.

Liu, S. and Marois, G. (2023), 'The effect of motherhood on the labour force participation of married women in China', *Asian Population Studies* pp. 1–17.

<https://doi.org/10.1080/17441730.2023.2193518>.

Liwiński, J. (2020), 'The impact of compulsory education on employment and wages in a transition economy', *Eastern European Economics* **58**(2), 137–173.

<https://doi.org/10.1080/00128775.2019.1689144>.

Maasoumi, E. and Wang, L. (2019), 'The gender gap between earnings distributions', *Journal of Political Economy* **127**(5), 2438–2504.

<https://doi.org/10.1086/701788>.

Marcenaro-Gutierrez, O. D. and Lopez-Agudo, L. A. (2021), 'Too late or too soon for school? The impact of school entry age', *Journal of research on educational effectiveness* **14**(2), 309–352.

<https://doi.org/10.1080/19345747.2020.1849479>.

Markussen, S. and Strøm, M. (2022), 'Children and labor market outcomes: separating the effects of the first three children', *Journal of Population Economics* **35**(1), 135–167.

<https://doi.org/10.1007/s00148-020-00807-0>.

Martins, P. S. and Pereira, P. T. (2004), 'Does education reduce wage inequality? quantile regression evidence from 16 countries', *Labour Economics* **11**(3), 355–371.

<https://doi.org/10.1016/j.labeco.2003.05.003>.

Matta, R., Ribas, R. P., Sampaio, B. and Sampaio, G. R. (2016), 'The effect of age at school entry on college admission and earnings: a regression-discontinuity approach', *IZA Journal of Labor Economics* **5**, 1–25.

<https://doi.org/10.1186/s40172-016-0049-5>.

McCrary, J. (2008), 'Manipulation of the running variable in the regression discontinuity design: A density test', *Journal of Econometrics* **142**(2), 698–714.

<https://doi.org/10.1016/j.jeconom.2007.05.005>.

Millea, M. J., Rezek, J. P., Shoup, B. and Pitts, J. (2017), 'Minimum wages in a segmented labor market: Evidence from South Africa', *Journal of Labor Research* **38**, 335–359.

<https://doi.org/10.1007/s12122-017-9241-z>.

Mincer, J. (1958), 'Investment in human capital and personal income distribution', *Journal of Political Economy* **66**(4), 281–302.

<https://doi.org/10.1086/258055>.

Mincer, J. (1974), 'Schooling, experience, and earnings. Human Behavior & Social Institutions No. 2.'

Mishra, V. and Smyth, R. (2015), 'Estimating returns to schooling in urban China using conventional and heteroskedasticity-based instruments', *Economic Modelling* **47**, 166–173.

<https://doi.org/10.1016/j.econmod.2015.02.002>.

Mitra, A. (2019), 'Returns to education in India: Capturing the heterogeneity', *Asia & the Pacific Policy Studies* **6**(2), 151–169.

<https://doi.org/10.1002/app5.271>.

Mlachila, M. M. and Moeletsi, T. (2019), 'Struggling to make the grade: A review of the causes and consequences of the weak outcomes of South Africa's education system', *IMF Working Papers* .

Moll, P. G. (1992), 'Quality of education and the rise in returns to schooling in South Africa, 1975–1985', *Economics of Education Review* **11**(1), 1–10.

[https://doi.org/10.1016/0272-7757\(92\)90016-V](https://doi.org/10.1016/0272-7757(92)90016-V).

Moll, P. G. (1998), 'Primary schooling, cognitive skills and wages in South Africa', *Economica* **65**(258), 263–284.

<https://doi.org/10.1111/1468-0335.00127>.

Mosomi, J. (2019a), 'Distributional changes in the gender wage gap in the post-apartheid South African labour market'.

<https://hdl.handle.net/10419/211247>.

Mosomi, J. (2019b), 'An empirical analysis of trends in female labour force participation and the gender wage gap in South Africa', *Agenda* **33**(4), 29–43.

<https://doi.org/10.1080/10130950.2019.1656090>.

Mosomi, J. W. et al. (2020), 'The labor market in South Africa, 2000–2017', *IZA World of Labor* .

<https://doi.org/10.15185/izawol.475>.

Msila, V. (2007), 'From apartheid education to the revised national curriculum statement: Pedagogy for identity formation and nation building in South Africa', *Nordic Journal of African Studies* **16**(2).

<https://doi.org/10.53228/njas.v16i2.63>.

Mtapuri, O. and Tinarwo, P. (2021), 'From apartheid to democracy', *Southern African Journal of Demography* **21**(1), 104–133.

<https://doi.org/10.2307/27125725>.

Muñoz, E. and Siravegna, M. (2021), 'Implementing quantile selection models in stata', *The Stata Journal* **21**(4), 952–971.

<https://doi.org/10.1177/1536867X211063148>.

Murray, J. and Van Walbeek, C. (2007), 'Impact of the sectoral determination for farm workers on the South African sugar industry: Case study of the KwaZulu-Natal North and South Coasts', *Agrekon* **46**(1), 94–112.

<https://doi.org/10.1080/03031853.2007.9523763>.

Mwabu, G. and Schultz, T. P. (1996), 'Education returns across quantiles of the wage function: alternative explanations for returns to education by race in South Africa', *The American Economic Review* **86**(2), 335–339.

<https://www.jstor.org/stable/2118148>.

- Mwabu, G. and 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.
<https://doi.org/10.1086/452460>.
- Nattrass, N. (2016), 'Setting minimum wages in South Africa', *International Labour Organization* **2**, 704.
- Naude, W. and Serumaga-Zake, P. (2001), 'An analysis of the determinants of labour force participation and unemployment in South Africa's North-West province', *Development Southern Africa* **18**(3), 261–278.
<https://doi.org/10.1080/03768350120041929>.
- Nikolov, P., Jimi, N. and Chang, J. (2020), 'The importance of cognitive domains and the returns to schooling in South Africa: Evidence from two labor surveys', *Labour Economics* **65**, 101849.
<https://doi.org/10.1016/j.labeco.2020.101849>.
- Ntuli, M. (2009), Exploring gender wage “discrimination” in South Africa, 1995–2004: a quantile regression approach, in 'Labor Markets and Economic Development', Routledge, pp. 498–525.
- Ntuli, M. and Wittenberg, M. (2013), 'Determinants of black women's labour force participation in post-apartheid South Africa', *Journal of African Economies* **22**(3), 347–374.
<https://doi.org/10.1093/jae/ejs039>.
- Oosterbeek, H., ter Meulen, S. and van Der Klaauw, B. (2021), 'Long-term effects of school-starting-age rules', *Economics of Education Review* **84**, 102144.
<https://doi.org/10.1016/j.econedurev.2021.102144>.
- Oreopoulos, P. (2006a), 'The compelling effects of compulsory schooling: Evidence from Canada', *Canadian Journal of Economics / Revue canadienne d'économie* **39**(1), 22–52.
<https://doi.org/10.1111/j.0008-4085.2006.00337.x>.
- Oreopoulos, P. (2006b), 'Estimating average and local average treatment effects of education when compulsory schooling laws really matter', *American Economic Review* **96**(1), 152–175.
<https://doi.org/10.1257/000282806776157641>.

- Oreopoulos, P. and Salvanes, K. G. (2011), 'Priceless: The nonpecuniary benefits of schooling', *Journal of Economic Perspectives* **25**(1), 159–184.
<https://doi.org/10.1257/jep.25.1.159>.
- Oyelere, R. U. (2010), 'Africa's education enigma? The Nigerian story', *Journal of Development Economics* **91**(1), 128–139.
<https://doi.org/10.1016/j.jdeveco.2009.04.001>.
- Oyenubi, A. and Mosomi, J. (2024), 'Utility of inequality sensitive measures of the gender wage gap: Evidence from South Africa', *Economic Analysis and Policy* **81**, 576–590.
<https://doi.org/10.1016/j.eap.2023.12.017>.
- Patrinos, H. A. (2016), 'Estimating the return to schooling using the mincer equation', *IZA World of Labor* .
- Patrinos, H. A., Psacharopoulos, G. and Tansel, A. (2021), 'Private and social returns to investment in education: the case of Turkey with alternative methods', *Applied Economics* **53**(14), 1638–1658.
<https://doi.org/10.1080/00036846.2020.1841086>.
- Patrinos, H. A., Ridao-Cano, C. and Sakellariou, C. (2006), *Estimating the returns to education: accounting for heterogeneity in ability*, The World Bank.
- Patrinos, H. A. and Sakellariou, C. (2006), 'Economic volatility and returns to education in Venezuela: 1992–2002', *Applied Economics* **38**(17), 1991–2005.
<https://doi.org/10.1080/00036840500427338>.
- Pehkonen, J., Viinikainen, J., Böckerman, P., Pulkki-Råback, L., Keltikangas-Järvinen, L. and Raitakari, O. (2015), 'Relative age at school entry, school performance and long-term labour market outcomes', *Applied Economics Letters* **22**(16), 1345–1348.
<http://dx.doi.org/10.1080/13504851.2015.1031864>.
- Peña, P. A. (2017), 'Creating winners and losers: Date of birth, relative age in school, and outcomes in childhood and adulthood', *Economics of Education Review* **56**, 152–176.

<https://doi.org/10.1016/j.econedurev.2016.12.001>.

Piek, M., von Fintel, D. and Kirsten, J. (2023), 'The impact of agricultural minimum wages on worker flows in South Africa', *South African Journal of Economics* .

<https://doi.org/10.1111/saje.12357>.

Posel, D. (2001), 'How do households work? migration, the household and remittance behaviour in South Africa', *Social Dynamics* **27**(1), 165–189.

<https://doi.org/10.1080/02533950108458709>.

Posel, D. (2010), 'Households and labour migration in post-apartheid South Africa', *Studies in Economics and Econometrics* **34**(3), 129–141.

<https://hdl.handle.net/10520/EJC21503>.

Posel, D. (2025), 'Who cares? gender differences in social reproduction and well-being in South Africa', *Gender, Work & Organization* .

<https://doi.org/10.1111/gwao.13249>.

Posel, D. and Bruce-Brand, J. (2021), "Only a housewife?" Subjective well-being and homemaking in South Africa', *Journal of Happiness Studies* **22**(1), 323–342.

<https://doi.org/10.1007/s10902-020-00232-w>.

Prieto-Rodriguez, J., Barros, C. P. and Vieira, J. A. (2008), 'What a quantile approach can tell us about returns to education in Europe', *Education Economics* **16**(4), 391–410.

<https://doi.org/10.1080/09645290801977033>.

Psacharopoulos, G. (2024), 'Returns to education: a brief history and an assessment', *Education Economics* **32**(5), 561–565.

<https://doi.org/10.1080/09645292.2024.2370119>.

Psacharopoulos, G. and Patrinos, H. A. (2018), 'Returns to investment in education: a decennial review of the global literature', *Education Economics* **26**(5), 445–458.

<https://doi.org/10.1080/09645292.2018.1484426>.

BIBLIOGRAPHY

- Rahman, M. and Al-Hasan, M. (2018), 'Returns to schooling in Bangladesh revisited', *The Bangladesh Development Studies* **41**(2), 27–42.
<https://www.jstor.org/stable/27031074>.
- Ranchhod, V. and Daniels, R. C. (2021), 'Labour market dynamics in South Africa at the onset of the covid-19 pandemic', *South African Journal of Economics* **89**(1), 44–62.
<https://doi.org/10.1111/saje.12283>.
- Republic of South Africa (1995), 'Labour relations act, 1995'.
- Republic of South Africa (1996a), 'Constitution of the Republic of South Africa', *Government Gazette* **378**(17678).
- Republic of South Africa (1996b), 'South African Schools Act, 1996', *Government Gazette* **377**(1867).
- Republic of South Africa (1997), 'Basic conditions of employment act, 1997', *Government Gazette* **390**(75).
- Republic of South Africa (1998), 'Employment equity act, 1998', *Government Gazette* **400**(55).
- Republic of South Africa (2004), 'Broad-based black economic empowerment act, 2003', *Government Gazette* **463**(46).
- Republic of South Africa (2014), 'Employment equity amendment act, 2013', *Government Gazette* **583**(37238).
- Republic of South Africa (2024), 'Basic Education Laws Amendment Act, 2024', *Government Gazette* **711**(51258).
- Riddell, W. C. and Song, X. (2011), 'The impact of education on unemployment incidence and re-employment success: Evidence from the US labour market', *Labour Economics* **18**(4), 453–463.
<https://doi.org/10.1016/j.labeco.2011.01.003>.
- Sakaue, K., Wokadala, J. and Ogawa, K. (2024), 'Updates on private returns to education in Uganda: evidence from universal primary education policy', *Education Economics* pp. 1–17.

<https://doi.org/10.1080/09645292.2024.2345708>.

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.

<https://doi.org/10.1016/j.ijedudev.2015.07.004>.

Schlotter, M., Schwerdt, G. and Woessmann, L. (2011), 'Econometric methods for causal evaluation of education policies and practices: a non-technical guide', *Education Economics* **19**(2), 109–137.

<https://doi.org/10.1080/09645292.2010.511821>.

Schultz, T. W. (1960), 'Capital formation by education', *Journal of Political Economy* **68**(6), 571–583.

<https://doi.org/10.1086/258393>.

Schultz, T. W. (1961), 'Investment in human capital', *The American Economic Review* **51**(1), 1–17.

<https://www.jstor.org/stable/1818907>.

Seekings, J. (2016), 'The quality of South African wage data: accounting for deductions, debt and disclosure', *Centre for Social Science Research Working Paper No. 397*.

Skovron, C. and Titunik, R. (2015), 'A practical guide to regression discontinuity designs in political science', *American Journal of Political Science* **2015**, 1–36.

Smith, M. (2023), 'Adam Smith on Growth and Economic Development', *History of Economics Review* **86**(1), 2–15.

<https://doi.org/10.1080/10370196.2023.2243741>.

Solarin, S. A. and Yen, Y. Y. (2016), 'A global analysis of the impact of research output on economic growth', *Scientometrics* **108**(2), 855–874.

<https://doi.org/10.1007/s11192-016-2002-6>.

Solli, I. F. (2017), 'Left behind by birth month', *Education Economics* **25**(4), 323–346.

<https://doi.org/10.1080/09645292.2017.1287881>.

- Somani, R. (2021), 'The returns to higher education and public employment', *World Development* **144**, 105471.
<https://doi.org/10.1016/j.worlddev.2021.105471>.
- Spaull, N. (2013), 'Poverty & privilege: Primary school inequality in South Africa', *International Journal of Educational Development* **33**(5), 436–447.
<https://doi.org/10.1016/j.ijedudev.2012.09.009>.
- Spaull, N. (2015a), Education quality in South Africa and sub-Saharan Africa: An economic approach, PhD thesis, Stellenbosch: Stellenbosch University.
- Spaull, N. (2015b), 'Schooling in South Africa: How low-quality education becomes a poverty trap', *South African Child Gauge* **12**(1), 34–41.
- Spence, M. (1973), 'Job market signaling', *The Quarterly Journal of Economics* **87**(3), 355–374.
<https://doi.org/10.1016/B978-0-12-214850-7.50025-5>.
- Staneva, A. V., Arabsheibani, G. R. and Murphy, P. (2022), Returns to education in four transition countries, in 'Critical Perspectives on Economics of Education', Routledge, pp. 182–212.
- Statistics South Africa (2016), 'General household survey 2015 [dataset]'.
Available from: <https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/715>.
- Statistics South Africa (2017a), 'General household survey 2014 [dataset]'.
Available from: <https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/526>.
- Statistics South Africa (2017b), 'General household survey 2016 [dataset]'.
Available from: <https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/621>.
- Statistics South Africa (2018), 'General household survey 2017 [dataset]'.
Available from: <https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/723>.
- Statistics South Africa (2019), 'General household survey 2018 [dataset]'.
Available from: <https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/801>.

- Statistics South Africa (2023), 'Census 2022 statistical release'.
Available from: https://census.statssa.gov.za/assets/documents/2022/P03014_Census_2022_Statistical_Release.pdf.
- Stefani, P., Biderman, C. et al. (2006), 'Returns to education and wage differentials in Brazil: a quantile approach', *Economics Bulletin* **9**(1), 1–6.
- Stefani, P. C. and Biderman, C. (2009), 'The evolution of the returns to education and wage differentials in Brazil: a quantile approach', *Applied Economics* **41**(11), 1453–1460.
<https://doi.org/10.1080/00036840601019299>.
- Stiglitz, J. E. (1975), 'The theory of "screening," education, and the distribution of income', *The American Economic Review* **65**(3), 283–300.
<https://www.jstor.org/stable/1804834>.
- Tansel, A. and Bodur, F. B. (2012), 'Wage inequality and returns to education in Turkey: A quantile regression analysis', *Review of Development Economics* **16**(1), 107–121.
<https://doi.org/10.1111/j.1467-9361.2011.00655.x>.
- The World Bank (2018), 'Overcoming Poverty and Inequality in South Africa: An Assessment of Drivers, Constraints and Opportunities', Technical report, World Bank.
- Thobejane, T. (2013), 'History of apartheid education and the problems of reconstruction in South Africa', *Sociology Study* **3**(1), 1–12.
- Timæus, I. M., Simelane, S. and Letsoalo, T. (2013), 'Poverty, race, and children's progress at school in South Africa', *The Journal of Development Studies* **49**(2), 270–284.
<https://doi.org/10.1080/00220388.2012.693168>.
- Torun, H. (2018), 'Compulsory schooling and early labor market outcomes in a middle-income country', *Journal of Labor Research* **39**(3), 277–305.
<https://doi.org/10.1007/s12122-018-9264-0>.
- UNESCO (2020), Global education monitoring report 2020: inclusion and education: all means all, Technical report, UNESCO, Paris, France.

- Valdés, M. T. and Requena, M. (2024), 'The effect of the age at school entry on educational attainment and field of study: an analysis using the Spanish census', *Higher Education* **87**(4), 1061–1083.
<https://doi.org/10.1007/s10734-023-01053-5>.
- van der Berg, S., van Wyk, C., Gustafsson, M., Clayton, R., Böhmer, B., Meyer, H., Qvist, J., Mkhombo, T., Euston-Bown, A., Claassen, J. et al. (2025), 'What SA-SAMS and Lurits data tells us about education: New insights from administrative data'.
<http://dx.doi.org/10.2139/ssrn.5181087>.
- van der Berg, S., Wills, G., Selkirk, R., Adams, C. and Van Wyk, C. (2019), The cost of repetition in South Africa, Technical report, Stellenbosch Economic Working Papers: WP13/2019.
- Viljoen, D. J. and Dunga, S. H. (2013), 'Determining the factors that influence female unemployment in a South African township', *International Journal of Social Sciences and Humanity Studies* **5**(1), 63–72.
- Vu, L. H., Tran, T. Q. and Phung, T. D. (2021), 'Children and female labor market outcomes in Vietnam', *Heliyon* **7**(7).
<https://doi.org/10.1016/j.heliyon.2021.e07508>.
- Wang, L. (2013), 'How does education affect the earnings distribution in urban China?', *Oxford Bulletin of Economics and Statistics* **75**(3), 435–454.
<https://doi.org/10.1111/j.1468-0084.2012.00697.x>.
- Wilcox, V. and Sahni, H. (2023), 'The effects on labor supply of living with older family members needing assistance with activities of daily living', *Journal of Family and Economic Issues* **44**(4), 900–918.
<https://doi.org/10.1007/s10834-022-09880-x>.
- Wittenberg, M. (2017a), 'Measurement of earnings: Comparing South African tax and survey data', Technical report, Southern Africa Labour and Development Research Unit, University of Cape Town.

BIBLIOGRAPHY

Wittenberg, M. (2017b), 'Wages and wage inequality in South Africa 1994–2011: Part 2–inequality measurement and trends', *South African Journal of Economics* **85**(2), 298–318.

<https://doi.org/10.1111/saje.12147>.

Zwane, T. T. (2020), 'The causal effect of education on earnings in urban and rural South Africa: A further update', *Demography and social economy* **39**(1), 79–94.