High School Dropout and Youth Labour Market outcomes in Sub-Saharan Africa: Evidence from South Africa

Godstime Osekhebhen Eigbiremolen

A thesis submitted for the degree of
Doctor of Philosophy

School of Economics
University of Kent
United Kingdom
September 2019

Dedication

To God and my family

Declaration

I declare that other than where a precise reference is made to the work of others, the contents of this thesis are original and have not been submitted in for consideration for any other degree or qualification in this university, or at any other university. This thesis is mine and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and acknowledgements.

Godstime Osekhebhen Eigbiremolen
September 2019

Acknowledgements

I would like to specially thank my supervisor, Andrey Launov for very useful feedback, comments, encouragement and expert guidance all through the course of this thesis. I could not have done this without him. I would also like to thank other faculty members at the School of Economics for helpful comments and suggestions at different stage of my thesis. Finally, I would like to thank the professional staff at the School of Economics for their help and support

Abstract summary

Chapter 2: Youth unemployment and high school dropout

The second chapter of this thesis provides a general description of the nature of high school dropout and youth unemployment in South Africa. This chapter is set as a motivational chapter for subsequent chapters. I defined a high school dropout in the data as an individual who was enrolled in high school at the beginning of the survey but failed to complete the last year of high school education (i.e., grade 12 in South Africa) before terminating their enrolment. A preliminary description of the data shows a high dropout rate of about 50% on the average. Upon estimation, I find that individuals who dropped out of high school have higher continuous spells of unemployment. In addition, I identify the following individuals as those more that are more vulnerable to high school dropout: black students, young students, students with record of poor academic performance, students who live close to school, and students with less educated fathers. On the other hand, factors that are associated with successful transition from school to employment include good academic performance, being white or coloured, age, and having an educated father.

Chapter 3: Assessing the Patterns of Self-selection into High School Drop-out and Graduation

The third chapter of my thesis focuses primarily on assessing the patterns of self-selection into high school dropout and high school graduation. Specifically, I adopt a flexible framework that allows me to quantify selection on unobservables between dropouts and graduates and establish to what extent this selection matters. To do this, I write down a flexible discrete-time conditional competing risks model that allows for a general correlated risks across destinations. I treat individ-

ual differences that may exist between dropouts and graduates as both observed and unobserved. I then use the structure of my model to estimate a reduced-form measure of self-selection into dropout and graduation. As a reduced-form measure of self-selection, I propose to consider the correlation coefficient of the bivariate distribution of unobserved heterogeneity across two possible exit states: high school dropout and high school graduation. I make no parametric assumption about the distribution of unobserved individual heterogeneity. Rather, I used a non-parametric approach, which relies on the data in estimating the distribution of unobserved differences. Results suggest that individuals who choose to drop out of high school are not systematically different from those who complete high school education in terms of unobservables. This runs contrary to typical findings in developed countries, which suggest that high school dropouts are individuals with relatively low ability, low expectations, and a set of negative preferences.

Chapter 4: High School Dropout and its Wage Consequences

The fourth chapter of my thesis focuses on the wage consequences of dropping out of high school. Specifically, I am interested in the severity and time pattern of the wage disadvantage dropouts face in the labour market, if any. This knowledge is important because lack of information or incomplete information about the worth of a high school diploma in the labour market could influence the decision to drop out of high school. Prior to estimating the wage equation, I check for self-selection both in the decision to drop out of high school and in the decision to participate in the labour market. Results from IV estimations show that OLS estimates are upward biased, even after controlling for measures of ability. Also, I find that dropouts earn less in the first year of labour market experience, and their wages progressively declined in subsequent years. This suggests that high school diploma may act as a signal of productivity, both immediately after graduation and in subsequent years.

Contents

List of Figures

List of Tables

Chapter 1

General Introduction

1.1 Overview

The youthful stage of life is an important phase of growth and development for every individual. Young people make decisions at this stage that could either ensure a meaningful and rewarding future or could potentially put them in a disadvantaged position for a life time. One of such decisions is the decision to either quit high school prior to graduation or complete high school education. Individuals who choose to dropout do not just lose the opportunity to get education today, they may also deprive themselves of further training in the future (Hill, 1979)

Thus, dropping out of high school comes at a cost, not just for the individual, but also for the society as whole. There is evidence in the literature, particularly in the context of developed countries, that individuals who dropped out of high school are more likely to experience a longer duration of unemployment, earn less in the labour market, less likely to contribute to national development and are more likely to be involved in social vices (Hill, 1979; Cawley et al., 2001; Lochner & Moretti, 2004). Thus, the benefits of addressing the problem of high school dropout are enormous. This is even more so in Sub-Saharan Africa, a region with a relatively high young population, high youth unemployment, socially and economically challenged and with cases of social vices. However, most of the existing evidence on high school dropout are in the context of developed countries, particularly the United States. Therefore, this thesis is set in the context of Sub-Saharan

Africa, using South Africa as a case study in order to bridge the gap in knowledge. The choice of South Africa is primarily because of data availability. South Africa has a unique longitudinal data (the Cape Area Panel Survey) that followed the educational and labour market outcomes of young people over a period of time. Most countries in the region do not have this kind of data.

In order to address the problem of high school dropout in a meaningful and sustainable way, it is important to first understand the nature of the problem and its possible implications in the context of the study. This is because it is difficult, if not impossible, to solve a problem without understanding the nature and underlying factors that drive such a problem. Therefore, the broad objective of this thesis is to examine the nature of high school dropout in South Africa and it's wage consequences; both immediately after dropping out of high school and in subsequent years.

1.1.1 Introduction to chapter 2

The second chapter of this thesis provides a general description of the nature of high school dropout and youth unemployment in South Africa. This chapter is set as a motivational chapter for subsequent chapters. I defined a high school dropout in the data as an individual who was enrolled in high school at the beginning of the survey but failed to complete the last year of high school education (i.e., grade 12 in South Africa) before terminating their enrolment. A preliminary description of the data shows a high dropout rate of about 50% on the average.

I then used transition probabilities to assess the duration of unemployment between high school dropouts and high school graduates. The purpose of this is to see whether high school dropouts spend longer time in the unemployment pool than individuals with high school qualifications. Furthermore, I used a set of logistic regressions to estimate observable individual and family characteristics that affect the likelihood of dropping out of high school prior to graduation and observable factors that facilitate a successful transition from school to work.

1.1.2 Introduction to chapter 3

To properly understand the nature of high school dropout in my data, it is important that I look beyond observable characteristics in explaining why young people quit high school prior to graduation. Thus, though useful for initial understanding, estimating just the correlates of high school dropout may not be sufficient. Put differently, if the decision to dropout from high school is not random, but based on an optimizing behaviour, observable characteristics alone become insufficient in explaining the dropout decision. Individuals may self-select into high school dropout or high school graduation based on unobserved factors. This knowledge is important to guide meaningful policies aimed at addressing the problem of high school dropout.

Therefore, the third chapter of my thesis focuses primarily on assessing the patterns of self-selection into high school dropout and high school graduation. Specifically, I adopt a flexible framework that allows me to quantify selection on unobservables between dropouts and graduates and establish to what extent this selection matters. To do this, I write down a flexible discrete-time conditional competing risks model that allows for a general correlated risks across destinations. I treat individual differences that may exist between dropouts and graduates as both observed and unobserved. I then use the structure of my model to estimate a reduced-form measure of self-selection into dropout and graduation. As a reduced-form measure of self-selection, I propose to consider the correlation coefficient of the bivariate distribution of unobserved heterogeneity across two possible exit states: high school dropout and high school graduation. I make no parametric assumption about the distribution of unobserved individual heterogeneity. Rather, I used a non-parametric approach, which relies on the data in estimating the distribution of unobserved differences.

1.1.3 Introduction to chapter 4

The fourth chapter of my thesis focuses on the wage consequences of dropping out of high school. Specifically, I am interested in the severity and time pattern of the wage disadvantage dropouts face in the labour market, if any. This knowledge is important because lack of information or incomplete information about the worth of a high school diploma in the labour market could influence the decision to drop out of high school. Prior to estimating the wage equation, I check for self-selection both in the decision to drop out of high school and in the decision to participate in the labour market. The method outlined in chapter 3 provides a way for me to check for self-selection in the dropout decision. To check for selection in labour market participation, I estimate a battery of Heckam selection model.

In order to identify the average return to high school diploma, I proceed in the following ways. First, I directly incorporate measures of ability into the wage equation. The measures used are IQ tests that were administered at the beginning of the survey. Secondly, I use parental education and proximity to school as instruments in IV estimations.

References

Cawley, J., Heckman, J., & Vytlacil, E. (2001). Three observations on wages and measured cognitive ability. *Labour Economics*, *8*(4), 419-442.

Hill, C. R. (1979). Capacities, opportunities and educational investments: The case of the high school dropout. *The Review of Economics and Statistics*, 9–20.

Lochner, L., & Moretti, E. (2004). The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports. *American Economic Review*, *94*(1), 155-189.

Chapter 2

Youth unemployment and high school dropout

2.1 Introduction

Youths are agents of change and development anywhere in the world when equipped with the right education, skills and opportunities. Irrespective of where they live or their socio-economic conditions, they have one thing in common; the desire to have a meaningful and rewarding future. Thus, policy makers, both in and outside Africa, are increasingly concerned about how best to address their needs (Filmer & Fox, 2014). According to the United Nation's Population Report, there will be about 1.4 billion young people (aged 15-24) in the world by 2020. For Sub-Saharan Africa, a substantial surge in the youth population has been observed over the years. In 2015, the number of youth stood at 226 million. This represents about 19% of the world youth population and their share in African's total population remains on the rise. In most parts of the world, national population seems to be ageing. But in Africa, especially Sub-Saharan Africa, youths are in the majority. This makes Africa one of the youngest continents in the world. Half of the Sub-Saharan Africa population is under 25 years of age and it is projected that between 2015 and 2035, there will be half a million more 15-year olds than the year before (Filmer & Fox, 2014).

The surge in youth population is not necessarily bad. In fact, the current demography could mean that youth is the most abundant asset possessed by Sub-Saharan Africa, which could be harnessed for the region's development. A large and educated youthful labour force with productive employment will have a significant and positive impact on the region's investment climate and growth prospects (Goldberg et al., 2009; Fuller, 2004). This is because young people, compared to the elderly, tend to work more and save less. That is, they are more likely to offer labour and capital to the economy, require less health care and rely less on social pensions than the elderly in the society (Bloom et al., 2010). However, as noted by Fuller (2004), if a society lacks the necessary social infrastructure to integrate, employ and care for a growing youth population, the potential demographic benefits could become a serious drain in the resources of the state and form a dangerously unstable element within society. In other words, youth education and gainful employment are essential if we must harness the opportunities for economic growth associated with perceived demographic dividend. While an educated and healthy work force bolster investment; unskilled and disillusioned youth make returns to investment low and uncertain. Thus, national governments and regional bodies have placed this issue in their policy agenda (Goldberg et al., 2009).

The young phase of life is essential. This is when young people attempt to exercise their financial independence and lay a basis for their future. Between the ages of 14 and 24, youths undergo several transitions. During this period, some young people transit from high school to colleges or universities, while others drop out of high school or complete high school education and join the labour market. The decisions young people make during this period have enormous consequences for their future prospects as well as those of the societies they live in (Goldberg et al., 2009). Across the world, successful transition from school to work is becoming increasing difficult for young people since unemployment among young people tends to be higher than in the remainder of the labour force. In addition, if youths are unable to develop the right skills needed to ensure gainful and productive employment, there could be a profound negative effects on a country's investment opportunities and growth prospects (Doiron & Gørgens, 2008; Goldberg et al., 2009).

Table E.1 shows that for Sub-Saharan Africa, the number of young people who were unemployed in 2015, 2016 and 2017 stood at 11.1 million, 11.3 million, and 11.6 million respectively. In terms of percentage, the region has a youth unemployment rates of 10.9%, 11.0%, and 11.1% in 2015, 2016, and 2017 respectively. The region has the highest labour force participation rate in the world for youth between the ages of 15-24 based on the 2016 estimates. The relatively high labour force participation rate in Sub-Saharan Africa may be due to the high number of adolescents (aged between 15 and 19) present in the labour force. This could be interpreted as an indication of poor economic and societal development in the region necessitating large shares of adolescents to work in order to earn a living, instead of continuing in education. These young people are often called upon to supplement household income, which in turn forces them to enter the labour force quite early. In developed regions with good economic and social development, large shares of adolescents do not necessarily need to work to earn a living. They have more opportunities to further their education and enhance their employment prospects in the future. There is however a possibility of overlap between education and labour force participation in some regions with high income.

Table 2.1: Youth unemployment trends and labour force participation rate

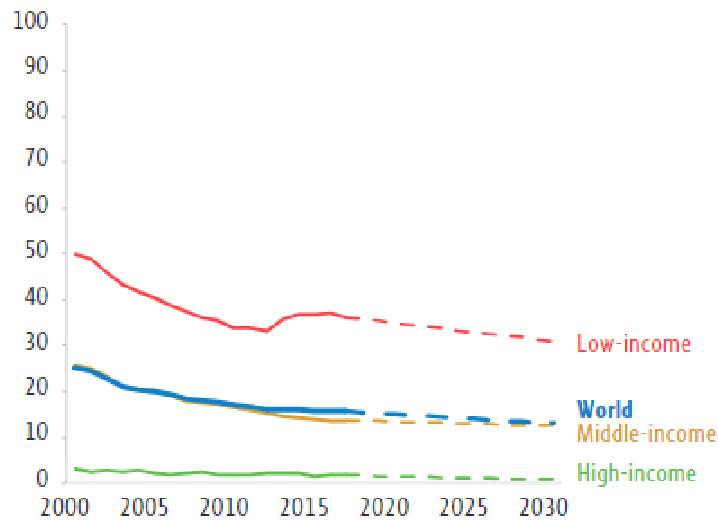
	Unemployment						Labour force part (2016)		
	Level (millions)			Percentages			Percentages by age		
Region	2015	2016	2017	2015	2016	2017	15-24	15-19	20-24
Sub-Saharan Africa	11.1	11.3	11.6	10.9	11.0	11.1	54.2	45.2	64.8
Northern Africa	3.7	3.7	3.7	29.4	29.3	29.2	31.9	18.9	44.9
Northern America	3.0	2.9	2.9	11.8	11.5	11.7	52.7	31.4	71.5
Latin America and the Caribbean	8.5	9.2	9.3	15.7	16.8	17.1	49.6	33.3	66.3
Arab States	2.6	2.7	2.6	30.6	30.6	29.7	30.4	17.3	44.1
Eastern Asia	11.9	11.4	11.0	10.6	10.7	10.9	52.5	29.4	70.6
South-Eastern Asia and the Pacific	7.4	7.7	8.0	12.4	13.0	13.6	51.3	32.6	70.1
Southern Asia	13.7	13.8	13.9	10.9	10.9	10.9	37.2	24.4	50.6
Central and Western Asia	2.1	2.1	2.2	16.6	17.1	17.5	37.2	24.4	50.6
Eastern Europe	2.0	1.8	1.7	17.1	16.6	16.2	36.3	9.3	57.6
Northern, Southern and Western Europe	4.5	4.3	4.1	20.6	19.7	18.9	44.4	24.3	63.1

Source: ILO calculations based on ILO Research Department?s Trends Econometric Models, April 2016.

The high level of youth unemployment in Sub-Saharan Africa is not showing any sign of decline. According to the Economic and Social Affairs Report of the United Nations 2015, the number of young people seeking employment is projected to increase by 60% between 2015 and 2030. The overall growth in the youth population tends to reinforce youth unemployment in the region. A large cohort of young people willing and able to work, all other things being equal, will reduce the employment prospects for the entire group. Empirical evidence from a large pool of developing countries indeed suggests a positive relationship between youth employment and youth cohort size (O'Higgins, 2003; Goldberg et al., 2009).

A related issue to the above discussion is high school dropout among young people. In the literature, a high school dropout refers to an individual who did not complete high school education for any reason after enrolment (Brooks-Gunn et al., 1993; Rumberger, 2001; Campbell, 2015). Young people dropping out of high school is increasingly becoming a source of concern, especially in developing countries. This is because dropping out of school could have far-reaching consequences in the future in terms of individual labour market outcomes and for the society as whole. For example, high school dropouts may stay longer in the unemployment pool and they may not get jobs that pay enough to keep them off public assistance. Higher rates of unemployment and lower earnings could mean lost productivity and reduced tax income. In addition, there is evidence in the literature that suggests that high school dropouts are more likely to have health problems, engage in criminal activities, and become dependent on welfare and other government programs than high school graduates (Rumberger, 1987).

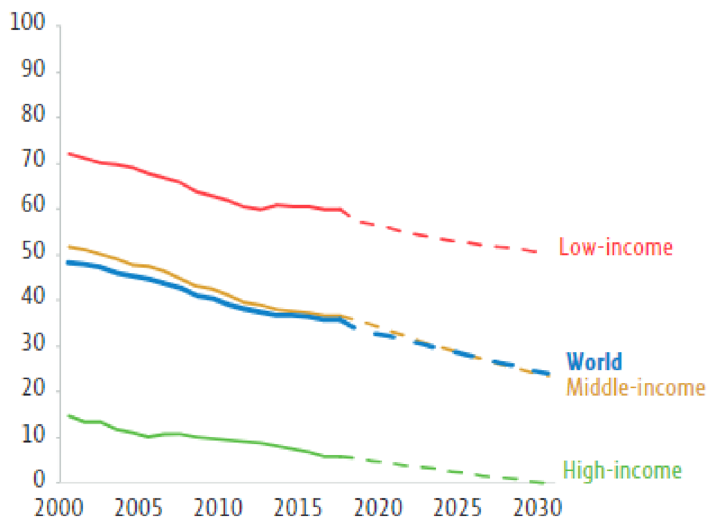
Figure 2.1: Out-of-school rates (%) for junior high school age (12-14)



Source: UNESCO (2019)

Data that specifically captures dropouts among young people across the world are not readily available. However, data from the United Nations Educational, Scientific and Cultural Organization (UNESCO) provide some evidence of young people that are out of of high school across the world. Although out-of-school does not necessarily represent school dropout, the available data helps to make some sense of the issue of access to education. A recent report by UNESCO (2019) reveal that, globally, 263 million or 18% of all children, adolescents and youth aged 6 to 17 years were out of school in 2017. More specifically, 61 million or 16% of adolescents of junior high school age (12-14 years) are out of school. Although the out-of-school rate for this group fell from 25% to 17% between 2000 and 2010, it has since remained relatively stagnant afterwards (see Figure 2.1). Furthermore, 138 million or 36% of senior high school age (15-17) are out of school. A fall in the rates from 48% to 37% (i.e., 11 percentage points gain) was observed between 2000 and 2013. However, the gain has slowed down considerably afterwards. While the out-of-school rate is as high as 60% in low income countries, it is only about 6% in high-income countries (see Figure 2.2). Table 2.2 provides out-of-school rates for upper secondary school age across different regions of the world. The out-of-school rates are far higher in Sub-Saharan Africa than in other regions.

Figure 2.2: Out-of-school rates (%) for senior high school age (15-17)



Source: UNESCO (2019)

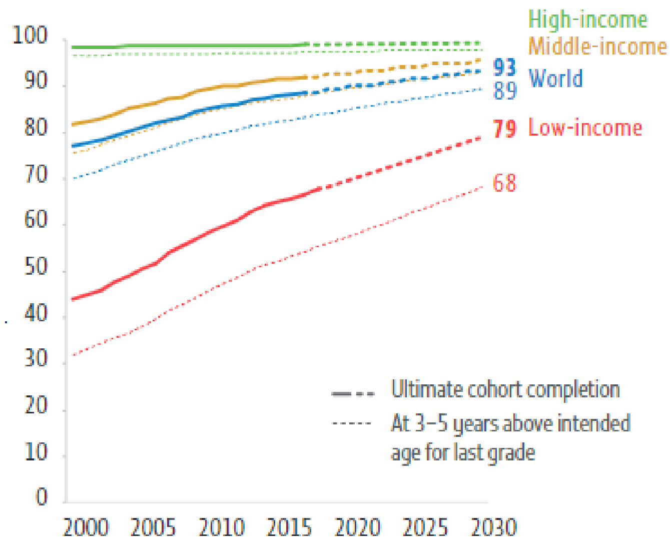
Table 2.2: Regional out-of-school rates for senior high school (2017)

	Out-of-school rate (%)		
Region	Both sexes	Male	Female
Sub-Saharan Africa	56.9	53.4	60.5
Eastern Asia	15.9	20.7	10.5
Latin America and the Caribbean	22.8	24.1	21.5
Northern Africa	33.3	32.3	34.4
South-Eastern Asia	25.7	27.6	23.8
Southern Asia	47.8	46.8	49.0
Central and Eastern Europe	12.1	12.2	12.0
North America and Western Europe	5.7	5.9	5.5
Arab States	38.3	36.1	40.6
World	35.8	35.6	36.1

Source: UNESCO Data for the Sustainable Development Goals (SDG), 2019

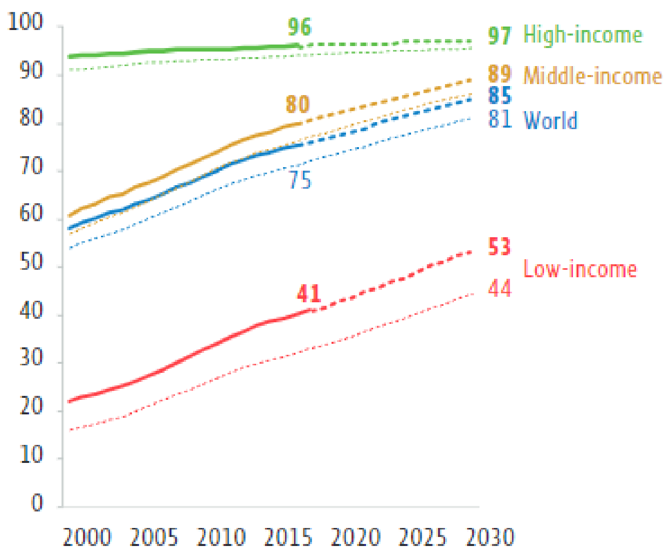
High school completion rate is another concept of interest. The conventional indicator for the completion rate measures school completion among students three to five years above the nominal age for the final grade. It is, therefore, a measure of reasonably timely completion of secondary school education (UNESCO, 2019). Figures 2.3 and 2.4 present completion rates for junior and senior high school

Figure 2.3: Completion rate (%) for junior high school age (12-14)



Source: UNESCO (2019)

Figure 2.4: Completion rate (%) for senior high school age (15-17)



Source: UNESCO (2019)

respectively. The figures show that for low-income countries, secondary school completion rate is about 72% for junior high school, but as low as 48% for senior high school. This indicates that dropping out mostly occur in the senior high school phase of education (i.e., between grades 10 and 12). There are different possible reasons for this observed outcome. One of such reasons could be that the introduction of fees from grade 10 onward in many countries with low-income could discourage some individuals, especially those from disadvantaged background, from continuing in education. Another possible explanation could be the increasing rigour of the curriculum, since this is the final phase of high school education

From the foregoing discussion, it is important to identify the correlates of high school drop out. In other words, why do young people quit high school education before completion? Also important is the issue of successful transition from school to work. What factors determine the successful transition from school to work? A lot of young people in Sub-Saharan African want to work after high school (some even drop out of high school to go into the labour market) in order to support themselves and their family and to possibly save for further education. In addition, it is also relevant to examine unemployment duration between high school dropouts and high school graduates. That is, do high school dropouts stay longer in the unemployment pool than high school graduates? These are the main questions that I seek to answer in this motivational chapter. The rest of the chapter is outlined as follows. Section 2.2 discusses the background of the study, section 2.3 reviews empirical literature, the data employed in the analysis is presented in section 2.4, section 2.5 discusses the transition and outcomes of high school dropouts and graduates, method of estimation and empirical results are presented in section 2.6, and section 2.7 concludes.

2.2 Background

Although South Africa has some elements of development, the country is still classified as a developing country like all other countries in Sub-Saharan Africa. In 1994, South Africa saw the end of Apartheid which started in 1948. The political freedom that was ushered in was viewed by many as an opportunity for economic growth and development that will be all-inclusive. However, the last 25 years have seen mixed results. Economic growth has been volatile. While inequalities in public services may have reduced, income inequality has increased, and poverty levels are not going down either (Bhorat et al., 2014). The presence of socio-economic inequalities has continued to threaten the long-term stability of the country. The bequests of apartheid, poor service delivery and widespread poverty have characterised current socio-political discourse with protests against service delivery widespread in various parts of the country. Unemployment remains high in South Africa, especially among the youths. The latest data from the International Labour Organization (ILO) shows that South Africa currently has the highest rate of youth unemployment in the word (percentage of total labor force ages 15-24). The rate is as high as 52.9% based on the 2018 data. High unemployment, especially among black South Africans, has been reported as one of the causes of widening income inequality (Kumo et al., 2016).

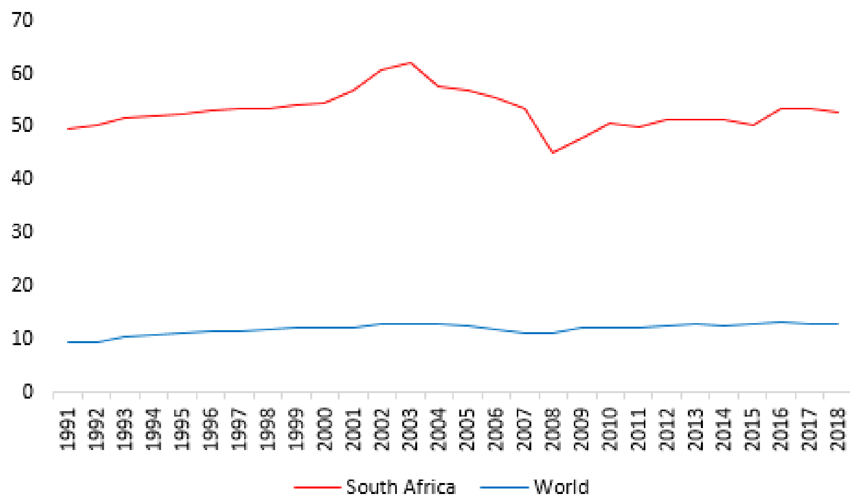
Table 2.3: Dropout rates among young people in South Africa

Province	Grade 10 (2014)	Grade 12 (2016)	Dropout rates
Northern Cape	22, 034	10, 041	54.4%
North West	67, 734	32, 045	52.7%
Free State	55, 293	26, 786	51.6%
Eastern Cape	154, 220	82, 902	46.2%
Limpopo	189, 170	101, 807	46.2%
KwaZulu Natal	264, 816	147, 648	44.2%
Mpumalanga	94, 528	54, 251	42.6%
Gauteng	174, 471	103, 829	40.5%
Western Cape	75, 791	50, 869	32.9%
South Africa	1, 100, 877	610, 178	44.6%

Source: National Department of Basic Education, South Africa

The available comprehensive data from the South Africa Department of Basic Education reveals a significant level of dropout among high school students. As shown in Table 2.3, drop out rate is as high as 54.4% in the Northern Cape province and stands at 45% for the country as a whole. High drop out rates are also observed across other provinces. Many of the the youths who dropped out of school are likely to go straight into the labour market with the hope of finding a job. However, due to low human capital accumulation, they may likely remain unemployed for a long time. This could further drive up the level of youth unemployment. Figure 2.5 shows that between 1991 and 2018, youth unemployment rates in South Africa stood way above the world average and do not show any significant sign of decline going forward. Given that youth unemployment is exacerbated by high drop out rate, addressing the latter may help to address, to some extent, the former.

Figure 2.5: Trends in Youth Unemployment (percentages)



Source: World Bank (2019)

The management of education in South Africa is shared between the Department of Basic Education (DBE) and the Department of Higher Education and Training (DHET). The DBE is responsible for all schools from grade R (i.e., grade 0 or the reception/preparatory year) to grade 12 (i.e., the final year of high school education, which is also known as the matriculation year), and adult literacy programmes. The DHET on the other hand deals with universities, and other post-school education and training, as well as coordinating the Human Resource Development Strategy for South Africa. The Basic Education schooling system is divided into a General Education and Training (GET) phase, which bridges primary schooling (grades 1 to 7) and junior secondary education or junior high school (grades 8 and 9), and a Further Education and Training (FET) phase, which represents senior secondary education or senior high school - grades 10 to

12 (Motala et al., 2009).

Thus, grades 1 to 7 represent Primary Education, while grades 8 to 12 capture High School Education (i.e., both junior and senior). The South African Schools Act (SASA) of 1996a provides legal backing for access to school and the National Education Policy Act of 1996 regulates admission into public and independent (private) schools. The School Act makes enrolment compulsory for all learners from the beginning of the year in which they turn six (i.e., turning six by 30 June in the year of admission) to the end of the year in which they turn 15, or the end of Grade 9, whichever occurs first. Parents are legally responsible for ensuring that those under their care are enrolled during the mandatory period. Compulsory education covers seven years of primary school and the first two years of high school education. (Motala, 2011).

With effect from 2007, some secondary schools became tuition-free or no-fee schools. This followed the amendment of the National Norms and Standard of School Funding in 2006. Under the new policy, schools classified as no-fee schools qualify for an increased funding by the State to balance out income initially obtained from school fees, and fee charging schools (previously all schools) may apply to their provincial Education Department to be declared a no-fee school as long as they meet the criteria (Sayed & Motala, 2012). Schools are classified as fee-free if they fall into the lowest two poverty quintiles. Poverty quintiles are determined nationally (this was previously determined at the provincial level), and the National Education Department establishes the amount that provinces are required to allocate per learner in each quintile.

The National Department also determines an adequacy benchmark (i.e., a minimum amount considered necessary for schools to provide adequate basic education). This amount was R527 in 2006 for non-salary expenditure, with the poorest quintiles receiving R703 per learner and the least poor received R117 per learner. Following this procedure, schools which receive adequate funding will be listed as no-fee schools. However, other costs of schooling like the costs of transport, school uniforms, books and stationery are not covered by this policy. Also, this selective fee exemption applies only from grade R to grade 9, which represents the compulsory education period (Motala, 2011). This means that students in grades 10 to 12 still have to pay tuition fee alongside other costs of education. This may partly explain why high drop out rates are observed mostly between grades 10 and 12 in the data (see Table 2.3).

2.3 Literature Review

Over the years, a number of studies have examined factors that explain high school dropout and successful transition from school to work. Before looking at both high school dropout and transition from school to work, I first examine the broader literature of human capital accumulation and earnings in the labour market using selected studies.

2.3.1 Human capital and labour market outcomes

In terms of data, many studies that examined human capital accumulation and labour market earnings employed panel (longitudinal) data in their analysis. Longitudinal data allows researchers to follow units of observation over a period of time and thus could aid identification. A few exceptions however exist. For example, Ge (2013) used simulated data to estimate the returns to schooling, comparing the estimates from a standard dynamic discrete choice model to those from ordinary least square (OLS) and instrumental variables (IV) estimates. In addition, few other papers used cross-sectional data to examine labour market outcome for young people (Angrist & Krueger, 1991; Card & Krueger, 1992a; Dorsett, 2006). However, Angrist & Krueger (1991) augmented their cross-sectional data with a natural experiment.

The literature has consistently identified the effect of racial differences in labour market outcomes; individuals who are black tend to earn less than their white counterparts. In addition, blacks are more likely to have fewer offers of employment; they have lower probability of transiting from school to work, and are more likely to lose their Jobs (Wolpin, 1992; Card & Krueger, 1992b; Eckstein & Wolpin, 1995, 1999; Keane & Wolpin, 2000; Bowlus et al., 2001). The difference in earnings between blacks and whites may not necessarily imply racial discrimination in the labour market as there could be other underlying factors driving the observed difference.

Keane & Wolpin (2000) carried out an interesting study about eliminating racial differences in school attainment and labour market success. They find that eliminating just wage discrimination would increase the expected present value (at age 16) of lifetime earnings for a typical black male from 73.7% to 83.8% of the average white wage. Also, it reduced the fraction of black males who drop out of high school by 14% (about 44% of the black-white gap) and increased the fraction who graduated from college by 20% (about 20% of the black-white gap).

The empirical literature also contains other results of interest. For example, years of schooling, higher school quality, skills enhancement programmes and better educated teachers are positively associated with higher earnings in the labour market (Card & Krueger, 1992a; Murphy & Peltzman, 2004). Furthermore, academic performance is positively associated with higher earnings (Andrews et al., 2002; Murphy & Peltzman, 2004). In addition, absence of tuition fees suggests higher private returns (Holzner & Launov, 2010). Interestingly, students are willing to stay in school with an increase in schooling cost provided the increase in cost is compensated for by the market in terms of high wages (Bilkic et al., 2012).

2.3.2 Transition from school to work

Many studies that examined school-to-work transition determinants employed the Cox proportional hazard model in their empirical evaluation. (Genda & Kurosawa, 2001; Corrales & Rodríguez, 2004; Nguyen et al., 2005; Jaunky & Khadaroo, 2008; Vanoverberghe et al., 2008; Sciulli & Signorelli, 2011). Examples of papers that used different methods include Biggeri et al. (2001) and Pugatch (2012) who employed three-step discrete time survival estimation method and a model of choice of enrollment respectively. Duration effect in youth unemployment is another area of interest apart from estimating the determinants of the successful transition from school to work. That is, researchers are interested in how long unemployment individuals stays in the unemployment pool (Biggeri et al., 2001; Böheim & Taylor, 2002; Doiron & Gørgens, 2008; Cockx & Picchio, 2012, 2013).

Many studies have identified educational qualification as an important determinant of successful transition fromschool to work (Biggeri et al., 2001; Salas-Velasco, 2007; Jaunky & Khadaroo, 2008; Sciulli & Signorelli, 2011). In addition, There seems to be some consensus in the literature on how the duration of unemployment affects subsequent employment opportunities. Many studies find that the probability of employment decreases with the duration of unemployment (Biggeri et al., 2001; Doiron & Gørgens, 2008; Cockx & Picchio, 2013). That is, the longer an individual stays in the unemployment pool, the harder it is to get a job. Furthermore, high reservation wage hampers the prospects of a successful transition from school to work Wolpin (1987). Individuals with relatively high reservation wage are, on the average, more likely to receive fewer job offers and thus spend more time in the unemployment pool. However as time passes, reservation wage tends to decline with the duration of employment (Kiefer & Neumann, 1979). This could be because an average individual knows that the longer they stay unemployed, the more difficult it will be to get a job. For example, Kiefer & Neumann (1979) find that the amount of decline in reservation wage as the duration of unemployment increases was 2.5% using data from the United States. However, Flinn (2006) find an opposite result, which is welfare maximizing: increasing minimum wage increases an agent's employment probability. This makes sense if we see high reservation wage as a possible signal of high productivity.

The literature contains other conflicting evidence about transition from school to work. There are other conflicting evidence in the literature that are worth noting. For example, while Salas-Velasco (2007) and Sciulli & Signorelli (2011) find area of residence to be an important determinant for a successful transition from school to work, Jaunky & Khadaroo (2008) find area of residence to be unimportant in explaining the transition process. There is also conflicting evidence with respect to gender effect. While Biggeri et al. (2001); Lassibille et al. (2001)find that females are less likely to transit from school to work, others report that gender has no effect on the transition process (Genda & Kurosawa, 2001; Jaunky & Khadaroo, 2008; Sciulli & Signorelli, 2011). The effect of educational achievement transition process does not also enjoy much consensus. While Biggeri et al. (2001) and Vanoverberghe et al. (2008) find that education attainment is positively associated with successful school to work transition, Sciulli & Signorelli (2011) find an opposite result: a negative association between high education achievement and transition from school work are. This could be because people with high education achievement may likely demand high reservation wage. Some existing studies find that family characteristics like parental education has no important effect on a successful transition from school to work (Lassibille et al., 2001; Nguyen et al., 2005).However, others find that family characteristics significantly determine the transition process (Salas-Velasco, 2007; Jaunky & Khadaroo, 2008).

2.3.3 School dropout

Most of the studies that used longitudinal data to analyse high school dropout were carried out in developed countries, especially the United States (Vallerand et al., 1997; Alexander et al., 1997; Rees & Mocan, 1997; Battin-Pearson et al., 2000; Farmer et al., 2003; Christle et al., 2007; South et al., 2007; Janosz et al., 2008; Neild et al., 2008; Plank et al., 2008; Archambault et al., 2009; Branson et al., 2014). The few studies on high school dropout in Sub-Saharan Africa are mostly based on cross-sectional data. Using cross-sectional data for this purpose is not very effective since it does not allow researchers to follow individuals over time and when and why they dropped out of school.

The methods of analysis adopted vary largely across studies: logistic regression and its variants (Alexander et al., 1997; Rees & Mocan, 1997; Sibanda, 2004; South et al., 2007; Neild et al., 2008), growth mixture model (Janosz et al., 2008; Archambault et al., 2009), hazard model (Plank et al., 2008), nested latent variable model (Battin-Pearson et al., 2000), ordinary least square (Branson et al., 2014), propensity score matching (Lee & Staff, 2007), and descriptive analysis (Vallerand et al., 1997; Christle et al., 2007; Motala et al., 2009; Ampiah & Adu-Yeboah, 2009; Motala, 2011; Ananga, 2011; Dunne & Ananga, 2013).

Different factors have been associated with high school dropout in the literature. Some of these include poor academic performance (Alexander et al., 1997; Battin-Pearson et al., 2000; South et al., 2007; Christle et al., 2007; Neild et al., 2008; Plank et al., 2008; Ampiah & Adu-Yeboah, 2009); low socio-economic status of parents (Battin-Pearson et al., 2000; Sibanda, 2004; Christle et al., 2007; Branson et al., 2014); lack of motivation (Vallerand et al., 1997; Wenger, 2002); mobility - recent change of both residences and schools (South et al., 2007); unstable pathways of school engagement (Janosz et al., 2008; Archambault et al., 2009); association with aggressive peers (Farmer et al., 2003; Neild et al., 2008); working intensively while in school (Lee & Staff, 2007); falling behind - older students (Neild et al., 2008; Branson et al., 2014); absenteeism and undesirable student behaviour (Christle et al., 2007; Ananga, 2011); death of parents (Ampiah & Adu-Yeboah, 2009); and large household size (Sibanda, 2004).

Some conflicting results exist in the literature. For example, while Sibanda (2004) and Christle et al. (2007) find that being black relative to being white increases school dropout rate, Neild et al. (2008) find that being black decreases the probability of high school dropout. Similarly, the literature contains conflicting evidence in terms of gender of effect. While Neild et al. (2008) find that being male increases the probability of dropping out of high school, Sibanda (2004) find that female students are more likely to quit high school education before completion.

2.3.4 Summary of literature review

The studies reviewed provide some insights into the nature of human capital accumulation and labour market outcomes for young people, especially as it relates to high school dropout and transition from school to work. However, there are still contradictory findings in the existing evidence, which makes it to be far from conclusive. More important, the review shows that existing results are mostly in developed countries, especially the United States and in Europe more recently. This is perhaps because the kind of data (i.e., longitudinal data) required for a meaningful analysis of high school dropout and school to work transition exist mostly in developed countries.

2.4 Data

The Cape Area Panel Study (CAPS) is the data used in this study. The CAPS is a five-wave longitudinal study of the lives of young people in Cape Town, South Africa. It is one of the very few data-set in the Sub-Saharan African region designed specifically to follow the lives of youths and young adults over a considerable time. The study's first wave collected interviews in August-December 2002 from about 4,800 randomly selected youth aged 14-22. Wave 1 also collected information on all members of the households of these young people, as well as a random sample of households not having members between 14 and 22 years of age. A third of the youth sample was re-interviewed in 2003 (wave 2a) and re-visited the remaining two-thirds in 2004 (wave 2b). In both 2005 (wave 3) and 2006 (wave 4), the full youth sample was then re-interviewed. For wave 5, full face-to-face interviews were carried out in 2009 with the sample comprising all respondents interviewed in any of Waves 2a, 3 or 4. A wide range of outcomes were covered in the study. These include: schooling, employment, health, family formation, individual life aspirations/expectations, intergenerational support systems, and social and political attitudes and behaviour.

A collaborative project between the Center for Population Studies at the University of Michigan Institute of Social Research and the Center for Social Science Research at Cape Town University in 2002 gave birth to the CAPS. UCT's Southern African Labour and Development Research Unit, and the Research Program in Development Studies at Princeton University joined in subsequent waves of the survey. The National Institute of Child Health and Human Development of the United State's National Institutes of Health (NIH) is the primary funder of the project. Additional funding were made available by the Office of AIDS Research, the Fogarty International Center, the National Institute of Aging of NIH, the Health Economics and HIV/AIDS Research Division at the University of KwaZulu-Natal, the European Union (through the Microcon research partnership on the microfoundations of violent conflict, via the CSSR) and by grants from the Andrew W. Mellon Foundation to the University of Michigan and the University of Cape Town.

2.5 Dropout rates and labour market outcomes

This section examines high school dropout rates in my data and also explore labour market outcomes for both dropouts and graduates. As noted before, the definition of high school dropout in this study is consistent with the definition broadly used in the literature - youths who enrolled but did not complete high school for any reason (Brooks-Gunn et al., 1993; Rumberger, 2001; Campbell, 2015). Specifically, dropouts in my data refer to young people who started high school but did not complete grade 12, which is the final year of high school in South Africa. For example, I classify individuals enrolled in wave 1, but no longer enrolled in wave 2 and have not completed grade 12 as dropouts only in wave 2. I then follow the same procedure in the other waves of the data.

Graduates on the other hand are those who have completed grade 12. The longitudinal nature of the CAPS allows me to see how individuals in my sample transit from one state to another over time. The starting state for all individuals (both dropouts and graduates) is enrolment in education in wave 1. High school drop out rates are presented in Table 2.4. The reported estimates show high dropout rates across waves or periods in the data. For example, periods, close to half of the total number of enrolled students dropped out of high school before completing

Table 2.4: Dropout rates

Period	Dropout rates (%)
First period	-
Second period	45.41
Third period	47.61
Fourth period	46.01
Fifth period	44.15

the final grade (an aggregation of these individuals formed my drop out sample). The observed high dropout rate in the data is worrisome due to the possible economic and socials costs associated with high school dropout (see Section 2.1 for discussion)

What are the possible drivers of the high dropout rates observed in the data? There are different factors that could affect the decision to drop out of high school. In Section 2.6, I explore this question further by examining observable determinants of dropout. In Table 2.5, I present labour market outcomes for dropouts and graduates in terms of unemployment duration and continuous employment spells (see Appendix A for a more detailed transition matrices across waves and multiple destination states).[1] The former measures the amount of time an individual spends in unemployment, while the latter measures the amount of time an individual spends on a given job.

The first three columns in Table 2.5 show the probability of continuous spell of unemployment for high school graduates relative to dropouts.[2] The probability estimates indicate, relative to high school graduates, high school dropouts have higher spells of continuous unemployment across the three periods examined.[3] This is intuitive: dropouts have low human capital accumulation which pushes them towards the bottom of the unemployment pool once they enter the labour force. Thus, they have lower chance of exiting the pool compared to individuals with high school qualification.

[1] Note that in the appendix tables E is education, J is job or employment, U is unemployment, and N is non-participation

[2] The transition matrix between wave 1 and wave 2 is excluded because wave 1 has only one state (i.e., enrollment in education), which makes transition across states impossible

[3] (i.e., waves 2 and 3, waves 3 and 4, and waves 4 and 5)

Table 2.5: Continuous spells of unemployment and employment

	Prob. of continuous spell of unemployment			Prob. of continuous spell of employment		
Period	Graduated (G)	Dropout (D)	G relative to D	Graduated (G)	Dropout (D)	G relative to D
Transition for one period	0.3929	0.4693	83.72	0.5149	0.5200	99.02
Transition for two periods	0.1533	0.2566	59.74	0.2483	0.2638	94.12
Transition for three periods	0.0366	0.1179	30.99	0.1273	0.1294	98.38

Source: Own estimation using CAPS

The last three columns of Table 2.5 show that the probability of continuous spell of employment for graduates, relative to dropout. The estimated probabilities are better interpreted as the likelihood of holding on to a particular job over time. For these estimates, I do not see much difference between dropouts and graduates. This suggests that although high school dropouts stay longer in the unemployment pool than high school graduates, dropouts are able to hold on to their jobs like their counterparts with high school qualification.

In Table 2.6, I compare raw monthly wage of high school dropouts to that of high school graduates. The estimates show that high school dropouts consistently earned less in the labour market than their counterparts with high school qualification. Combining the relatively low wage high school dropouts earn in the labour

Table 2.6: Monthly wage (Rand)

Period	Graduates	Dropouts
First period	2,058	1,191
Second period	2,627	1,700
Third period	2,627	1,755
Fourth period	2,975	1,969
Fifth period	4,433	2,619

Source: Own estimation using CAPS

with the evidence that they have a relatively long spells of continuous unemployment paints a somewhat gloomy picture for their future and the society they live in.

2.6 Method

In this section, I estimate observable characteristics that influence the decision to drop out of high school. Observable factors that determine successful transition from school to work are also estimated. This section also include a simple framework that could be used to motivate the estimation of high school dropout.

2.6.1 Statistical model of school choice

Cameron & Heckman (1998) provides a simple statistical model of school choice. The explained variable equals the ex-post observed school attainment, and the independent variables includes individual and household's socio-economic characteristics of interest. Formally, let s denotes the grade attained in school (i.e., completed schooling). Furthermore, let $D_s = 1$ if an individual completes grade s, and $D_s = 0$ otherwise. Independent variables that explain transition from $s-1$ to s can be presented as $\mathbf{X}_s$=$\mathbf{x}_s$. Therefore, the probability of transiting from $s-1$ to s can be written as:

$$\Pr(D_s = 1|\mathbf{X}_s = \mathbf{x}_s, D_{s-1}) = P_{s-1,s}(\mathbf{x}_s) \tag{2.1}$$

The probability that an individual drops out at grade $s-1$ is the event, $D_{s-1} = 0$. In such an occurrence, we do not have any meaningful event in congruence with the outcome, $D_s = 1$ and $D_{s-1} = 0$. These probabilities explain the schooling-transition model, which become a fully specified Markov process when an initial condition D_1 is assumed.

For practical purpose, a logistic transition probability model can be used for estimation. See Mare (1980), and Allison (1982) for an excellent motivation and detailed discussion. A summary, in the context of Equation (2.1), can be represented as

$$P_{s-1,s}(\mathbf{x}_s) = \frac{\exp(\mathbf{x}_s\beta_s)}{1 + \exp(\mathbf{x}_s\beta_s)} \tag{2.2}$$

Apart from logistics regressions, linear regressions have also been used in the literature to study the relationship between total completed education and socio-economic characteristics (Sewell & Hauser, 1975; Featherman & Hauser, 1975). A summary of this approach, following Cameron & Heckman (1998), is presented below, using notations from Equation (2.1). Completed schooling is the sum of the D_s. This can be written more formally as:

$$\text{completed schooling} = \sum_{s=1}^{\bar{S}} D_s,$$

where the highest attainable grade is represented by $\bar{S}$. I can let $\mathbf{X} = (\mathbf{X}_1, ..., \mathbf{X}_{\bar{S}})$. Then the probability of completing $S{=}s$ years of schooling can be written as

$$\Pr\left(\sum_{j=1}^{\bar{S}} D_j = s|\mathbf{X} = \mathbf{x}\right) = \left[\prod_{l=1}^{s} P_{l-1,l}\left(\mathbf{x}_l\right)\right]\left[1 - P_{s,s-1}(\mathbf{x}_s)\right], \tag{2.3a}$$

where $P_{\bar{S},\bar{S}+1} = 0$. Expected schooling therefore writes as

$$E\left(\sum_{s=1}^{\bar{S}} D_s|\mathbf{X} = \mathbf{x}\right) = \sum_{s=1}^{\bar{S}} s\Pr\left(\sum_{j=1}^{\bar{S}} D_j = s|\mathbf{X} = \mathbf{x}\right). \tag{2.3b}$$

In terms of the schooling-transition model, linear regression approximation to (2.3a) and (2.3b) combine the β_s in an uninterpretable manner (Cameron & Heckman, 1998). This is perhaps the reason why logistic regressions are preferred to linear regression in estimating simple models of school choice.

2.6.2 Estimating dropout and transition to work

Cox proportional hazard (CPH) model, which is due to Cox (1972), is often used in the literature to estimate the correlates of interest in an time-to-event analysis (transitions to work and high school dropout in my case). However, the CPH model is designed primary to admit continuous-time data. Different extensions of the CPH model to accommodate discrete-time data have been proposed by different authors. In this motivational chapter, I follow closely the method outlined in Allison (1982) due to its simplicity.

In the case of discrete-time data, the assumption is that time can only take positive integer values ($t = 1, 2, 3, ...$) and the researcher observe a total of n independent individuals ($i = 1$, ...,n) beginning at some natural starting point $t = 1$. This will continue until time t_i. At this point, there are two possibilities; either an event of interest occurs (e.g, drop out from high school or successfully transit to work) or the observation is censored. Most models used in analysing transition data assume that censoring is independent of the occurrence of events (i.e., censoring does not depend on the hazard rate). Put differently, there is no selective withdrawal of individuals because they are more or less likely to experience an event (Allison, 1982). This is known as the *non informative censoring* assumption. This assumption is maintained throughout this chapter.

To begin, I define a discrete-time hazard rate as follows

$$P_{it} = Pr[T_i = t | T_i \geq t, \mathbf{x}_{it}] \tag{2.4}$$

Where T is the discrete random variable given the uncensored time an event of interest occurred, and $\mathbf{x}_{it}$ is a K x 1 vector of explanatory variables. The descriptive statistics of selected variables used in my estimation is presented in Table B.1. Equation (2.4) is the conditional probability that an individual drop out of school or transit to work at time t, given that neither event has already occurred. The next thing is to establish how the hazard rate defined in Equation (2.4) depends on time and explanatory variables. A simple method often used in the literature to estimate the effect of observable characteristics on hazard rate (or the probability of transition) is the logistic regression function (Cox, 1972; Myers et al., 1973; Byar & Mantel, 1975; Brown, 1975; Thompson, 1977; Mantel & Hankey, 1978). I can write the logistic regression function as

$$P_{it} = \frac{1}{1 + \exp(-\alpha_t - \beta' \mathbf{x}_{it})} \tag{2.5}$$

The logit form of Equation (2.5) can be represented as

$$\log\left[\frac{P_{it}}{(1 - P_{it})}\right] = \alpha_t + \beta^{'} \mathbf{x}_{it} \tag{2.6}$$

As it is the case in proportional hazard model, $\alpha_t (t = 1, 2, ...)$ is just a set of constants that are left unspecified. For Equation (2.5), the likelihood of the data can be written as

$$L = \prod_{i=1}^{n} [\Pr(T_i = t_i)]^{\delta_i} [\Pr(T_i > t_i)]^{1-\delta_i} \tag{2.7}$$

where the variable δ_i is set equal to 1 if i is uncensored and zero otherwise. Each of the probabilities in Equation (2.7) can be expressed as a function of the hazard rate. Employing basic properties of conditional probabilities, it can be shown that

$$\Pr(T_i = t) = P_{it} \prod_{j=1}^{t-1} (1 - P_{ij}) \tag{2.8}$$

$$\Pr(T_i > t) = \prod_{j=1}^{t} (1 - P_{ij}) \tag{2.9}$$

If I substitute Equations (2.8) and (2.9) into Equation (2.7) and take the logarithm, I get the likelihood function

$$\begin{aligned} \log L = & \sum_{i=1}^{n} \delta_i \log\left[\frac{P_{it_i}}{(1 - P_{it_i})}\right] \\ & + \sum_{i=1}^{n} \sum_{j=1}^{t_i} \log(1 - P_{ij}) \end{aligned} \tag{2.10}$$

At this point, I can simply substitute Equation (2.5) into Equation (2.10) and then proceed to maximize log L with respect to β.

The estimates of β for both high school dropout and transition from school to work are presented in Table 2.7. The results in column 1 of Table 2.7 show that age is negatively associated with the likelihood of dropping out of high school. In addition, the results suggest that being a female increases the likelihood of dropping out og high school. Possible explanations for this outcome may include early marriage and teenage pregnancy; gender bias in human capital investment;

Table 2.7: Logistic regression for dropout and transition to work

Covariates	(1) Dropout	(2) Work
Siblings	0.011	-0.011
	(0.02)	(.012)
Age	-0.185***	0.222***
	(0.03)	(0.03)
Sex (female)	0.417**	-0.313**
	(.17)	(.14)
Distance to School	-0.009*	0.006
	(0.01)	(0.01)
Numeracy test score	-0.042**	0.035**
	(0.02)	(0.02)
Race (white)	-2.053***	1.631***
	(0.56)	(0.44)
Above average in Sec. Sch. (Average/below)	-0.167	0.287*
	(0.20)	(0.16)
Teacher absenteeism (not absent)	0.084	0.021
	(0.22)	(0.20)
Time spent on household work	0.014	-0.020**
	(0.01)	(0.01)
Father's years of schooling	-0.083***	0.078***
	(0.03)	(0.02)
Mother's years of schooling	0.010	0.012
	(0.03)	(0.02)
Observations	1,161	1,398

Source: Own estimation using CAPS. Robust standard errors are in parentheses at individual level.

and excessive household chores. Distance to school is negatively associated with high school dropout. This outcome seems counter-intuitive, as one would have expected an opposite result. However, it could be that the farther the school, the more valuable it is to students. As expected, test score is negatively associated with high school drop pout; good academic performance could serve as a motivation to continue in education. Furthermore, being black increases the likelihood of dropping out of high school prior to graduation. Inequality of opportunity between whites and black still exists in South Africa even after many years of apartheid. Results in Table 2.7 further show that parental education (i.e., father's years of education) increases the likelihood of completing high school education.

Column 2 of Table 2.7 provides estimates of the determinants of successful transition from school to work. Results suggest that age is positively associated with school to employment transition. That is, older students are more likely to successfully transit from school to work than younger students. In addition, being a female reduces the likelihood of a successful transition from school to work. Furthermore, good academic performance, as measured by test scores, also increases the likelihood that an average student will successfully transit from school to work. Good academic performance is an indicator of high ability, which could also act as a signal of productivity in the labour market. Similarly, father's education is positively associated with a successful transition from school to employment. Results in Table 2.7 also indicate that being white increases the likelihood of successfully transiting from school to work. Spending time on household work seems to reduces the likelihood of a successful transiting to work.

2.7 Conclusion

Using the Cape Area Panel Study (CAPS) data, this paper examines the labour market outcomes of school dropouts and graduates and also explore possible observable factors that determine high school dropout and successful transition from school to employment. First, a high dropout rate is observed in the data. In addition, estimates from transition probabilities show that school dropouts have higher continuous spells of unemployment in relation to high school graduates. Furthermore, high school dropouts consistently earned less in raw wages than high school graduates.

Estimates from panel logistic regression identity the following as those who are more vulnerable to high school dropout: black students, young students, students with record of poor academic performance, students who live close to school, and students with less educated fathers. On the other hand, factors that increase the chance of a successful transition to work include good academic performance, being white or coloured, age, and having an educated father. These results should be interpreted as correlates of high school dropout and transition to employment, and not necessary causations.

References

Alexander, K. L., Entwisle, D. R., & Horsey, C. S. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of education*, 87–107.

Allison, P. D. (1982). Discrete-time methods for the analysis of event histories. *Sociological methodology*, *13*, 61–98.

Ampiah, J. G., & Adu-Yeboah, C. (2009). Mapping the incidence of school dropouts: a case study of communities in northern ghana. *Comparative Education*, *45*(2), 219–232.

Ananga, E. D. (2011). Typology of school dropout: The dimensions and dynamics of dropout in ghana. *International Journal of Educational Development*, *31*(4), 374–381.

Andrews, M. J., Bradley, S., & Stott, D. (2002). Matching the demand for and supply of training in the school-to-work transition. *The Economic Journal*, *112*(478), C201–C219.

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

Archambault, I., Janosz, M., Morizot, J., & Pagani, L. (2009). Adolescent behavioral, affective, and cognitive engagement in school: Relationship to dropout. *Journal of school Health*, *79*(9), 408–415.

Battin-Pearson, S., Newcomb, M. D., Abbott, R. D., Hill, K. G., Catalano, R. F., & Hawkins, J. D. (2000). Predictors of early high school dropout: A test of five theories. *Journal of educational psychology*, *92*(3), 568.

Bhorat, H., Hirsch, A., Kanbur, R., & Ncube, M. (2014). *Economic policy in south africa, past, present and future.* Development Policy Research Unit, University of Cape Town.

Biggeri, L., Bini, M., & Grilli, L. (2001). The transition from university to work: a multilevel approach to the analysis of the time to obtain the first job. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *164*(2), 293–305.

Bilkic, N., Gries, T., & Pilichowski, M. (2012). Stay in school or start working? the human capital investment decision under uncertainty and irreversibility. *Labour Economics*, *19*(5), 706–717.

Bloom, D. E., Canning, D., & Fink, G. (2010). Implications of population ageing for economic growth. *Oxford review of economic policy*, *26*(4), 583–612.

Böheim, R., & Taylor, M. P. (2002). The search for success: do the unemployed find stable employment? *Labour Economics*, *9*(6), 717–735.

Bowlus, A. J., Kiefer, N. M., & Neumann, G. R. (2001). Equilibrium search models and the transition from school to work. *International Economic Review*, *42*(2), 317–343.

Branson, N., Hofmeyr, C., & Lam, D. (2014). Progress through school and the determinants of school dropout in south africa. *Development Southern Africa*, *31*(1), 106–126.

Brooks-Gunn, J., Guo, G., & Furstenberg Jr, F. F. (1993). Who drops out of and who continues beyond high school? a 20-year follow-up of black urban youth. *Journal of Research on Adolescence*, *3*(3), 271–294.

Brown, C. C. (1975). On the use of indicator variables for studying the time-dependence of parameters in a response-time model. *Biometrics*, *31*(4), 863–872.

Byar, D. P., & Mantel, N. (1975). 393: Some interrelationships among the regression coefficient estimates arising in a class of models appropriate to response-time data. *Biometrics*, 943–947.

Cameron, S. V., & Heckman, J. J. (1998). Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of american males. *Journal of Political economy*, *106*(2), 262–333.

Campbell, C. (2015). High school dropouts after they exit school: Challenges and directions for sociological research. *Sociology Compass*, *9*(7), 619–629.

Card, D., & Krueger, A. B. (1992a). Does school quality matter? returns to education and the characteristics of public schools in the united states. *Journal of political Economy*, *100*(1), 1–40.

Card, D., & Krueger, A. B. (1992b). School quality and black-white relative earnings: A direct assessment. *The Quarterly Journal of Economics*, *107*(1), 151–200.

Christle, C. A., Jolivette, K., & Nelson, C. M. (2007). School characteristics related to high school dropout rates. *Remedial and Special Education*, *28*(6), 325–339.

Cockx, B., & Picchio, M. (2012). Are short-lived jobs stepping stones to long-lasting jobs? *Oxford Bulletin of Economics and Statistics*, *74*(5), 646–675.

Cockx, B., & Picchio, M. (2013). Scarring effects of remaining unemployed for long-term unemployed school-leavers. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *176*(4), 951–980.

Corrales, H., & Rodríguez, B. (2004). Transition from education to first significant job in spain: the influence of educational attainment. In *The 58th international atlantic economic conference. chicago, illinois* (pp. 7–10).

Cox, D. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society, Series B*, *34*, 187–202.

Doiron, D., & Gørgens, T. (2008). State dependence in youth labor market experiences, and the evaluation of policy interventions. *Journal of Econometrics*, *145*(1), 81–97.

Dorsett, R. (2006). The new deal for young people: effect on the labour market status of young men. *Labour Economics*, *13*(3), 405–422.

Dunne, M., & Ananga, E. D. (2013). Dropping out: Identity conflict in and out of school in ghana. *International Journal of Educational Development*, *33*(2), 196–205.

Eckstein, Z., & Wolpin, K. I. (1995). Duration to first job and the return to schooling: Estimates from a search-matching model. *The Review of Economic Studies*, *62*(2), 263–286.

Eckstein, Z., & Wolpin, K. I. (1999). Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica*, *67*(6), 1295–1339.

Farmer, T. W., Estell, D. B., Leung, M.-C., Trott, H., Bishop, J., & Cairns, B. D. (2003). Individual characteristics, early adolescent peer affiliations, and school dropout: An examination of aggressive and popular group types. *Journal of School Psychology*, *41*(3), 217–232.

Featherman, D. L., & Hauser, R. M. (1975). Design for a replicate study of social mobility in the. *Social indicator models*, 219.

Filmer, D., & Fox, L. (2014). *Youth employment in sub-saharan africa.* World Bank Publications.

Flinn, C. (2006). Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates. *Econometrica*, *74*(4), 1013–1062.

Fuller, G. E. (2004). The youth crisis in middle eastern society. *Brief Paper. Clinton Township: Institute of Social Policy and Understanding*.

Ge, S. (2013). Estimating the returns to schooling: Implications from a dynamic discrete choice model. *Labour Economics*, *20*, 92–105.

Genda, Y., & Kurosawa, M. (2001). Transition from school to work in japan. *Journal of the Japanese and International Economies*, *15*(4), 465–488.

Goldberg, J., Garcia, M., & Fares, J. (2009). *Youth in africa's labor market.* World Bank Publications.

Holzner, C., & Launov, A. (2010). Search equilibrium and social and private returns to education. *European Economic Review*, *54*(1), 39–59.

Janosz, M., Archambault, I., Morizot, J., & Pagani, L. S. (2008). School engagement trajectories and their differential predictive relations to dropout. *Journal of social Issues*, *64*(1), 21–40.

Jaunky, V. C., & Khadaroo, A. J. (2008). The school-to-work transition for university graduates in mauritius: A duration model approach. *The Empirical Economics Letters*, *7*(1), 25–36.

Keane, M. P., & Wolpin, K. I. (2000). Eliminating race differences in school attainment and labor market success. *Journal of Labor Economics*, *18*(4), 614–652.

Kiefer, N. M., & Neumann, G. R. (1979). An empirical job-search model, with a test of the constant reservation-wage hypothesis. *Journal of political Economy*, *87*(1), 89–107.

Kumo, W., Chulu, O., & Minsat, A. (2016). *Economic outlook: South africa.* African Development Bank, Organisation for Economic Co-operation and Development, United Nations Development Programme.

Lassibille, G., Gómez, L. N., Ramos, I. A., & de la O Sánchez, C. (2001). Youth transition from school to work in spain. *Economics of Education Review*, *20*(2), 139–149.

Lee, J. C., & Staff, J. (2007). When work matters: The varying impact of work intensity on high school dropout. *Sociology of Education*, *80*(2), 158–178.

Mantel, N., & Hankey, B. F. (1978). A logistic regression analysis of response-time data where the hazard function is time dependent. *Communications in Statistics-Theory and Methods*, *7*(4), 333–347.

Mare, R. D. (1980). Social background and school continuation decisions. *Journal of the american statistical association*, *75*(370), 295–305.

Motala, S. (2011). Educational access in south africa. *Journal of Educational Studies*, *2011*(Special issue 1), 84–103.

Motala, S., Dieltiens, V., & Sayed, Y. (2009). Physical access to schooling in south africa: mapping dropout, repetition and age-grade progression in two districts. *Comparative Education*, *45*(2), 251–263.

Murphy, K. M., & Peltzman, S. (2004). School performance and the youth labor market. *Journal of labor Economics*, *22*(2), 299–327.

Myers, M. H., Hankey, B. F., & Mantel, N. (1973). A logistic-exponential model for use with response-time data involving regressor variables. *Biometrics*, 257–269.

Neild, R. C., Stoner-Eby, S., & Furstenberg, F. (2008). Connecting entrance and departure: The transition to ninth grade and high school dropout. *Education and Urban Society*, *40*(5), 543–569.

Nguyen, A., Taylor, J., et al. (2005). *From school to first job: a longitudinal analysis* (Tech. Rep.).

O'Higgins, N. (2003). Trends in the youth labour market in developing and transition countries.

Plank, S. B., DeLuca, S., & Estacion, A. (2008). High school dropout and the role of career and technical education: A survival analysis of surviving high school. *Sociology of Education*, *81*(4), 345–370.

Pugatch, T. (2012). Bumpy rides: School to work transitions in south africa.

Rees, D. I., & Mocan, H. N. (1997). Labor market conditions and the high school dropout rate: Evidence from new york state. *Economics of Education Review*, *16*(2), 103–109.

Rumberger, R. W. (1987). High school dropouts: A review of issues and evidence. *Review of educational research*, *57*(2), 101–121.

Rumberger, R. W. (2001). Who drops out of school and why. *Understanding Dropouts: Statistics, Strategies, and High-Stakes Testing. Washington*.

Salas-Velasco, M. (2007). The transition from higher education to employment in europe: the analysis of the time to obtain the first job. *Higher Education*, *54*(3), 333.

Sayed, Y., & Motala, S. (2012). Equity and no feeschools in south africa: Challenges and prospects. *Social Policy & Administration*, *46*(6), 672–687.

Sciulli, D., & Signorelli, M. (2011). University-to-work transitions: an empirical analysis on perugia graduates. *European journal of higher education*, *1*(1), 39–65.

Sewell, W. H., & Hauser, R. M. (1975). Education, occupation, and earnings. achievement in the early career.

Sibanda, A. (2004). Who drops out of school in south africa? the influence of individual and household characteristics.

South, S. J., Haynie, D. L., & Bose, S. (2007). Student mobility and school dropout. *Social Science Research*, *36*(1), 68–94.

Thompson, W. (1977). On the treatment of grouped observations in life studies. *Biometrics*, 463–470.

UNESCO. (2019). *Meeting commitments: Are countries on track to meet sdg 4?* The United Nations Educational, Scientific and Cultural Organization.

Vallerand, R. J., Fortier, M. S., & Guay, F. (1997). Self-determination and persistence in a real-life setting: toward a motivational model of high school dropout. *Journal of Personality and Social psychology*, *72*(5), 1161.

Vanoverberghe, J., Verhaest, D., Verhofstadt, E., & Omey, E. (2008). The transition from school to work in flanders: a duration analysis. *Journal of education and work*, *21*(4), 317–331.

Wenger, J. W. (2002). Does the dropout age matter? how mandatory schooling laws impact high school completion and school choice. *Public Finance & Management*, *2*(4).

Wolpin, K. I. (1987). Estimating a structural search model: the transition from school to work. *Econometrica: Journal of the Econometric Society*, 801–817.

Wolpin, K. I. (1992). The determinants of black-white differences in early employment careers: Search, layoffs, quits, and endogenous wage growth. *Journal of Political Economy*, *100*(3), 535–560.

Appendix A

Transition matrices

Table A.1: Wave 1 $\longrightarrow$ Wave 2 for graduates

	E	J	U	N	*cens.*	total obs.
E	65.44	8.27	7.14	19.08	0.07	1,499
Total obs.	981	124	107	286	1	1,499

Table A.2: Wave 2 $\longrightarrow$ Wave 3 for graduates

	E	J	U	N	*cens.*	total obs.
E	72.10	10.90	5.50	7.33	4.17	981
J	14.52	40.32	22.58	11.29	11.29	124
U	12.15	23.36	19.63	35.51	9.35	107
N	12.59	19.23	15.40	18.18	34.62	286
Total obs.	774	237	147	176	165	1,499

Table A.3: Wave 3 ⟶ Wave 4 for graduates

	E	J	U	N	*cens.*	total obs.
E	67.70	11.50	4.01	10.59	6.20	774
J	9.28	41.77	30.38	8.02	10.55	237
U	6.80	35.37	23.13	23.81	10.88	147
N	10.80	45.45	10.23	30.68	2.84	176
Total obs.	575	320	155	190	259	1,499

Table A.4: Wave 4 ⟶ Wave 5 for graduates

	E	J	U	N	*cens.*	total obs.
E	31.65	33.04	15.48	0.35	19.48	575
J	7.50	51.25	12.50	0.31	28.44	320
U	7.10	47.10	19.35	0.65	25.81	155
N	12.11	50.00	27.89	0.00	10.00	190
Total obs.	240	522	212	14	511	1,499

Table A.5: Wave 1 ⟶ Wave 2 for dropouts

	E	J	U	N	*cens.*	total obs.
E	62.26	5.43	15.76	16.55	0.00	1,142
Total obs.	771	62	180	189	0	1,142

Table A.6: Wave 2 ⟶ Wave 3 for dropouts

	E	J	U	N	*cens.*	total obs.
E	64.98	7.59	7.59	19.69	0.14	771
J	3.23	51.61	35.48	1.61	8.06	62
U	5.00	26.67	21.67	38.89	7.78	180
N	5.82	22.22	21.16	48.68	2.12	189
Total obs.	484	176	155	303	24	1,142

Table A.7: Wave 3 $\longrightarrow$ Wave 4 for dropouts

	E	J	U	N	*cens.*	total obs.
E	55.99	14.05	5.79	23.76	0.41	484
J	3.98	45.45	40.91	4.55	5.11	176
U	3.23	35.48	21.94	33.55	5.81	155
N	9.90	21.12	8.58	55.12	5.28	303
Total obs.	313	267	160	345	57	1,142

Table A.8: Wave 4 $\longrightarrow$ Wave 5 for dropouts

	E	J	U	N	*cens.*	total obs.
E	0.00	36.42	58.79	0.00	4.79	313
J	0.00	47.94	31.09	0.00	20.97	267
U	0.00	39.38	40.63	0.00	20.00	160
N	0.00	28.99	52.75	0.00	18.26	345
Total obs.	0	405	514	0	223	1,142

Appendix B

Descriptive statistics

Table B.1: Descriptive statistics of selected variables

Variable	Mean	St. Dev.
School to employment transition	0.44	0.50
School to drop-out	0.67	0.46
Age	19.60	2.89
Siblings	2.24	2.95
Distance to School	19.96	15.70
Numeracy test score	9.74	5.87
Sex (female)	0.549	0.50
Performance in Sec. Sch.	0.291	0.45
Time on household work	10.53	64.68
Father?s schooling (in years)	8.90	3.90
Mother?s schooling (in years)	8.85	3.32

Source: Own estimation using CAPS

Chapter 3

Assessing the patterns of self-selection into high school dropout and high school graduation

3.1 Introduction

Dropping out of high school has both economic and social costs. Individuals who drop out may have difficulties finding "good" jobs due to low human capital accumulation that persists into the future in the form of low wages. Limited work opportunities and low earnings in the labour market may lead to inter-generational cycle of poverty. Furthermore, dropouts may not be able to fully contribute to national economic growth and development, and may also be susceptible to social vices (Hill, 1979; Cawley et al., 2001; Lochner & Moretti, 2004). These possible costs of high school dropout have been acknowledged by different countries, multinational agencies and relevant stakeholders.

The economic and social costs associated with high school drop-out are particularly severe in South Africa where high school drop-out rate has reached a point of national crisis. Approximately 60% of first graders will ultimately drop out rather than complete 12th Grade. Likewise, by Grade 12, only 52% of the age appropriate population remain enrolled (Weybright et al., 2017). As a result of low human capital accumulation, many young people who drop out remain unemployed for a long time. South Africa currently has the highest rate of youth unemployment (percent of total labor force ages 15-24) in the world (ILO, 2019).

The first step in addressing the problem of high school drop-out is to identify possible factors that drive selection into dropout and graduation. There is an understanding in the literature that apart from observable characteristics, individuals may choose to drop out or complete high school education due to unobserved factors. Individuals may vary in motivation, innate ability, preferences and expectations. Thus, the patterns of dropout and graduation that we observe may be driven largely by unobservable characteristics. This knowledge is important to guide policies aimed at addressing the problem of high school drop-out. For example, if those who choose to drop out are systematically different from those who graduates, "incentives" policies may not keep them in school.

In order to assess the possible differences in unobservables between high school dropouts and graduates, I propose to simultaneously estimate the joint distribution of unobserved differences between dropouts and graduates using a competing risks model. I allow for a general correlation of unobserved differences in my estimation since the unobserved factors that drive selection into dropout could be related to the unobserved factors that drive selection into high school graduation in different ways.

My estimation technique imposes no stringent parametric assumption on the distribution of unobserved differences. Rather, I take the unobserved differences to be a mixture of finite types. That is, each destination (dropout or graduation) is made up of different types of people with varying level of unobservables. One of the advantages of this approach is that it relies less on parametric assumptions and depend more on the distribution in data during estimation. After estimating the competing risks model, I then write an analytical expression for the correlation of unobservable differences between dropouts and graduates. The estimated correlation coefficient is my proposed measure of selection on unobservables as it allows me to directly examine the relationship between the unobserved factors that drive selection into high school dropout and high school graduation in a flexible way. Although the competing risks model, which I estimated is an existing model, using its structure to write an analytical expression that allows for the direct assessment of selection on unobservables in education choices is novel in the economics of education literature. This method is not only flexible, but it is also data-driven and it does not rely on sundry assumptions.[1]

The estimated correlation coefficient can either be positive, negative or zero. A statistically significant correlation indicates the presence of selection on unobservables. A positive and significant correlation estimate suggests a positive selection between dropping out and graduating. That is, the same sets of unobserved factors drive selection into high school dropout and high school graduation. A negative and significant correlation estimate suggests a negative selection. That is, different sets of unobserved factors drive selection into dropout and graduation. A statistically insignificant correlation implies the absence of selection on unobservables. This method that I have proposed to check for selection on unobservables in the decision to dropout of high school or complete high school education is similar to the Heckman selection model in principle. The bivariate (dropout and graduation) distribution of unobserved heterogeneity in my approach is not fundamentally different from the bivariate distribution of errors in the Heckman model. While the method I propose has a discrete bivariate distribution with 2 points of support, Heckman model has a continuous bivariate normal distribution.

My estimation returns a non significant measure of selection. That is, the estimated correlation coefficient is not statistically different from zero at standard level of significance. This suggests the absence of systematic sorting or selection on unobservables between high school dropouts and graduates in my data. The rest of the chapter is adumbrated as follows. Section 3.2 reviews the literature; section 3.3 discusses the data and context of the study, econometric specification

[1] The only assumption I make in is on the number of types in my mixture when estimating the competing risks model. Specifically, I take the unobserved heterogeneity in each destination (dropout or graduation) to have a mixture of two or a 2-point support. That is, each destination is a mixture of people with either high or low unobservables like innate ability, expectation, motivation, etc. This categorization is intuitive and also make estimation more manageable

is presented in section 3.4; section 3.5 describes the correlation coefficient between high school dropouts and graduates; empirical results are discussed in section 3.6; while section 3.7 concludes.

3.2 Literature review

3.2.1 Self-selection

There are studies in the education and self-selection literature that are close to my work in this chapter, although my estimation approach is different from what exists in the literature. Roy (1951) provides one of the earliest discussions of self-selection, albeit in the context of occupational choices. In a very simple setting, Roy's theoretical model assumes that an individual has the choice of only two occupation: hunting and fishing. There is freedom to move across occupations and considerable degree of association may exist between individual performances as hunters and fishermen. He argued that this association could either be positive (the best hunter is also the best fisherman) or negative association (the best hunter makes the worst hunter and vice-versa).

The paper by Willis & Rosen (1979a) was the first to extend Roy's self-selection model and analyze schooling decision in a non-hierarchical (multiple skills) structure. The application of the self-selection model to the schooling decision process is considered a major advance as it brings us into a framework where we no longer think of the return to schooling as a single parameter (Belzil, 2007). Willis and Rosen used a structural equation model, where the schooling decision is a single latent index equation, and optimal schooling decision is obtained by assuming that individuals maximize life-time earning. They applied their model to a sample of World War II veterans and find evidence that supports positive sorting between high school and college decisions. That is, those who choose to go to college have higher lifetime earnings in college-required jobs than those who did not and those who did not enrol in college have higher lifetime earnings in high school-required jobs than college enrollees, had they decided to work in high school-required jobs.

The contribution made by Willis & Rosen (1979a) is particularly remarkable in helping researchers to understand that ability bias is a complex phenomenon in the presence of multiple skills (two skills or dimensions in their case). A hierarchical (single dimension) notion of ability bias is a simplification of this process. However, the authors could not obtain an estimate of the correlation between unobservables across schooling choices (i.e., they did not quantify the nature of self-selection) because they were unable to identify the distributions of counterfactual choices. n this study, I am able to quantify the nature of self-selection and show to what extent selection matters.

Carneiro et al. (2011) extend Willis and Rosen's self-selection model and and also apply it to high school and college education choices. Their first extension is to account for the uncertainty in the returns to education. They also differentiate between present value income-maximizing and utility-maximizing evaluation of schooling choices. Unlike Willis & Rosen (1979b), the authors were able to identify the distributions of counterfactual choices in their structural equation model and quantify the nature of selection using the National Longitudinal Survey of Youth (NLSY) and Panel study of income Dynamics (PSID). Like Willis and Rosen, they find evidence that supports positive sorting between high school and college decision. My study is most closely related to Carneiro et al. (2011) in overall objective (i.e., estimation of self-selection in schooling choices). What I do differently is to show an alternative estimation strategy and apply my analysis to the decision to drop out of high school or complete high school education.

The paper by Dolton & van der Klaauw (1999) is technically close to my estimation approach. The authors used a dependent competing risks framework that allows for a flexible semi-parametric specification of both the duration dependence and unobserved individual heterogeneity. While they analyze teacher's decision to quit their job into either non-teaching career or non-working state, I employed a similar model to assess the pattern of self-selection into high drop-out and graduation.

3.2.2 Conditioning on unobserved heterogeneity

There is also a body of work in the literature that describes observable factors that determine high school dropout, conditional on unobserved heterogeneity. Angrist & Krueger (1991) used the season of birth as an instrumental variable for educational attainment in the United States. They find that the birth season influences the decision to leave high school. In particular, people born at the beginning of the year start school at an older age and can therefore drop out after less schooling than people born at the end of the year. The authors also find that compulsory schooling laws reduces high school drop out by about 25%.

Using a longitudinal data from the United States, Eckstein & Wolpin (1999) structurally estimated a sequential model work decision and high school attendance. They find that youths who drop out of high school possess different attributes than those who complete high school education. High school dropouts have lower school ability or motivation, they have lower expectations about the worth of a high school diploma in the labour market, they are better at jobs that are done by non-graduates, they value leisure and do not place much value on school attendance.

Educational sorting hypothesis (signaling model) predicts that an environment where individuals are constrained from entering university will see a decrease in high school dropout rates, compared to an environment with greater university access. On the other hand, human capital theory predicts higher increase in university enrolment and an identical high school dropout rates in an environment with greater university access. Bedard (2001) used the United State's National Longitudinal Survey of Young Men (NLSYM) and the National Longitudinal Survey of Young Women (NLSYW) to test this contradiction. The author finds that labour markets containing universities have greater dropout rates for high schools. This outcome is compatible with the signalling model.

Using the the United State's 1954 Supreme Court judgement that called separate schools for black and white children inherently unequal as an identification strategy, Guryan (2004) examined whether the desegregation benefited black and white students in desegregated school districts. Census data from the 1970s and 1980s reveals that desegregation pushed down high school dropout rates of blacks by two by three percentage points. However, no significant effect is observed among whites. The results remained the same after accounting for family income, parental education, migration selection and state and regional trends.

PROGRESA is an antipoverty programme in Mexico that provides monetary transfers to families contingent upon their children's regular attendance at school. The benefits provided by the programme are intended to offset the opportunity costs of not sending children to school. The first phase of the program was implemented as a randomized social experiment. Behrman et al. (2005) applied a Markov schooling transition model to the experimental data and assessed the impact of the anti-poverty program on schooling attainment and on the underlying behaviours that determine schooling attainment. They find that the program efficiently improves educational achievement by decreasing dropout rates and promoting grade development. In addition, the authors mimicked the program's long-term impacts and discover that if children were to engage in the program between the ages of 6 and 14, the average instructional level would rise by 0.7 years and the percentage of children attending junior high school would rise by 21%. Brazil has a similar anti-poverty programme called *Bolsa Escola/Familia* and findings by Glewwe & Kassouf (2012) show a relatively similar effect.

Montmarquette et al. (2007) developed an econometric model that examines the determinants of working while in school, academic performance and the decision to drop out of high school. The likelihood function of their model allows for a heterogeneous preferences for schooling. Using a Canadian micro data of high school students and high school dropouts, the authors find that being a female student, attending a private school, and living with educated parents are associated with a strong preference for schooling over the labor market. They further find that an individual's decision to drop out of high school is affected by the legal age to access the labour market, high minimum wages, and low unemployment rates.

Using a rich Canadian data-set, Foley et al. (2014) examined how the decision to drop out of high school varies with parental education, while incorporating effects from cognitive and non-cognitive ability and parental valuation of education (PVE). The authors find that cognitive ability and PVE have substantial impacts on the decision to drop out of high school and that parental education has little direct effect on dropping out after controlling for these factors. Their results confirm the importance of determinants of ability by age 15 but also indicate an important role for PVE during teenage years.

Becoming a Man (BAM) is a randomized control trial developed by Youth Guidance (a non-for-profit organization) in the United States in 2009–2010 and 2013–2015. BAM provides youth with weekly group sessions during the school day and utilizes cognitive behavioral therapy to encourage youth to slow down in high-stakes situations. Analyzing the experimental data, Heller et al. (2017) find that in the two studies, participants increased school engagement and a follow-up data reveal that graduation rates increased by 12–19%. In addition, the authors find that the effectiveness of the programme lies in its ability to help youth slow down and reflect on whether their thoughts and behaviors are well suited to the situation they are in, or whether the situation could be differently construed.

3.2.3 Summary of literature review

The literature review indicates that existing evidence on self-selection in education choice is still quite limited, compared to other aspects of education economics like returns to schooling. Thus, this study will add to knowledge in this regard. In addition, almost all the studies reviewed (both the self-selection studies and studies that conditioned on unobserved heterogeneity) are in developed country context, particularly in the United States. My study provides evidence in a developing country context, a setting that is socially and economically different from developed countries.

3.3 Data and context

This study is based on the Cape Area Panel Study (CAPS). The data is a 5 wave longitudinal study of the lives of youths and young adults in metropolitan Cape Town, South Africa. The CAPS is the first data in Sub-Saharan African region designed to follow the lives of youths and young adults over a considerable time. The study's first wave collected interviews in August-December 2002 from about 4,800 randomly selected youth aged 14-22. Wave 1 also collected information on all members of the households of these young people, as well as a random sample of households not having members between 14 and 22 years of age. A third of the youth sample was re-interviewed in 2003 (wave 2a) and re-visited the remaining two-thirds in 2004 (wave 2b). In both 2005 (wave 3) and 2006 (wave 4), the full youth sample was then re-interviewed. For wave 5, full face-to-face interviews was carried out in 2009 with the sample comprising all respondents interviewed in any of Waves 2a, 3 or 4.

Table 3.1: Descriptive statistics

	Dropouts		Graduates	
Variables	Mean	St.Dev	Mean	St. Dev.
Sex (male)	0.470	0.499	0.454	0.498
Local (urban)	0.787	0.409	0.811	0.391
Reservation wage	1726	2700	1406	1777
Time in household work	5.878	5.956	5.574	5.603
Test scores	8.330	5.558	10.735	5.532
Distance to school	19.729	14.590	19.773	14.833
Teacher absenteeism	0.142	0.350	0.106	0.308
Mother's education	8.642	3.258	9.530	3.212
Mother helped with homework	0.200	0.399	0.224	0.417
Censored spell	37%		44%	
Observations	1,539		1,314	

Notes: Reservation wage is in South African Rand.

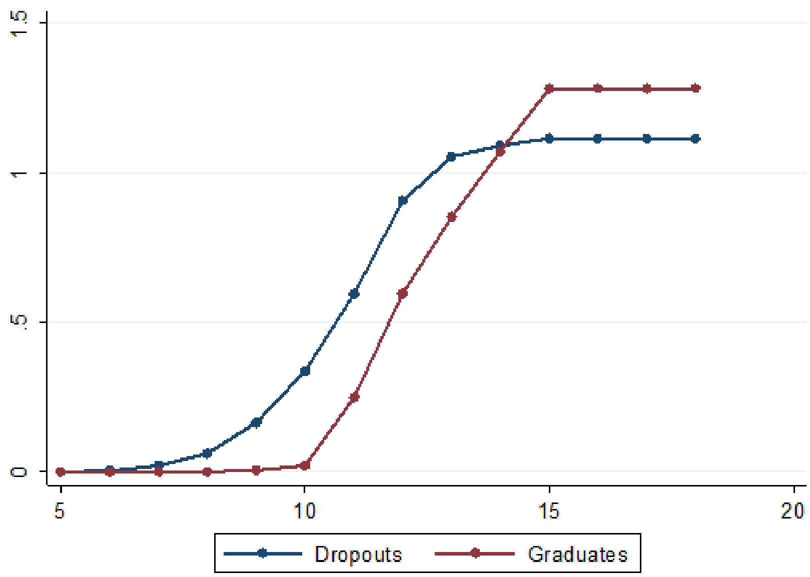
The sample for analysis compromises individuals that were enrolled in high school at the beginning of the survey (2002). These individuals are then followed over time until they either drop-out or graduate. Drop-out is defined as leaving high school without completing the final grade (i.e., grade 12 in South Africa). Individuals who were still enrolled in high school at the end of the survey (i.e., the *survivals*) are right-censored. These individuals are however retained in the sample to avoid selection bias. Nonparametric hazard estimates presented in Figure 3.1 show that the risk of dropping out of high school is higher than the risk of graduating during the first 13 years of curriculum life.[2] The risks of dropping out and graduating are the same in the 14th year. Figure 3.1 further shows that after the 14th year, the risk of graduating becomes higher than the risk of dropping out of high school.[3]

Most developed countries in the world tend to have well structured and regulated curriculum that allow students to complete their schooling within the stipulated time. However, in a developing country setting like South Africa, the curriculum may not be well structured and strictly regulated. This means that students may spend more than the required time before completing their education. For example, more than 50% of graduates in my sample spent above the stipulated 12 years in education before obtaining a high school diploma (see Figure 3.2).

[2]Given that curriculum is not strictly regulated in a developing country like South Africa, it is possible for a student to still drop out or graduate in the 13th year (or beyond) of curricula life. See Figure 3.2 for example.

[3]The idea here is that anyone who has been able to *survive* in education up to the 14th year would rather want to push to graduate, instead of quitting at this stage.

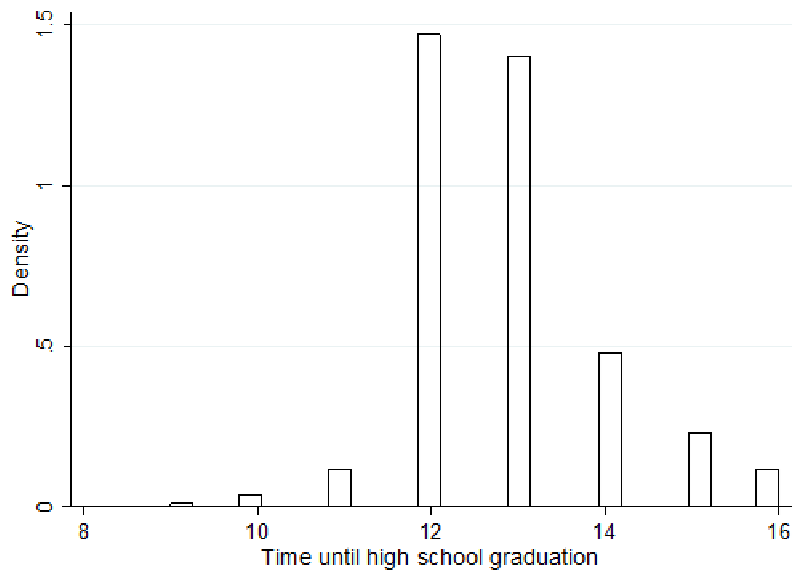
Figure 3.1: Nonparametric hazard estimates



This means that the stock sample of the first wave of the CAPS data potentially over-samples individuals with too long duration of curriculum, leading to possible length bias in estimation. This requires some methodological adjustment to the standard competing risks model in order to account for possible bias. In section 3.3, I show how I make this adjustment. The distribution in Figure 3.2 shows only complete spells; censored spells are excluded.

The management of high school education in South Africa rests with the Department of Basic Education. The DBE is responsible for all schools from Grade R (i.e., grade 0 or the reception/preparatory year) to Grade 12 (the final year of high school). The South African Schools Act (SASA) of 1996a provides legal backing for access to school and the National Education Policy Act of 1996 regulates admission into public and independent (private) schools (Motala et al., 2009).

Figure 3.2: curriculum distribution



Under apartheid, the Bantu system of Education (i.e., the official education system for black South Africans) very much limited the quality of education, and the apartheid government in power regularly under-fund black schools. During the 1980s, there was increase in high schools for black learners, and by 1994 South Africa was able to provide broad access to basic education, although in a system seen as racial and offering poor quality to most learners (Soudien, 2004). Although education policy during apartheid made provision for separate education across racial groups, the emphasis of the post-apartheid democratic government has been to pursue equitable access to education for all based on the Constitution and Bill of Rights (Motala et al., 2009).

3.4 Econometric specification

To assess the pattern of self-selection into high school dropout and graduation, I set up a conditional discrete-time multivariate mixed proportional hazard (competing risks) model. Let t_e (elapsed duration) denote the number of years an individual has spent in education prior to sampling, and t_r (residual duration) the number of years spent in education after sampling. Let T denote the entire duration spent in education and t denote its realization, such that $t = t_e + t_r$. An individual can exit to a $\bar{y}$ mutually exclusive and exhaustive destination states.[4] Let the random variable $Y, (Y = 1, ..., \bar{y})$ represent destination states. For ease of notation, $t_e + t_r$ is replaced by t. Thus, the exit process in terms of $\bar{y}$ transition intensities at each point in time can be described as

$$\theta_y(t) = P(T = t, Y = y | T \geq t), \tag{3.1}$$

We can think of a competing risks as a model in which the transition intensities are hazard functions of $\bar{y}$ independent destination-specific latent duration or survival times (Dolton & van der Klaauw, 1999). More generally, the actual exit time and destination can be represented as

$$T = \min(T_y : y = 1, ..., \bar{y})$$

$$Y = \operatorname{argmin}_y(T_y : y = 1, ..., \bar{y})$$

Each independent random variable $T_y, y = 1, ..., \bar{y}$, is a latent duration that captures the length of stay before an exit of type y occurs in the absence of all other types of exit risks. The distribution $f_y(t)$, conditional on observed and unobserved characteristics, can be parameterized with the proportional hazard

$$\theta_i^y(t|\mathbf{x}, v) = \lambda_y(t) \cdot \exp[\mathbf{x}_i^{'}\boldsymbol{\beta}_y] \cdot v_y, \tag{3.2}$$

where $\lambda_y(t)$ is the risk-specific baseline hazard rate, $\mathbf{x}_i$ represents a set of explanatory variables (see Table 3.1), $\boldsymbol{\beta}_y$ is a vector of unknown parameters, while v_y is a destination-specific realization of the draw from the multivariate distribution of unobserved heterogeneity.

Let time be grouped into interval $K_i^y = [k_{y0}, k_{y1}); [k_{y1}, k_{y1})$ ... $[k_{j-1}^y, \infty)$. The researcher is only able to observe time $T \in \{1, ..., j\}$, where $T = t$ denotes an exit within the interval $[k_{yt-1}, k_{yt})$. Thus, the survival function with the above interval

[4] Dropout or graduate in my case. A student who drops out is no longer at risk of graduating because the competing event (dropout) hinders the occurrence of the event of interest (graduation). I observe continuous exits not just into dropouts, but also into graduation

(i.e., an average student survives in education until k_{yt}, given that no transition has occurred prior to $k_{yt-1}, t = 1, 2...j-1$) can be represented as

$$\exp\left(-\int_{k_{yt-1}}^{k_{yt}} \theta_y(s_y; \beta_y, v_y) ds_y\right) \tag{3.3}$$

Therefore, the probability that a risk-specific transition has taken place within the interval $[k_{yt-1}, k_{yt})$, given that no transition occurred before k_{yt-1} can be presented as

$$h_i^y(t) = 1 - \exp\left(-\int_{k_{yt-1}}^{k_{yt}} \theta_y(s_y; \beta_y, v_y) ds_y\right) \tag{3.4}$$

In Equation (3.4), k represents an indicator of grouped hazard within the interval $[k_{yt-1}, k_{yt})$, and s_y is used to capture the underlying continuous time. The covariates and unobserved heterogeneity are specified in exponential for the purpose of empirical application. Using Equation (3.3), I can rewrite Equation (3.4) as

$$h_i^y(t) = 1 - \exp\left[-\exp\left(\log\left(\int_{k_{yt-1}}^{k_{yt}} \lambda_y(s_y) dt\right) + \mathbf{x}_i^{'}\boldsymbol{\beta}_y + \log(v_y)\right)\right] \tag{3.5}$$

Equation (3.5) specifies the risk-specific group hazard rate in the form of complementary log-log, with unobserved heterogeneity and baseline hazards having unknown distributions. The flexibility of the complementary log-log specification in (3.5) makes it possible for me to specify both the baseline hazard and unobserved heterogeneity non-parametrically. In order to specify the baseline hazard non-parametrically, I use a set of dummies λ_{yk}, to characterize the continuous baseline as

$$\lambda_k^y = \log\left(\int_{k_{yt-1}}^{k_{yt}} \lambda_y(s_y) ds_y\right)$$

To further simplify notation, I define $\log(v_y) = \mu_y$ and then rewrite Equation (3.5) as

$$h_i^y(t) = 1 - \exp\left\{-\exp\left\{\lambda_{yk} + \mathbf{x}_i^{'}\boldsymbol{\beta}_y + \mu_y\right\}\right\} \tag{3.6}$$

Let g_i denote the censoring indicator for each destination state, where $g_i = 1$ if the spell is not censored and $g_i = 0$ otherwise. The overall destination-specific likelihood for student i who has a spell duration of k_{yt}, conditional on the observed component x_i and the unobserved component v_y can be presented as

$$l_i = (h_i^y(k_y; \beta_y; \mu_y))^{g_i} \cdot \prod_{s_y=1}^{k_{yt-gi}} (1 - h_i^y(s_y; \beta_y; \mu_y)) \tag{3.7}$$

There are two major ways to estimate μ (i.e., the unobserved component): parametrically and non-parametrically. The formal approach requires making parametric assumption about the distribution μ. The challenge with this approach is that I do not know the true distribution of μ ex-ante. Consequently, choosing a particular parametric distribution for μ that is not consistent with the data may lead to bias in estimation. The non-parametric approach on the other hand does not make any parametric assumption about the distribution of μ. Rather, it relies on the data in order to estimate the distribution. This is the approach I follow in this study due to its flexibility.

More specifically, I assume a discrete multivariate distribution of unobserved components. For this distribution I suggest that destination-specific marginal distributions have identical supports but potentially different probability mass values. Let there be Z mass points in the support, forming a set $\Omega_\mu = \{\mu^{[1]}, \mu^{[2]}, ..., \mu^{[Z]}\}$. Then for any individual a draw from the distribution of unobserved heterogeneity is a set of destination specific probability mass points $\{\mu_y\}_{y=1,...,\bar{y}}$, where $\mu_y \in \Omega_\mu$ for any y, and $\{p_{yz}\}_{y=1,...,\bar{y},\ z=1,...,Z}$ is the corresponding set of probability mass values, $\sum_{y=1}^{\bar{y}} \sum_{z=1}^{Z} p_{yz} = 1$. Since $\mathbf{x}_i$ includes an intercept, normalization requires setting one element in Ω_μ to zero. Conditioning on the elapsed duration in order to avoid length bias (see Figure 3.2), the total conditional log-likelihood to be maximized is presented as

$$L_i(k_y, \beta_y, \mu_y) = \sum_{i=1}^{N} \log \frac{\left[\sum_{y=1}^{\bar{y}} \sum_{z=1}^{Z} P_{yz} \left((h_i(k_y, \beta_y, \mu_y))^{g_{i,y}} \cdot \prod_{s=1}^{k_{yt-gi}} (1 - h_i(k_y, \beta_y, \mu_y)) \right)\right]}{\left[\sum_{y=1}^{\bar{y}} \sum_{z=1}^{Z} P_{yz} \left(\prod_{s_y=1}^{k_{yte}} (1 - h_i(s_y; \beta_y, \mu_y)) \right)\right]} \tag{3.8}$$

Heckman & Honoré (1989) and Abbring & Berg (2003) show that competing risks models are identified by explanatory variables. More specifically, Abbring & Berg (2003) established that it is sufficient to have one covariate that has continuous variation within an interval to achieve identification. One of such variables in my case is reservation wage, which is sufficient for identification based on the results in Abbring & Berg (2003).

3.5 Correlation coefficient and self-selection

3.5.1 Interpretation

The correlation coefficient represents a measure of self-selection, which I compute using the structure of the competing risks model in Equation (3.8). It allows for a direct assessment of the pattern of self-selection into drop-out and graduation. The computed value can be positive, negative, or zero. Each of these outcomes affect the selection pattern differently.

I have a bivariate distribution of unobserved individual heterogeneity, which is either 0 or μ across risks. There are four possible pairs that can be drawn from this distribution: $(0,0)$; $(0,\mu)$; $(\mu,0)$; (μ,μ). Since both destination states are allowed to share the same support point in my specification; let the first element in the pairs be dropout-specific draw, and the second element in the pairs be graduation-specific draw. Thus, the pattern of selection individuals exhibit depends on the combination of μ and 0 that they drew, say from birth.

Negative correlation

If the measure of selection is negative, that means I predominantly draw $(0, \mu)$ or $(\mu, 0)$ from all possible pairs. For example, if an individual is graduating faster and dropping out slower $(0, \mu)$, it means they are high-ability type. High ability increases the odds of "'good" academic performance and accelerates high school completion. Because graduation and drop-out are mutually exclusive (i.e., they are competing risks), such an individual may never drop out from high school. On the other hand, if an individual is graduating slower and dropping out faster (i.e., they draw 0 for graduation and μ for dropout), it suggests low-ability type. Although an individual with this kind of draw may eventually graduate if they push themselves through the curriculum, they will do so much slower than their peers. Dropping out faster is a sort of reality check; the possibility of earning a high school diploma is low. So its better they drop out of high school and explore alternative prospects.

This is consistent with the theory of return to human capital (Lindsay, 1971; Dominitz & Manski, 1996; Goldin, 1999; Carneiro et al., 2011). If an individual knows that there is a positive return to high ability (e.g., high wages in the labour market) and they are high-ability type, they will choose to complete high school education instead of dropping out. On the other hand, if they know know about the positive return to human capital, but are low-ability type, they will rather drop out fast, than wait to go into the labour market with a qualification they may never achieve.

Positive correlation

If the measure if selection is positive, I predominantly draw $(0, 0)$ or (μ, μ) from all possible pairs. Positive correlations suggests that if an individual is "good" (i.e., high ability for example), they will be "good" in both dimensions (i.e., dropout and graduation), and if they are "bad" (i.e., low ability for example), they will be "bad" in both dimensions. For instance, if an individual is graduating faster and dropping out faster (i.e., they draw μ for for both graduation and dropout), it means they can quickly complete high school education and could also drop out very quickly. Dropping out of high school in this case could be because the individual knows that they have such a high return in the labour market that they do not need education, despite that fact that they have the ability to do it.[5] Positive sorting suggests that individuals who self-select into select dropout will earn a relatively high lifetime wages in jobs that do not require high school diploma, while they while they will also earn high lifetime wages in jobs that require high school diploma if they had chosen to complete high school education (Willis & Rosen, 1979b; Carneiro et al., 2011).

Zero correlation

If the measure of selection is not statistically different from zero, it means that I cannot establish any pattern of selection into dropout or graduation. This implies lack of any systematic difference between those who choose to dropout and those who choose to complete high school education.

3.5.2 Analytical expression

Let G and D by two (discrete) random variables representing graduates and dropouts respectively. Thus, I am interested in the correlation that exists between G and D. If I assume that all marginal distributions have identical mass points at the support, I can show (see Appendix C) that the expression for the correlation coefficient is given as

$$\text{Corr}(G, D) = \frac{p_{11} - p_1^G p_1^D}{[p_1^G(1 - p_1^G)p_1^D(1 - p_1^D)]^{0.5}} \tag{3.9}$$

where

$$p_1^G = p_{10} + p_{11}$$

$$p_1^D = p_{01} + p_{11}$$

The three parameters (p_{01}, p_{10}, p_{11}) required to compute the measure of selection are estimated in the competing risks model specified in Equation (3.8). In order

[5] For these individuals, education is considered as "bad" or "useless"

to test whether the estimated measure of selection is statistically significant, I use the delta method to compute the variance of the correlation coefficient (see Appendix D for derivation).

3.6 Estimation results

Empirical results from the competing risks specification in Equation (3.8) are summarized in Table 3.2. Estimates in columns 1 and 2 do not account for the role of unobserved individual heterogeneity. The assumption here is that only observable characteristics are relevant in predicting high school drop-out and graduation. The results in columns 3 and 4 are accounts for a univariate distribution of unobserved heterogeneity, while estimates in columns 5 and 6 allow for a generalization of unobserved heterogeneity.

The results indicate that being a male increases the likelihood of dropping out of high school, *ceteris paribus*. This makes sense in the context of a developing country like South Africa where male children are often called upon to take up menial jobs in order to supplement low family income. This result is consistent with the findings in Oreopoulos & Page (2006). Moving from the model without unobserved heterogeneity (column 2) to the model with a univariate distribution of unobserved differences (column 4) leads to substantial upward change in age effect, whereas generalizing unobserved heterogeneity (column 6) does change the age estimate significantly. However, I find no gender effect on graduation. That is, all things being equal, it will take the same amount of time for a man and a woman to graduate. Similarly, I find no location (urban or rural) or proximity-to-school effects across destination states.

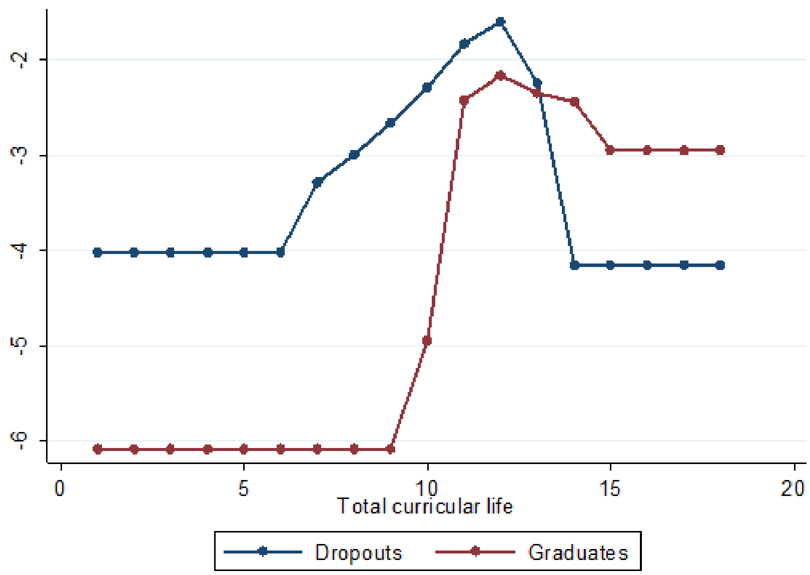
As expected mother's education and higher test scores reduce the likelihood of high school dropout. This is also consistent with previous studies in the literature (Eckstein & Wolpin, 1999; Li et al., 2004; Foley et al., 2014). On the other hand, individuals with high test scores and whose mothers are more educated take longer time to graduate. This seemingly counter-intuitive result could be as a result of higher extracurricular activity related to education. This is especially so if higher-educated mothers are also correlated with wealthier families. Comparing columns 2, 4, and 6 for dropouts shows that the inclusion of a univariate unobserved heterogeneity does significantly change the results, while the generalization of unobserved factors leads to minor changes in the estimates. The same is true for graduates when I compare results in columns 1, 3 and 5.Furthermore, teacher's availability reduces the likelihood of high school drop-out. As before, the introduction of a univariate distribution of unobserved heterogeneity leads to a substantial increase in the estimated coefficient, whereas a generalization of unobserved differences only cause a small change. However, I find no absenteeism effect on the duration until graduation. The time spent on household work and lack of help with homework both increase the likelihood of high school drop-out and the duration until graduation for those who have graduated. In addition, while reservation wage is positively associated with the likelihood of dropping out of high school, I find no effect on the likelihood of graduation.

Table 3.2: Competing Risks Model by Destination States

	No UH		Univariate Dist. of UH		General UH	
	Graduates	Dropouts	Graduates	Dropouts	Graduates	Dropouts
Sex (male)	0.110	0.284**	0.155	0.457**	0.182	0.460**
	(0.076)	(0.067)	(0.086)	(0.097)	(0.105)	(0.096)
Location (urban)	-0.168	-0.244**	-0.174	-0.252	-0.179**	-0.233**
	(0.107)	(0.084)	(0.115)	(0.154)	(0.029)	(0.110)
Mother's education	-0.043**	-0.050**	-0.054**	-0.111**	-0.060**	-0.116**
	(0.009)	(0.008)	(0.012)	(0.019)	(0.019)	(0.018)
Test scores	-0.030**	-0.081**	-0.042**	-0.113**	-0.045**	-0.110**
	(0.008)	(0.007)	(0.010)	(0.011)	(0.013)	(0.010)
Distance to school	0.000	0.004	-0.000	0.007	-0.000	0.006
	(0.002)	(0.002)	(0.002)	(0.003)	(0.000)	0.004
Teacher not absent	0.119	-0.250**	0.130	-0.377**	0.151	-0.388**
	(0.120)	(0.087)	(0.143)	(0.133)	(0.086)	(0.124)
Time on household work	0.021**	0.025**	0.021**	0.035**	0.022	0.033**
	(0.006)	(0.005)	(0.008)	(0.012)	(0.016)	(0.010)
Reservation wage	0.062	0.124**	0.063	0.126**	0.062	0.127**
	(0.041)	(0.033)	(0.051)	(0.057)	(0.036)	(0.035)
No help with homework	0.300**	0.353**	0.336**	0.505**	0.345	0.493**
	(0.100)	(0.094)	(0.111)	(0.116)	(0.193)	(0.125)
Unobserved component			1.448**		1.477**	
			(0.149)		(0.098)	
Probability Type 1			0.1067**		0.0381	
			(0.038)		(0.046)	
Probability Type 2			0.893**		0.043	
			(0.0385)		(0.043)	
Probability Type 3					0.147**	
					(0.073)	
Probability Type 4					0.770**	
					(0.066)	
Correlation coefficient					**0.215**	
					(**0.155**)	
Log-likelihood at maximum	-3414.87		-3402.23		-3401.09	
Sample size	2,877		2,877		2,877	

Notes: Standard errors are in parenthesis and ** represents statistical significance at 5%. UH: unobserved heterogeneity.

Figure 3.3: Estimated flexible baseline hazards



Overall, the inclusion of a univariate unobserved heterogeneity changes the estimated coefficients substantially, while generalization from a univariate distribution to a bivariate distribution of unobserved individual heterogeneity does not seem to add much.

In Figure 3.3, I plot the estimated flexible baseline hazards. The result indicates that the risk of dropping out of high school is higher than the risk of completing high school education in the first 12 years of schooling. The risks of dropping out and graduating are the same in the 13th year of education. Beyond this the 13th year, graduating from high school becomes more likely than dropping out. Therefore, conditioning on observed and unobserved individual unobserved heterogeneity, the drop out and graduation pattern is consistent with the pattern observed in the nonparametric hazard estimates (see Figure 3.1).

In order to assess the patterns of self-selection, I turn to the competing risks estimates in columns 5 and 6 of Table 3.2. Particular interest is in the estimated correlation coefficient (a measure of self-selection). Unlike the results in columns 3 and 4, the correlation coefficient incorporates the bivariate distribution of unobserved heterogeneity between dropouts and graduates. This allows me to check for systematic patterns of self-selection that may exist across destination states. The measure of selection is estimated as 0.215. Ignoring statistical significance for a moment, the estimated measure of self-selection suggests the presence of positive sorting. That is, individuals in my sample could either graduate faster or drop out faster based on unobserved factors. If an individual is "good" (e.g, high ability), they will be "good" in both dimensions, and if "bad" (e.g, low ability), they will be "bad" in both dimensions (see Section 3.5.1 for more discussion on this).

However, the estimated measure of self-selection is not statistically different from zero. This is important because it suggests that there is no systematic pattern of self-selection into high school dropout and graduation. That is, individuals who choose to drop out of high school are not systematically different from those who choose to complete high school education in terms of unobservables. It is typically found in developed country data that individuals who drop out of high school are those with relatively low ability, motivation, low expectations, and a set of negative preferences (Eckstein & Wolpin, 1999; Li et al., 2004; Foley et al., 2014). My results show that this is not necessarily the case in South Africa, a developing country in Sub-Saharan Africa. Individuals who drop out in my data do not seem to do so because they perceive or understand that they do not stand a chance in education. Rather, individuals who drop out of high school could potentially complete high school education like their counterparts with high school qualifications.

What can possibly explain this finding? South Africa as a country still lives with the legacy of unequal opportunity after many years of apartheid. Inequality of opportunity implies that individuals with disadvantaged background may drop out of school more because the monetary and opportunity cost of completing high school education is higher for them. Also, South Africa, like many other developing countries, lacks viable social protection programmes or other forms of government support that could keep disadvantaged youths in school until graduation. Competing needs for the little resources available to many poor households may not leave enough to support high school education. For example, the results in column 6 of Table 4.2 shows that low level of parental education (a variable that is positively correlated with family income) increases the likelihood of high school dropout. In addition, individuals with disadvantaged background are sometimes called upon to supplement low family income. This often implies dropping out of high school and taking up menial jobs.

This is not the kind of result one would expect to see in most developed countries where government support and social protection is relatively high. In such countries, we would expect to see some sort of negative sorting: individuals who dropout are not potential graduates. They are different, either in their motivation, ability or expectations.

3.7 Conclusion

This chapter assess the pattern of self-selection into high school drop-out and graduation using longitudinal data from South Africa. Using the structure of competing risks model, I quantify the nature of self-selection and establish to what extent selection matters. Results show that there is no systematic pattern of self-selection into high school dropout and graduation. That is, individuals who choose to drop out of high school are not systematically different from those who choose to complete high school education in terms of unobservables. Individuals who drop out of high school do not do so because they are weak and lack any future ambition. Rather, individuals who drop out of high school could potentially earn a high school diploma.

Therefore, my results show that there is scope for government intervention in order to address high school drop-out in South Africa, and possibly other Sub-Saharan African countries . One way to do this is through a well coordinated cash transfer targeted at individuals with disadvantaged background. There is evidence in the literature that supports the effectiveness of this kind of policy intervention Behrman et al. (2005); Glewwe & Kassouf (2012). In addition, Individuals may not be aware of the benefit of high school diploma in the labour market. Thus, the provision of accurate information may incentivize an individual to complete high school education.

The results in this chapter, to the best of my knowledge, provide the first evidence of self-selection into high school dropout and high school graduation in the literature. I use a very flexible and data-driven approach to quantify selection on unobservables in schooling choices and establishing to what extent selection matters. This approach requires no sundry assumptions for identification. Rather, it simply uses the structure of a well-known model to estimate a measure of self-selection and document an empirical fact.

References

Abbring, J. H., & Berg, G. J. V. D. (2003, August). The identifiability of the mixed proportional hazards competing risks model. *Journal of the Royal Statistical Society Series B*, *65*(3), 701-710.

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

Bedard, K. (2001, August). Human Capital versus Signaling Models: University Access and High School Dropouts. *Journal of Political Economy*, *109*(4), 749-775.

Behrman, J. R., Sengupta, P., & Todd, P. (2005, October). Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment in Rural Mexico. *Economic Development and Cultural Change*, *54*(1), 237-275.

Belzil, C. (2007, July). The return to schooling in structural dynamic models: a survey. *European Economic Review*, *51*(5), 1059-1105.

Carneiro, P., Heckman, J. J., & Vytlacil, E. J. (2011, October). Estimating Marginal Returns to Education. *American Economic Review*, *101*(6), 2754-2781.

Cawley, J., Heckman, J., & Vytlacil, E. (2001). Three observations on wages and measured cognitive ability. *Labour Economics*, *8*(4), 419-442.

Dolton, P., & van der Klaauw, W. (1999, August). The Turnover of Teachers: A Competing Risks Explanation. *The Review of Economics and Statistics*, *81*(3), 543-550.

Dominitz, J., & Manski, C. F. (1996, Winter). Eliciting Student Expectations of the Returns to Schooling. *Journal of Human Resources*, *31*(1), 1-26.

Eckstein, Z., & Wolpin, K. I. (1999, November). Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities. *Econometrica*, *67*(6), 1295-1340.

Foley, K., Gallipoli, G., & Green, D. A. (2014, September). Ability, Parental Valuation of Education, and the High School Dropout Decision. *Journal of Human Resources*, *49*(4), 906-944.

Glewwe, P., & Kassouf, A. L. (2012, March). The impact of the Bolsa Escola/Familia conditional cash transfer program on enrollment, dropout rates and grade promotion in Brazil. *Journal of Development Economics*, *97*(2), 505-517.

Goldin, C. (1999). Egalitarianism and the Returns to Education during the Great Transformation of American Education. *Journal of Political Economy*, *107*(S6), 65-94.

Guryan, J. (2004, September). Desegregation and Black Dropout Rates. *American Economic Review*, *94*(4), 919-943.

Heckman, & Honoré, B. E. (1989, June). The Identifiability of the Competing Risks Model. *Biometrika*, *76*(2), 325–330.

Heller, S. B., Shah, A. K., Guryan, J., Ludwig, J., Mullainathan, S., & Pollack, H. A. (2017, February). Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago. *The Quarterly Journal of Economics*, *132*(1), 1-54.

Hill, C. R. (1979). Capacities, opportunities and educational investments: The case of the high school dropout. *The Review of Economics and Statistics*, 9–20.

ILO. (2019). *Global unemployment data.* International Labour Organization.

Li, M., Poirier, D. J., & Tobias, J. L. (2004, March). Do dropouts suffer from dropping out? Estimation and prediction of outcome gains in generalized selection models. *Journal of Applied Econometrics*, *19*(2), 203-225.

Lindsay, C. M. (1971, December). Measuring Human Capital Returns. *Journal of Political Economy*, *79*(6), 1195-1215.

Lochner, L., & Moretti, E. (2004). The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports. *American Economic Review*, *94*(1), 155-189.

Montmarquette, C., Viennot-Briot, N., & Dagenais, M. (2007, November). Dropout, School Performance, and Working while in School. *The Review of Economics and Statistics*, *89*(4), 752-760.

Motala, S., Dieltiens, V., & Sayed, Y. (2009). Physical access to schooling in south africa: mapping dropout, repetition and age-grade progression in two districts. *Comparative Education*, *45*(2), 251–263.

Oreopoulos, P., & Page, M. E. (2006, October). The Intergenerational Effects of Compulsory Schooling. *Journal of Labor Economics*, *24*(4), 729-760.

Roy, A. D. (1951, June). Some Thoughts On The Distribution Of Earnings. *Oxford Economic Papers*, *3*(2), 135-146.

Soudien, C. (2004). Constituting the class': An analysis of the process of 'integration'in south african schools. *Changing class: Education and social change in post-apartheid South Africa*, 89–114.

Weybright, E. H., Caldwell, L. L., Xie, H., Wegner, L., & Smith, E. A. (2017). Predicting secondary school dropout among south african adolescents: A survival analysis approach. *South African journal of education*, *37*(2).

Willis, R. J., & Rosen, S. (1979a). Education and self-selection. *Journal of political Economy*, *87*(5, Part 2), S7–S36.

Willis, R. J., & Rosen, S. (1979b, October). Education and Self-Selection. *Journal of Political Economy*, *87*(5), 7-36.

Appendix C

Correlation coefficient derivation

Consider

$$Corr(G,D) = \frac{Cov(G,D)}{[(Var(G)Var(D))]^{0.5}}$$

so that

$$Corr(G,D) = \frac{E(GD) - E(G)E(D)}{[(Var(G)Var(D))]^{0.5}}$$

and

$$Corr(G,D) = \frac{E(GD) - E(G)E(D)}{[(E(G^2) - E(G)^2)(E(D^2) - E(D)^2)]^{0.5}}$$

Given the bivariate distribution of unobserved heterogeneity (2-points of support) below

$$\begin{bmatrix} G,D & \mu_0 & \mu_1 \\ \mu_0 & p_{00} & p_{01} \\ \mu_1 & p_{10} & p_{11} \end{bmatrix}$$

the marginal distribution for graduates can be represented as

$$\mu_0 :\ = p_{00} + p_{01}$$

$$\mu_1 :\ = p_{10} + p_{11}$$

and that of dropouts is given as

$$\mu_0 : = p_{00} + p_{10}$$

$$\mu_1 : = p_{01} + p_{11}$$

If I assume that all marginal distributions have identical mass points at the support, $Corr(G, D)$ can be presented in a more general form as follows:

$$Corr(G, D) = \frac{\sum_i \sum_j \mu_i \mu_j p_{ij} - \sum_i \mu_i p_i^G \sum_i \mu_i p_i^D}{[(\sum_i (\mu_i)^2 p_i^G - (\sum_i \mu_i p_i^G)^2)(\sum_i (\mu_i)^2 p_i^D - (\sum_i \mu_i p_i^D)^2)]^{0.5}}$$

Expanding, using the bivariate distribution of unobserved heterogeneity, I have that $Corr(G, D)$ is given as

$$\frac{(\mu_0\mu_0 p_{00} + \mu_0\mu_1 p_{01}) + (\mu_1\mu_0 p_{10} + \mu_1\mu_1 p_{11}) - (\mu_0 p_0^G + \mu_1 p_1^G)(\mu_0 p_0^D + \mu_1 p_1^D)}{\left[\left((\mu_0^2 p_0^G + \mu_1^2 p_1^G) - (\mu_0 p_0^G + \mu_1 p_1^G)^2\right)\left((\mu_0^2 p_0^D + \mu_1^2 p_1^D) - (\mu_0 p_0^D + \mu_1 p_1^D)^2\right)\right]^{0.5}}$$

Setting $\mu_0 = 0$ across destinations states (normalization assumption for identification purpose), I have that

$$Corr(G, D) = \frac{(\mu_1\mu_1 p_{11}) - (\mu_1 p_1^G)(\mu_1 p_1^D)}{\left[\left(\mu_1^2 p_1^G - (\mu_1 p_1^G)^2\right)\left(\mu_1^2 p_1^D - (\mu_1 p_1^D)^2\right)\right]^{0.5}}$$

Rearranging and collecting terms, I get

$$Corr(G, D) = \frac{\mu_1^2 (p_{11} - p_1^G p_1^D)}{\left[\mu_1^2\left(p_1^G - (p_1^G)^2\right)\mu_1^2\left(p_1^D - (p_1^D)^2\right)\right]^{0.5}}$$

And

$$Corr(G, D) = \frac{\mu_1^2 (p_{11} - p_1^G p_1^D)}{\mu_1^2\left[\left(p_1^G - (p_1^G)^2\right)\left(p_1^D - (p_1^D)^2\right)\right]^{0.5}}$$

Ultimately, the correlation coefficient between G and D writes as:

$$Corr(G, D) = \frac{p_{11} - p_1^G p_1^D}{[p_1^G(1 - p_1^G) p_1^D (1 - p_1^D)]^{0.5}}$$

Appendix D

Delta method and variance estimation

Recall the formula for correlation coefficient between dropouts and graduates is given as

$$r = \frac{p_{11} - p_1^G p_1^D}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{0.5}}$$

where $p_1^G = p_{10} + p_{11}$, and $p_1^D = p_{01} + p_{11}$.

- Rearrangement of the expression for r

I estimate a parameter vector $[p_{00}, p_{01}, p_{10}]$, so for the purpose of computing the confidence interval of $\hat{r}$ it is more convenient to consider the following substitution

$$p_{11} = 1 - (p_{00} + p_{01} + p_{10})$$

such that

$$\begin{aligned} p_1^G &= 1 - p_{00} - p_{01} \\ p_1^D &= 1 - p_{00} - p_{10} \end{aligned}$$

and

$$r = \frac{1 - (p_{00} + p_{01} + p_{10}) - p_1^G p_1^D}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{0.5}}.$$

This can be further rearranged as follows:
Consider

$$\begin{aligned}
p_1^G p_1^D &= (1 - p_{00} - p_{01})\, p_1^D = p_1^D - p_{00} p_1^D - p_{01} p_1^D \\
&= [1 - p_{00} - p_{10}] - p_{00} p_1^D - p_{01} p_1^D.
\end{aligned}$$

Inserting this into r I have

$$\begin{aligned}
r &= \frac{1 - (p_{00} + p_{01} + p_{10}) - \left\{[1 - p_{00} - p_{10}] - p_{00}p_1^D - p_{01}p_1^D\right\}}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{0.5}} \\
&= \frac{-p_{01} + p_{00}p_1^D + p_{01}p_1^D}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{0.5}} = \frac{p_{00}p_1^D - p_{01}\left[1 - p_1^D\right]}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{0.5}} \\
&= \frac{p_{00}\left[1 - 1 + p_1^D\right] - p_{01}\left[1 - p_1^D\right]}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{0.5}} = \frac{p_{00} - p_{00}\left(1 - p_1^D\right) - p_{01}\left[1 - p_1^D\right]}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{0.5}} \\
&= \frac{p_{00} - \left(1 - p_1^D\right)\{p_{00} + p_{01}\}}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{0.5}} = \frac{p_{00} - \left(1 - p_1^D\right)\left\{1 - p_1^G\right\}}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{0.5}} \\
&= \frac{p_{00}}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{\frac{1}{2}}} - \left[\frac{\left(1 - p_1^D\right)\left(1 - p_1^G\right)}{p_1^G p_1^D}\right]^{\frac{1}{2}}.
\end{aligned}$$

- Derivatives of r w.r.t p_{00}

  First

$$\begin{aligned}
&\frac{\partial}{\partial p_{00}}\left(\frac{p_{00}}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{\frac{1}{2}}}\right) \\
&= \frac{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{\frac{1}{2}} - p_{00}\frac{1}{2}[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{-\frac{1}{2}}\frac{\partial\left(p_1^G(1-p_1^G)p_1^D(1-p_1^D)\right)}{\partial p_{00}}}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{\frac{1}{2}}} \\
&= \frac{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{\frac{1}{2}}}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]} - \frac{p_{00}\frac{1}{2}[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{-\frac{1}{2}}\frac{\partial\left(p_1^G(1-p_1^G)p_1^D(1-p_1^D)\right)}{\partial p_{00}}}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]} \\
&= \frac{1}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{\frac{1}{2}}} - \frac{p_{00}}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{\frac{3}{2}}}\frac{1}{2}\frac{\partial\left(p_1^G(1-p_1^G)p_1^D(1-p_1^D)\right)}{\partial p_{00}} \\
&= \frac{p_{00}}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{\frac{3}{2}}}\left\{\frac{p_1^G(1-p_1^G)p_1^D(1-p_1^D)}{p_{00}} - \frac{1}{2}\frac{\partial\left(p_1^G(1-p_1^G)p_1^D(1-p_1^D)\right)}{\partial p_{00}}\right\}
\end{aligned}$$

and

$$\frac{\partial}{\partial p_{00}}\left(\left[\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right]^{\frac{1}{2}}\right)$$

$$= \frac{1}{2}\left[\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right]^{-\frac{1}{2}}\frac{\partial}{\partial p_{00}}\left(\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right)$$

$$= \frac{\left[p_1^G p_1^D\right]^{\frac{1}{2}}}{\left[\left(1-p_1^D\right)\left(1-p_1^G\right)\right]^{\frac{1}{2}}}\frac{1}{2}\frac{\partial}{\partial p_{00}}\left(\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right)$$

Now, since $r = \dfrac{p_{11} - p_1^G p_1^D}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{0.5}}$ both partial derivatives can be written more compactly as

$$\frac{\partial}{\partial p_{00}}\left(\frac{p_{00}}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{\frac{1}{2}}}\right)$$

$$= \frac{r^3 p_{00}}{\left(p_{11} - p_1^G p_1^D\right)^3}\left\{\frac{\left(p_{11} - p_1^G p_1^D\right)^2}{r^2 p_{00}} - \frac{1}{2}\frac{\partial\left(p_1^G(1-p_1^G)p_1^D(1-p_1^D)\right)}{\partial p_{00}}\right\}$$

and

$$\frac{\partial}{\partial p_{00}}\left(\left[\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right]^{\frac{1}{2}}\right)$$

$$= \frac{r[p_1^G p_1^D]}{p_{11} - p_1^G p_1^D}\frac{1}{2}\frac{\partial}{\partial p_{00}}\left(\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right)$$

so that

$$
\begin{aligned}
\frac{\partial r}{\partial p_{00}} &= \frac{r^3 p_{00}}{\left(p_{11} - p_1^G p_1^D\right)^3} \left\{ \frac{\left(p_{11} - p_1^G p_1^D\right)^2}{r^2 p_{00}} - \frac{1}{2} \frac{\partial \left(p_1^G (1-p_1^G) p_1^D (1-p_1^D)\right)}{\partial p_{00}} \right\} \\
&\quad - \frac{r[p_1^G p_1^D]}{p_{11} - p_1^G p_1^D} \frac{1}{2} \frac{\partial}{\partial p_{00}} \left(\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D} \right) \\
&= \frac{r}{p_{11} - p_1^G p_1^D} \left[1 - \frac{r^2 p_{00}}{2\left(p_{11} - p_1^G p_1^D\right)^2} \frac{\partial \left(p_1^G (1-p_1^G) p_1^D (1-p_1^D)\right)}{\partial p_{00}} \right. \\
&\quad \left. - \frac{p_1^G p_1^D}{2} \frac{\partial}{\partial p_{00}} \left(\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D} \right) \right]
\end{aligned}
$$

Consider first

$$
\begin{aligned}
& \frac{\partial \left(p_1^G (1-p_1^G) p_1^D (1-p_1^D)\right)}{\partial p_{00}} \\
&= \frac{\partial}{\partial p_{00}} \left((p_1^G - \left[p_1^G\right]^2)(p_1^D - \left[p_1^D\right]^2) \right) \\
&= (p_1^D - \left[p_1^D\right]^2) \frac{\partial}{\partial p_{00}} \left(p_1^G - \left[p_1^G\right]^2 \right) + (p_1^G - \left[p_1^G\right]^2) \frac{\partial}{\partial p_{00}} \left(p_1^D - \left[p_1^D\right]^2 \right) \\
&= (p_1^D - \left[p_1^D\right]^2) \left(2p_1^G - 1\right) + (p_1^G - \left[p_1^G\right]^2) \left(2p_1^D - 1\right) \\
&= p_1^D (1-p_1^D) \left(2p_1^G - 1\right) + p_1^G (1-p_1^G) \left(2p_1^D - 1\right) \\
&= (1-p_1^D)(1-p_1^G) \left[p_1^D \frac{2p_1^G - 1}{(1-p_1^G)} + p_1^G \frac{2p_1^D - 1}{(1-p_1^D)} \right] \\
&= \frac{\left(p_{11} - p_1^G p_1^D\right)^2}{r^2} \left[p_1^D \frac{2p_1^G - 1}{1-p_1^G} + p_1^G \frac{2p_1^D - 1}{1-p_1^D} \right]
\end{aligned}
$$

so that

$$\begin{aligned}
\frac{\partial r}{\partial p_{00}} &= \frac{r}{p_{11}-p_1^G p_1^D}\left[1-\frac{r^2 p_{00}}{2\left(p_{11}-p_1^G p_1^D\right)^2}\frac{\left(p_{11}-p_1^G p_1^D\right)^2}{r^2}\left[p_1^D\frac{2p_1^G-1}{1-p_1^G}+p_1^G\frac{2p_1^D-1}{1-p_1^D}\right]\right. \\
&\left.\quad -\frac{p_1^G p_1^D}{2}\frac{\partial}{\partial p_{00}}\left(\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right)\right] \\
&= \frac{r}{p_{11}-p_1^G p_1^D}\left[1-\frac{p_{00}}{2}\left[p_1^D\frac{2p_1^G-1}{1-p_1^G}+p_1^G\frac{2p_1^D-1}{1-p_1^D}\right]\right. \\
&\left.\quad -\frac{p_1^G p_1^D}{2}\frac{\partial}{\partial p_{00}}\left(\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right)\right]
\end{aligned}$$

Finally consider

$$\begin{aligned}
&\frac{\partial}{\partial p_{00}}\left(\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right) \\
&= \frac{p_1^G p_1^D \frac{\partial}{\partial p_{00}}\left(1-p_1^D-p_1^G+p_1^G p_1^D\right)-\left(1-p_1^D\right)\left(1-p_1^G\right)\frac{\partial}{\partial p_{00}}\left(p_1^G p_1^D\right)}{\left[p_1^G p_1^D\right]^2} \\
&= \frac{p_1^G p_1^D\left(2+\frac{\partial}{\partial p_{00}}\left(p_1^G p_1^D\right)\right)-\left(1-p_1^D\right)\left(1-p_1^G\right)\frac{\partial}{\partial p_{00}}\left(p_1^G p_1^D\right)}{\left[p_1^G p_1^D\right]^2} \\
&= \frac{2p_1^G p_1^D+\left[p_1^G p_1^D-\left(1-p_1^D-p_1^G+p_1^G p_1^D\right)\right]\frac{\partial}{\partial p_{00}}\left(p_1^G p_1^D\right)}{\left[p_1^G p_1^D\right]^2}
\end{aligned}$$

$$\begin{aligned}
&= \frac{2p_1^G p_1^D+\left[p_1^D+p_1^G-1\right]\frac{\partial}{\partial p_{00}}\left(p_1^G p_1^D\right)}{\left[p_1^G p_1^D\right]^2} \\
&= \frac{2p_1^G p_1^D+\left[p_1^D+p_1^G-1\right]\left\{-p_1^D-p_1^G\right\}}{\left[p_1^G p_1^D\right]^2} \\
&= \frac{2p_1^G p_1^D-\left[\left(p_1^D+p_1^G\right)^2-\left(p_1^D+p_1^G\right)\right]}{\left[p_1^G p_1^D\right]^2} \\
&= \frac{p_1^D-\left[p_1^D\right]^2+p_1^G-\left[p_1^G\right]^2}{\left[p_1^G p_1^D\right]^2} \\
&= \frac{p_1^D\left(1-p_1^D\right)+p_1^G\left(1-p_1^G\right)}{\left[p_1^G p_1^D\right]^2}
\end{aligned}$$

such that

$$
\begin{aligned}
\frac{\partial r}{\partial p_{00}} &= \frac{r}{p_{11}-p_1^G p_1^D}\left[1-\frac{p_{00}}{2}\left[p_1^D\frac{2p_1^G-1}{1-p_1^G}+p_1^G\frac{2p_1^D-1}{1-p_1^D}\right]-\frac{p_1^G p_1^D}{2}\frac{p_1^D\left(1-p_1^D\right)+p_1^G\left(1-p_1^G\right)}{[p_1^G p_1^D]^2}\right] \\
&= \frac{r}{p_{11}-p_1^G p_1^D}\left[1-\frac{p_{00}}{2}\left[p_1^D\frac{2p_1^G-1}{1-p_1^G}+p_1^G\frac{2p_1^D-1}{1-p_1^D}\right]-\frac{1}{2}\frac{p_1^D\left(1-p_1^D\right)+p_1^G\left(1-p_1^G\right)}{p_1^G p_1^D}\right] \\
&= \frac{r}{p_{11}-p_1^G p_1^D}\left[1-\frac{p_{00}}{2}\left[p_1^D\left(\frac{p_1^G-\left(1-p_1^G\right)}{1-p_1^G}\right)+p_1^G\frac{p_1^D-\left(1-p_1^D\right)}{1-p_1^D}\right]\right. \\
&\quad \left.-\frac{1}{2}\left[\frac{p_1^D\left(1-p_1^D\right)}{p_1^G p_1^D}+\frac{p_1^G\left(1-p_1^G\right)}{p_1^G p_1^D}\right]\right] \\
&= \frac{r}{p_{11}-p_1^G p_1^D}\left[1-\frac{p_{00}}{2}\left[p_1^D\left(\frac{p_1^G}{1-p_1^G}-1\right)+p_1^G\left(\frac{p_1^D}{1-p_1^D}-1\right)\right]\right. \\
&\quad \left.-\frac{1}{2}\left[\frac{1-p_1^D}{p_1^G}+\frac{1-p_1^G}{p_1^D}\right]\right]
\end{aligned}
$$

- Derivatives of r w.r.t p_{01}

First

$$
\begin{aligned}
&\frac{\partial}{\partial p_{01}}\left(\frac{p_{00}}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{\frac{1}{2}}}\right) \\
&= -\frac{p_{00}p_1^D(1-p_1^D)}{2[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{\frac{3}{2}}}\frac{\partial}{\partial p_{01}}\left(p_1^G-\left[p_1^G\right]^2\right) \\
&= -\frac{p_{00}p_1^D(1-p_1^D)}{2[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{\frac{3}{2}}}\left(2p_1^G-1\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{p_{00}}{[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]^{\frac{1}{2}}}\frac{p_1^D(1-p_1^D)}{2[p_1^G(1-p_1^G)p_1^D(1-p_1^D)]}\left(2p_1^G-1\right) \\
&= \left(r+\left[\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right]^{\frac{1}{2}}\right)\frac{2p_1^G-1}{2p_1^G(1-p_1^G)}
\end{aligned}
$$

and

$$\begin{aligned}
&\frac{\partial}{\partial p_{01}}\left(\left[\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right]^{\frac{1}{2}}\right)\\
&= \frac{1}{2}\left[\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right]^{-\frac{1}{2}}\frac{1-p_1^D}{p_1^D}\frac{\partial}{\partial p_{01}}\left(\frac{1-p_1^G}{p_1^G}\right)\\
&= \frac{1}{2}\left[\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right]^{-\frac{1}{2}}\frac{1-p_1^D}{p_1^D}\left(-\frac{1}{[p_1^G]^2}\right)\\
&= -\frac{1}{2}\frac{1}{\left[p_1^D\left(1-p_1^D\right)p_1^G\left(1-p_1^G\right)\right]^{\frac{1}{2}}}\frac{1-p_1^D}{p_1^G}\\
&= -\frac{1}{2}\frac{r}{p_{11}-p_1^G p_1^D}\frac{1-p_1^D}{p_1^G}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{\partial r}{\partial p_{01}} &= \left(r+\left[\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right]^{\frac{1}{2}}\right)\frac{2p_1^G-1}{2p_1^G(1-p_1^G)}+\frac{1}{2}\frac{r}{p_{11}-p_1^G p_1^D}\frac{1-p_1^D}{p_1^G}\\
&= \frac{r}{2p_1^G}\left[\left(1+\frac{1}{r}\left[\frac{\left(1-p_1^D\right)\left(1-p_1^G\right)}{p_1^G p_1^D}\right]^{\frac{1}{2}}\right)\left[\frac{p_1^G}{1-p_1^G}-1\right]+\frac{1-p_1^D}{p_{11}-p_1^G p_1^D}\right]\\
&= \frac{r}{2p_1^G}\left[\left(\frac{(1-p_1^G)(1-p_1^D)}{p_{11}-p_1^G p_1^D}+1\right)\left[\frac{p_1^G}{1-p_1^G}-1\right]+\frac{1-p_1^D}{p_{11}-p_1^G p_1^D}\right]\\
&= \frac{r}{2p_1^G}\left[\left(\frac{1-p_1^G-p_1^D+p_{11}}{p_{11}-p_1^G p_1^D}\right)\left[\frac{p_1^G}{1-p_1^G}-1\right]+\frac{1-p_1^D}{p_{11}-p_1^G p_1^D}\right]\\
&= \frac{r}{2p_1^G}\left[\left(\frac{p_{00}+p_{01}+p_{11}-p_1^D}{p_{11}-p_1^G p_1^D}\right)\left[\frac{p_1^G}{1-p_1^G}-1\right]+\frac{1-p_1^D}{p_{11}-p_1^G p_1^D}\right]\\
&= \frac{r}{2p_1^G}\left[\left(\frac{1-p_{10}-[1-p_{00}-p_{10}]}{p_{11}-p_1^G p_1^D}\right)\left[\frac{p_1^G}{1-p_1^G}-1\right]+\frac{1-p_1^D}{p_{11}-p_1^G p_1^D}\right]\\
\tfrac{\partial r}{\partial p_{01}} &= \frac{r}{2p_1^G\left(p_{11}-p_1^G p_1^D\right)}\left[p_{00}\left[\frac{p_1^G}{1-p_1^G}-1\right]+\left(1-p_1^D\right)\right].
\end{aligned}$$

Derivative of r with respect to p_{10} is analogous to that of p_{01}. The only difference is the indexing. With these partial deviates, I compute the variance of the correlation coefficient using the values of $p_{00}, p_{01}, p_{10}, p_{11}$ obtained from the competing risks model specified in Equation (3.8).

Chapter 4

High school dropout and its wage consequences

4.1 Introduction

Every youth enrolled in high school is faced with an economic decision: to either continue in education or terminate their educational investment before completing high school. This is an important decision because young people who choose to drop out do no just lose the opportunity to get education today, they may also deprive themselves of further training in the future (Cameron & Heckman, 2001). Thus, individuals who drop out of high school may be disadvantaged in the labour market in terms of earnings. However, little is known about the severity and time pattern of this disadvantage, especially in a developing country context. My aim in this chapter is to examine the immediate and future wage consequences of high school drop out in South Africa. This kind of knowledge is important because incomplete information about the worth of high school diploma in the labour market could influence the decision to drop out of high school (Hill, 1979).

Although high school dropouts have acquired more educational training than those who terminated their schooling after primary education, failure to stay in high school until graduation could make all the difference in terms of future earnings in the labour market. The difference in outcomes is particularly likely in developing countries due to relatively low level of skilled labour force. Another thing that is different about developing countries is the high monetary and opportunity cost of staying in school until graduation. Young people are often called upon to supplement low household income. This often means dropping out of school and taking up menial jobs.

One key issue in estimating the wage consequences of high school dropout is the possibility of self-selection; both in the drop out decision and the decision to participate in the labour market. First, the decision to drop out may not be random. Individuals who choose to drop out of high school may be systematically different from those who choose to complete high school education in terms of unobservables. For example, using a United States longitudinal data, Eckstein & Wolpin (1999) find that individuals who choose to drop out of high school have lower academic ability, have lower motivation, and have lower expectations about returns in the labour market. Therefore, simply comparing the labour market earnings for dropouts with that of high school graduates may not necessarily show how high school dropouts would have performed in the labour market if they had completed high school education or how high school graduates would have performed if they had dropped out (Bjerk, 2012).

Secondly, the decision to participate in the labour market or not may not be random for both high school dropouts and graduates. For example, individuals who choose not to participate in the labour market may be systematically different from those who participate. Since we do not have any information on earnings for non-participants, the sample used in the wage equation becomes a non-random or a selected sample. Therefore, the observed earnings for participants my be overstated or understated depending on the nature of selection.

However, it is important not to assume selection in a given data as it's presence or absence could be context-specific. In order to ascertain whether high school dropouts are systematically different from graduates in terms of unobservables, I estimate a competing risks model with a discrete bivariate distribution of unobserved heterogeneity.[1] To ascertain whether non-participants in the labour force are systematically different from participants in my data, I estimate a battery of the well-known Heckman selection model. I find no systematic pattern of self-selection in the decision to drop out of high school and in the decision to participate in the labour market. Empirical results from different specifications show that dropouts earn less in the first year of labour market experience, and their wages progressively declined over time.

4.2 Literature

There are few studies in the literature that specifically examined the wage consequences of high school dropout. One of the earliest of such studies is Hill (1979). Using the 1966 National Longitudinal Survey of Young Men (NLSYM), the author estimates a recursive four equation model. The ordinary least squared (OLS) technique is applied to each equation. Empirical results show that high school dropouts earn less than those with high school diploma in the time periods immediately following the completion of high school education.

Blakemore & Low (1984) also investigated high school dropout decision and its wage consequences in the United States using the the National Longitudinal Study of the High School Class of 1972 (NLSHS). The authors modelled high school drop out decision as interdependent events that involve both the dropout decision and expected wage in the labour market and also account for possible selection bias in the drop out decision. They find, unlike Hill (1979), that high school dropouts earn more than graduates in the labour market. One possible explanation that has been offered in the literature for this kind of unexpected positive relationship between dropping out and wage is the possibility that dropouts attained additional work experience.

[1] The method I have proposed to check for self-selection in the decision to drop out of high school is similar to the Heckman selection model in principle. The bivariate distribution of unobserved heterogeneity in my approach is not fundamentally different from the bivariate distribution of errors in the Heckman model. The main difference is that while the model I used has a discrete bivariate distribution with 2 points of support, Heckman model has a continuous bivariate normal distribution.

In his paper, Oreopoulos (2006) exploited the history of high school drop out rates in the United Kingdom to estimate the local average treatment effect (LATE) of high school education. The author took advantage of the change in minimum high school leaving age from 14 to 15 in 1947 to achieve identification. He use both instrumental variable and regression discontinuity techniques for estimations. Empirical result show that high school dropouts earned considerably less than high school graduates in the labour market.

Using the Canadian Youth in Transition Survey (YITS), Campolieti et al. (2010) estimated the effect dropping out of high school has on different outcomes, including wage. They employed the two-stage least squared (2SLS) estimation approach in order to adjust for the endogeneity problem that may arise due to selection bias in the drop out decision. The authors find that dropouts have poorer wages and employment outcomes, and that they do not make up for their lack of education through additional skill acquisition and training.

Bjerk (2012) examined the impact of dropping out of high school on labour market and criminal outcomes. The author explored for important heterogeneities within dropouts that may correlates with their future outcomes. The essence of this is to see whether the reason an individual gave for dropping out of high school is correlated with labour market and criminal outcomes many years after dropping out. Using the National Longitudinal Survey of Youth 1997 (NLSY97), he find that individuals who drop out because they feel "pushed" out of high school earn lower wages than otherwise similar high school graduates. However, those who drop out of high school because they feel "pulled" out do not do worse in their labour market earnings compared to otherwise similar high school completers.

4.2.1 Summary

Overall, there seems to be paucity of empirical studies that specifically examined the wage consequences of dropping out of high school and the few studies that exist are mostly in the context of developed countries: the United States, the United

Kingdom and Canada. In addition, many of the existing studies are analyzed in a static setting; they do not tell us anything about the time pattern of the disadvantage dropouts face in the labour market. This chapter adds to the existing evidence in the literature, especially in a developing country context. The rest of the chapter is outlined as follows: section 4.3 presents the data and context, the theory of human capital is discussed in section 4.4, methods of estimation are described in section 4.5, section 4.6 discusses empirical results, while section 4.7 concludes.

4.3 Data and context

Like the previous chapters, this chapter is based on the Cape Area Panel Study (CAPS). The CAPS is a five-wave longitudinal study of the lives of young people in Cape Town, South Africa. The CAPS is one of the few data in Sub-Saharan African region designed to follow the lives of youths and young adults over a considerable time. The study's first wave collected interviews in August-December 2002 from about 4,800 randomly selected youth aged 14-22. Wave 1 also collected information on all members of the households of these young people, as well as a random sample of households not having members between 14 and 22 years of age. A third of the youth sample was re-interviewed in 2003 (wave 2a) and re-visited the remaining two-thirds in 2004 (wave 2b). In both 2005 (wave 3) and 2006 (wave 4), the full youth sample was then re-interviewed. For wave 5, full face-to-face interviews were carried out in 2009 with the sample comprising all respondents interviewed in any of Waves 2a, 3 or 4.

The sample for analysis compromises of individuals that were enrolled in high school at the beginning of the survey (2002). These individuals are then followed over time until they either dropout or graduate. Drop out is defined as leaving high school without completing the final grade (i.e., grade 12 in South Africa). Individuals who were still enrolled in high school at the end of the survey (i.e., the *survivals*) are right-censored. These individuals are retained in the sample in order to avoid possible selection bias.

South Africa provides a good context for analysing high school drop out among young people and its wage consequences. This is because the social and economic costs associated with high school dropout are particularly severe in South Africa where high school drop out has reached a point of national crisis. Approximately 60% of first graders will ultimately drop out rather than complete 12th Grade. Likewise, by Grade 12, only 52% of the age appropriate population remain enrolled (Weybright et al., 2017). In addition, South Africa currently has the highest rate of youth unemployment (percent of total labor force ages 15-24) in the world (ILO, 2019).

4.4 Theory of human capital investment

There are plethora of studies that have examined the relationship between years of schooling (often used as a measure of human capital investment and) and earnings in the labour market. However, if the decision to invest in education is not random, but a result of an optimizing behaviour, a typical earning regression may overstate the the contribution of schooling to earnings (Hill, 1979).

Becker (1964) developed a simple theoretical model that explains interpersonal differences in educational investment and returns.[2] Becker termed this model "optimal investment in human capital". Basically, the model sees an average individual as maximizing the present value of his net earnings over the life cycle by investing in formal education up to the point at which the marginal rate of return from the investment is the same as the cost of the investment. This can be expressed more formally in a standard demand and supply framework. The demand for schooling can be explained by two factors: the expectation of the worth of a particular level of schooling in the labour market and the probability that the individual will actually achieve that level of education (Griliches, 1975).

The first factor (i.e., the expectation of returns) is determined by the exogenous forces of the labour market to a great extent and individual differences arise due

[2]More formal extensions have been provided by other authors over the years.

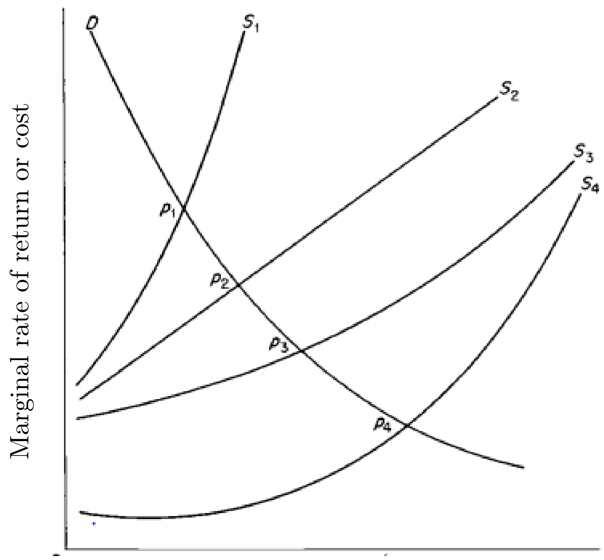
to imperfect knowledge. The second factor (i,e., the probability of success in education) is largely determined by individual ability (capacities), the schooling environment and the extent to which the individual believes the school curriculum will actually led to an increase in human capital. On the other hand, the supply side of the model captures environmental factors and investment opportunities which are mostly determined by the availability of funds (Hill, 1979).

In essence, Becker's model implies that the total amount invested in human capital and earnings differ among individuals because of differences in either demand or supply conditions. In what follows, I discuss two special cases of Becker's model and how they both relate to each other

4.4.1 Egalitarian model

Here, Becker's model assume that demand conditions are the same for everyone. Thus, the differences observed in human capital investment and earnings are only due to supply conditions. That is, every individual has the same ability or capacity to benefit from human capital investment. Investment in human capital and earnings differ because of differences in luck, subsidies, family wealth and other socioeconomic characteristics which give some individuals the opportunity to invest more than others. The idea here is that eliminating these differences would remove the differences in opportunities, which in turn would eliminate the differences in human capital investment and earnings. In essence, the model assumes that most individuals are intrinsically equally capable of high school or college education; only poverty, ignorance, and prejudice prevents some from doing so. Becker noted that availability of fund is the most important cause of differences in opportunity. Accessibility to funds differ for individuals for different reasons. This may be due to the area or country where one lives, the kind of family one is born into, political contact, etc.

Figure 4.1: Equilibrium level of human capital investment at different opportunities



Source: Becker (1964)

If demand remains the same (ability held constant), and only supply conditions are allowed to vary, Figure 4.1 shows that the equilibrium level of human capital investment would be by the intersection of the different supply curves with the given demand curve. These points are represented by p_1, p_2, p_3, and p_4 in Figure 4.1. Becker noted that full knowledge of these points, combined with the marginal rate of return of each amount of human capital invested would help to identify the common demand curve. Individuals that have the means (i.e., favourable supply conditions) will invest relatively more in human capital (the more the equilibrium level). Becker's model relates earnings to the amount of capital invested by the average rate of return of on investment. Thus, the distribution of earnings depends on the the distribution of capital investment. If the demand curve for human capital was to be horizontal, then the average rate of return on investment would be the same for each individual, and earnings and investment will have the same distribution.[3] See Becker (1964) for a more detailed discussion.

[3]The only difference will be a difference in the units of the average rate of return that depended on the aggregate supply of and demand for human capital.

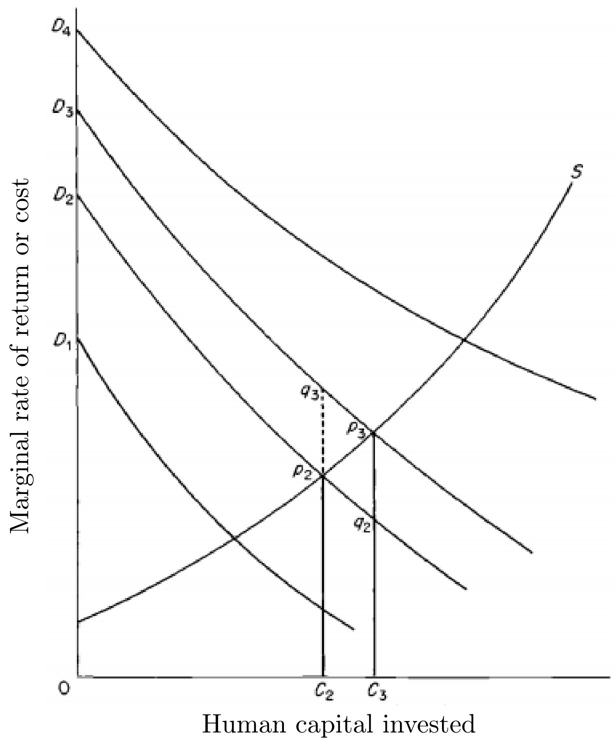
4.4.2 Elite model

This approach is the other end of the egalitarian approach. It assumes that supply conditions are the same for everyone and only the demand conditions vary. Put differently, every individual has more or less effectively the same opportunity to invest in human capital and get a good return in the labour market. Any observed difference in human capital investment and earnings occur because of differences in the capacity or ability to benefit from human investment in human capital. Thus, individuals with high capacity form the "elite" of the society.

As noted before, capacity or ability is measured in Becker's model primarily by demand curve, and supply curves measure opportunities. Thus, for a given amount of money invested in human capital, an individual with a higher demand curve receive higher returns than others. Given that a higher demand curve implies more returns from a given investment, in effect, ability is being measured indirectly in essence. This is reflected in the earnings received when investment in human capital is held constant. Becker argued that represents a compromise on how ability is defined or presented. Specifically, it is a compromise between the definition of ability in terms tests scores (IQ, personality and motivation), but disregarding the effect on earnings in the labour market and the definition in terns of earnings that does not take opportunities into account.

Figure 4.2 shows that if the supply curve (i.e, opportunity) remains the same, and only the demand curve (i.e., ability or capacity) is allowed to vary, the marginal rate of return to each individual's investment in human capital (equilibrium point) will be at the intersections of the different demand curves with the singular supply curve. Figure 4.2 also shows a positive relationship between the height of a demand curve (i.e., the amount of capital invested by an average individual) and the marginal rate of return. The knowledge of capital investments and the associated marginal rate of returns for different individuals help helps to identify the common supply curve. The intuition is akin to that of Figure 4.1.

Figure 4.2: Equilibrium level of human capital investment
at different abilities



Source: Becker (1964)

The model in Figure 4.2 implies that the marginal rate of return for every level of human capital investment vary by ability. To see this, consider the marginal rate of return to investing $0C_3$ instead of $0C_2$. For an individual with the capacity or ability level D_2, this would be proportional to the area represented by $p_1C_2C_3q_2$. For an individual with the ability level D_3, the marginal rate of return would correspond to the larger area represented by $q_3C_2C_{3p}3$. However, if we had simply estimated the marginal rate of return by taking the difference in earnings between individuals investing $0C_2$ and $0C_3$ respectively, the estimate would have been proportional to the area represented by $D_2p_2C_2C_3p_3D_3$, which exceeds both true estimates. Becker noted that correct estimates could be obtained by either adjusting upward the earnings of the individual investing $0C_2$ by the area $D_3D_2q_2p_3$, or by adjusting downward the earnings of the individual investing $0C_3$ by the area $D_3D_2q_2p_3$. See Becker (1964) for discussions on elasticity and distributions.

4.4.3 A comparison of both models

One of the key implications of Becker's "egalitarian" model or approach is that the higher the amount of human capital invested, the lower will be the marginal rate of return. This suggests a negative relationship between capital investments and earnings (see Figure 4.1). This outcome is a direct opposite of what is obtainable in the "elite" approach; the higher the amount of human capital invested, the higher the marginal rate of return. This implies a positive relationship between capital investments and earnings (see Figure 4.2).

As schooling increases, the returns to schooling tend to decrease in some countries, albeit marginally. This is somewhat consistent with Becker's "egalitarian" model. For example, using global data, the returns associated with an extra year of schooling are shown to have declined significantly over time. However, the decrease in returns is at a more slower rate than has occurred with the expansion in education, especially since the late 1990s (Patrinos, 2016). Schooling has expanded by almost 50 percent since 1980. Montenegro & Patrinos (2014) noted that over a 30 year period the returns to education have declined by 3.5 percentage points, or 0.1 percent a year. Within the same period, schooling increased by 2 percent a year. On average, another year of schooling causes a reduction of the returns to schooling by one percentage point.

Another basis of comparison is in how both model explained observed differences in returns to human capital investment. To explain inequality in earnings, the "egalitarian" model presume or emphasises greater inequality in opportunity, while the "elite" model emphasises greater inequality in ability or capacity. Inequality in earnings seems to be more of a problem under the ""egalitarian"" model than in the "elite" model. In addition, inequality of opportunity, which is the key feature of the "'egalitarian" model, tends to be more of a problem in developing countries than in developing countries on the average. This may be explained by the weakness of State's institutions that allows corruption to thrives as well as the lack of social protection programmes. In such countries, individuals with the required ability or capacity to invest in human capital may not engage in such investment because of lack of opportunity. On the other hand, the equality of opportunity assumed by the “elite" model may find more relevance in developed countries, especially Scandinavian European Coventries. Although the opportunity to invest in human capital may be relatively the same, lack of capacity may may hinder human capital investment.

4.5 Method of analysis

4.5.1 Selection into high school drop out and graduation

To check for possible selection bias in the decision to drop out of high school in my data, I write down a flexible discrete-time competing risks model with a discrete bivariate distribution of unobserved heterogeneity, and allow for a general correlated risks across destinations. I treat individual differences that may exist between dropouts and graduates as both observed and unobserved and then use the structure of my model to estimate a reduced-form measure of self-selection into dropout and graduation. The results show that the estimated measure of self-selection is not statistically different from zero. This suggests that there is no systematic pattern of self-selection into dropout and graduation in my data. Put differently, individuals who choose to drop out may not be systematically different from high school graduates in terms of unobservables. See section 3.4 for econometric specification and section 3.6 for result.

4.5.2 Selection into the labour market

To check if there is self-selection in the decision to participate in the labour market, I follow the method outlined in Heckman (1979). To formalize notations, consider a labour market participation equation defined as

$$z_i^* = \mathbf{w}_i\gamma + \epsilon_i \tag{4.1}$$

where

$$z_i = \begin{cases} 1 & \text{if } z_i^* > 0 \\ 0 & \text{if } z_i^* < 0 \end{cases}$$

and z_i^* represents a latent variable measuring the underlying propensity to participate in the labour market. Thus, the wage equation can be represented as

$$y_i = \begin{cases} y_i = \mathbf{x}_i\beta + \mu_i & \text{if } z_i^* > 0 \\ - & \text{if } z_i^* < 0 \end{cases} \tag{4.2}$$

where y_i is log hourly wage and $\mathbf{x}_i$ is observed characteristics including a measure of education. The variables in $\mathbf{w}_i$ and $\mathbf{x}_i$ are identical except for an additional variable in w_i that satisfies the exclusion restriction. That is, the additional variable should affect the outcome variable in the selection equation, but not the wage equation.

In my case, I use job referral as an additional variable in the selection equation. Job referral is defined in my data as when family and friends refer or helped individuals in my sample to get a job at their place of work. This recommendation or referral could determine or influence eventual employment in the labour market. Thus, job referral as a variable belongs to the selection or first stage regression. On the other hand, job referral could be legitimately excluded from the second stage or wage regression as the amount an individual earn in the labour market is not directly determined by whether they were referred for the job or not. Employers have established criteria for determining wage like the level of education, experience, productivity, etc.

Equation (4.2) says that we only observe wage if and only if individuals participate in the labour market. Thus, if the non-participants are systematically different from participants, we may understate or overstate the true returns to high school diploma. The Heckman selection model typically makes the following assumption about the distribution of the error terms in the selection and wage equations:

$$\epsilon_i \sim N(0,1)$$
$$\mu_i \sim N(0,\sigma^2)$$
$$\text{corr}(\mu_i,\epsilon_i) = \rho$$

A statistically significant ρ indicates the presence of self-selection. I estimate a battery of Equations (4.1) and (4.2) using the maximum likelihood estimation approach. The results presented in Table **??** (Appendix **??**) show a statistically insignificant ρ across specifications. This suggests the absence of self-selection in the decision to participate in the labour market.

4.5.3 Empirical wage equation

Although there is no self-selection into high school dropout and labour market participation in my data (see sections 4.5.1 and 4.5.2), the estimation of the returns to education in Equation (4.2) may still be biased. To see this, consider a simple two-equation system that describes schooling (S_i) and and log wages (y_i) for individual i:

$$S_i = \mathbf{x}_i\beta + \xi_i \tag{4.3}$$

$$y_i = \mathbf{x}_i\beta + S_i\delta + \upsilon_i \tag{4.4}$$

Where $\mathbf{x}_i$ is a vector of observed characteristics, and δ captures the true average return to education. There are variety of reasons why the measure of schooling may be correlated with the unmeasured or unobserved component in Equation (4.4).[4] One that has received considerable attention in the literature is ability bias. See for examples Griliches (1977) and (Card, 2001). Let us assume some individuals in my sample have innate ability that helps them to earn higher wages in the labour market at different level of education. If these individuals acquire a higher-than-average schooling, then OLS estimates of δ will be biased upward (Card, 1995).[5] Thus, ability represents a missing variable in Equation (4.4). Consequently, the wage equation can be correctly represented as

$$y_i = \mathbf{x}_i\beta + S_i\delta + A_i\alpha + \eta_i \tag{4.5}$$

[4]Measurement error is another possible reason

[5]This bias is based on the assumption that ability has an independent positive effect on earnings above and beyond the amount of schooling accumulated and that the relationship between the excluded ability and included measure of schooling is positive (Griliches, 1977). Measurement errors will likely attenuate this upward bias.

Where A_i is a measure of unobserved ability. The endogeneity problem we face in estimating the true average return to schooling is simply a consequence of missing A_i, which is a classical omitted variable problem.

Thousands of studies have been carried in the literature in a bid to adequately capture omitted ability and estimate the true returns to education. Two approaches have been predominately used in the literature. The first approach is to find a measure of innate ability and plug it into the wage equation (Griliches & Mason, 1972; Griliches, 1977; Blackburn & Neumark, 1993; Blundell et al., 2000; Cawley et al., 2001). This seems to be an obvious way of dealing with the problem. However, this approach has been criticized on the ground that it is difficult to get a variable or a set of variables that completely capture unobserved ability.

The second approach is the use of instrumental variable or variables to identify the true returns to education (Angrist & Keueger, 1991; Kling, 2001; Chen, 2008; Jensen, 2010). The idea is to get an instrument or a sets of instrument that affects schooling, but not earnings (relevance and "excludability" conditions). This approach has also been criticized on that ground that it is very difficult to get an instrument or a sets of instruments that could be legitimately excluded from the wage equation.

In this study, I use both approaches in estimating the wage consequences of high school dropout. As I measure of omitted ability, I used measures of IQ tests that were administered to individuals in my sample at the beginning of the survey in 2002. Even though the measures I used may not completely capture innate ability, they should provide at least a lower bound for the severity of ability bias. For the instrumental variable estimation, I use mother's education as instrument for schooling. It is plausible to argue that mother's education is correlated with children's education. This is evident in my data (see first stage regression in Table F.1). Also, mother's education is not expected to have a direct effect on wage as the labour market does not directly pay for parental education. Thus, I can legitimately exclude this variable from the wage equation. As a robustness check, I use the interaction of high school proximity and mother's education as an additional instrumental variable.

4.6 Empirical results

I start by presenting raw differences in wages between dropouts and graduates in Table 4.1. The results show that dropouts consistently earn less than counterparts with high school diploma. The observed premium associated with high school diploma increased progressively between 2005 and 2009, suggesting that high school dropouts do relatively worse-off over time. However, these estimates do not control for individual and family characteristics that could affect earnings in the labour market. They also do not account for possible differences in unobserved factors.

Table 4.1: Raw differences in wages

	Graduates	Dropouts	Premium	N
First-year (2005)	1824	1373	451***	511
Second-year (2006)	2380	1687	692***	552
Fourth-year (2009)	3891	2283	1608***	508

Note: Wages are in South Africa Rand.

*** represents statistical significance at 1% level

In Table 4.2, I present OLS estimates that control for individual and family characteristics, but ignore unobserved ability. Again, results show that individuals with high school diploma consistently have significantly higher wages than dropouts. For example, high school dropouts earned about 20% less than high school graduates in the first year of labour market experience. A follow-up in the next year show an increase in the wage premium by about 3 percentage points. Four years after dropping out of school (second follow-up), the premium associated with high school diploma increased to about 41% (i.e., a 15 percentage points increment between 2005 and 2009).

Since OLS results in Table 4.2 do not account for the possible role of unobserved ability, I present estimates that incorporate measures of ability in Table 4.3.

Table 4.2: OLS Estimates

	(1) First-year	(2) Second-year	(3) Fourth-year
Education (dropouts)	-0.216***	-0.251***	-0.411***
	(0.069)	(0.063)	(0.062)
Sex (male)	0.295***	0.237***	0.112*
	(0.067)	(0.061)	(0.060)
Age	0.013	-0.002	-0.045**
	(0.020)	(0.018)	(0.018)
Area (urban)	0.211**	0.205**	0.303***
	(0.097)	(0.086)	(0.091)
Group (black)	0.051	-0.253*	-0.446***
	(0.132)	(0.134)	(0.171)
Experience	0.196***	0.226**	0.126
	(0.072)	(0.094)	(0.173)
Experience^2	-0.068***	-0.038**	-0.008
	(0.017)	(0.015)	(0.014)
Number of hours worked	0.010	0.060***	0.063***
	(0.019)	(0.018)	(0.022)
Constant	6.502***	6.832***	8.360***
	(0.469)	(0.460)	(0.736)
Observations	464	356	303

Notes: Standard errors are in parentheses

Measures of ability are literacy and numeracy test scores administered at the start of the CAPS survey in 2002. Like the OLS results in Table 4.2, these results indicate that high school dropouts earn less than comparable high school graduates in the first year of labour market experience. Also, the earnings of high school dropouts progressively declined over time. However, accounting for innate ability consistently reduces the wage premium associated with high school diploma. This show that OLS estimates indeed overstates the true returns to high

Table 4.3: OLS estimates with measures of ability

	(1) First-year	(2) Second-year	(3) Fourth-year
Education (dropouts)	-0.162**	-0.174**	-0.253***
	(0.078)	(0.073)	(0.069)
Sex (male)	0.288***	0.234***	0.100*
	(0.067)	(0.061)	(0.058)
Age	0.011	-0.002	-0.044***
	(0.020)	(0.018)	(0.017)
Area (urban)	0.177*	0.154*	0.215**
	(0.099)	(0.089)	(0.090)
Group (black)	0.105	-0.169	-0.279
	(0.137)	(0.138)	(0.169)
Experience	0.184**	0.208**	0.063
	(0.072)	(0.093)	(0.168)
Experience^2	-0.067***	-0.037**	-0.004
	(0.017)	(0.015)	(0.014)
Numbers of hours work	0.015	0.063***	0.054**
	(0.019)	(0.018)	(0.022)
IQ test	0.008	0.011**	0.021***
	(0.005)	(0.005)	(0.005)
Constant	6.251***	6.470***	7.912***
	(0.500)	(0.488)	(0.718)
Observations	464	356	303

Notes: Literacy and numeracy tests scores administered at the start of the survey are used as measures of ability. Standard errors are in parentheses.

school education (i.e., OLS estimates are upward biased). Table 4.4 presents results from instrumental variable (IV) estimation. The results follow the same pattern like the estimates in Tables 4.2 and 4.3: high school dropouts earn less than high school graduates in the first year of labour market experience and the

Table 4.4: IV Estimates

	(1) First-year	(2) Second-year	(3) Fourth-year
Education (dropouts)	-0.103***	-0.123***	-0.202***
	(0.040)	(0.034)	(0.034)
Sex (male)	0.309***	0.258***	0.154**
	(0.068)	(0.062)	(0.062)
Age	0.011	-0.002	-0.045**
	(0.020)	(0.018)	(0.018)
Area (urban)	0.187*	0.183**	0.242***
	(0.097)	(0.086)	(0.093)
Group (black)	0.105	-0.177	-0.282
	(0.140)	(0.141)	(0.182)
Experience	0.189***	0.234**	0.096
	(0.072)	(0.094)	(0.175)
Experience^2	-0.065***	-0.039**	-0.004
	(0.017)	(0.015)	(0.014)
Number of hours worked	0.015	0.065***	0.065***
	(0.019)	(0.018)	(0.023)
Constant	6.362***	6.580***	8.097***
	(0.472)	(0.470)	(0.752)
Observations	464	356	303

Notes: Mother's education is used as instrument for schooling. Standard errors are in parentheses

earnings of dropouts progressively declined over time. However, results in 4.4 show that compared to OLS estimates that include measures of ability (see Table

4.3), IV estimates provide lower wage consequences of high school dropout, both in the first and subsequent years of labour market experience. This suggests that OLS estimates may still be upward bias even after controlling for measures of ability.

Table 4.5: Summary of the wage consequences of high school dropout

	(1)	(2)	(3)	(4)	N
First year	-0.216***	-0.162**	-0.103***	-0.119***	464
	(0.069)	(0.078)	(0.040)	(0.041)	
Second year	-0.251***	-0.174**	-0.123***	-0.121***	356
	(0.063)	(0.073)	(0.034)	(0.034)	
Fourth year	-0.411***	-0.253***	-0.202***	-0.200***	303
	(0.062)	(0.069)	(0.034)	(0.035)	

Notes: (1) OLS estimates (2) OLS estimates with measures of ability (3) IV estimates with mother's education as instrument for schooling (4) IV estimates with the interaction of proximity to school and mother's education as instrument for schooling. Standard errors are in parentheses.

As a form of robustness check for my IV estimates, I use an interaction of proximity to high school and mother's education as an additional instrument. Although a number of studies have used proximity to school as an instrumental variable on its own, the possibility of endogenous placement could mean that proximity to school can not be legitimately excluded from the wage equation. Interacting high school proximity with family characteristics (mother's education in my case) is one way to reduce the problem of endogenous placement. See Card (1995) for a detailed discussion. I present the additional estimates in column 4 of Table 4.5, which provides a summary of the key results from different specifications. The results in column 4 are largely similar to the results in column 3, which used only mother's education as instrumental variable.

Table 4.5 shows that, irrespective of the model specifications, individuals who choose to drop out of high school earned significantly less than individuals with high school diploma in the first year of labour market experience. The wages of high school dropouts also progressively declined over time compared to that of graduates. The first thing to note about these results is that the earnings disadvantage of high school dropout seem, on the average, to be more severe than what is reported in developed countries like the United States and Canada. For example, high school dropouts earn about 10% less than graduates in the United States on the average (Bjerk, 2012), and about 8% less in Canada (Campolieti et al., 2010). This difference may be due to the relative scarcity of skilled labour in developing countries.

Also, given that the estimates in Table 3.2 suggest that individuals who choose to drop out of high school may not be systematically different from high school graduates in terms of unobservables, it may be argued that high school diploma acts as a signal of productivity in the labour market, both immediately after graduation and in subsequent years. Thus, for deciding to terminate education today, high school dropouts may face a life cycle of poverty.

4.7 Conclusion

In this chapter, I examine the wage consequences of leaving high school prior to graduation using longitudinal data from South Africa. I show that there is no systematic pattern of self-selection into high school dropout and graduation in my data, and that those who participate in the labour market are not different from non-participants. I use two approaches to try to account for the endogeneity of education attainment. First, I use IQ tests results as measures of ability in an OLS regression. Secondly, I use parental education and proximity to school as instruments in IV estimations. Results from IV estimations show that OLS estimates are upward biased, even after controlling for measures of ability. Although not systematically different from high school graduates in terms of unobservables, high school dropouts earn less in the first year of labour market experience and their wages progressively declined in subsequent years.

References

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

Becker, G. S. (1964). Human capital: A theoretical and empirical analysis, with special reference to education. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*.

Bjerk, D. (2012). Re-examining the impact of dropping out on criminal and labor outcomes in early adulthood. *Economics of education review*, *31*(1), 110–122.

Blackburn, M. L., & Neumark, D. (1993). Omitted-ability bias and the increase in the return to schooling. *Journal of labor economics*, *11*(3), 521–544.

Blakemore, A. E., & Low, S. A. (1984). The high-school dropout decision and its wage consequences. *Economics of Education Review*, *3*(2), 111–119.

Blundell, R., Dearden, L., Goodman, A., & Reed, H. (2000). The returns to higher education in britain: evidence from a british cohort. *The Economic Journal*, *110*(461), F82–F99.

Cameron, S. V., & Heckman, J. J. (2001). The dynamics of educational attainment for black, hispanic, and white males. *Journal of political Economy*, *109*(3), 455–499.

Campolieti, M., Fang, T., & Gunderson, M. (2010). Labour market outcomes and skill acquisition of high-school dropouts. *Journal of Labor Research*, *31*(1), 39–52.

Card, D. (1995). Using geographic variation in college proximity to estimate the return to schooling, aspects of labour market behaviour: essays in honour of john vanderkamp. ed. *LN Christofides, EK Grant, and R. Swidinsky*.

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

Cawley, J., Heckman, J., & Vytlacil, E. (2001). Three observations on wages and measured cognitive ability. *Labour economics*, *8*(4), 419–442.

Chen, S. H. (2008). Estimating the variance of wages in the presence of selection and unobserved heterogeneity. *The Review of Economics and Statistics*, *90*(2), 275–289.

Eckstein, Z., & Wolpin, K. I. (1999). Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica*, *67*(6), 1295–1339.

Griliches, Z. (1975). Economic problems of youth', in (js coleman et al.). *Youth: Transition to Adulthood, Report of the Panel on Youth of the President's Science Advisory Committee, Washington, DC: Government Printing OfŽce, pp. 64q75*.

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica: Journal of the Econometric Society*, 1–22.

Griliches, Z., & Mason, W. M. (1972). Education, income, and ability. *Journal of political Economy*, *80*(3, Part 2), S74–S103.

Heckman. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153–161.

Hill, C. R. (1979). Capacities, opportunities and educational investments: The case of the high school dropout. *The Review of Economics and Statistics*, 9–20.

ILO. (2019). *Global unemployment data.* International Labour Organization.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, *125*(2), 515–548.

Kling, J. R. (2001). Interpreting instrumental variables estimates of the returns to schooling. *Journal of Business & Economic Statistics*, *19*(3), 358–364.

Montenegro, C. E., & Patrinos, H. A. (2014). *Comparable estimates of returns to schooling around the world.* The World Bank.

Oreopoulos, P. (2006). Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review*, *96*(1), 152–175.

Patrinos, H. A. (2016). Estimating the return to schooling using the mincer equation. *IZA World of Labor*.

Weybright, E. H., Caldwell, L. L., Xie, H., Wegner, L., & Smith, E. A. (2017). Predicting secondary school dropout among south african adolescents: A survival analysis approach. *South African journal of education*, *37*(2).

Appendix E

Self-selection into the labour market

Table E.1: Heckman selection estimates (first and second stage estimates)

	(1)		(2)		(3)		(4)		(5)	
	First-stage	Second	First-stage	Second	First-stage	Second	First-stage	Second	First-stage	Second
Education (dropout)	-0.153	-0.294***	-0.150	-0.286***	-0.148	-0.288***	-0.154	-0.274***	-0.0353	-0.189**
	(0.117)	(0.0919)	(0.120)	(0.0871)	(0.120)	(0.0873)	(0.123)	(0.0870)	(0.141)	(0.0762)
Sex (male)	0.167	0.356***	0.167	0.348***	0.168	0.351***	0.163	0.345***	0.157	0.357***
	(0.117)	(0.0831)	(0.118)	(0.0823)	(0.118)	(0.0829)	(0.119)	(0.0817)	(0.121)	(0.0926)
Hours worked	-0.0326	-0.000470	-0.0328	0.00128	-0.0394	-0.00313	-0.0410	-0.00169	-0.0277	0.00397
	(0.0322)	(0.0327)	(0.0326)	(0.0334)	(0.0330)	(0.0337)	(0.0330)	(0.0335)	(0.0335)	(0.0343)
Area (urban)			-0.0971	0.129*	0.0277	0.197**	0.0322	0.213***	-0.0121	0.170**
			(0.164)	(0.0762)	(0.172)	(0.0767)	(0.172)	(0.0774)	(0.181)	(0.0769)
Group (black)			-0.0652	0.0444	-0.101	0.0247	-0.0462	0.0238	0.0596	0.118
			(0.224)	(0.263)	(0.225)	(0.268)	(0.230)	(0.269)	(0.248)	(0.270)
age					0.0864**	0.0477	0.0936***	0.0458	0.0968***	0.0577
					(0.0357)	(0.0368)	(0.0359)	(0.0382)	(0.0360)	(0.0493)
Experience							-0.197	0.130	-0.265	0.0581
							(0.164)	(0.104)	(0.173)	(0.149)
Experience^2							0.0756	-0.0425	0.0817	-0.0286
							(0.054)	(0.032)	(0.0564)	(0.042)
Job search (yes)									0.432***	0.251
									(0.121)	(0.218)
IQ test									0.0155*	0.0142
									(0.009)	(0.008)
Job referral	0.140**		0.137**		0.134**		0.135**		0.102*	
	(0.062)		(0.062)		(0.063)		(0.062)		(0.059)	
ρ		-0.380		-0.346		-0.363		-0.334		-0.459
		(0.340)		(0.346)		(0.353)		(0.346)		(0.457)
Observations	532	466	532	466	532	466	528	463	528	463

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix F

Preliminary results

Table F.1: First stage regression results

	(1) First-year	(2) Second-year	(3) Fourth-year
Mother's education	-0.0979***	-0.0960***	-0.0960***
	(0.0231)	(0.0256)	(0.0285)
Sex (male)	0.344***	0.358**	0.394***
	(0.121)	(0.141)	(0.149)
Age	-0.00878	0.00598	0.0120
	(0.0377)	(0.0414)	(0.0452)
Area (urban)	-0.201	-0.152	-0.195
	(0.178)	(0.206)	(0.236)
Group (black)	0.930***	1.066***	1.289**
	(0.281)	(0.404)	(0.515)
Experience	-0.00613	0.182	0.282
	(0.132)	(0.217)	(0.438)
Experience^2	0.0128	-0.0153	-0.0135
	(0.0329)	(0.0365)	(0.0367)
Hours worked	0.0615*	0.0837**	0.0452
	(0.0336)	(0.0408)	(0.0537)
Constant	-0.384	-1.366	-2.202
	(0.933)	(1.141)	(1.921)
F-statistics	17.6	14.1	11.3
Observations	480	365	315

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Chapter 5

Conclusion

5.1 Summary

This thesis sets out to examine the nature of high school dropout and youth labour market outcomes in a Sub-Saharan Africa context, using South Africa as case study. I use the Cape Area Panel Survey data in my analysis. I find a substantially high rate of high school dropout in the data (about 50% on the average). Using transition probabilities, I find that high school dropouts have longer duration of unemployment than high school graduates. Furthermore, I use a discrete-time proportional hazard model and a discrete-time mixed proportional hazard model to estimate the determinants of high school dropout and graduation. While the proportional hazard model estimations ignore the role of unobserved individual heterogeneity, I account for the role of unobserved individual differences in the mixed proportional hazard model. I find that, compared to the mixed proportional hazard model, the model that ignores unobserved differences among individuals consistently understate the likelihood of dropping out or graduating from high school. However, when I generalize from a univariate distribution of unobserved heterogeneity to a bivariate distribution, the estimated likelihood of dropping out or graduating from high school remains largely the same.

In order to further understand the nature of high school dropout, I adopt a flexible framework that allows me to quantify selection on unobservables between dropouts and graduates and establish to what extent this selection matters. Specifically, I estimate a discrete-time conditional competing risks model that that allows for a general correlated risks across destinations. I treat individual differences that may exist between dropouts and graduates as both observed and unobserved. I then use the structure of my model to estimate a reduced-form measure of self-selection into drop-out and graduation. As a reduced-form measure of self-selection, I propose to consider the correlation coefficient of the bivariate distribution of unobserved heterogeneity across two possible exit states: graduation and drop-out.

It is typically found in developed country data that individuals who drop out of high school are those with relatively low ability, motivation, low expectations, and a set of negative preferences (Eckstein & Wolpin, 1999; Li et al., 2004; Foley et al., 2014). My results indicate that this is not necessarily the case in a developing country context, using South Africa as a case study. Specifically, I find no systematic difference between dropouts and graduates in terms of unobservables. This suggest that individuals who dropped out could potential complete high school education. One possible explanation of this finding is that unlike most developed countries, developing countries lack viable social protection programmes or other forms of government support that could keep disadvantaged youths in school until graduation. Thus, inequality of opportunity means that individuals with disadvantaged background may drop out of school more because the monetary and opportunity cost of completing high school education is higher for them.

Having understood, to some extent, the nature of high school dropout in my data, I estimate a wage equation in order to determine the severity and time pattern of the wage consequences of leaving high school prior to graduation. Prior to estimation, I checked for selection, both in the decision to dropout from high school and in the decision to participate in the labour market. The estimated measure of selection from the competing risks model shows no selection in the dropout decision. Also, estimates from a battery of Heckman selection model show no selection in the decision to participate in the labour market. Results from IV estimations show that OLS estimates are upward biased, even after controlling for measures of ability. Although estimates from a selection model suggests that individuals who choose to drop out of high school may not be systematically different from high school graduates in terms of unobservables, my results show that dropouts earn less in the first year of labour market experience, and their wages progressively declined in subsequent years. This suggests that high school diploma may act as a signal of productivity, both immediately after graduation and in subsequent years. The earning disadvantage of high school dropouts reported in this study is higher, on the average, than what is reported in developed countries like the United States and Canada.

5.2 Contribution

To the best of my knowledge, this study presents the first evidence of self-selection into high school dropout and high school graduation in the literature. Previous attempts to account for unobserved differences in the decision to dropout of high school took a univariate approach. That is, researchers specify their model with a univariate distribution of unobserved heterogeneity. Given the restrictive nature of this approach, it does not allow for the quantification of self-selection in terms of unobservables. By allowing for a bivariate distribution of unobserved heterogeneity in my model, I am able to estimate a measure of selection.

In addition, the method I proposed in this chapter provides an alternative, flexible and data-driven approach to analyze self-selection in education choices. The method requires no sundry assumptions for identification. Rather, it simply uses the structure of a well-known model to estimate a measure of self-selection and document an empirical evidence.

5.3 Recommendation

The results from this study indicate that individuals who drop out of high school are not systematically different from high school graduates in terms of unobservables (i.e., expectations, motivation, preferences, etc). This suggests that individuals who drop out do not do so because they perceive or understand that they do not stand a chance in education. Thus, dropouts could potentially complete high school education. This provides a scope for government to intervene in order to address the problem of high school dropout. Disadvantaged students could be supported with social protection programme like conditional cash transfers and other related incentive-based policies. This study has also shown that the worth of a high school diploma in the labour market is not trivial and long-lasting. Providing accurate information of what students stand to gain by completing high school education is another possible area of intervention.

5.4 Limitation and future work

The Cape Area Panel survey data used in this study collects data only from Cape Town South Africa. The data collection process did not cover the entire country. This non-representative nature of the data is a limitation of this study as it constrains the generalization of my findings. However, it is important to mention that Cape Town is not very different from other cities in South Africa and some cities in Sub-Saharan Africa in terms of socioeconomic characteristics. In terms of future work, it would be useful and insightful to write down a structural model of high school dropout and then simulate different government policies and see how they may affect high school dropout.

References

Eckstein, Z., & Wolpin, K. I. (1999). Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica*, *67*(6), 1295–1339.

Foley, K., Gallipoli, G., & Green, D. A. (2014). Ability, parental valuation of education, and the high school dropout decision. *Journal of Human Resources*, *49*(4), 906–944.

Li, M., Poirier, D. J., & Tobias, J. L. (2004, March). Do dropouts suffer from dropping out? Estimation and prediction of outcome gains in generalized selection models. *Journal of Applied Econometrics*, *19*(2), 203-225.