Beyond enrolment:
Academic incentives, outcomes and performance in higher education

Leigh Neethling

School of Economics

Faculty of Commerce

University of Cape Town

Thesis presented for the degree of

Doctor of Philosophy

October 2022

Supervisor: Associate Professor Anthony Leiman

Plagiarism Declaration

I, Leigh Neethling, hereby declare that the work on which this thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree at this or any other university.

I authorise the University to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signed: Date: October 2022

Abstract

This study examined incentives, academic outcomes and student performance in the South African higher education (HE) sector using the University of Cape Town as a case study. The analysis was conducted using a dataset that stacked three cross-sections of first-year entering students and tracked these students over time.

The thesis comprises six chapters. The introductory chapter provides background content on the research. Chapter 2 presents an exploratory and descriptive analysis of the South African HE sector over the period 2004–2015. It focuses on a descriptive analysis of key South African HE indicators and outcomes. The first objective was to evaluate access to HE by identifying the pool of potential entrants through an examination of the school-leaving cohort of each year. The second objective was to identify racial enrolment, progression, and completion patterns to observe whether significant changes occurred over the period. The author finds that the differentials in performance between racial subgroups have narrowed over time. White students are shown to have the highest student success rates at above 80%, and although other subgroups show some improvement, they do not catch up to these rates. A trend analysis of the data, however, provided support for a fall in the dropout rate for all students. This chapter also provides evidence for persistence in but slower progression through HE. Chapter 3 presents a way to consider and evaluate the Dean's Merit List (DML) incentive system in the context of an African economy. The author evaluated the impact of academic recognition policies, specifically the DML, on student outcomes. Using a regression discontinuity approach, the chapter shows that the DML as an academic incentive policy, has largely negative rather than the intended positive effects over the short- and long-term on academic performance in the South African context. The results indicate that the DML has an unfavourable impact on subsequent academic performance. Students who received the award tended to earn lower grade point averages in subsequent years than expected.

The findings suggest that the DML does not reinforce academic achievement. These results appear to be counterintuitive but support Bénabou and Tirole's theoretical expectations regarding extrinsic motivation in a situation of asymmetric information between an agent and principal. Chapter 4 investigates student performance over time by introducing a ranking variable of student achievement. The main finding is that race, gender and performance on final school-leaving examinations are important determinants of academic achievement. Female students outperform male students across the distribution of grade point average, and this finding is consistent with the growing international literature. Chapter 5 presents detailed evidence on the determinants of academic outcomes using discrete-time methods for competing risks survival analysis. An important contribution of this chapter is studying the determinants of dropout and graduation in HE in the context of an African country. While graduation is the preferred route of exit, voluntary and involuntary exit before completion remain prominent for a significant number of students. Interestingly, and contrary to other international studies, the author did not find support for financial aid status contributing to either voluntary dropout or graduation, even after controlling for academic and socio-economic background factors. Students on academic programmes are shown to be more likely to be involuntarily excluded and less likely to graduate or voluntarily exit HE than mainstream students. This is a cause for concern as these programmes are an initiative intended to address transformation and equity in HE, attracting significant resources from within and outside universities. Chapter 6 summarises and presents policy discussions.

Overall, the study shows that one-size-fits-all policies within the same institution applied across heterogenous faculties do not achieve their desired outcomes in the South African HE setting. Considerable thought should be given to the nature of recognition policies as other basic requirements, such as course progression criteria, tend to crowd out the desired incentive effects of recognition policies. In addition, academic administrators should consider programmes that promote a decrease in outcome disparities in HE, including establishing more and expanding academic development programmes.

Dedication

To Sean, my incredible husband, whose love and support made it possible for me to complete this thesis, and to our amazing daughters Belen and Isla, who are patient and kind.

I love you.

This is for you.

Acknowledgements

The road to completion has been long and not without difficulties. This thesis would not have been possible without the support of my family, who have encouraged me every day since I embarked on this long, often lonely journey. Thank you for always being there and stepping in when needed.

My biggest supporter is my husband. He is the best partner I could have hoped for. And he has given me the best gifts in life, two beautiful daughters. They are strong, amazing little girls who have shouldered more than their fair share of mom not being around to play.

I would like to acknowledge financial support from the UCT Research Office and Emerging Researcher Fund.

A very big thank you to my supervisor Associate Professor Anthony Leiman. He has been supportive and encouraging throughout this long process. His long chats about anything and everything related to education will not be forgotten.

A special thanks to two of the best colleagues one could hope for: Lindokuhle Njozela and Katherine Eyal made this process easier by easing the load at work, encouraging me when I am down, and helping me to keep moving forward. They are inspiring.

The staff at UCT's Institutional Planning Department were patient through all my many requests for student data. They did not complain when I persisted with data requests over the years.

I would like to thank the participants at the ESSA conferences in 2015 and 2019 for providing valuable feedback on Chapters 3 and 5. Chapter 4 was presented at the 4th Annual International Conference on the Scholarship of Teaching and Learning in Higher Education at Central University of Technology in 2018. I thank the participants for their feedback.

To the random person reading this, don't give up. Don't let anyone tell you it isn't possible.

Table of Contents

List of Tables

List of Figures

Acronyms

AYOS	Academic Year of Study
CIF	Cumulative Incidence Function
CJM	Cattaneo, Jansson and Ma
DHET	Department of Higher Education and Training
DML	Dean's Merit List
GER	Gross Enrolment Rate
GPA	Grade Point Average
HE	Higher Education
HEI	Higher Education Institution
IPD	Institutional Planning Department
NMAR	Not Missing at Random
NSC	National Senior Certificate
RD	Regression discontinuity
SHR	Sub-distribution Hazard Ratio
UCT	University of Cape Town
US	United States

Chapter 1: Introduction

“Education is the most powerful weapon which you can use to change the world”
Nelson Mandela

1.1 Introduction and Motivation

The benefits of post-secondary schooling at an individual and societal level are well established in the literature (McMahon, 2009). The literature shows quite clearly that the returns to education in South Africa rise with the level of education completed and that returns to education in South Africa are amongst the highest in the world (Keswell & Poswell, 2004; Oreopoulos & Petronijevic, 2013). At the individual level, the labour market opportunities for individuals with HE qualifications are numerous with associated benefits such as improved health and more engaged citizens (Milligan, Moretti, & Oreopoulos, 2004).

Numerous socio-economic issues confront South Africa, such as excessive income and wealth inequality, high unemployment rates, and underwhelming basic and secondary education systems. Given these difficulties, the HE system can play a significant role in eliminating social disparities by enhancing access and outcomes and by producing graduates with the necessary skills to ensure the nation’s economic success. Along with encouraging access to opportunities that increase labour market participation, this also increases social mobility by increasing income and wealth mobility (Statistics South Africa, 2019).

The demand for tertiary education in South Africa remains high and robust. Research for South Africa shows that individuals with tertiary qualifications are significantly less likely to be unemployed, face shorter periods of unemployment (Nonyana and Njuho, 2018), are more likely to be formally employed (Branson and Leibbrandt, 2013), and are twice as likely to obtain employment than individuals who only completed Grade 12; and are significantly more likely to be self-employed than unemployed (Branson, Leibbrandt & Zuze, 2009).[1] In addition to the labour market participation effects, individuals with tertiary education qualifications are expected to earn significantly higher lifetime earnings compared to individuals with incomplete high school or incomplete matric. Relative to individuals with incomplete high school, individuals with a degree or certificate earn between 170% and 220% more, and individuals who completed a degree earn between 250% and 400% more than individuals who had incomplete matric between 2000 and 2007 (Branson, Leibbrandt & Zuze, 2009).

One motivator for this study is expenditure on education. Expenditure on the South African HE sector has increased over the last few years. It is envisioned by the government that over the period from 2017 to 2024, government expenditure on the HE sector will reach 7% of consolidated government expenditure and 2.2% of domestic GDP. In addition, the real per capita expenditure per student in university has consistently exceeded R80 000 per year since 2010 and reached a high of R90 000 in 2017. The rand amount remains just below R90 000 in current expenditure periods.[2] Expenditure per pupil at technikons is about half the rand amount for every year since 2010 (Khuluvhe, & Netshifhefhe, 2021) Overall, expenditure per full-time student is non-trivial in domestic currency terms and absorbs a large part of the government budget.

A second motivator for this thesis is the contestation between the racial demographics of the country and educational outcomes. South Africa's demographic profile is 80% African, 9% Coloured, 8% White and 3% Indian (Statistics South Africa, 2022). However, the labour market profile of post-secondary educated workers still does not mirror the country's national

---

[1] In this thesis, the term 'matric' is used interchangeably with 'Grade 12', which is the final year of high school, in keeping with everyday language in South Africa. When referring to complete Grade 12 or matric, completion of the final or matriculation examinations are referenced.

[2] The dollar amount is approximately $4250 when converted at a rand-dollar exchange rate of R18/$.

racial profile 25 years post-democracy. The majority population group of African individuals experience worse educational outcomes compared to all other race groups (Statistics South Africa, 2019). Investigating what happens when individuals are enrolled in HE adds to the stock of knowledge on human capital formation in South Africa.

In this context it is important to understand the microeconomics of HE as it relates to participation, success, and student outcomes.[3] It is not only necessary to understand how individuals perform in the HE system but also to tease out these patterns of participation, performance, and student success outcomes across observable characteristics, including race and gender. Currently, the policies that are in place to address these inequalities in HE are ineffective.

Research on access and success has grown over the past few years, albeit at relatively slow rates because of shortages of individual-level data from the HE sector. This includes a lack of panel data linking students from high school into tertiary education and a lack of detailed data on students once enrolled in HE. The research in this area has increased as the volume of HE data has grown. However, there remains a significant shortage of microeconomic analysis to show how students progress through HE, with particular reference to how and when exit decisions are taken by students.

This study examined human capital formation at the HE level using the University of Cape Town (UCT) as a case study. This study contributes to the economics of education literature on incentives, performance, and outcomes in HE by contributing three papers on the South African experience. The first of these explores the use of incentives as a motivator for student achievement. The linkages between incentive and recognition policies and student motivation

[3] The term ‘success’ is loosely defined as completion of a task, qualification, or activity. The term is defined in each of the contexts within which it is used throughout the thesis.

have not been fully explored in a developing country context, especially in sub-Saharan Africa. The second examines the pathways through HE. Policy makers and university administrators may not fully comprehend the nature, duration, and complexity of pathways through HE, and this hinders the design and implementation of good academic policies. The third considers how students exit from HE in South Africa and investigates the competing nature of academic outcomes. The uncertainty of what happens once students are enrolled remains a key obstacle to isolating and identifying those factors that hinder or foster performance and success in HE.

Data was drawn from cohorts entering between 2006 and 2008, which were the last few cohorts before new entrants into HE wrote the National Senior Certificate (NSC) exams.[4] Each of the three areas investigated focuses on distinct but interrelated aspects of the HE system in the South African context. Moreover, given the dearth of literature emanating from developing economies, this research aimed to describe and explain the HE decisions made by individuals once they have enrolled in HE.

1.2 Research Questions

The overarching objective of this research was to expand the knowledge and understanding of the economics of HE with a particular focus on issues relating to student performance and outcomes in a developing country setting by conducting a South African case study. Accordingly, this thesis explored the following questions:

- What is the impact of student recognition policies as incentives?
- Do student recognition policies deliver the intended outcomes for students?

[4] Prior to 2009 students wrote the matric exams called the Senior Certificate. The entering cohort from 2009 onward wrote the NSC exams.

- Are the effects of student recognition policies homogenous across degree duration?
- Do racial and gender pathways through HE differ?
- How does student performance differ between the first and last year of study?
- How different are student outcomes among the student body?
- How do exit outcomes differ by race and gender?

As this research is located within the South African HE sector, it was important to have a working understanding of the background and historical context of the sector. These are discussed in Chapter 2.

1.3 Contribution and Structure of the Thesis

This study sought to contribute to the literature that explains student pathways through HE. Most writing on the topic has been in the context of developed countries, where success rates and the resulting study design requirements are meaningfully different from those to developing countries. This makes it difficult to directly compare outcomes across these two contexts. The relatively higher dropout rates for developing countries like South Africa need to be accounted for in any methodology selected to study issues where loss of participation occurs due to non-random reasons.

The methodologies adopted in this study were chosen to correct for a range of potentially biasing effects because of the nature of HE participation and outcomes in the South African HE system. This allowed for better identification of the relationships that are explored in each chapter and were expected to give rise to more accurate predictors of academic and student outcomes. In the South African context, there is a high dropout rate in HE, and to account for this, all methodologies chosen should consider if the selection bias arising as the characteristics of those who remain enrolled may differ from the characteristics of those who drop out.

Chapter 2 gives the contextual basis for the understanding and interrogating the HE sector in a developing country, and specifically in an African context. HE in developing countries is characterised by issues around access and equity, which can lead to challenges relating to participation, progression, performance, and exit. The chapter explores the aggregate measures of progression by examining student success rates by race group to observe differential patterns. This is subsequently followed by an exploration of graduation rates and dropout by race group to identify consistency in the patterns observed in the progression data.

Chapter 3 explores the short- and long-term effects of a student incentive programme known as the Dean's Merit List (DML), commonly referred to as the Dean's List, at many international universities on student outcomes using UCT as a case study. The DML is an academic award given to high-achieving students who achieve a minimum grade point average (GPA) each year. This accolade is meant to incentivise students to maintain their solid academic performance and to continue excelling in subsequent years.

Chapter 4 investigates student performance once enrolled in HE by exploiting the longitudinal nature of HE data. The chapter examines how students perform over time by following the same students as they progress through the system from enrolment to exit. For the purposes of this chapter, exit is defined by graduation. The primary contributions of this chapter include tracking cohorts from entry to graduation, a detailed examination of annual academic performance by faculty and degree programme, the development of a ranking variable that allows for the control of differences in degree programme difficulty across the institution, and the development of a methodology designed to overcome the dropout bias present in HE data in countries such as South Africa.

Chapter 5 examines the determinants of academic outcomes using a competing risks model. The contribution of this approach is that it controls for the different outcomes observable in the data. An important contribution of this chapter is studying the determinants of dropout and graduation in HE in the context of an African country.

The concluding chapter of this study draws on the results of Chapters 2 to 5 to provide a summary of the main findings and highlights several potentially important areas for policy. The limitations of the work are presented and thoughts for further work and extensions in this area are discussed.

Chapter 2: An Overview of Higher Education in South Africa: 2004–2015

2.1 Introduction and Background

Many great documents begin with an assertion that what follows is obvious. Probably the two best known are, ‘*It is a truth universally acknowledged*’ and ‘*We hold these truths to be self-evident*’. Either phrase could be used to introduce the phrase, ‘that education matters in developing countries’. Despite this, the process and pitfalls associated with HE in these countries remain under-researched. This study investigated what happens once students are enrolled in HE in a developing country context. It contributes to the literature by presenting an African case study because relatively few of these exist. This is an important topic as success or completion rates in developing countries are lower than in developed countries, and many developing countries are investing aggressively in HE without getting the desired improvements in outcomes. The impact of these outcomes extend into individual’s labour market experiences. The economics literature provides evidence that individuals with completed HE have better labour market outcomes, including labour market participation rates, shortened unemployment durations, higher wages, and more secure jobs (Branson, Leibbrandt & Zuze, 2009; Branson and Leibbrandt, 2013).

Before exploring the impact of policy and investigating academic outcomes, it is useful to describe and discuss the national South African HE landscape. This is relevant because it provides the background and context for the subsequent chapters and findings and shows the entrenched disparities in educational and economic outcomes that permeate the South African economy.

The primary objective of this chapter is to present an exploratory and descriptive analysis of the South African HE sector over the period 2004–2015. This chapter focuses on a descriptive analysis of key South African HE indicators and outcomes to emphasise the differentiated set of academic outcomes that are observed across the student body and the heterogeneity of students' experiences. The first objective is to evaluate access to HE: the pool of potential entrants is identified by examining the school-leaving cohort of each year. The second objective is to identify racial enrolment, progression, and completion patterns and observe whether significant changes have occurred over the period. This provides a background and context for the findings in the subsequent chapters of this study.

Since the end of apartheid, higher education institutions (HEIs) have been open to all South Africans. The South African government broke down and removed structural blockages that prevented students from attending HEIs of their choice and implemented policies to expand the HE sector to encourage participation. This included creating special units within universities that focused on increased affirmative action admissions into dedicated programmes designed exclusively for previously disadvantaged students only. It also entailed updating admission criteria to take into account the possible under-preparedness of students from government or public sector high schools; and setting targets or benchmarks in the admission profile of new entering first-year students. Despite these interventions, the HE sector is still categorised by low participation and high dropout rates (Fisher & Scott, 2011). Sections 2.2 and 2.3 provide a detailed discussion on the South African landscape of access and enrolment to present a background on the current state of HE and to draw attention to the differences that may exist between different groups.

It is widely acknowledged that challenges in the basic education system hamper access to and participation in HE.[5] Branson, Hofmeyr and Lam (2014) conducted one of the most thorough reviews of progress through school (primary and high school) using data from South Africa's only longitudinal survey the National Income Dynamics Study, and they found that only about 50% of students who start Grade 1 reach Grade 12. They tracked individuals over time and followed young people through their schooling careers and found that most dropout occurs in high school after Grade 7, the largest numbers being between Grades 9 and 11. Their findings support those of the Department of Basic Education that almost 40% of students have dropped out by the time they reach Grade 11, leaving a smaller pool of students to write the school exit examinations and even fewer who qualify for access to HE.

South Africa has one of the lowest rates of HE relative to adult population in the world. Despite some improvements in participation in the HE sector, only 2.2% of South Africans between the ages of 25 and 64 have completed a HE qualification (Statistics South Africa, 2019). This compares unfavourably with other middle-income or developing countries and leaves South Africa trailing developed economies significantly. In Brazil, at least 5% of the adult population have a bachelor's degree, and the proportion is much higher in the US at almost 30% (OECD, 2021).

Completion rates in the universities are also problematic. Only 45% of those who register for university degrees in South Africa ever complete their degrees (DHET, 2018). In addition, there are considerable differences in graduation rates by race, gender, and qualification type (Fisher & Scott, 2011). Not only do race and gender affect HE entry and graduation rates, they

[5] Basic education refers to the 12 years of primary (7 years) and high (5 years) school in South Africa.

also affect the grades obtained by students and hence their prospects for entry to higher levels of tertiary study. This is an issue that has received little attention to date.

Before proceeding to the detailed discussion in the following sections, it is prudent to briefly define key concepts that recur within the discussion. The first concept that requires a specific definition is ‘success’. The term success, as used in this chapter, refers to individuals who have completed the requirements for the award of the qualification.

The racial classifications commonly applied in South Africa may be problematic for those unfamiliar with the country. The apartheid government divided the population into four primary population groups or classifications, namely White, Black (African), Coloured and Indian (including individuals of Asian descent), and these classifications were enforced via identity numbers and identity documents. The Coloured group represents individuals of mixed-race and served as a catch-all for those who were not included in the other three groups. For the purposes of this paper, the terms ‘Black’ and ‘African’ are used interchangeably, with African being the preferred term for Black South Africans.

The next sections in this chapter discuss in detail some of the specific issues facing South African HE today, looking at who qualifies for entry and student success rates. It is important to place this discussion in context because while South Africa has improved access to HE, the same improvements have not been noted in success or throughput rates for each identified population racial subgroup. The three issues mentioned above, namely access, enrolment and progression, are discussed in further detail in sections 2.2 to 2.4.

2.2 Access

Access to HE in South Africa is determined almost entirely on the basis of student performance in the Grade 12 school-leaving final examinations.[6] These standardised examinations are written by approximately 400 000–550 000 scholars per year and are overseen by the Department of Basic Education in conjunction with the provincial departments of basic education.[7] Exams typically run over an eight-week period from October to early December each year.[8] Students' final results determine whether they receive firm offers for a place of study in HE in the next academic year. Approximately 30% of Grade 12 learners who write the school-leaving exams achieve bachelor passes.[9] About 70% of those achieving bachelor passes go on to enter HE, but only 45% of these students go on to graduate with an undergraduate qualification (DHET, 2018).

The South African education system has experienced notable growth in recent years, especially in HE (Statistics South Africa, 2019). Compared to a decade ago, a larger proportion of young people enter HE today. Growth in this sector comes from two major sources. Firstly, there has been a substantial increase in the number of scholars completing basic education each year. Many more South Africans are completing high school with a bachelors pass and thereby qualifying for entry to HE today compared to 10 years ago.[10] Secondly, there have been modest

---

[6] In recent years, universities have required students to write the National Benchmark Tests. The results from the National Benchmark Tests and the student's Grade 12 examination results determine whether an offer for a place of study is made. However, the National Benchmark Tests are not relevant for the period covered by this thesis.

[7] The term 'scholar' indicates an individual in the basic education system. 'Student' refers to an individual in the HE sector.

[8] In South Africa, the academic and calendar years coincide.

[9] A bachelor pass refers to a grade above the minimum needed to qualify for admission to bachelor-level studies. Importantly, it does not guarantee an offer of admission.

[10] Grade 12 students writing their final examinations must achieve government-mandated pass criteria called a 'bachelors pass' that is a minimum prerequisite for entry to HE. Many universities set entrance requirements far above bachelors pass.

increases in the number of places available. Between 2011 and 2015 the number of study places available increased by 47 011. This translates into an annual average growth rate of 1.2% over the same period, significantly down from an annual average rate of 4% over the previous five-year period. There is a much higher demand for HE today, but the supply of places has plateaued in recent years (Council on Higher Education, 2018).

Figure 1 shows a summary of the school-leaving exit examinations between 2004 and 2015. The number of students writing the school exit examinations has fluctuated over the period under consideration but has shown a general upward trend over the period. The proportion of students passing and meeting some minimum criterion for entry to HE, shown by the proportion of bachelor passes, has consistently increased. This increased eligibility is one indicator of the rising potential demand for HE.

Figure 1 Summary of high school exit examinations: 2004–2015

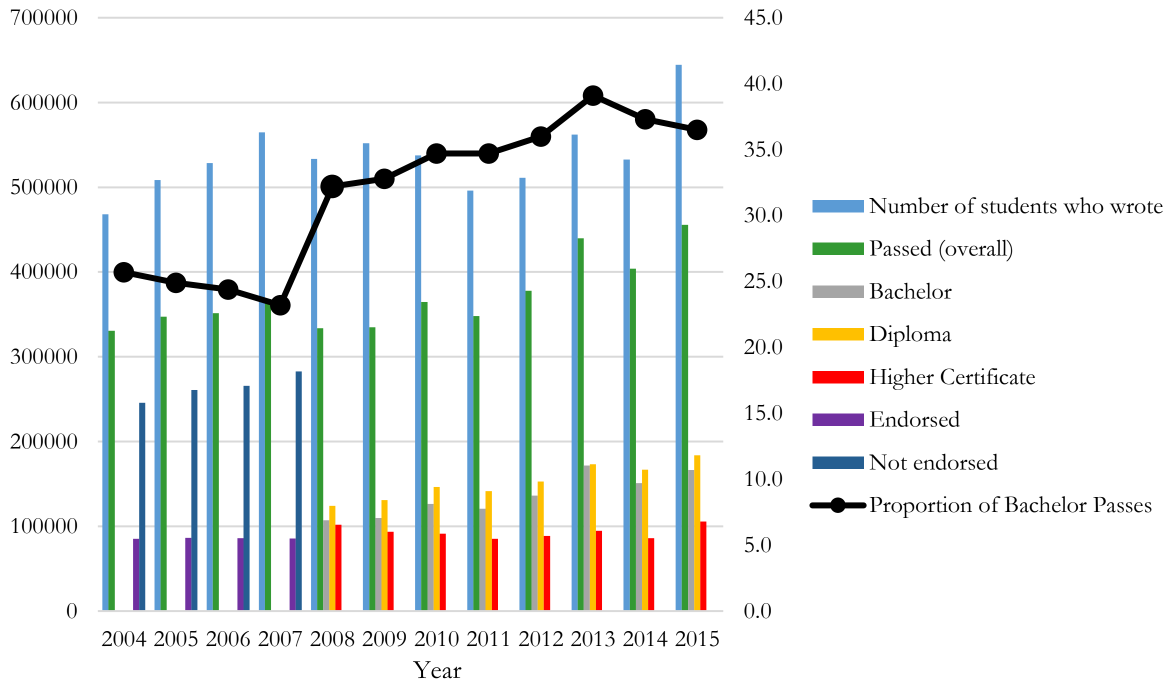

Figure 1 shows a summary of the number of scholars who wrote the school-leaving examinations and the classification of passes by category. The figure also shows the proportion of scholars achieving the minimum criteria to be eligible for HE (proportion of bachelor passes read from the right-hand side).
Data: Department of Basic Education (2005–2015) and Statistics South Africa (2018).

Prior to 2008, scholars wrote the old Senior Certificate curriculum examinations in which they were required to achieve an endorsed pass on their final school-leaving examinations to be eligible to enrol at a HEI. In an effort to improve the quality of schooling and cognitive outcomes for school-leavers, the new NSC was introduced, and the first cohort wrote these school exit-level exams at the end of 2008. The NSC ranked eligibility to HE on a three-tiered system, bachelor (university) entry, diploma, and higher certificates (at the college level), from highest to lowest, and therefore, did away with the endorsed and not endorsed categories. To be eligible for degree-level (bachelor) studies, students must achieve on the NSC a minimum of a 2, or 30% final grade, for their chosen home language and at least a 4, or 50%, in four chosen subjects.[11] The matric achievement criteria is in the Appendix, Table 35. The entrance requirements for admission to degree-level (bachelor) studies are highest, followed by diplomas and higher certificates.

The two school-leaving exit systems are not directly comparable as the government did not set strict matching criteria between the old and new systems. Research conducted in 2010 reveals that NSC results map to Senior Certificate but about 20% higher, depending on the student's location in the distribution, meaning that a student who would have achieved 60% on the old Senior Certificate exams could expect to achieve 80% on the NSC exams (Schoer et al., 2010). Further research in this area and across cohorts of students leaving school with an NSC showed similar results.

This grade inflation is one reason for the significant jump in bachelor passes between 2007 and 2008. Figure 1 shows that the number of students who are eligible for entry to bachelor's

[11] Details on these chosen subjects and the matric grading system are available in the Appendix, sections A1–A4.

study between 2007 and 2008 increased by 21 820 students or nine percentage points, from 23.2% in 2007 to 32.2% in 2008.

Between 2008 and 2015 there was a 55% increase in the number of scholars who passed their school-leaving examinations with a bachelor's pass. During this period no additional new universities opened in South Africa nor was there any significant expansion in the availability of study places. The next section examines enrolment patterns and trends by racial composition to trace any changes from qualifiers to enrolments.

2.3 Enrolment

The South African HE landscape is deeply fragmented. Prior to 1994, universities were not open to all students and most access was based on race. For example, there were specific universities for Africans, Coloureds and Indians respectively. Institutions like UCT did their best to open themselves where gaps in government regulations allowed, e.g., where specific courses were not available at Historically Black institutions, the students wanting to do them could relocate to UCT. Once the democratic government took over in 1994, the process began to de-racialise HEIs. Historical issues meant that enrolments at the different HEIs are still skewed according to race and income where income represented affordability, creating a highly divided HE sector.

The HE sector has undergone a significant transformation over the past few years. In an effort to create a more efficient HE sector, the government consolidated the number of institutions while re-purposing others. Overall, this has resulted in a decrease in the number HEIs from 36 to 26. Between 2003 and 2005, the HE sector underwent significant transformation and the number of HEIs was consolidated from 36 to 23 (DHET, 2019), to which three new ones were then added This consolidation had multiple objectives, such as efficiency in university administration through economies of scale and scope, which was achieved. In addition, some institutions were regarded as ‘second rate’, and by merging these less well received institutions with others of better academic standing, the government hoped to increase the value of their qualifications. Two new universities opened in 2014, namely the University of Mpumalanga and Sol Plaatje University, and the Sefako Makagatho Health Sciences University commenced operation in 2015, enrolling its first students for the 2015 academic year. This takes the present-day number of public universities to 26, and these universities enrol most students in the HE sector. About 15% of university students enrol in private universities but very little is known about students’ performance in private HEs, and therefore, this analysis focused on public universities.

Identifying the capacity constraints in public universities has been rather difficult as many universities do not disclose their enrolment capacity constraints. Many universities face capacity constraints because of national health and safety regulations in terms of limited physical space on campus and building size.

Initial investigations revealed that the number of places for new first years has remained static over the period under consideration. Universities such as the UCT, the University of Johannesburg, Witwatersrand University, and Stellenbosch University have not significantly increased their enrolment capacity over the period of investigation. As universities are not obliged to report publicly on annual admissions or enrolment capacity, the next best option was to evaluate the limited enrolment statistics over the period. Figure 2 shows the enrolment in universities by race from 2004 to 2015.

Figure 2 South African HE enrolment: 2004–2015

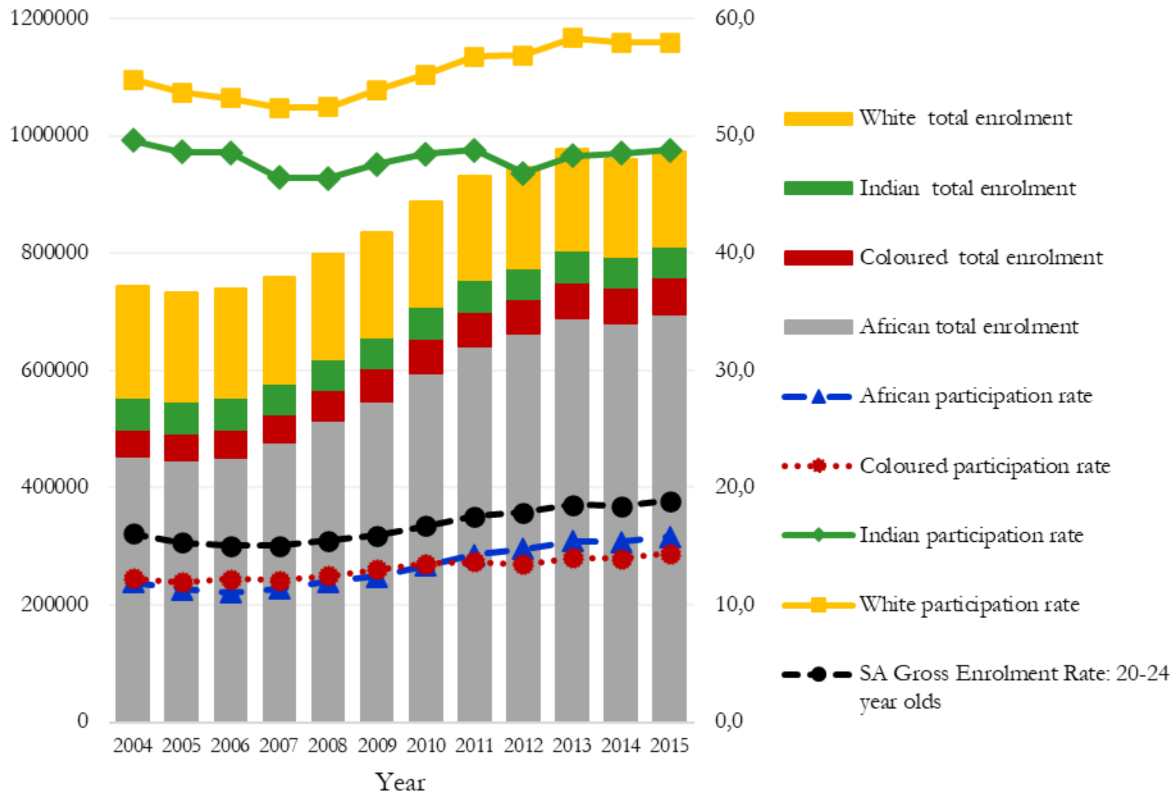

Figure 2 shows headcount enrolments and participation rates in South African HE from 2004 to 2015.
Data: Council on Higher Education (2018) and Statistics South Africa (2019)

The first indicator of significance is the gross enrolment rate (GER). The GER is a participation rate measuring the number of 20–24 years olds in the general population who are enrolled in HE.[12] The National Plan for Higher Education (Ministry of Education, 2001) sets out a post-secondary participation target of 20%, as defined by the GER (Ministry of Education, 2001). In 2014, North America had a GER above 84%, compared to sub-Saharan Africa that had a GER of approximately 8% (Roser & Ortiz-Ospina, 2013). The long-term GER for South Africa is 15%, which is above the sub-Saharan average but compares poorly with other middle-income countries. This new target was set to promote and achieve equity in access and success, to improve transformation in the HE sector, and to increase the number of

[12] The South African government and the DHET use the terms participation rate and GER interchangeably.

graduates in South Africa. Importantly, from a transformational perspective, the DHET set out to transform enrolment in HE the sector to more closely replicate the demographic profile of the country within HE. Their aim was to provide the country with a steady supply of non-white graduates who would transform the labour market in the future and help equalise racial participation in the market for skilled labour.

At an aggregate level an improvement in HE participation is noted.[13] Figure 2 shows that, by 2015, South Africa had not yet reached one million students participating in HE. However, it is evident from Figure 2 that the HE system shows continuous increases in enrolment over the period, with significant gains between 2008 and 2012.

To date, the 20% GER target has not yet been achieved. Between 2008 and 2013 the South African participation rate increased by 2.6 percentage points, from 16.6% to 19.2% (Statistics South Africa, 2018). The aggregate percentages mask significant racial differences observed in the data. Figure 2 shows the actual headcount for South African HE and the participation rates by race. It is evident from Figure 2 that the participation rate for white students has increased quite strongly over the period, especially from 2008, while the participation rates for Coloured and African students display much lower growth. Indian/Asian students show a decline in participation rate over the same period. Note that both the Coloured and African overall participation rates are low and remain below the national target of 20%. This pattern has persisted despite affirmative action policies at institutions that were effectively reserved for White students until the late 1980s.

Another significant indicator is a measure of the ratio of the number of first-time entering students to the number of bachelor's passes in the previous academic year. Figure 3 graphically

[13] The discussions that follow in this chapter relate to information on public universities in South Africa.

presents this capacity indicator. This ratio is a measure of the capacity constraints of HEIs. The ratio shows that not all students who qualify enter HE in the year immediately following their Grade 12 year due to various factors. It also shows that some sort of HE capacity constraint was reached by 2010. After 2010, the ratio exhibits a downward trend. The ratio may fall for two reasons. Firstly, as Figure 1 shows, there have been consistent increases in the number of students writing exit examinations and achieving bachelor-level passes but neither the numerator nor the denominator in this ratio remains constant, implying that relative changes in each variable may be driving the pattern observed in Figure 3. In this instance, a faster rise in the number of students achieving bachelor passes relative to the number of students entering HE is indicative of the capacity of institutions reaching a ceiling. Secondly, without additional capacity being created by HEIs through the creation of new HEIs or existing HEIs enrolling more students, matriculants are left with fewer post-school options.[14] This observation is consistent with trends observed in labour market statistics where more young people are filtering into the labour market each year. The lack of employment opportunities means that many youths do not enter employment instead of education, causing a rise in the number of young people in the 'Not in Education, Employment or Training' category.

Despite African participation rates increasing by 31% over the period, the increase comes off a relatively small base. The increase in nominal terms amounts to almost a quarter of a million students. Even in the face of such large nominal increases, it is still not enough to meet the targeted participation rate of 20% or to create equitable participation in HE as desired by the government.

[14] 'Matriculants' is the colloquial term used in South Africa to describe an individual who has completed the 12 years of basic education.

Figure 3 Ratio of first-time entering students to the number of bachelors passes of the previous academic year: 2004–2015

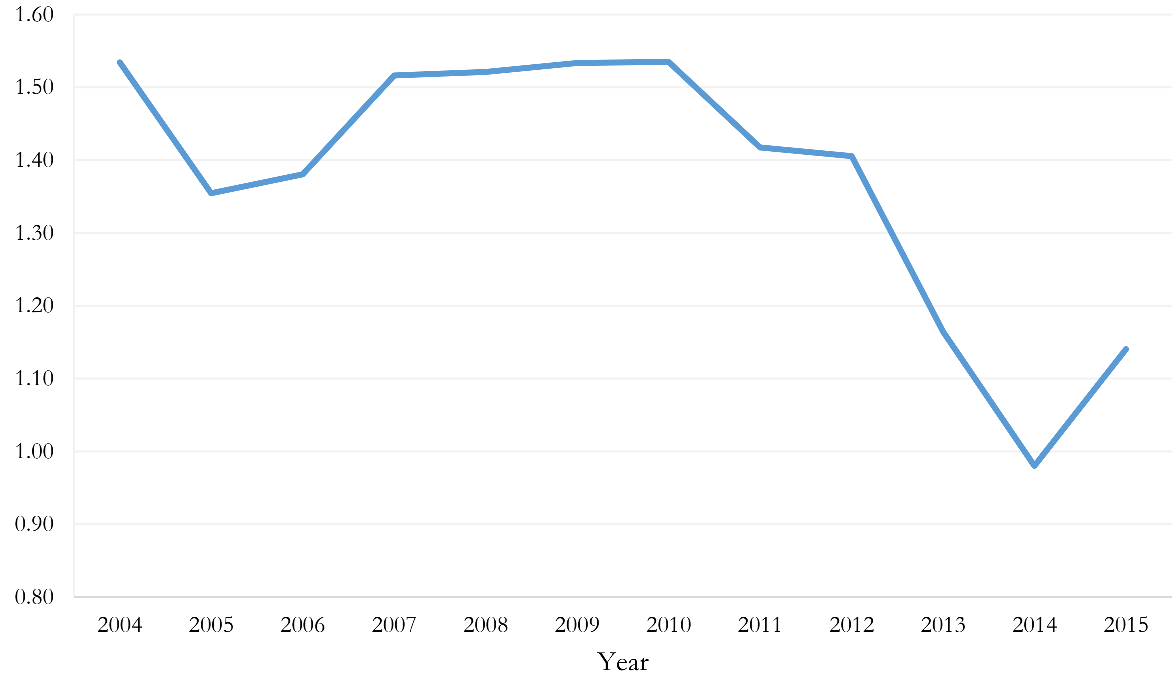

Figure 3 shows the ratio of first-time entering students to the number of bachelor passes of the previous academic year.
Data: Statistics South Africa (2018).

Despite the evident growth from the demand side in the HE sector, only 2.2% of South Africans between the ages of 25 and 64 have completed a HE qualification (Statistics South Africa, 2018). When broken down into narrower age categories, HE acquisition shows significantly more variation. For individuals between the ages of 25 and 34 years, South Africa's share of individuals with a HE qualification was 5.6% in 2018, compared to 21% for Brazil and 51.8% for the USA at the same point. Russia, Canada, and Korea exceed the 60% share of individuals aged 25–34 years with completed HE. South Africa's performance does not compare favourably with other middle-income or developing countries, and leaves South Africa trailing developed economies quite significantly.

2.4 Progression and Dropout

There are notable gains in terms of success in the HE sector, even though the sector remains highly fragmented. The previous section highlighted the increasing headcount enrolments in HE over the period under investigation, which were largely driven by a sharp increase in the number of African students, which increased by over 30% for the period under consideration. The next subsection breaks down the throughput and success categories and examines outcomes by race and gender to gain greater insight into the differences between groups. The breakdown by race largely aligns with the commonly held belief that race is a strong proxy for income levels in South Africa, where income levels are also often a proxy for the quality of school education obtainable as South African state schools are not all free. The gender breakdown aligns with commonly reported statistics on performance and correlates with an appreciation of the gender gap in education.

2.4.1 Progression

Success in HE can be viewed from a number of perspectives as there is no single defined and universally accepted measure of student success. Universities use a variety of indicators as they can interrogate their own data at the individual student level. By using multiple measures one can cut across a static representation of the data to obtain a more detailed and nuanced picture of the student experience over time.

The most common indicator used worldwide is the number of graduating students. This is important because it allows South Africa to be compared to other countries (as noted when commenting on the low historical completion rate of students in South Africa. While this is significant, it remains critical to know how students move through the system. This leads to the next measure, success rates.

2.4.1.1 Course success rates

South African universities are required by law to measure and release indicators of success (Council on Higher Education, 2018). Course or student success rates are one such measure of success and an established primary indicator of student progression.[15] Course or student success rates measure the percentage of courses that a student has successfully completed in any given year. Nationally, a benchmark of 80% was set by the DHET, indicating that students should pass a minimum of 80% of all enrolled courses in any given year of study to maintain the highest chance of meeting the minimum requirements for award of the qualification within the $n + 2$ period. The $n+2$ 'rule' states that if a degree programme is listed as a minimum of 3 years, students are expected to complete the requirements for the award of the degree within 5 years. Similarly for 4-year degree programmes, expected completion is within 6 years of first registration. Universities in turn set progression rules to encourage students to remain on track to meet this government-set criteria. Figure 4 shows a breakdown of course success rates by race to illustrate how students have fared relative to the benchmark.

The data in Figure 4 raises three important points. Firstly, at the national level, no race group exceeded the 80% benchmark prior to 2008. Since 2008, only White students have exceeded the national benchmark. Other race groups show upward trends in course success rates over time but have not exceeded the benchmark. Unfortunately, the graph does not allow for analytical breakdown of success rates by race across institutions. This analysis is very important in the context of highly unequal post-secondary institutions in South Africa.

The second important observation from Figure 4 is that African students show the largest improvements over the period when compared to the student body as a whole. White students

[15] The terms 'course success rates' and 'student success rates' are used interchangeably.

show the second-highest improvement over the period, but start from a much higher base. There was an almost 12% improvement in course success rates for African students between 2005 and 2017. Coloured students had the lowest improvement over time at approximately 8%. There was also a marginal narrowing of the differential between White and African students over the period. From a course success perspective, this narrowing in progression performance is small and amounts to approximately two percentage points. Of the three observations made about Figure 4, this is the least encouraging.

Figure 4 Student/course success rates by race: 2005–2017

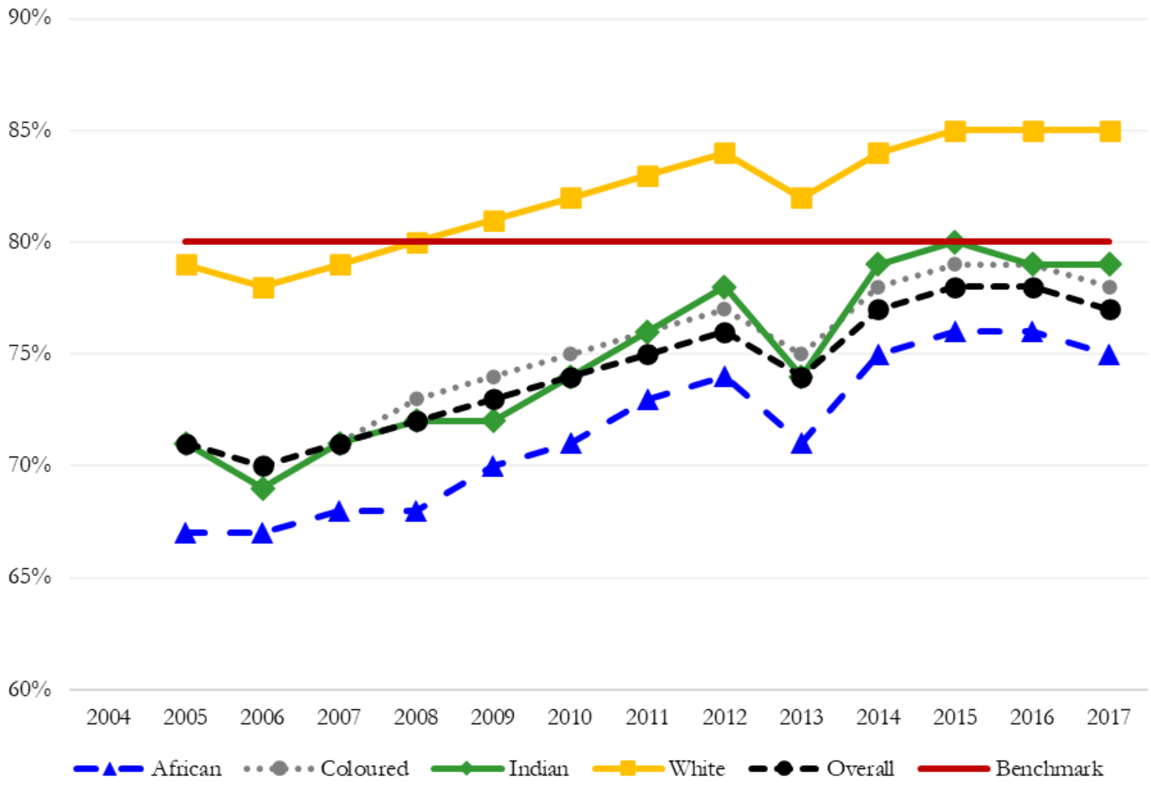

Figure 4 shows student/course success rates by race from 2005 to 2017. Data for 2004 was unavailable.
Data: Council on Higher Education (2012, 2015 and 2018)

The third observation from Figure 4 that is worth noting is that all race groups showed an upward trend over time. At first glance, this improvement in course success is indicative of an improvement in performance. On the other hand, it could be that standards have dropped at course level, translating into higher pass rates and better apparent performance. More cannot be said about the drivers of this improvement without further information.

2.4.1.2 Aggregate completions

A general measure of success is the aggregate number of qualifications awarded over time. Figure 5 shows the aggregate number of diplomas and degrees awarded between 2004 and 2018.

Figure 5 Degree and diploma level undergraduate completions (2004–2018)

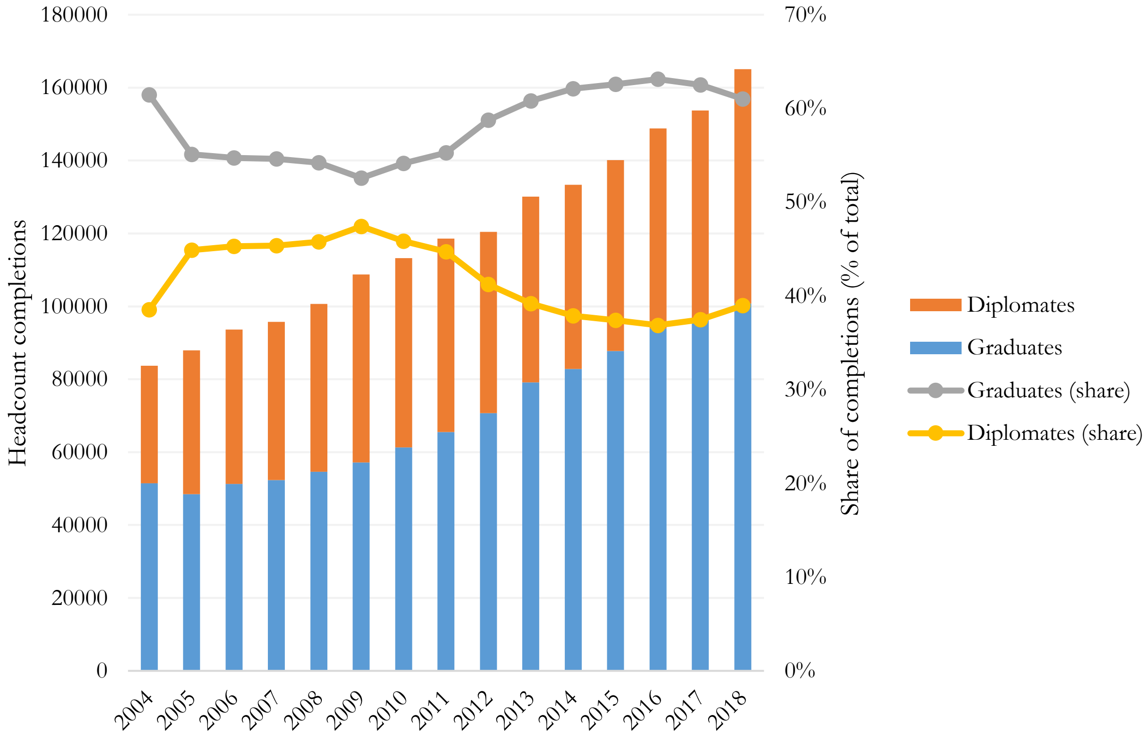

Figure 5 shows headcount degree and diploma undergraduate completions and shares of degree and diploma completions from 2004 to 2018.
Data: Council on Higher Education (2018) and Statistics South Africa (2018)

It is evident from Figure 5 that most qualifications awarded in South Africa remain university degrees rather than technikon-based diplomas.

While the previous section emphasised first-time enrolling students in the HE sector, this section focuses on students completing undergraduate degree or diploma qualifications. Figure 5 shows the number and share of degree and diploma graduates in South Africa between 2004 and 2015. Over this period, the number of students successfully exiting HE with a first qualification almost doubled, increasing from 83 665 total completions in 2004 to 165 086 completions in 2018, an annual growth rate of approximately 5%. Throughout the period, the number of graduates from universities exceed the number of diplomates from technikons or universities of technology.[16] Graduates include students completing 3-year undergraduate degrees, 4-year undergraduate degrees, or 4-year professional degrees.[17] Diplomates include students completing diploma level 1-, 2- or 3-year qualifications at either a university or technikon. As few students in South Africa complete qualifications in the minimum time allotted, the above graph is extended to 2018 to capture students who may still be in the system.

Between 2004 and 2009, the share of degree graduations steadily declined while the share of diploma graduations increased. From 2010 to 2016, this trend reversed, with a rising share of degree graduations and a falling share of diplomates. Figure 5 also confirms that the completion or graduation rate increased faster than the enrolment rate over the period. The information in Figure 5 supports the patterns observed in Figure 6, which shows that African students experienced relatively large improvements over the period but not enough to unequivocally catch up with any other race group.

While aggregate information like that given in Figure 5 is illuminating, it is much more informative to dig deeper into the data to evaluate other measures of success. One of the most revealing measures is completion rates by race group.

[16] Diplomates are individuals who meet the requirements and qualify for the award of a diploma-level qualification. While graduates tend to come from universities only, diplomates may come from either a university or a technikon.

[17] Professional degrees are qualifications in the fields of Engineering and Actuarial Science.

2.4.1.3 Completion by race

Student completion or graduation rates in South Africa remain extremely low, especially when broken down by race. The numbers also do not fare well when compared to similar middle-income countries (Roser & Ortiz-Ospina, 2013).

Figure 6 shows cumulative graduation rates by cohort, broken down by race. The information was adapted from a key 2019 report by the DHET. The graphs show the cumulative graduation rates for each cohort between 2004 and 2015. Each graph has the same y-axis to facilitate ease of reading and interpretation. Students are tracked for up to 10 years when calculating completion rates. This is the DHET's preferred time frame for following students, especially when it can be shown that many students do not enter HE in the year immediately after completing high school.

Figure 6 clearly shows that completion rates increase at relatively rapid rates until year 4 or 5 and that this may differ by race group. Thereafter, the rate of increase in completion rates slows down, indicating that the $n + 2$ rule has some impact from year 6 onward. This is true across all race groups. It is also evident that white students experience higher graduation rates from year 3 relative to the other three race groups.

The most fascinating observations in this data lie within race groups rather than between. On average, African students experience the lowest completion rates in undergraduate studies, and White students experience the highest. This has been attributed to many things, the predominant amongst these being a lack of academic preparedness that hampers performance from the beginning of African students' academic careers. For each subsequent cohort in the data, the completion rate rises steadily. This is clear when looking at later cohorts data, for example, when comparing the 2011 cohort to the 2004 cohort.

Figure 6 Cohort cumulative graduation rates by race

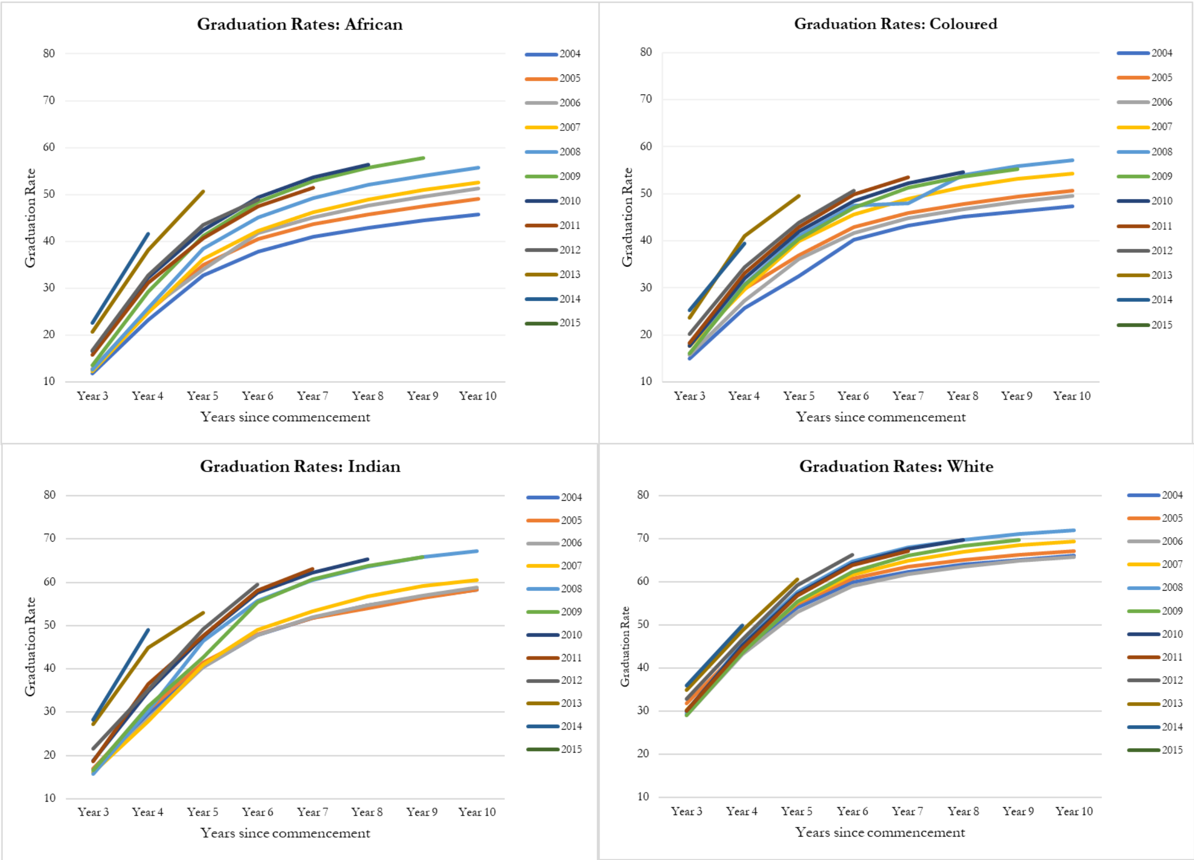

Figure 6 shows cumulative graduation rates by race group. Each line per chart represents a cohort from 2004 to 2015.
Data: Adapted from DHET (2019)

Most South African students register for a 3-year undergraduate qualification. This makes year 5 the first important point at which to analyse the data in Figure 6. Many, if not all, institutions impose an *n + 2* rule on student progression. Students are required to ensure that they complete any qualification in no longer than the minimum formal time or duration of the qualification plus two additional years. Since 4-year qualifications comprise a much smaller percentage of all qualifications, the most useful metric remains year 5. An analysis of intra-group performance in year 5 shows that African students experienced the highest increase in success rates, at approximately 54%, compared to white students, who experienced a 12% increase over the same period. The improvement for Coloured students is similar to that of African students at 51%. However, when the success rates are broken down further, the impressive improvement in graduation rates is not sufficient to raise the 10-year graduation rates for Coloured and African students above the year 5 graduation rate for white students.

Without additional information, it is difficult to pinpoint the reasons for the observed convergence of success rates. A few potential reasons have been suggested, such as increased extension of student funding (including financial aid), grade inflation across all institutions, and institutions admitting significantly stronger students compared to earlier cohorts.

The key finding from the data in Figure 6 is the convergence of graduation rates between African and White students over time. By 2013, the African 5-year graduation rate is 50.6%, compared to 54.1% for white students in 2004 and 60.6% in 2013. The substantial catch-up exhibited by African students is indicative of a decline in another, arguably more important academic outcome, that of dropout.

2.4.2 Dropout

In line with the general patterns displayed in the previous section, Figure 7 shows that there has been a steady decline in the number of dropouts over the period. This is shown by the dropout curve for each cohort successively shifting upward over time. This is also partly mirrored by steady increases in the number of students graduating each year. Evidence for this is shown in Figures 6 and 7 via the different outcome measures that are tracked by the DHET.

An interesting standout from Figure 7 is that African students do not necessarily experience the worst dropout rates from HE in South Africa over an extended period of tracking. At an aggregate level, the data suggests that Coloured students experience the highest dropout rates when the same cohort is tracked over time or the general trend evaluated for each subsequent cohort. However, many HEIs monitor the 6-year dropout rate as this is correlated with the 6-year graduation rate. On this measure, African students have the lowest graduation rate and the highest dropout rate after six years. In the period 2004 to 2015, there was a noticeable 4.6% decline in the differential between the dropout rates of White and African students.

Figure 7 Cohort cumulative dropout rates by race

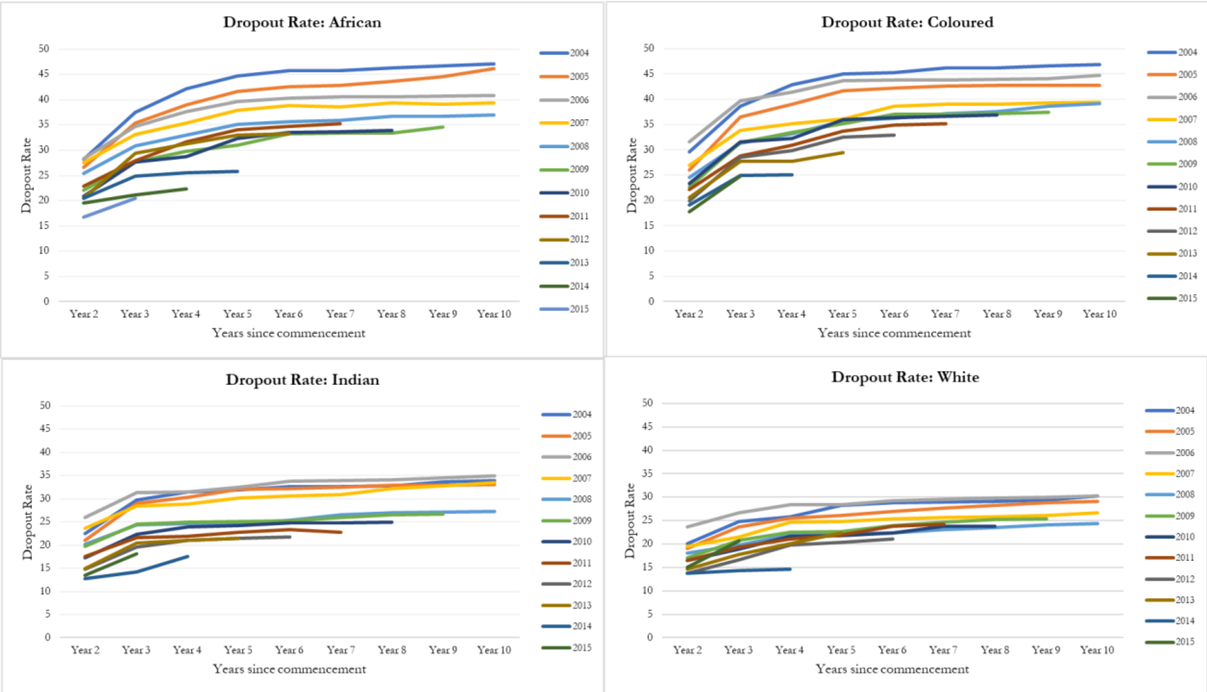

Figure 7 shows cumulative dropout rates by race group. Each line per chart represents a cohort from 2004 to 2015. Data: Adapted from DHET (2019).

Importantly, the 6-year dropout rate of African students fell by 12.5 percentage-points while the 10-year dropout rate fell by 10 percentage-points. This is an impressive outcome for a system that continues to display inequities in admission and performance. Coloured students displayed a similar trend for the 6-year dropout rate but had slightly higher 10-year dropout rates than African students.

An important distinction, absent from the aggregate dropout data is the separation of dropout statuses into voluntary and involuntary dropout. HEIs have set rules against which student progression is measured each year. When students fail to meet these minimum requirements, they are involuntarily excluded from that institution. Detailed figures concerning students refused re-entry to South African universities on academic grounds is not publicly available but could shed considerable light on the South African HE dropout problem.

2.5 Conclusion

A large body of research shows that the benefits of HE accrue both to the individual and to society at large. For a country like South Africa, defined by high levels of poverty, inequality, and unemployment, the unequal nature of access to and participation in HE ensures that unequal access to more lucrative jobs and income opportunities persists until greater equity is achieved in both access and success in HE. The South African HE landscape is characterised by low participation and high dropout rates. The DHET shows that approximately 45% of all students who enrol in HE exit the system with a qualification. This is below regional and international standards and ensures that South Africa continues to lag behind its peers.

This discussion has highlighted the problem that the capacity constraints of the HE sector have prohibited the expansion needed to allow significantly more students access to HE. In addition, the negligible increases in enrolment over the period, despite the addition of three new universities, show a sector that is struggling to expand and provide opportunities to more individuals.

Once students are in the system, the differentials in performance between subgroups have narrowed over time. White students have the highest student success rates, exceeding 80% over the period. Other subgroups do not catch up to these rates but do show some improvement. When graduation rates are evaluated, White students again display the highest rates but other subgroups, especially African students, show improvements in performance over time. Lastly, the decline in dropout rates is discernible in the data, with a trend analysis showing a fall in the dropout rate for all students.

This chapter provided evidence for persistence in but slower progression through HE. Successively higher course success rates lead to improved graduation and dropout outcomes for all students. The rate at which the identified subgroups progress through the system is different, but the evidence suggests that there is a decrease in the African-White differential at the aggregate level.

Chapter 3: Miss(ing) the Mark: Recognition Incentives and Student Outcomes

3.1 Introduction

Many universities implement policies of reward or reprimand as a means to ensure students meet their academic continuation requirements from year to year. These policies can take the form of static recognition (warning) during or at the end of the academic year, or dynamic warning systems that alert students to a below-expected level of performance throughout the academic year. Some university administrators, without knowing the true impact of these policies, suggest that such incentive schemes generate an effort response, and by implication a performance response.

One such tool is the DML, a key measure used by HEIs worldwide to recognise and incentivise the academic performance of high-achieving students. The DML is an academic honour that is recorded on a student's official academic transcript. It is also known as the Dean's List in countries like the USA. In its simplest form, students may be recognised by the university or HEI for outstanding academic performance, usually above some specified threshold. Some universities advertise these honour rolls to enhance and encourage future student performance.

There are two potential explanations for its implementation. The first is simply to reward high-achieving students whose performance exceeds a given standard. A recognition of academic excellence is added to the academic transcripts of such students and noted officially via a letter to the awardee. This list is updated annually. Students may appear on the list one year but not meet the requirements in subsequent years. They may reappear in later years, depending on course and overall performance.

The second reason is as an aspirational goal for students who are good, but still below the merit list cutoff. The hypothesis is that such students, informed by classmates of the recognition that occurs just above the threshold, may be incentivised to increase their effort in subsequent years to gain access to the group of students recognised by the university.

Despite the popularity of the DML system around the world, the literature on its effects on student performance and academic outcomes remains sparse. There are two reasons for the lack of evaluation. Firstly, the lack of data in HE is widespread. Many HEIs do not have sufficiently detailed individual-level records due to poor record keeping practices. Within this, data records are often incomplete, and missing data is a key challenge to evaluating policies and student outcomes. Secondly, many studies are largely descriptive or correlation-based in nature, failing to account for group differences or selection into the award of the DML. Fortunately, with the emergence of techniques that assist in identifying causal effects, the volume of research in this area has started to increase.

This chapter uses a sharp regression discontinuity (RD) design where students who earn GPAs just above a given threshold are compared to students whose GPAs were just below it. Therefore, a good counterfactual is presented for a student just below the threshold who did not receive academic recognition and a student just above the threshold who has received the DML academic recognition. Given the level at which the reward is located, both groups of students are expected to perform relatively well, and the RD design allows the evaluation of any incremental impact of the policy in addition to the already good performance demonstrated by the students (Fletcher & Tokmouline, 2010).

This chapter finds significant negative effects on short-term performance but considerable heterogeneity across students for long-term effects. As the university studied in this chapter offers 3- and 4-year degree options, the results are split to show the differences between the two different degree durations. Humanities, male, 3-year degree and low entry score students responded the most negatively to the policy. An important aspect of these findings is that these effects do not fade over time but increase in intensity between the second year of enrolment and the graduation year, depending on the programme of study. For longer-term outcomes such as graduation there are very small but statistically significant negative effects. Students above the threshold in their first year are more likely to be involuntarily excluded from the university in subsequent years, which is an interesting observation given that these are high-achieving students in their first year. Overall, the findings suggest heterogenous effects with no consistent patterns across observable characteristics. None of the heterogeneous effects was consistent across faculties, suggesting that the application of strict cutoffs or absolute scores as incentives is limited in scope and impact in the context in which they are examined.

This chapter makes several contributions to the literature. Firstly, this chapter contributes to the literature on gender and racial differences in response to educational incentives and recognition. In countries where student performance is less understood, especially at the post-secondary level, this model provides some insight into causal mechanisms that might (dis)incentivise students to increase their academic performance. This information may be of value to academic administrators who are uncertain about performance thresholds for short-term recognition. This chapter also contributes to work on the effects of reward or recognition programmes on developing country HE students, and specifically South African students. In the only such developing country study in the education literature, Wright (2018) studied students in Jamaica and finds that incentive policies had significantly positive effects on student performance. However, most of students in developed countries, and there is little evidence available for developing country contexts where education completion rates are much lower.

The rest of the chapter is organised as follows: Section 3.2 reviews background literature on academic probation and reward; section 3.3 describes the data and institutional setting, including a detailed discussion about the DML programme; section 3.4 describes the empirical strategy; section 3.5 presents the results; and section 3.8 concludes with a discussion of the research implications.

3.2 Background

As previously mentioned, despite the popularity of the DML system around the world, the literature relating to its effects on student performance and academic outcomes is not well developed. The lack of data in HE is widespread as many institutions do not have exemplary data management systems. Many studies are also largely descriptive or correlational in nature, and fail to account for group differences or selection into award of the DML, and are thus unable to identify causal effects of such policies. There is, however, much literature examining the effects of university policies, such as academic probation (Fletcher & Tokmouline, 2010), affirmative action (Massey & Mooney, 2007), financial aid (Goldrick-Rab et al., 2012; Dynarski, 2000; Dynarski & Scott-Clayton, 2013), merit aid (Leeds & DesJardins, 2015; Dynarski, 2002; Dynarski, 2004) and student advising (Bettinger & Baker, 2014), on student performance and outcomes. While the literature focused on these topics has continued to grow, the literature on policies that recognise annual academic performance has stagnated, largely due to a lack of data or to the manner in which such policies have been implemented. The research examining the effects of student reward or recognition that does exist is mostly correlational, comparing student outcomes for those who have received academic reward and those who have not (Crandall & McGhee, 1968). Comparing two groups whose observed and unobserved characteristics vary significantly leads to biased results as the two groups are not directly comparable along any given dimension.

Thistlethwaite and Campbell (1960) was one of the earliest papers to conduct a study of the impact of merit awards on future academic outcomes using a regression discontinuity design. Examining the effect of increased public recognition, the authors found that recipients were more likely to receive future academic scholarships and pursue post-graduate studies but that no effect was found for impact on career plans.

Seaver and Quarton (1973) also conducted a causal study on the effect of the DML on subsequent academic performance. Using an RD design the authors find that DML students achieved higher than expected GPAs in the subsequent term relative to non-DML students. This effect held true for annual GPA and cumulative GPA in the semester immediately following the award of recognition as well as the term after that. While they found the DML award had positive short-term and long-term effects on student performance, their study did not investigate any heterogeneity of outcomes.

Wright (2018) presents another comprehensive analysis of the impact of recognition and probationary policies on student performance. Using data from a large Jamaican university, the author finds that the outcomes are sensitive to the design and intensity of the programmes' implementation. Specifically, academic recognition policy effects are heterogeneous across the student body with students in social sciences responding differently than students in the natural sciences. The author finds statistically significant, strong short- and long-term responses to the DML policy in the social sciences for almost all student outcomes but does not find many significant effects in the medical, pure or applied sciences.[18] To explain these differences, the author evaluated students' course selection. The findings suggest that this significant result is driven by both strategic course taking and an increase in effort as the improved results could lead to tangible benefits such as access to financial assistance. The author does not show if there are any gender effects.

[18] This includes GPA, cumulative GPA, dropout and graduation.

3.3 Conceptual framework

The conceptual framework follows Lindo, Sanders and Oreopoulos (2010), in applying Bénabou and Tirole's (2000) model of individual responses to incentives. The model is implemented as an interaction between a principal (manager, teacher) and an agent (worker, student). The principal has the ability to set performance standards for the agent. Similar to Lindo, Sanders and Oreopoulos (2010), the model is presented from the perspective of the agent (student). The predicted outcomes of the model are then related to academic rewards.

Academic awards for excellence may be viewed as powerful social reinforcers of behavior with large, positive impacts due to the impact it has on recipients. It has been noted that an award may work through a number of channels. These include, improving students' self-evaluations, encouraging students to increase effort in subsequent terms, and improving the social-academic climate (Seaver and Quarton, 1973).

Consider an agent facing a given set of two options. Each option i incurs a private cost c_i and if completed successfully generates benefit V_i to the agent. If the agent fails or chooses not to engage, the net payoff is 0. Option 1 is the easy option with low potential benefits. Option 2 is the difficult option with high potential benefits. Benefits and costs may be summarized as

$$0 < V_1 < V_2 \qquad\qquad 0 < c_1 < c_2 \tag{3.1}$$

It is assumed that the probability of success is the same for both options and given by θ. It is further assumed that

$$\theta_1 \equiv \frac{c_1}{V_1} < \theta_2 \equiv \frac{c_2 - c_1}{V_2 - V_1} < 1 \tag{3.2}$$

The agent is fully informed and knows θ. The agent therefore solves

$$max\ \{0, \theta V_1 - c_1, \theta V_2 - c_2\}.$$

The agent therefore shirks or chooses neither option if $0 \leq \theta < \theta_1$, chooses option 1 if $\theta_1 \leq \theta < \theta_2$ and option 2 if $\theta_2 < \theta \leq 1$. Consider the situation where the principal makes option 1 less attractive by imposing additional costs on the agent, resulting in the agent choosing option 2 if

$$\theta \geq \frac{c_2}{V_2} \equiv \theta^*, \tag{3.3}$$

with $\theta_1 < \theta^* < \theta_2$. If option 1 remained a possibility for agents who would otherwise choose option 1, the standard either makes agents of the ambitious type $[\theta^*, \theta_2]$ work harder or agents of the weaker type $[\theta_1, \theta^*]$ give up.

The model can also be used to show the effect of positive incentives, such as student responses to academic rewards. Consider two students whose first year non-cumulative annual GPAs are near the DML threshold with one above it and on below it. The student just above the threshold faces what the university considers to be academic encouragement, rewarding the students with a positive incentive. Similar to Lindo, Sanders and Oreopoulos (2010), in the following year each of the two students has three options: return to university aiming to achieve a low GPA (option 1), return to university aiming to achieve a high GPA (option 2), or not return at all (neither option 1 nor 2).

The model therefore allows for the following testable outcomes:

Weaker students are more likely to drop out, independent of option 1

Stronger students return and continue to excel academically (option 2)

Bénabou and Tirole (2000) argue that incentives are weak reinforcers in the short run and negative reinforcers in the long run. The short run in this application of the model is the academic year immediately after the year of recognition or reward, i.e Year 2. The long-run is the cumulative GPA over time, and exit outcomes of graduation, voluntary and involuntary dropout.

Within this theoretical framework Bénabou and Tirole (2000), the authors discuss the notion of coasting. Students who develop high self-confidence from the receipt of an award may reduce effort in subsequent periods due to the reward. In this instance, effort and ability become substitutes, rather than complements. It may also be true that students take the signal from the award and direct their efforts along new and risky paths. In the context of this chapter, the student may change programmes towards easier programmes due to elevated feelings of self-confidence.

3.4 Data and Institutional Setting

3.4.1 Data

The data for this chapter comes from administrative records at UCT in South Africa. For each year a student is enrolled, data is available on their demographic characteristics, programme and duration of registration, individual course enrolment, academic performance, financial aid, and student residence status. The analysis was restricted to the entering cohorts of 2006 to 2008 to take advantage of the uniformity of entry requirements during that time period. This allows them to be cleanly tracked through to their exit from the system. The Health Sciences (medical) faculty was specifically excluded from this analysis as its entry requirements are significantly different from other faculties. Students who completed their high schooling abroad and those for whom there are missing variables of interest were also excluded. The analysis focuses on those individuals within 6 years of their initial entry. This accords with the national norm in South Africa, which focuses on academic performance within six years of registration.[19] This also facilitates comparability with international studies, which focus on a similar period following enrolment. Lastly, the sample was limited to those within the optimal bandwidth of the DML requirement. This effectively excluded any students who failed a course in the year in which the DML is awarded and students who missed the DML requirement by significant margins as these students would not be eligible for the DML, and more than likely, differ significantly on observed characteristics from those who do meet the requirements for the award of the DML.

First-time entering students are exposed to a detailed orientation programme before commencing their formal studies. This orientation programme usually includes the provision of important university documentation such as handbooks and the actual registration process. Students apply to and are accepted into degree programmes that are typically very structured and largely pre-determined. The least structured degrees are in the Humanities faculty where students have significant freedom when selecting their majors and subjects while meeting less formal programme restrictions within the overall degree structure. Most students register for 3-year degrees. Their course loads tend to vary with their faculty of registration as each faculty is free to determine the level of work contained in each course and degrees they offer subject to national standards. The Commerce faculty has the highest course count requirements for the award of the degree, and the Science faculty has the lowest course requirements. However, the work intensity tends to be higher in Science than in other faculties. The Humanities faculty is somewhere between the Commerce and Science faculties, depending on the programme of choice. Other popular degrees are typically professional 4-year degrees, which usually include a fourth year of study that is accessed as an honours-level degree for students completing 3-year degree programmes. For many degrees offering a 4th year of study, access and progress

---

[19] This assumption is similar to Casey et al (2018) where graduation is defined as ‘within 6 years of entry’.

from 3rd year to 4th year is not automatic and there are minimum grade requirements being required before students are admitted to the final year. This progression incentive means that for many disciplines a jump in performance may be observed in the 3rd year, on the part of those students attempting to gain entry to 4th year programmes.

Students' GPAs are recorded on their academic transcripts for each year of registration. In South Africa, grades are awarded on a scale of 0–100. Students' annual non-cumulative GPA is calculated as the average of the grades for a given academic year of study (AYOS) weighted by the credit weighting of the courses registered for in a particular year. Credit weightings differ by faculty due to faculties having autonomy to set their own credit requirements, but they ostensibly imply the same required amount of hours by the level of study.

At the University of Cape Town, students with an unrounded weighted average non-cumulative annual GPA of 70% or more are recognised for their academically excellent performance by having the DML award added to their academic record.[20] The key criteria to be recognised with the DML are not having failed a course in the year of interest; completing a full course load, unless students are in their final year that does not require a full load; and completing all courses during the standard term.[21] Students receive notification that they have been placed on the DML from faculty offices via official email correspondence. Because many courses span the entire year, students' academic records are usually evaluated at the end of the academic year and not on a semester basis as may be the case elsewhere.

[20] A 2/1 aggregate or better

[21] This means that students who enroll for summer or winter term courses to spread their academic loads may not be rewarded with the DML.

3.4.2 Descriptive statistics

As previously indicated, the data used in this analysis is from an administrative dataset from UCT that included race, gender, age, home language, programme of registration, home faculty, financial aid, and residence status.

The main sample in this analysis comprises students whose annual GPAs are two percentage points above and below the cutoff of 70%. In addition, for completeness and accuracy of results, the sample was expanded to include a number of different bandwidths. The widest of these was the data-driven bandwidth based on Cattaneo, Jansson and Ma, 2018. The CJM method is chosen for the bandwidth selection as it is data driven and does not assume any functional form of the data. The recommended bandwidth based on the CJM method includes students 5.179 percentage points below the cutoff and 6.643 percent above the cutoff. The students in the three cohorts (2006–2008 entering years) were followed for a maximum of n + 2 years for this analysis, which is in line with national tracking. Students were omitted if their data were missing for any variables of interest.

Descriptive statistics of student characteristics are shown in Table 1. The first column of Table 1 shows the descriptive statistics for the full sample. Approximately 56% are female, 61% are white, 35% are from the Commerce faculty, and 57% are enrolled in 3-year degrees. The average entrance score, the proportion of students with financial aid, and whether students live in student residences or not are fairly consistent across all bandwidths. The same pattern is noted for students enrolled in extended academic programmes.

To ensure that the results from this analysis are statistically sound, a number of robustness checks were performed through the chapter. The analysis proceeded in a step-wise manner by first examining the effects of first-year DML recognition on exit outcomes. Once exit outcomes have been examined, annual and cumulative GPAs are evaluated as outcomes. The long-term effects of the year 1 DML award on year 3 and 4 performance (where available) are also investigated.

Table 1 Descriptive statistics by bandwidth

	(1)	(2)	(3)	(4)	(5)
Covariates	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Female	0.53	0.53	0.53	0.54	0.56
	(0.49)	(0.49)	(0.49)	(0.50)	(0.50)
African	0.17	0.17	0.18	0.18	0.19
	(0.38)	(0.38)	(0.38)	(0.39)	(0.39)
White	0.62	0.62	0.62	0.62	0.61
	(0.48)	(0.48)	(0.48)	(0.48)	(0.48)
Coloured	0.12	0.12	0.12	0.12	0.11
	(0.33)	(0.33)	(0.32)	(0.32)	(0.32)
Indian/Asian	0.06	0.06	0.06	0.06	0.07
	(0.24)	(0.24)	(0.24)	(0.24)	(0.26)
Commerce	0.34	0.35	0.35	0.35	0.35
	(0.47)	(0.47)	(0.48)	(0.48)	(0.48)
Humanities	0.35	0.35	0.35	0.35	0.35
	(0.47)	(0.47)	(0.48)	(0.48)	(0.48)
Science	0.11	0.11	0.11	0.11	0.12
	(0.32)	(0.32)	(0.32)	(0.32)	(0.32)
Engineering	0.17	0.17	0.16	0.16	0.16
	(0.37)	(0.37)	(0.37)	(0.37)	(0.37)
3-year degree	0.57	0.57	0.58	0.58	0.57
	(0.49)	(0.49)	(0.49)	(0.49)	(0.50)
4-year degree	0.42	0.42	0.42	0.42	0.43
	(0.49)	(0.49)	(0.49)	(0.49)	(0.49)
Private school	0.38	0.38	0.38	0.38	0.38
	(0.48)	(0.48)	(0.48)	(0.49)	(0.49)
House of Delegates	0.02	0.02	0.02	0.02	0.02
	(0.15)	(0.15)	(0.15)	(0.14)	(0.14)
House of Representatives	0.05	0.05	0.05	0.05	0.05
	(0.21)	(0.21)	(0.22)	(0.22)	(0.22)
Department of Education and Training	0.06	0.06	0.06	0.06	0.06
	(0.23)	(0.23)	(0.24)	(0.24)	(0.25)
Cape Education	0.54	0.54	0.54	0.53	0.53
	(0.49)	(0.49)	(0.50)	(0.49)	(0.50)
Entance Score	42.30	42.17	42.2	42.38	42.29
	(8.36)	(8.51)	(8.56)	(8.43)	(8.76)
Financial Aid status	0.08	0.08	0.08	0.08	0.08
	(0.27)	(0.27)	(0.27)	(0.27)	(0.27)
Residence Status	0.31	0.31	0.31	0.31	0.28
	(0.46)	(0.46)	(0.46)	(0.46)	(0.45)
Extended Programme	0.09	0.09	0.09	0.09	0.09
	(0.28)	(0.28)	(0.29)	(0.29)	(0.29)
Observations	2153	1989	1686	1264	858

Notes: Columns 1-5 show student characterisitcs for the different bandwidths applied in this chapter. For each variable, the mean is presented with the standard deviation presented in parantheses below.

3.5 Empirical Design

3.5.1 Validity of the regression discontinuity design

The main empirical strategy for the analysis in this chapter was RD. The RD approach provides causal estimates of the policy's impact on student outcomes. RD approaches allow for the estimation of local average treatment effects (LATE), where observations on either side of the threshold or cutoff are most similar. This facilitates the comparison of treatment and control groups around the cutoff. Imbens and Lemieux (2008) provide a detailed discussion on the practical implementation of the RD approach for causal identification and argued that three criteria must be met for any RD study to be valid. The first condition that must be satisfied is there should be a discontinuous change in the allocation of the assignment of treatment. To qualify for the DML, a student must satisfy the raw, unrounded GPA requirement of 70% or more, with no failures. The second criterion to be satisfied is that of local randomisation (Lee & Lemieux, 2009). Put simply, any differences in observable characteristics within the neighbourhood of the discontinuity must be continuous through the cutoff. To evaluate this assumption, the balance on covariates was evaluated to check for discontinuities in their distribution around the cutoff. The results for this test are presented in Table 2. The results show no significant differences for any variable through the cutoff, supporting the validity of the RD design. The third assumption to be met is that there should be no manipulation of the running variable in the region of the discontinuity, in this instance, the GPA around the award grade. Fortunately, this assumption seemed to be satisfied. Courses at the university are graded separately, and there is no convention of setting grades to fit a curve. More importantly, 70% is not a target in any course listed in the university handbook, although 75% (the lower boundary of a first-class pass) may be since it is the highest broad class of pass students can achieve at UCT.

Given the 'sharp' nature of the discontinuity, the following equation was used to estimate the impact of the DML award after the first year on student outcomes:

$$Y_i = \theta \, GPA_i^{year1} + \delta_1 \left(GPA_i^{year1} > GPAMIN \right) + u_i \tag{3.4}$$

where Y_i is an outcome for student i. In this analysis, Y_i is one of two measures for GPA: non-cumulative or annual GPA and cumulative GPA, or one of three exit outcomes namely graduation, voluntary dropout or non-voluntary dropout. $\theta(GPA_i^{year1})$ is a continuous function of students' non-cumulative first-year GPA; $1\left(GPA_i^{year1} > GPAMIN\right)$ is an indicator variable equal to one if the student's non-cumulative GPA is above the award grade; and u_i is a random error term. The coefficient of interest is δ_1, which is the estimated impact of being recognised by being placed on the DML.

The issue of bandwidth remains pertinent for this type of analysis. The data-driven methodology suggested by Cattaneo, Jansson and Ma (CJM) (2018) provides bandwidths that are very wide, their suggested upper bound crossing into the highest class of pass. The key reason for choosing bandwidths so much narrower than those derived from the CJM methodology, is the boundaries that the university uses for its classes of pass. The upper limit of the CJM bandwidth crosses over into a new grade category that is related to the cumulative performance of students and is recognised at the point of graduation.[22] Intuitively, this would create confounding effects that could be difficult to disentangle from the DML effect.

To test the sensitivity of the findings to the bandwidth selected, and to ensure that robustness checks were simultaneously carried out, the analysis was replicated using four bandwidths, all of which are narrower than the recommended CJM bandwidth. In choosing the

[22] A first-class pass is awarded to students who achieve an aggregate GPA of 75% upon graduation.

final bandwidths, points from the cutoff were evaluated, and students falling below the GPA value of 65.2 and above a GPA value of 74.8 were removed. As a further robustness check, estimates for bandwidths of 4, 3 and 2 percentage points around the cutoff were shown with the main results to demonstrate that the results are not sensitive to the choice of bandwidth. This robustness check was built into the analysis and presented with the main findings where possible.

Importantly for RD analysis, sample sizes should be sufficiently large around the cutoff to allow RD approaches to closely approximate randomized experiments (Cattaneo et al, 2019). In calculating the optimal data-driven bandwidths, the CJM method yields a bandwidth that is wider than expected, but this width ensures large enough sample sizes to warrant the validity of the results of this analysis. Therefore, while the remaining narrower bandwidths are presented as robustness checks, not every specification within the remaining four bandwidths meet the criteria for inference based on RD.

Further extensions present the results for the long-term impact of a DML recognition in first year. This is because a DML might have been awarded in first year only, with non-cumulative GPAs in later years falling below the cutoff of 70% for DML recognition. In addition to the long-term effects of the DML award, the impact of *ever* being awarded a DML was of interest.[23]

A key concern in any RD study is manipulation of the running variable of interest since this could result in significant discontinuities at the cutoff or threshold of interest. In the context of this study, this might have occurred if students were able to manipulate their GPAs through strategic course-taking or if they could forecast their GPA with some certainty (Casey et al.,

[23] Due to data unavailability, the impact of ever being awarded a DML is not considered in this study. It is however an area for future research.

2018). If students were able to manipulate their GPA, it is likely that clumping would have been observed in the running variable at the threshold (Lee & Lemieux, 2009). The impact of this would have been extreme discontinuity at the threshold and consequently biased the estimated effects of the policy on student outcomes. Fortunately, given that the first-year curriculum in most faculties is fairly fixed, required courses for a chosen major or specialisation are pre-determined and the largest contribution to final course grades is performance in final course examinations. In such faculties, students have few opportunities to strategically choose courses to maximise their GPAs. The university also has significant systems in place to prevent fraudulent manipulation of grades (and hence GPAs). Students use an assigned student number rather than a name on examination scripts. This means that should instructors mark examinations, they will not be able to easily identify specific students and thereby award grades that do not reflect the actual examination performance. In addition, a sample of each course's examination scripts are externally examined to ensure that all marking has been completed fairly and consistently, and to ensure that standards are maintained over time.

Casey et al. (2018) similarly argue that if students can manipulate their GPA through strategic course selection, it will be observable in the density of GPAs around the threshold. If students are able to engage in strategic course taking, the distribution of GPAs would be discontinuous just below and just above the threshold, with very few observations below the threshold and many observations just above. When interpreting this discontinuous allocation of GPAs, it would have been simple to assume that students just above the threshold are significantly more motivated than students just below the threshold. However, the inability to exert significant influence over the choice of subjects largely prevents this from occurring at UCT.

3.5.2 Satisfying the regression discontinuity assumptions

To confirm that GPA manipulation was not a feature of the results, manipulation tests were conducted to search for evidence of a discontinuity in the density of the running variable, which was in this instance the GPA at the 70% cutoff. A statistically significant discontinuity in annual GPA at the cutoff would have implied GPA manipulation by students in the form of non-random selection or self-selection into courses, and therefore, into the control and treatment groups in the analysis (Cattaneo, Jansson & Ma, 2018).

To test for manipulation of annual GPA, the procedure outlined by Cattaneo, Jansson and Ma (2018) was followed. The test for manipulation involves estimating the discontinuity in the density function of the running variable (McCrary, 2008). The results of the manipulation tests were based on data-driven optimal bandwidths using local polynomial density estimation for RD analysis. Figure 8 shows the local polynomial density function plot, including confidence intervals. Plotting it allowed the data to be visually and statistically examined for evidence of discontinuity or heaping of the annual GPA variable. At the DML cutoff of 70% there is a small decrease in the density of GPA. The estimated discontinuity of the density is not statistically significant. All manipulation test results indicated that the null hypothesis of no manipulation could not be rejected[24].

[24] The manipulation tests are robust to the bandwidth selected or recommended via the data-driven process.

Figure 8 DML density test and probability of treatment: Whole sample

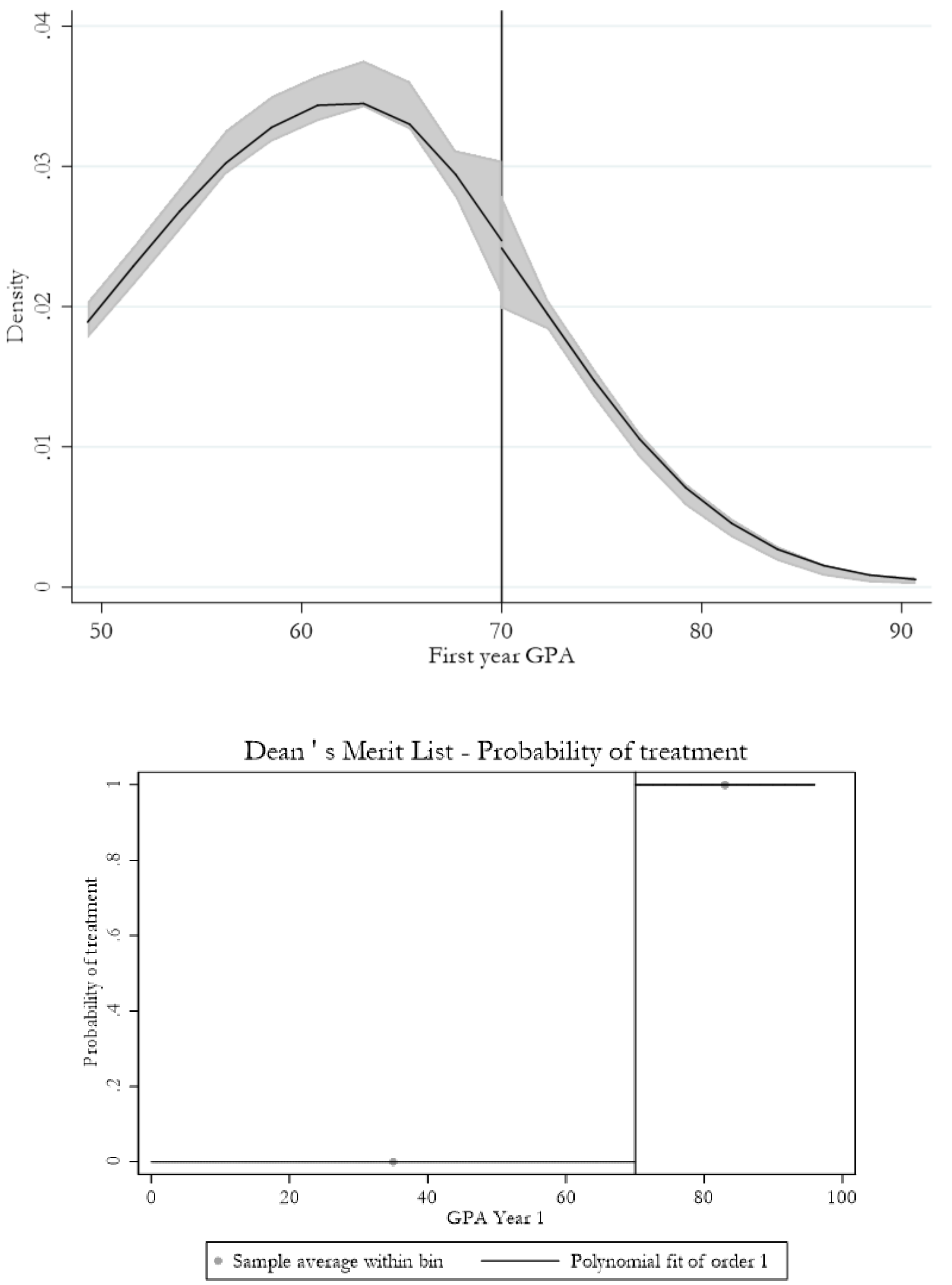

Figure 8 shows the density distribution of first-year GPA through the 70% DML cutoff. The bottom panel shows the probability of treatment.

The density tests presented in Figures 8 and 9 support the causal interpretation of the RD results. The statistics showed no significant discontinuity, and eyeballing Figure 8 confirms this. Significant consideration was given to the choice of bandwidth as very broad bandwidths are likely to distort the analysis and produce less credible results while too narrow bandwidths will provide too few data points for valid estimation. The bottom panel of Figure 8 shows the probability of treatment. The jump from 0 to 1 is very strong at the GPA threshold of 70, aligning with the sharp RD framework.

Figure 9 shows the density tests by degree duration. The university studied in this thesis offers both three- and four-year general bachelor's degree. The characteristics of students who enroll in four-year degrees vary in terms of a few key characteristics.

From the manipulation tests, the optimal bandwidth for usage in the study was derived using the CJM methodology. The CJM methodology uses simple local polynomial density estimators. This avoids pre-binning the data which leads to improvements in size properties while also allowing for restrictions on other features, which leads to improvements in power properties (Cattaneo, Jansson and Ma, 2018). In addition, the CJM method yields sufficiently large population samples for estimation and inference purposes. The data-driven CJM method indicated a lower bandwidth of 5.179% below the cutoff and 6.643% as the upper bandwidth above the cutoff value of 70%. The student GPA is given as a number between 0 and 100, with the average student achieving a GPA in the mid-50s to mid-60s. Based on the grading system at UCT, students who achieve an average of 70% to 74.5% are inherently different from those who achieve GPAs above 75% as 70–74% represents an upper-second class of pass (also known as a 2+) and 75% and above represents a first-class pass (also known as a first). The RD manipulation tests also indicated that it would be sensible to conduct the analysis with varying bandwidth sizes to ensure the robustness of results as the size of the coefficients of the manipulation tests varied quite substantially across the five bandwidths (see Table 2).

Therefore all results presented commence with the CJM optimal bandwidth followed by researcher-selected bandwidths of BW(2) = 4.8, BW(3) = 4, BW(4) = 3 and BW(5) = 2.[25]

Figure 9 Density tests by degree duration

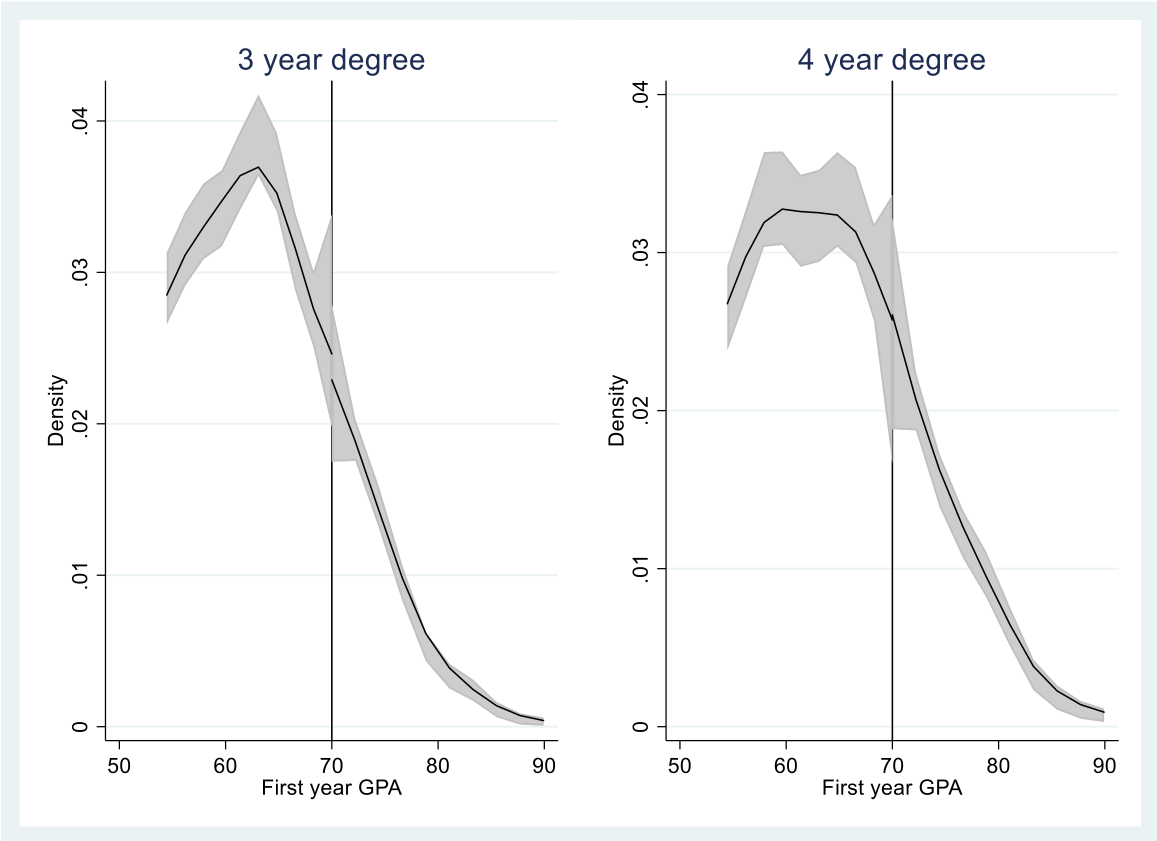

Figure 9 shows the distribution of first-year GPAs by degree duration.

A potential complicating issue that was affected by the choice of bandwidth was the explanatory power of the RD model. The model relies on a large number of observations around the cutoff to ensure validity of the results. Using individual cohorts leads to small sample sizes. To sidestep the risk of too small a sample size, the data was collected over three years and pooled. Combining the three annual cohorts of students provided a large enough sample in even the narrowest bandwidths.

[25] CJM is representative of the optimal bandwidth given by the Cattaneo, Jansson and Ma (2018) stata programme.

If there are sufficient observations around the cutoff, RD approaches may closely approximate randomized experiments. This is true when the treatment and control groups are equal in terms of all observed and unobserved characteristics. Table 2 shows the characteristics of students for each of the bandwidth choices. All characteristics, except female, entrance score, engineering faculty and African are statistically equal across the treatment and control groups. Note that for these characteristics, only one bandwidth per characteristics tends to show some discontinuity. Table 2 provides evidence to show that almost all of these observable characteristics are continuous through the threshold, except for the few differences mentioned above. Taken together, the different bandwidths and the power analysis support the presentation of the CJM results in all instances. The power analysis supports the use of all bandwidths when using the full sample, but not on the smaller of 3-year and 4-year degree. The power analysis is available in the Appendix, Tables 37-39.

As mentioned, substantial discontinuities at the DML threshold could have invalidated the research design, indicating that individuals with certain characteristics were driving the observed results. To test that covariates were continuous through the threshold, the procedure adopted by Casey et al. (2018) was followed to regress pre-determined student characteristics on GPA. These results were estimated via the following equation:

$$\mathrm{X_i} = \theta\ \mathrm{GPA_i^{year1}} + \delta_1\left(\mathrm{GPA_i^{year1}} > \mathrm{GPAMIN}\right) + u_i \tag{3.5}$$

The results are given in Table 2. A set of pre-determined characteristics were used and, there were no statistically significant differences across the DML threshold. Put plainly, there are no significant jumps in pre-determined characteristics across the DML thresholds, for all student characteristics, except in 3 different cases. Importantly, performance in the school leaving matriculation examinations (entrance score), which is a strong predictor of performance, was not discontinuous across the DML threshold. At the 10% level, at a bandwidth of ± 4.8 above and below the threshold, there is a significant difference by gender, females being more likely to be above the threshold. While a similar observation is noted with respect to the engineering faculty for the CJM bandwidth, residence status (bandwidth of ± 4.8) and entrance score (bandwidth of ± 2), for the remaining student characteristics the covariates are continuous through the threshold. However, none of the above invalidates the RD analysis and the main bandwidth of interpretation is the CJM bandwidth. As noted by Casey et al (2018), some of the point estimates in Table 2 may be small in magnitude implying that small imbalances across covariates may have an impact on potential outcomes when covariates are included in the analysis.

Another potential complicating factor could have been the re-enrolment of students in subsequent years. This could have affected the RD estimates if students above the threshold dropped out at significantly different rates than those with GPAs below the threshold. However, evidence suggested this is not the case, and it did not present a significant threat to the validity of the results.

Table 2 Validity of RD design: Balance on covariates

	(1)	(2)	(3)	(4)	(5)
Covariates	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Female	1.2395	3.1436*	2.1884	1.7085	0.2944
	[0.8724]	[1.1618]	[1.4714]	[2.2429]	[3.9321]
African	0.8432	1.4720*	0.3765	-0.7055	-1.0617
	[0.6651]	[0.8888]	[1.1450]	1.7554	[3.0769]
Coloured	-0.6899	-0.4033	0.0937	0.5944	2.6837
	[0.5767]	[0.7654]	[0.9529]	[1.4451]	[2.5719]
Indian/Asian	0.2094	0.2976	0.8915	2.6527	1.0064
	[0.4267]	[0.5699]	[0.7133]	[1.0821]	[2.0664]
Humanities	0.5144	1.4253	1.4157	1.0441	2.0929
	[0.8331]	[1.1105]	[1.4095]	[2.1584]	[3.7905]
Science	-0.0613	0.43773	0.0014	1.0886	4.1048
	[0.5614]	[0.7411]	[0.9509]	[1.4298]	[2.5511]
Engineering	-1.2748*	-1.2347	-0.0643	-0.5248	-0.4901
	[0.6607]	[0.8739]	[1.0863]	[1.6608]	[2.9108]
3-year degree	0.2154	0.2227	-0.0460	0.5189	3.8801
	[0.8649]	[1.1529]	[1.4604]	[2.2301]	[3.9233]
Private school	-0.2187	0.0661	0.76914	-0.5522	2.5633
	[0.8472]	[1.1277]	[1.4279]	[2.1896]	[3.8311]
House of Delegates	-0.2736	-0.2621	-0.2405	0.9573	1.7376
	[0.2934]	[0.3804]	[0.4755]	[0.6772]	[1.2365]
House of Representatives	0.1768	0.4874	0.3289	-0.3173	-0.0941
	[0.4104]	[0.5481]	[0.7088]	[1.1065]	[1.9182]
Department of Education and Training	0.1803	0.7376]	0.2456	0.1603	-0.2895
	[0.4538]	[0.5964]	[0.7764]	[1.1875]	[2.1223]
Entance Score	7.6278	6.0291	20.8213	33.8186	139.2332*
	[14.6206]	[19.8623]	[25.3932]	[38.3141]	[69.5935]
Financial Aid status	-0.00415	0.3966	-0.0780	-0.4964	-2.298
	[0.4749]	[0.6344]	[0.8112]	[1.2379]	[2.2085]
Residence Status	-1.0923	-1.7822*	-2.0648	-2.2567	1.6271
	[0.8093]	[1.0745]	[1.3609]	[2.0695]	[3.5892]
Extended Programme	0.7341	0.4132	0.1251	0.4137	-0.4502
	[0.4963]	[0.6721]	[0.8609]	[1.3167]	[2.3688]
Observations	2153	1989	1686	1264	858

Notes: This table presents estimates of the above-cutoff indicator using predetermined covariates as the dependent variable. Column (1) uses observations within the optimal CJM bandwidth. Column (2) uses our preferred bandwidth based on the grade structure at the institution under consideration. Columns (3), (4) and (5) each use a fixed bandwidth of 4, 3 and 2 grade points around the cutoff, respectively. All estimates are clustered on the running variable. * implies p value<0.1, ** implies p < 0.05, and *** implies p < 0.01.

3.6 Results

The evidence presented in Section 3.5 indicates the validity of the RD design in this context. In addition, running tests of the pre-determined characteristics on the treatment variables showed that they too were not discontinuous through the DML threshold, further supporting the RD design used in this analysis. As a result, any discontinuities in subsequent years between academic performance of students who were on the DMLs and those who were not, can be attributed to the DML policy.

All the results that follow in sections 3.6.1 to 3.6.3 are estimated on the basis of standard RD estimations that do not include covariates. Section 3.6.4 presents estimations with covariates. Lee and Lemieux (2009) contend that if the distribution of covariates around the cutoff are not discontinuous then the requirement for local randomisation has been met and the RD approach yields consistent estimates of the treatment effect of the DML on student outcomes.

The main student outcomes evaluated are non-cumulative GPA in subsequent years, cumulative GPA in subsequent years, graduation, and academic dropout (both voluntary and involuntary). Voluntary and involuntary dropout are measures of persistence, allowing us to compare South African results with equivalent results from institutions in developed countries. All estimations show results for conventional and robust RD estimates. Robust estimates are calculated using a robust variance estimator while conventional estimates are calculated using a conventional variance estimator. The results for exit outcomes are presented first, followed by the results for the GPA outcomes.

All results presented in this section are considered to be local average treatment effects (LATE) as treatment effects are evaluated around the cutoff point where treatment and control groups are most similar. In keeping with standard practice, results are shown for the cutoff centered around 0.

3.6.1 Exit outcomes

The results shown in Table 3 are encouraging as the size of the estimated coefficients are within the expected range. Panel C and Panel A in Table 3 are linked through the relationship of dropout on graduation. It is important to acknowledge that the exit outcomes in Table 3 are linked – if a student graduates, they cannot drop out. If a student drops out, they may still subsequently successfully complete their degrees.

For graduation, the estimated effects of the DML policy are negative. This is rather unexpected. Students who appeared on the DML were 4.4% less likely to graduate than those in the control group. This may be illustrative of the dropout or attrition problem discussed below. Unsurprisingly, a student that is more likely to exit involuntarily, is also less likely to graduate

Dropout is divided into two categories – voluntary and involuntary. Voluntary dropout occurs when students withdraw from their studies of their own volition through their own actions or circumstances. Involuntary dropout describes academic exclusion from the university after a student fails to meet progression requirements, e.g. not passing the required minimum number of courses per year, or failing a ‘major’ course twice.

The DML policy has a statistically significant and positive impact on involuntary dropout across all the specified bandwidths. Based on the optimal bandwidth, students appearing on the DML in their first year are 2.3% more likely to involuntarily dropout relative to the control group Voluntary dropout is not significant at the widest bandwidth but is significant and positive for the three narrower bandwidths, with the policy effect strongest for students closest to the cutoff. As is the case with other results in this chapter, the size of the coefficients increase as the bandwidth narrows. Schochet (2008) asserts that for a given sample size, narrower bandwidths may yield less precise estimates, and that could be key rising for the sharp rise in the size of the coefficient between the CJM bandwidth and the narrowest bandwidth.

Table 3 Impact of the DML on exit outcomes

	(1)	(2)	(3)	(4)	(5)
Panel A: Graduation	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Robust	-0.044**	-0.057***	-0.074***	-0.085***	-0.112***
	[0.017]	[0.019]	[0.020]	[0.021]	[0.021]
Conventional	-0.007	-0.015]	-0.024	-0.045***	-0.062***
	[0.012]	[0.014]	[0.015]	[0.017]	[0.019]
Observations	2153	1989	1686	1264	858
Panel B: Voluntary Dropout					
Robust	0.020	0.023	0.034*	0.037*	0.052***
	[0.016]	[0.017]	[0.018]	[0.019]	[0.019]
Conventional	0.002	0.005	0.008	0.020	0.024
	[0.011]	[0.013]	[0.014]	[0.015]	[0.017]
Observations	2153	1989	1686	1264	858
Panel C: Involuntary Dropout					
Robust	0.023***	0.033***	0.040***	0.047***	0.061***
	[0.006]	[0.007]	[0.008]	[0.008]	[0.008]
Conventional	0.005	0.009**	0.015***	0.025***	0.037***
	[0.004]	[0.004]	[0.005]	[0.006]	[0.007]
Observations	2153	1989	1686	1264	858

Table 3: Impact of the Dean's Merit List Policy: Exit outcomes. Estimates are presented for the full sample. Standard errors are clustered along the running variable. * implies p value<0.1, ** implies p < 0.05, and *** implies p < 0.01. This specification does not include covariates. The cutoff has been recentered on zero.

The DML policy has a statistically significant and positive impact on involuntary dropout across all the specified bandwidths. Voluntary dropout was significant and positive for the three narrowest bandwidths only, with the policy effect strongest for students closest to the cutoff. Similar to other results related to the DML policy, the size of the coefficients increased as the bandwidth narrowed. The DML policy estimates on exit outcomes with second order polynomials show similar results to the linear estimations in Table 3. See Appendix, tables 40 and 41 for second order polynomial estimates for estimations with and without covariates.

3.6.2 Non-cumulative GPA

Table 4 shows the impact of academic recognition on students' annual non-cumulative GPA for the entire sample. On average, the effect of the DML policy is negative, with the impact being significantly more negative in the third year of study than in the second. The DML policy appears to depress subsequent GPA scores; an unexpected result. The university's intention is to reward and encourage studious behaviour, which should appear as higher second year GPA scores. However, in general, students that appeared on the DML in their first year at the university appear to experience less stable academic pathways; the impact of the policy is negative and lasting.

The robustness checks are included in each set of results in this section and all yield similar results. Irrespective of the chosen bandwidth, the current implementation of the DML policy leads to lasting, negative impacts on performance. The results also show that the policy effect is strongest for students within the narrowest bandwidth, further supporting the validity of the RD design.

Table 4 Impact of the DML policy: full sample

	(1)	(2)	(3)	(4)	(5)
Annual non-cumulative GPA Year2	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Robust	-1.058*	-1.388**	-1.668***	-2.387***	-2.884***
	[0.544]	[0.575]	[0.615]	[0.669]	[0.718]
Conventional	-1.047***	-0.781*	-0.908**	-1.154**	-1.851***
	[0.384]	[0.415]	[0.447]	[0.505]	[0.586]
Observations	2153	1864	1539	1148	753
Annual non-cumulative GPA Year3					
Robust	-2.690***	-3.153***	-3.517***	-4.484***	-4.868***
	[0.911]	[0.988]	[1.06]	[1.172]	[1.298]
Conventional	-1.735***	-2.041***	-2.274***	-2.697***	-3.786***
	[0.632]	[0.707]	[0.769]	[0.875]	[1.034]
Observations	2144	1857	1537	1148	751

Table 4: Estimates are presented for the full sample. Standard errors are clustered along the running variable. * implies p value <0.1, ** implies p < 0.05, and *** implies p < 0.01. This specification does not include covariates. The cutoff has been recentered on zero.

The results were split into the two main degree types, namely 3-year and 4-year degrees. This is an important distinction as the students who tend to enrol in 3-year qualifications tend to differ from students who enrol in 4-year qualifications. On average, 3-year qualifications have lower entry requirements than 4-year or professional degrees, suggesting weaker students enrol more frequently in the shorter duration degree programmes.

For the 3-year degree options, the DML has a significantly negative effect both when measured using the CJM bandwidth and the other, smaller bandwidth options. All results indicate that students appearing on the DML are likely to experience a significantly lower annual GPA in subsequent years compared to their first year of enrolment. This is true for the full sample and in particular for students registered for a 3-year degree. This might be indicative that second and third year results depend less on high school achievement, like first-year performance does. Instead, later years of study may require more critical thinking skills. This holds across all potential bandwidths and is especially pronounced in columns (4) and (5) of Table 4. The treated students do significantly worse compared to similar students who were not treated.

Table 5 Impact of the DML policy: 3-year degree

	(1)	(2)	(3)	(4)	(5)
Annual non-cumulative GPA Year2	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Robust	-2.066**	-2.389**	-2.797**	-4.285***	-5.0118***
	[0.958]	[1.061]	[1.150]	[1.280]	[1.419]
Conventional	-1.405**	-1.358*	-1.5667*	-1.914**	-3.096***
	[0.642]	[0.735]	[0.802]	[0.923]	[1.127]
Observations	1229	1066	881	652	422
Annual non-cumulative GPA Year3					
Robust	-3.472**	-3.926**	-4.279**	-5.331***	-5.543***
	[1.385]	[1.542]	[1.697]	[1.920]	[2.085]
Conventional	-2.229**	-2.701***	-2.97***	-3.430***	-4.475***
	[0.883]	[1.004]	[1.115]	[1.329]	[1.679]
Observations	1220	1059	877	648	418

Table 5: Impact of the Dean's Merit List Policy: 3-year degree. Estimates are presented for the full sample. Standard errors are clustered along the running variable. * implies p value<0.1, ** implies P < 0.05, and *** implies P < 0.01. This specification does not include covariates. The cutoff has been recentered on zero.

The results are also presented for selected subgroups. The results for the 4-year degree students differed substantially from those of students in 3-year degrees. Among students registered for 4-year degrees, the DML did not have a statistically significant impact on subsequent non-cumulative GPA. A potential reason for the lack of effect is that the progression requirements are notably different between the two degree types, even if they occur within the same faculty. In particular, the effect is about 50% larger in the third year of study relative to the second. It is only in the final year of the degree that the effect is particularly pronounced, and the effect is negative and statistically significant at the 5% level of significance, but only for the narrowest bandwidth.

A key concern with respect to the GPA results is the size of the coefficients. Given that the running variable has been recentered on zero, coefficients are expected to be smaller and quite close to zero in most estimations. The coefficients for the GPA estimations are all greater than one which implies extremely largely policy effects not supported by the data. Results for the GPA outcomes should therefore be interpreted with more caution.

Table 6 shows the results for the 4-year degree students. This significant negative effect (column 5 of Table 6) that only occurs in the final year may be related to the strand of literature showing that students in their final year of registration are potentially more concerned with securing a job than giving all their attention to their academic studies, leading to a lower academic performance shown through a lower annual GPA. However, results in these subgroups should be interpreted with caution due to smaller sample sizes.

Table 6 Impact of the DML policy: 4-year degree

	(1)	(2)	(3)	(4)	(5)
Annual non-cumulative GPA Year2	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Robust	0.486	0.213	0.110	0.597	1.107
	[1.061]	[1.130]	[1.209]	[1.312]	[1.466]
Conventional	-0.494	0.109	0.131	0.053	0.056
	[0.749]	[0.812]	[0.875]	[0.994]	[1.146]
Observations	924	798	658	496	331
Annual non-cumulative GPA Year3					
Robust	-1.411	-1.798	-2.084	-2.523	-2.316
	[1.334]	[1.401]	[1.460]	[1.580]	[1.856]
Conventional	-1.002	-0.9777	-1.179	-1.500	-2.362*
	[0.999]	[1.091]	[1.159]	[1.267]	[1.397]
Observations	924	798	658	496	331
Annual non-cumulative GPA Year4					
Robust	-1.555	-1.251	-1.482	-1.249	-3.159**
	[1.26]	[1.331]	[1.371]	[1.408]	[1.421]
Conventional	-2.299	-2.020**	-1.731	-1.713	-1.505
	[0.926]	[1.012]	[1.089]	[1.202]	[1.297]
Observations	800	687	567	433	287

Table 6: Impact of the Dean's Merit List Policy: 4-year degree. Estimates are presented for the full sample. Standard errors are clustered along the running variable. * implies p value<0.1, ** implies P < 0.05, and *** implies P < 0.01. This specification does not include covariates. The cutoff has been recentered on zero.

The more stringent entry requirements of four year degrees may account for the observed differences in the effect of the DML on student performance. Students who enrol for the 4-year degrees tend to have higher school-leaving exam scores, complete more subjects during their secondary education, and are more likely to sign up for math-based disciplines such as actuarial science and engineering. There may also be unobservable differences between the two cohorts across both degree types that can explain the very different impacts on non-cumulative GPA.

Surprisingly, no heterogeneous effects across gender, race, schooling background or home language were found when non-cumulative GPA was evaluated. In the South African context, academic results vary significantly with race, gender and schooling background (Council on Higher Education, 2018).[26]

There is very little evidence to suggest that the DML policy is implemented differently in the faculty of Commerce than in other faculties. All faculties rely on the undergraduate handbooks to explain the DML process to students, and no additional information is distributed via any student communication channels. There are no awards ceremonies at which students who achieve GPAs above the DML threshold are recognised publicly.

3.6.3 Cumulative GPA

Cumulative GPA is considered a superior measure of a student's academic performance as it represents the long-term performance of the student. It is the preferred measurement of GPA as it removes some of the variance that a single poor set of grades add to the GPA.

Table 7 shows the impact of the DML on cumulative GPA. The cumulative results suggested that the DML reward does not have the desired impact on high-performing students. Across all specifications the results showed that being awarded the DML in first year has a negative effect on cumulative GPA, providing evidence that while there is some dynamic adjustment occurring, this effect is also negative on average across the student population. This result indicated that incentives crowd out the potential benefit of the DML in the long term.

There were small changes in sample size moving from one year to another, especially when examining the GPA outcomes. This is due to dropout or attrition examined in the

[26] Academic results refer to both Grade 1-12 and post-secondary schooling.

previous chapter, which results in the loss of a few students in each subsequent year. However, the loss per bandwidth from one year to another is not statistically significant and does not influence the results. For example, moving from cumulative GPA in Year 2 to Year 3 resulted in the loss of 9 students due to dropout.

The rigid degree requirements across the university make it unlikely that students start engaging in strategic course selection behaviour from year 2 onward to maintain their DML status. While first-year curricula are fairly fixed, second-year curricula are marginally more relaxed and allow students to select at most two of their courses out of a maximum of nine. This does not give students enough room to manipulate their GPA as the elective courses are typically lower weighted than the compulsory courses for either the degree or specialisation in a discipline.

Table 7 Impact of the DML on cumulative GPA: full sample

	(1)	(2)	(3)	(4)	(5)
Cumulative GPA Year2	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Robust	-0.540*	-0.756**	-0.959**	-1.458***	-1.815***
	[0.303]	[0.321]	[0.345]	[0.379]	[0.408]
Conventional	-0.396*	-0.294	-0.391	-0.575**	-1.079***
	[0.211]	[0.228]	[0.247]	[0.282]	[0.333]
Observations	2153	1864	1539	1148	753
Cumulative GPA Year3					
Robust	-1.228**	-1.544***	-1.812***	-2.489***	-2.820***
	[0.518]	[0.555]	[0.597]	[0.656]	[0.699]
Conventional	-0.885**	-0.902**	-1.025**	-1.279***	-1.997***
	[0.351]	[0.387]	[0.425]	[0.489]	[0.582]
Observations	2144	1857	1537	1148	751
Cumulative GPA Year4					
Robust	-1.869**	-1.854**	-2.076**	-2.508***	-3.0346***
	[0.761]	[0.803]	[0.858]	[0.918]	[0.972]
Conventional	-1.398***	-1.382**	-1.466**	-1.694**	-2.045**
	[0.526]	[0.571]	[0.622]	[0.714]	[0.821]
Observations	1298	1115	919	691	455

Table 7: Impact of the Dean's Merit List Policy: Cumulative GPA. Estimates are presented for the full sample. Standard errors are clustered along the running variable. * implies p value<0.1, ** implies p < 0.05, and *** implies p < 0.01. This specification does not include covariates. The cutoff has been recentered on zero

3.6.4 Covariate-adjusted results

In an analysis of randomised experiments, researchers often include covariates as means to increase precision of the treatment effect estimation (Frölich and Huber, 2019). Lee (2008) points out that when the baseline covariates are independent of treatment status, estimates should be insensitive to the inclusion of these covariates and that estimates should be somewhat stable. The estimations including covariates are presented below.

Table 8 presents the covariate adjusted estimates for the three exit outcomes; graduation, voluntary dropout and involuntary dropout. The covariate-adjusted estimations for exit outcomes are robust to the inclusion of covariates, yielding similar results to Table 3. The stability of the size of the estimates between the non-covariate and covariate-adjusted estimations for the exit outcomes meet the principles described by Lee (2008). The size of the coefficients is within the expected range and resembles the pattern of significance seen in the estimations without covariates. On average, the inclusion of covariates decreased the size of estimates by approximately 25%. The impact of the DML policy on the treated group remains negative with respect to graduation, compared to the control group. For the CJM optimal bandwidth, DML students are 3.6% less likely to graduate than students whose GPAs were just below the DML cutoff.

With respect to involuntary dropout, the expected sign on the coefficient was negative but like the estimations without covariates, the sign on the coefficient was positive. This is evidence of the DML policy failing to provide the expected incentive to enhance academic excellence.

The estimated effects for exit outcomes remain stable with the inclusion of baseline covariates, as would be expected if the treatment is locally independent of all pre-determined characteristics. This lends credence to the findings of negative impacts of the DML policy on the treated group.

Table 8 Impact of the DML on exit outcomes: full sample

	(1)	(2)	(3)	(4)	(5)
Panel A: Graduation	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Robust	-0.036**	-0.046**	-0.063***	-0.067***	-0.087***
	[0.017]	[0.019]	[0.020]	[0.021]	[0.023]
Conventional	-0.001	-0.009	-0.017	-0.036**	-0.046**
	[0.012]	[0.014]	[0.015]	[0.016]	[0.018]
Observations	2153	1989	1686	1264	858
Panel B: Voluntary Dropout					
Robust	0.019	0.021	0.031*	0.037**	0.054**
	[0.015]	[0.017]	[0.017]	[0.019]	[0.0219]
Conventional	0.002	0.003	0.007	0.018	0.024
	[0.011]	[0.013]	[0.014]	[0.015]	[0.016]
Observations	2153	1989	1686	1264	858
Panel C: Involuntary Dropout					
Robust	0.016**	0.025***	0.031*	0.030***	0.033***
	[0.006]	[0.007]	[0.008]	[0.009]	[0.008]
Conventional	0.003	0.007	0.010***	0.018***	0.022***
	[0.005]	[0.005]	[0.005]	[0.006]	[0.007]
Observations	2153	1989	1686	1264	858

Table 8: Impact of the Dean's Merit List Policy: Exit outcomes. Covariate-adjusted estimates are presented for the full sample. Standard errors are clustered along the running variable. * implies p value<0.1, ** implies p < 0.05, and *** implies p < 0.01. The cutoff has been recentered on zero.

Table 9 shows the results for the addition of baseline covariates to the annual non-cumulative GPA analysis. The results suggest that adding covariates to the model virtually eliminates the effect of the DML on annual non-cumulative GPA in year 2 and year 3 at the optimal bandwidth of BW(CJM). The stability of the results for the GPA outcome variables is brought into question when considering the roughly 50% reduction in the magnitude of the coefficients for the non-cumulative GPA in year 3.

Table 9 Impact of the DML on non-cumulative GPA: full sample

	(1)	(2)	(3)	(4)	(5)
Annual non-cumulative GPA Year2	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Robust	0.094	-0.242	-0.327	-0.618	-0.695
	[0.557]	[0.596]	[0.636]	[0.697]	[0.794]
Conventional	-0.262	0.041	-0.036	-0.096	-0.462
	[0.404]	[0.435]	[0.464]	[0.516]	[0.594]
Observations	2153	1864	1539	1148	753
Annual non-cumulative GPA Year3					
Robust	-1.506	-1.857	-2.197*	-2.723**	-2.413
	[1.067]	[1.165]	[1.242]	[1.373	[1.514]
Conventional	-1.237*	-1.338	-1.366	-1.606	-2.320**
	[0.740]	[0.834]	[0.902]	[1.005]	[1.177]
Observations	2144	1857	1537	1148	751

Table 9: Covariate-adjusted estimates are presented for the full sample. Standard errors are clustered along the running variable. * implies p value <0.1, ** implies p < 0.05, and *** implies p < 0.01. The cutoff has been recentered on zero.

Table 10 below shows the results with the addition of covariates to the model. The addition of covariates has decreased the size of most of the coefficients and rendered most coefficients statistically insignificant. However, the coefficients remain larger than expected, and the magnitude of coefficients increase quite dramatically as the bandwidths narrow.

Table 10 Impact of the DML on cumulative GPA: full sample

Cumulative GPA Year2	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Robust	0.089	-0.146	-0.270	-0.529	-0.671
	[0.306]	[0.327]	[0.351]	[0.388]	[0.437]
Conventional	0.035	0.164	0.097	-0.009	-0.345
	[0.217]	[0.235]	[0.253]	[0.283]	[0.332]
Observations	2153	1864	1539	1148	753
Cumulative GPA Year3					
Robust	-0.349	-0.638	-0.858	-1.216*	-1.193
	[0.548]	[0.588]	[0.626]	[0.683]	[0.751]
Conventional	-0.369	-0.294	-0.333	-0.474	-0.947
	[0.376]	[0.415]	[0.452]	[0.508]	[0.588]
Observations	2144	1857	1537	1148	751
Cumulative GPA Year4					
Robust	-1.106	-1.043	-1.056	-1.057	-1.400
	[0.791]	[0.835]	[0.868]	[0.911]	[1.085]
Conventional	-1.640***	-1.434**	-1.268*	-1.125	-1.374*
	[0.572]	[0.623]	[0.666]	[0.739]	[0.819]
Observations	1298	1115	919	691	455

Table 10: Impact of the Dean's Merit List Policy: Cumulative GPA. Covariate-adjusted estimates are presented for the full sample. Standard errors are clustered along the running variable. * implies p value<0.1, ** implies p < 0.05, and *** implies p < 0.01. The cutoff has been recentered on zero.

Table 11 shows the estimation results for the impact of the DML policy by the 3-year degree students. All estimated coefficients for the covariate-adjusted estimations are smaller than the no-covariate results. However, coefficients are still quite large relative to expectations, and all coefficients are not statistically significant anymore.

Table 11 Impact of the DML policy: 3-year degrees

	(1)	(2)	(3)	(4)	(5)
Annual non-cumulative GPA Year2	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Robust	0.126	-0.214	-0.528	-1.486	-2.315
	[0.938]	[1.014]	[1.098]	[1.245]	[1.509]
Conventional	0.429	0.499	0.335	0.128	-0.826
	[0.631]	[0.708]	[0.766]	[0.855]	[1.022]
Observations	1229	1066	881	652	422
Annual non-cumulative GPA Year3					
Robust	-0.227	-0.582	-0.753	-1.109	-0.644
	[1.718]	1.945	[2.104]	[2.334]	[2.487]
Conventional	-0.351	-0.527	-0.458	-0.551	-0.771
	[1.106]	[1.304]	[1.436]	[1.660]	[2.038]
Observations	1220	1059	877	648	418

Table 11: Impact of the Dean's Merit List Policy: 3-year degree. Covariate-adjusted estimates are presented. Standard errors are clustered along the running variable. * implies p value<0.1, ** implies P < 0.05, and *** implies P < 0.01. the cutoff has been recentered on zero

Table 12 presents the results for the students enrolled in 4-year degrees. Similar to the results for the 3-year degree students, all coefficients in the covariate-adjusted estimations are not statistically significant, including the CJM bandwidth which has the highest numbers of observations, and the results are not as informative as would be expected, even when evaluated at the data-driven optimal bandwidth in Column (1) of Table 12.

Table 12 Impact of the DML policy: 4-year degree

	(1)	(2)	(3)	(4)	(5)
Annual non-cumulative GPA Year2	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Robust	0.673	0.591	0.792	1.198	1.756
	[1.008]	[1.077]	[1.137]	[1.188]	[1.302]
Conventional	-0.765	-0.191	-0.063	0.292	0.350
	[0.728]	[0.770]	[0.817]	[0.920]	[1.020]
Observations	924	798	658	496	331
Annual non-cumulative GPA Year3					
Robust	-2.082	-2.117	-2.317	-2.500	-2.463
	[1.405]	[1.446]	[1.507]	[1.612]	[1.977]
Conventional	-1.733*	-1.700	-1.674	-1.712	-2.569*
	[1.044]	[1.118]	[1.179]	[1.264]	[1.381]
Observations	924	798	658	496	331
Annual non-cumulative GPA Year4					
Robust	-1.881	-1.747	-1.757	-1.500	-3.795**
	[1.222]	[1.304]	[1.348]	[1.424]	[1.586]
Conventional	-2.417**	-2.197**	-1.963**	-1.917*	-2.001
	[0.848]	[0.918]	[0.999]	[1.121]	[1.219]
Observations	800	687	567	433	287

Table 12: Impact of the Dean's Merit List Policy: 4-year degree. Covariate-adjusted estimates are presented. Standard errors are clustered along the running variable. * implies p value<0.1, ** implies P < 0.05, and *** implies P < 0.01. the

3.7 Threats to Validity

3.7.1 Robustness

The results presented above were found to be robust to a number of falsification checks. Throughout this chapter, all estimation results are presented with varying bandwidths as robustness checks for ease of reference, including the bandwidths and the bias-corrected estimates and bandwidths described in Cattaneo, Jansson and Ma (2018).

When the different bandwidths used in this analysis was selected, consideration was given to the structure of the grading system as the DML cutoff is near the grade for a first-class pass. Given the nature of the results presented in section 3.5, readers may be sceptical of the outcomes as many of the results are counterintuitive, such as the large negative effect of the DML on subsequent performance and the small but negative impact on graduation outcomes. However, a number of robustness tests were performed to support these results.

Column 5 of Tables 2–6 shows the results for the narrowest bandwidth around the DML threshold. It is worth stressing that this was the preferred bandwidth for this analysis as it represents the students closest to the cutoff on either side. The key results should be interoperated from the narrowest band around the actual DML cutoff, and all other bandwidths should be viewed as falsification tests or robustness checks. All bandwidth results are presented in the same table for ease of access but also to show what a logical widening of the bandwidth interval may be.

Columns 1–4 of Tables 2–6 should be considered the robustness checks at each stage of estimation. Column 1 of Tables 2–6 show the results where the reported effects were determined by a data-driven process underpinned by the CJM method. This process also resulted in the widest bandwidth of all the estimations. All results were insensitive to the choice of bandwidth, as is shown by the consistent direction of sign and magnitude of the estimated coefficients. In addition, coefficients were largest for column 5 in Tables 2–6, which suggests that effects are strongest for those students just above and below the cutoff.

The finding that academic performance decreased or was marginally weaker in subsequent years remained statistically significant. As discussed above, the results were negative and significant irrespective of the specification. The size of the coefficients increased as the bandwidth narrowed, increasing by approximately 170% from the CJM data-driven bandwidth to the narrowest choice of two percentage points above and below the threshold. This was likely due to converging characteristics of students just above and below the eligibility threshold, which is a key assumption of RD-techniques.

Another method to check the robustness of the results is to implement higher-order global polynomials in the analysis. Gelman and Imbens (2019) strongly advise against the use of quadratic polynomials. The use of global polynomials applies larger weighting to the individuals with extreme values of the forcing variable, but local linear estimations would attach zero weighting to these individuals. Gelman and Imbens state that the implementation of high order polynomials results in significant variation in the estimated coefficients and very wide confidence intervals. Instead, they recommend that researchers implement local polynomial approximations, as were implemented in this analysis.

Implementation of the higher-order polynomials as a robustness check yields results that remain consistent (magnitude and sign) but with much wider confidence intervals (Gelman & Imbens, 2019; see Appendix, Tables 40 and 41). The estimates based on the third, fourth and fifth order local polynomials ranged from -2.88 to -12.2. At the same time, the increase in the size of the standard errors was exponential, and the widening of the confidence intervals shed doubt on the use of higher-order polynomials in a setting such as this study. This supports the findings of Gelman and Imbens (2019). Overall, the higher-order polynomials showed a similar result to the main results.

Overall, the results were not sensitive to the choice of bandwidth or estimation technique, yielding results of similar magnitude and statistical significance in almost all specifications.

3.7.2 Attrition

Longitudinal studies should take cognisance of any potential tendency to attrition bias present in their data as this may affect the results. The historically high levels of dropout in the South African HE sector made the possibility of attrition bias a source of concern. Depending on the source of data and whether attrition is evaluated at institutional or national level, published HE attrition rates vary between 25% and 40% (Council on Higher Education, 2018).

This chapter focuses on a subset of students who perform relatively well compared to the average student. Research shows that students who perform better are less likely to drop out or stop out and should be observed in the data for longer periods than weaker students who are more at risk (Meggiolaro, Giraldo & Clerici, 2017). This was also the case in this study.

The overall attrition rate in the sample was approximately 5.8% in the widest bandwidth and increased to 6.3% for the narrowest bandwidth. However, the distribution of attrition between 3- and 4-year degrees was significantly different. The attrition rate for students enrolled in 3-year degree programmes was less than 1%, irrespective of the bandwidth used. The attrition rate for students enrolled in 4-year programmes was 13.4% for the widest bandwidth and 13.2% for the narrowest.

While the attrition rates remained somewhat stable over the entire sample in this analysis, the number and proportion of students who exited the sample increased as the bandwidth widened. This was expected. However, the reason for attrition is not observable in the data and the absolute number of students exiting the sample is very small for the narrowest bandwidth. Based on the pattern of results from the sub-sample analysis (3- versus 4-year degrees) from the narrowest to the widest bandwidths, it is reasonable to assume that the attrition may be viewed as missing at random, and therefore, no data imputation was necessary to correct for it.

By proceeding with the missing at random assumption that missingness depends on observed but not unobserved data, there was little justification to impute missing data. This is especially understandable when observing the narrowest bandwidth that experienced the smallest proportion of dropouts from the sample and that the coefficients were largest and significant in all specifications in the narrowest bandwidth. In addition, the use of re-weighting and the implementation of imputation did not seem appropriate for such low levels of overall attrition in the data. Given the consistency of the findings across the different sub-samples, the results did not appear to have been biased by the level of attrition.

3.8 Potential Mechanisms – programme selection

The negative impacts of the DML on subsequent performance necessitates the search for the most plausible explanations through which the DML policy impacts on subsequent academic performance. The negative impact of the disincentive effect was an unexpected finding that merits additional investigation. However, to truly uncover the factors driving this outcome, additional data on behaviour and effort is required, most of which is currently unavailable.

The full sample results suggest that students' performance decreases in subsequent years based on the GPA outcomes. One plausible explanation for this is that students treated with the DML award select into more difficult programmes. At the university under consideration, students in many degree programmes are not wholly free to select their individual courses. Instead, courses are prescribed based on their chosen degree programme. Therefore, instead of looking at the level of difficulty at the course level, we look at the level of difficulty at programme level. Based on the pass rate for each programme, we divide the existing 55 programmes into 5 quintiles, ranging from hardest 20% (11 programmes per quintile) to easiest 20% of programmes.

The results for the programme selection analysis are presented in Table 13 below. No covariates are included in the first programme selection analysis (Panel A of Table 8), in line with Cattaneo et al (2021). Columns 1-5 show the different programmes ranked from top 20% most difficult programmes (given by the lowest pass rate) to the programmes with the highest pass rates called bottom 20%. All results are presented for the CJM optimal bandwidth only due to power considerations. Column 3 of Panel A shows the middle 20% of programmes, while column 4 shows the second easiest 20% of programmes by pass-rate. The results for programme selection show that only one set of estimations is statistically significant. Students who are treated with the DML in first year are between 3.5% and 6.1% more likely to select into the second hardest programmes by pass rate in the university when covariates are included in the estimations, compared to the control group.

Table 13 Impact of the DML policy on programme selection

	Top 20%	Next 20%	Middle 20%	Next 20%	Bottom 20%
Panel A: Programme selection (no covariates)	**BW=CJM**	**BW=CJM**	**BW=CJM**	**BW=CJM**	**BW=CJM**
Robust	0.002	0.045	0.010	-0.051	-0.006
	[0.016]	[0.020]	[0.027]	[0.028]	[0.034]
Conventional	-0.003	0.035	0.009	-0.004	-0.038
	[0.012]	[0.014]	[0.019]	[0.020]	[0.023]
Observations	2153	2153	2153	2153	2153
Panel B: Programme selection (inc covariates)					
Robust	-0.000	0.061**	-0.043	-0.007	-0.009
	[0.018]	[0.025]	[0.031]	[0.034]	[0.034]
Conventional	-0.008	0.035**	-0.023	0.031	-0.034
	[0.013]	[0.017]	[0.022]	[0.024]	[0.024]
Observations	2153	2153	2153	2153	2153

Table 13: Impact of the Dean's Merit List Policy on programme selection. Estimates are presented for the full sample. Standard errors are clustered along the running variable. * implies p value<0.1, ** implies P < 0.05, and *** implies P < 0.01. The top specification does not include covariates. The cutoff has been recentered on zero.

The results across both the no covariate and covariate-adjusted estimations show that, except for the 21-40% hardest programmes coefficient, none of the estimated coefficients are statistically significant. This provides evidence that students treated with the DML on average do not select into easier or more difficult programmes, indicating that programme selection is not driving the disincentive effect. As such, the idea that programme selection may be driving the negative impact of the DML is not supported by the estimations in Table 13.

One of the main reasons posited by Bénabou and Tirole (2003) for a weaker subsequent effort and performance in the presence of rewards is that of coasting. When individuals have high self-confidence, as some students would after an excellent academic performance in first year, they may tend to exert less effort, independent of ability. In the case of the DML, once treated students achieve the DML benchmark after their first academic year of study, they may believe that they do not need to exert as much effort in order to meet the cutoff in subsequent years, and thus reduce effort. The increase in self-esteem helps up to a certain point, after which the agent (student) starts to exhibit poor academic decision-making regarding effort allocation, yielding poorer academic results in subsequent years of study. The correlation between first year and second year DML is very low at 32%, suggesting that lower effort may partly explain the impact on subsequent performance.

These results provide some support for the Bénabou and Tirole (2003) coasting argument. They do not suggest that switching to easier programmes, akin to enrolling in courses which award higher grades, is driving the negative policy impact. The results permit a number of remaining potential explanations to explain the negative outcomes of DML on treated students, including that of over-confidence or coasting.

3.9 Conclusion

This chapter presented a way to consider and evaluate the DML incentive system in the context of an African economy. Many colleges and universities have policies that recognise good student performance, but very little is known about the impact these policies have on student outcomes. Only a few papers in the economics literature examines this policy, and to the best of the researcher's knowledge, the international literature in both developing and developed economies is not especially well developed. It is difficult to locate the results of this study in the context of the broader literature or HE systems around the world.

This chapter evaluated the impact of academic recognition policies, specifically the DML, on student outcomes. A RD approach was used to show that the DML as an academic incentive policy has negative (rather than the intended positive) effects over the short- and long-term on academic performance in the South African context. However, the effects of the policy on exit outcomes, such as graduation and dropout, appear to be smaller in impact but vary across the specified bandwidths.

The results indicate that the DML has an unfavourable impact on subsequent academic performance. Students who received the award tend to earn lower GPAs in subsequent years than expected. The findings suggest that the DML does not reinforce academic achievement. These results seem particularly counterintuitive relative to the findings of international empirical studies that show that DML reward policies have significant and positive impacts in the long term. However, the results are in line with Bénabou and Tirole's (2003) theoretical expectations regarding extrinsic motivation in a situation of asymmetric information between an agent and principal. A few reasons were suggested for the difference in findings between this study and that of international investigations.

The effects of the DML incentive are more mixed than expected when analysed by subgroup (results not shown). Female response was effectively unchanged by the DML policy. Intuitively, this makes sense as females tend to be the better performers than males and also tend to have higher annual and cumulative GPAs. On the other hand, sizeable negative effects were found for males. This was disappointing as many administrators look to incentives to improve male performance and/or outcomes.

The findings of this study may be of use to academic administrators because the results suggest that the DML does not achieve its stated goals. It may make more sense for administrators to set recognition policies by faculty, given the heterogeneity across the university, or to increase the cutoff for recognition. By focusing on the above two options, administrators are likely to improve their chances of achieving targeted policy outcomes and their desired effects. In addition, it may remove the band of GPA grades over which mixed signals could occur. While the results presented in this chapter are largely discouraging, it does represent an opportunity for university administrators to revise, rethink and update the student incentive structure.

Chapter 4: Mind the Gap: An Analysis of Time Path Gender and Racial Differences in Student Achievement

4.1 Introduction

Successful participation, and by extension completion, in HE is correlated with a host of benefits, including better labour market outcomes and job security (Fisher & Scott, 2011). In South Africa, the rate of return to education for completed Grade 12 and above is rising across the distribution of education and is higher for matric than for lower levels of education (Keswell & Poswell, 2004) and significantly higher for post-secondary education (Branson, Leibbrandt & Zuze, 2009), making education an attractive investment option.

There is a large body of research within the economics discipline that identify achievement or attainment gaps by race and gender from young ages.[27] Racial differences in educational attainment have been established as an empirical regularity (Aucejo, 2013). There is a concentration of literature that focusses exclusively on primary or secondary education in developed country contexts (Fryer and Levitt, 2004). Within this literature, it has been established, using almost all available metrics, that White individuals significantly outperform non-white individuals (Branson, Leibbrandt & Zuze, 2009). This includes reading or literacy scores and mathematics or numeracy (Spaull, 2013a). Research on South African third grade children indicate significant differences in performance by race and gender, extending to larger

[27] Achievement gaps are differences in performance on tests, access to opportunities, and educational attainment.

differences in educational attainment by Grade 9 (Spaull, 2015a). These differences persist throughout high school and are also evident in HE internationally (Bowen & Bok,1998).

Student performance and academic outcomes are increasingly important policy issues in the HE debate and understanding how students perform once enrolled in HE is very important to education policy makers. With competitive and globalised labour markets, secondary education is no longer sufficient to gain entry to the best paying jobs. HE has become the standard, and how students perform during their degrees is important to employers when making employment decisions (Fisher & Scott, 2011). It has become common practice among employers in South Africa to ask for student transcripts when applying for job vacancies.

However, once enrolled in HE, and specifically university-level education, there are still differences in performance. These differences have been noted by race and gender These differences, commonly referred to as the 'achievement gap', are large. The achievement gap is an important and divisive issue in the South African educational sector as educational outcomes and opportunities are still observed along racial lines. Considerably less research has been undertaken on performance at the post-secondary level compared to lower levels of education.

Internationally at the tertiary or post-secondary level, Black and Hispanic students take longer to complete their degrees (Bowen, Chingos, & McPherson, 2009), have lower GPAs (Betts & Morell, 1999), and experience lower graduation rates than white or Indian students (Bowen & Bok, 1998).[28] Reducing these gaps will decrease racial and gender differences in educational attainment (Bowen & Bok, 1998), earnings (Reardon, 2013), labour market participation, and representation at senior levels of leadership, including businesses and government (Oosthuizen & Bhorat, 2005).

[28] In this context, Black refers to individuals of African descent.

A key motivator for this study was the current debate on equity in HE. HEIs are facing pressure from government and civil society organisations to increase equity and diversity within entering cohorts in order to create a more diverse pool of university graduates. Admission alone will not address the inequitable allocation of education across the population. It is important to know and understand the potential academic pathways that a student may encounter within the HE system to come up with viable solutions to address completion rates and time-to-degree patterns. Therefore, detailed knowledge of student performance is essential. Students from previously disadvantaged schools perform poorly on the Grade 12 school-leaving exam. This performance extends into HE where significant differences in outcomes are noted (Branson, Leibbrandt & Zuze, 2009). Evidence suggests that academic performance affects post-HE earnings. One potential implication is that students whose performance is significantly lower than the median fares significantly worse in terms of future earnings.

Studies find a significant and positive link between college or university GPA and subsequent labour market earnings. Loury and Garman (1995) show that weekly earnings of white males are strongly correlated with increases in college GPA, and for Black males the effect is still positive but weaker. Other studies had similar results (Betts & Morell, 1999; Loury & Garman, 1995, Keswell, 2004).

Given that educational performance in South Africa is largely characterised along racial lines, this places African, Coloured and Indian students at a significant disadvantage in the labour market from the outset. A primary focus of this chapter is to evaluate the extent of gender and racial differences in student performance. The racial aspect is especially important given that socio-economic background coincides very strongly with race in South Africa.

While there has been a proliferation of studies focusing on student learning, retention and success, these studies typically use cross-sectional analysis to identify the determinants of educational outcomes. Another shortcoming of studies in these areas is the focus on the early years in HE, with many specifically focusing on first year, rather than a complete overview of all years of study. To the best of the researcher's knowledge, there is a lack of research that focuses on the post-secondary racial/gender gap as it relates to students' progression through the HE system, especially in a developing country context.

The evolution of non-cumulative GPA was examined by focusing on racial and gender differences. A ranking approach is introduced to measure student performance through the conversion of non-cumulative GPA into a ranking variable. The path of student grades over time and how these grades evolve as students move through their academic studies are of primary interest. While the standard approach in the literature has been to measure performance by observing students' GPA, this research focused on relative performance rather than absolute. Relative measures of programme rank allowed for observed distributional changes in one student's rank relative to another. Thus, we measured students' programme rank within their degree programmes over time. This measurement of performance was preferred to the GPA measurement for a few reasons. First, it allowed for the measurement of relative gains of students. If it is true that students who are high achievers in year 1 remain high achievers through to graduation, there will be little or no changes in their relative rank. However, a student might experience a poor start to their studies and be ranked much lower in the distribution. If this student improves their performance between their first year and year of graduation, the student will improve their rank quite substantially, relative to all other students in the class. The key dimension of student performance investigated was therefore student rank.

Two key concerns arose from the use of GPA. The first is that GPAs are determined differently across departments and faculties, making GPA a heterogenous variable of interest. This issue was overcome by converting annual GPA variable into a ranking variable to allow the focus to be on relative movements rather than absolute movements. This also compensated for some programmes with extremely high GPAs and others with very low GPAs due to the nature of courses within the programme. The second concern was the relevance of this study to other HEIs. Focusing on one HEI only ensured the results are not driven by differences in grading patterns across institutions or unobserved heterogeneity.

The aim of this analysis was two-fold. Firstly, this work will help university administrators better understand the factors that play a role in predicting student success. This is especially important where the schooling system varies significantly in quality between the public and private spheres of the primary and secondary education sector (Spaull, 2013a), which impacts on student preparedness for HE. Moreover, as universities are adapting admission criteria to ensure greater access and equity in student enrolments, understanding which factors contribute to successful participation, retention and completion is critical.

Secondly, this research contributes to the body of knowledge on performance in HE by examining the variation in performance over time and specifically identifying student trajectories. This will enable university administrators to identify key bottlenecks in student progression and allow greater targeting in terms of academic interventions. This work may also help inform the types of interventions that may be necessary at different levels of study by faculty or discipline.

The chapter is structured as follows: Section 4.2 presents the related literature, including possible linkages between pre-university experiences and post-secondary outcomes; section 4.4 describes the data used in this study; section 4.65 discusses the trends in academic performance; section 4.8 discusses the empirical results; and the final section considers conclusions and the implications of the results.

4.2 Related Literature and Background

Achievement gaps, described as differences in performance between groups on identifiable characteristics, have been noted as early as kindergarten (Fryer & Levitt, 2004).

Existing criteria for which differences in academic performance have been found to include racial background, gender, home language, socio-economic status, and disability (Fryer & Levitt, 2004). A key reason for concern on this topic is that achievement gaps at early ages lead to longer-term gaps that persist into high school and post-school studies (Bowen & Bok, 1998). This has been documented quite extensively in the US where there are significant differences in student performance by race and gender across the education distribution. Longer-term impacts include the types of jobs that students can access once they enter the labour market, and by implication, the income earned from these jobs.

A wide range of theories have been examined to explain achievement gaps in education. Based on the theory of catch-up, Fryer and Levitt (2004) use the Early Childhood Longitudinal Study (ECLS-K) to investigate the Black-white test score gap for kids in their first two years of school. The authors were the first to use this informative data set, which included significantly more covariates than existing studies had available at the time. This allowed the authors to theorise a few potential influences on standardised test scores and use the longitudinal data to follow students on their academic journeys. A key finding from this study is that Black students enter the schooling system with an equivalent performance to white students when controlling for as many observable characteristics as possible. Importantly, once in the education system, Black students perform significantly worse over time compared to observationally equivalent white students with similar characteristics. If the trend continues, by Grade 5, Black students are 0.50 of a standard deviation behind their white counterparts, a gap that is similar in size to those in the studies by Phillips (2000) and Phillips, Crouse, and Ralph (1998). Another interesting finding by Fryer and Levitt is that Black females outperform Black males. This is similar to this study's finding that Black males are the worst performers relative to White females. Fryer and Levitt (2004) also conducted the analysis on a within-race basis to investigate the possibility that not all covariates affect students in the same way and to check that the estimates are not biased. For example, it may be that an improvement in socio-economic status of Black children does not have the same effect as an improvement in socio-economic status of white children. When Black kids had the characteristics of average white kids in their analysis, Black kids still perform below a comparable white kid by as much as 0.21 standard deviations on both math and reading tests. This result highlights the importance of policy targeting by interrogating the potential gains that could be achieved for a given policy implementation.

For South Africa there is a growing body of research analysing performance of school-going age students (see Spaull, 2013). This includes interrogating performance on literacy and numeracy tests. The evidence for children of school-going age in South Africa is similar to that of the USA. Spaull (2013a) finds that there are significant differences in performance between students who attend private and public schools. The gaps in performance are found as early as Grade 3 and tend to grow wider until Grade 7, showing little or no sign of catch-up or convergence (Spaull, 2013a).

Using data from South Africa's only nationally representative educational panel dataset, the National School Effectiveness Study, Spaull (2015b) evaluates learning deficits in Grades 3, 4 and 5. While they do not directly examine achievement gaps, it highlights the disparity in performance by school type. The South African schooling system is divided into five quintiles with 1 being the poorest quintile in any province and 5 being the richest.[29] This is an important measure by which to evaluate schooling outcomes as most South African children, the majority of whom are African, attend quintile 1–4 schools. Most children attending private schools in most provinces are white. An important finding by Spaull (2015b) is that 26% of Grade 5

[29] A province is similar to a state in the US.

students still function at a Grade 3 level and 45% of the quintile 5 students, the wealthiest students, are not yet operating at a Grade 3 level by the end of Grade 5. This type of information is useful as it presents a snapshot of performance at a point in time, but it does not allow for an evaluation of learning trajectories over time. Spaull (2015b) identifies South Africa's participation in a few international benchmarking tests that allow for the cross-country comparison of performance. Results show that South African youth have large literacy and numeracy gaps by the age of eight and that these gaps do not decrease as students' progress through the educational grades. In South Africa's case, socio-economic disparities are distributed along racial lines, and therefore, race and socio-economic status can be used interchangeably.

While the body of literature on lower levels of schooling has expanded over time, it has not translated into a growing body of research on student performance in HE, largely due to data and other information constraints. The evidence on performance in HE is much sparser. Arcidiacono, Aucejo and Spenner (2012) use a ranking variable derived from GPA and finds that the gap between white and Black students' GPA narrows between freshman and senior years, irrespective of the performance measure used. The authors argue that the observed grade conversion is not driven by Black students catching up but rather by differences in grading patterns and course selection. Given these differences in grading patterns, weaker students are more likely to switch away from the natural sciences and engineering, which are viewed as academically more challenging, into the social sciences where courses are viewed as less technical and the grades awarded may be higher.

Arcidiacono, Aucejo and Spenner (2012) studied student performance over time using, focusing on the Black/White achievement gap in HE. The authors found that major choice is an important determinant of performance, and the closing of a Black-White achievement gap. Their analysis used a ranking of class GPA as the main dependent variable. We follow their adoption of the ranking variable in the chapter but focus on a ranking at the programme rather than the class level.

Martin, Spenner and Mustillo (2017) conducted a longitudinal study of undergraduate performance over time, and test for leading explanations of the post-secondary achievement gap. They focus on the racial/ethnic achievement gap at a highly selective, private university in the US. They find that almost half of the observed gap is attributable to family background characteristics and pre-college preparation. While personal resources and patterns of campus involvement significantly impact GPA, they do not account for the achievement gap. Importantly, they identify two key areas for targeted and strategic interventions: the major or specialisation selection process, and campus climate. Regarding the first intervention, achievement gaps are largest in the first year when students are the most uncertain about future studies and are yet to settle fully into university life. This implies that more and better information at either the application stage or when making course decisions can play a major role in determining student success for future years. White students are less likely to experience a major change. This is largely due to two effects. The first is that many White students are legacy students, meaning that they have an immediate family member, such as a parent, who has attended and completed HE from the same college or university, and therefore, able to give suitable guidance in terms of choosing a major or field of specialisation.[30] The second effect is that White students assimilate more easily into student life (Oates, 2009). A growing body of research highlights the impact that patterns of assimilation into an environment have on the functioning of individuals in that environment. Gallop and Bastien 92016) show that when institutions make concerted efforts to create welcoming and inclusive environments, more

[30] Legacy students are defined as individuals whose parents attended and graduated from the same college or university.

students thrive in those circumstances. This is supported by work from Quaye, Harper & Pendakur (2019) who show that creating welcoming spaces for all students, at all levels of study, leads to increased engagement, persistence and improved academic outcomes.

Very few studies provide evidence for catch-up between students as they progress through the system or as they become more familiar with their educational environment and resources. Arcidiacono, Aucejo and Spenner (2012) is one of the few who attempt to illustrate catch-up and investigate the drivers of observed catch-up that is evidenced. This chapter draws on some of the work of Arcidiacono, Aucejo and Spenner (2012).

4.3 Education production function

The conceptual approach most used for studies on academic outcomes in the education production function. In this framework educational outcomes are assumed to be related to a set of inputs such as individual, family and school levels. Following Caviglia-Harris and Maier (2020), the conceptual approach models outcomes using a reduced-form education production function given by:

$$y_{it} = \alpha + \beta_1 Z_i + \beta_2 K_{it} + \beta_3 X_{it} + \varepsilon_{it} \tag{4.1}$$

where y_{it} represents an academic outcome of student i in time t, $Z_i$is their personal characteristics, $K_{it}$is socio-economic characteristics and X_{it}represents factors such as academic characteristics and ε_{it} represents unobserved determinants of academic outcomes

4.4 Data

The data used for analysis in this chapter came from UCT's Institutional Planning Department (IPD). It was drawn from the application and admissions process where every applicant submits information on their schooling background and some basic personal characteristics. As a historically white institution, UCT attracts a significant application pool from historically white secondary and independent or private schools, but there is a growing pool of applicants from historically disadvantaged secondary schools or lower quintile public schools. The key variables contained in the undergraduate records can be classified as (i) personal information, including gender, race, home province, financial aid status, and residence status; (ii) academic background, including name of high school attended, Grade 12 results, and programmes applied for; and (iii) annual history, including GPA per year, class ranking per year, and changes in class ranking per year.

The final sample excluded international students because the primary motive of this research was to isolate the impacts for students who have been through the South African education system. Students from the medical faculty were also excluded as the admission process is fundamentally different to the rest of the university. Finally, transfer-in students were excluded as they exhibit a significantly different academic career than the first-time entering cohort.

Among the variables used in the analysis were controls for gender, race, Grade 12 examination score (matric score), faculty of first enrolment, degree type, first-year GPA, financial aid status, residence status, number of courses failed in the first year, home province, and high school type. Gender is a binary variable with the reference group set to female. The race variable comprised a set of four binary variables, including African, Coloured, white and Indian. White was used as the reference group.

A key drawback of the UCT IPD data was that it did not contain information on family or parental background because this information was not required at the time of application. Similar studies show that parental education and occupation account for a large proportion of the variation in student performance, and the absence of this information could result in omitted variable bias.

Table 8 shows the descriptive statistics for the analysis sample. The sample is 49% female with an even spread of students across the three entering cohorts. Twenty-eight percent of students live in residence while 14% receive financial aid. Less than half of the students register for a 4-year degree. The two largest faculties are Commerce and Humanities, while Science is approximately half the size of Commerce.

The racial breakdown of the sample reveals that slightly less than half is White while just more than one quarter is African. Approximately 68% of the student body speak English as their first language.

Table 14 Descriptive statistics for full sample

		Full Sample	Graduates	Dropouts
Gender	*(Female)*	0.49	0.51	0.42
Race	*African*	0.26	0.21	0.41
	Coloured	0.17	0.16	0.19
	Indian	0.08	0.08	0.08
	White	0.46	0.51	0.29
Affluence		0.88	0.92	0.79
Matric score		81.5	83.2	75.7
3-year degree		0.58	0.58	0.58
Faculty	*Commerce*	0.32	0.34	0.24
	Humanities	0.31	0.32	0.3
	Engineering	0.19	0.17	0.22
	Science	0.16	0.14	0.2
English Home Language		0.68	0.74	0.53
Financial Aid		0.14	0.11	0.23
Residence		0.28	0.26	0.34
2006		0.31	0.31	0.3
2007		0.32	0.31	0.33
2008		0.36	0.36	0.36
Observations		8452	6310	2142

Data source: UCT 2006, 2007 and 2008.
Author's own calculations

Because of the nature of school enrolment policies in South Africa, an affluence variable was created based on the status of high school attended. This variable, from the health economics literature, was based on the high school classification. The intention was for this variable to capture some of the information about household income, even imperfectly.

Retention or attrition bias is discussed in the next subsection, as this informs the selection equation in Section 4.7.1.

4.4.1 Retention or attrition bias

A critical issue faced in many HE studies is that of selection bias: The students who graduate and are observed for longer periods (evidence of persistence or retention) are characteristically different to the students who dropout. Failure to control for these differences results in biased results, which is typically upward or overstating the impact of a particular variable on the outcome. Concerningly, academic administrators or policy makers in countries where dropout is high may well be informing education policy based on these biased results, which could impact students' academic careers or outcomes.

Graduation rates at HEIs in many developing countries developed economies are lower than that for developed countries. Survivorship or attrition bias (selection bias) is therefore a significant problem in education studies. Institutional graduation rates in South Africa vary extensively between institutions. At the institution in question, graduation rates hover at 75% and exhibits substantial variation between race groups.

Retention should be considered as it relates to student persistence. Some students will move through the system much slower compared to others and in the South African context they could be just meeting the minimum progression requirements each year. Attrition, or retention, the probability that a student remains enrolled in university, plays a much larger part earlier in students' academic careers as students either fail to meet progression requirements moving from first to second year or fail to meet the requirements moving from year 2 to 3. By the start of the third academic year, students would have met most, if not all, of the progression requirements to proceed to the final year of study. Attrition, or retention, therefore plays a significant role in who completes their degrees and overall student performance and should be considered in subsequent analyses of the data.

4.5 Time Path of Differences in Performance

Universities use GPAs almost exclusively to evaluate student success as practically all courses are required to award grades. These grades are averaged across all courses to create an aggregate measure of academic performance. GPAs are typically measured on a scale from 0 to 100, and the bulk of students locate just above the 50% GPA across the student population.[31]

GPA data is inherently longitudinal in composition (Grove & Wasserman, 2004). The collection of non-cumulative annual GPA is an indication of students' performance over time and represents the life cycle pattern of student GPA. This is in contrast to the Grade 12 examination or once-off HE entrance exams such as the National Benchmark Tests, which are cross-sectional data and present difficulties of their own. Grove and Wasserman's (2004) study is one of the better-known studies that examine the time-series dynamics of student performance. They examined the life cycle of undergraduate GPA for six cohorts over eight semesters of enrolment at a large private university in the US. The data reveals a 'check-mark' pattern for GPA. Grades typically fall after the first semester but rise again until the second-last semester of study. In the final semester, students again experience a decline in GPA, commonly referred to as the 'senior slump', which is largely explained by most students' change in focus from academics to the labour market.

Figure 10 shows the non-cumulative GPA by gender and race. It is evident from Figure 10 that female students outperform male students for each race group for every year of available data. The aggregated data does not display the check-mark pattern of performance, even where the sample might be restricted to students who complete their degrees in minimum time.[32] It also evident that white females are the strongest performing students from enrolment through to graduation. Similar to Grove and Wasserman's (2004) observation based on US data, Figure 10 shows that on average males experience a senior slump either between years 4 and 5 or

[31] The class of pass for grades vary at different universities but 50% remains the minimum passing grade in HE in South Africa.

[32] Students who graduate in the advertised time for a given degree (shortest allowed time).

years 5 and 6, supporting the argument that the senior slump is somewhat a male phenomenon. Figure 10 also suggests a gender gap for a developing-economy context.

Figure 10 Annual non-cumulative GPA by race and gender

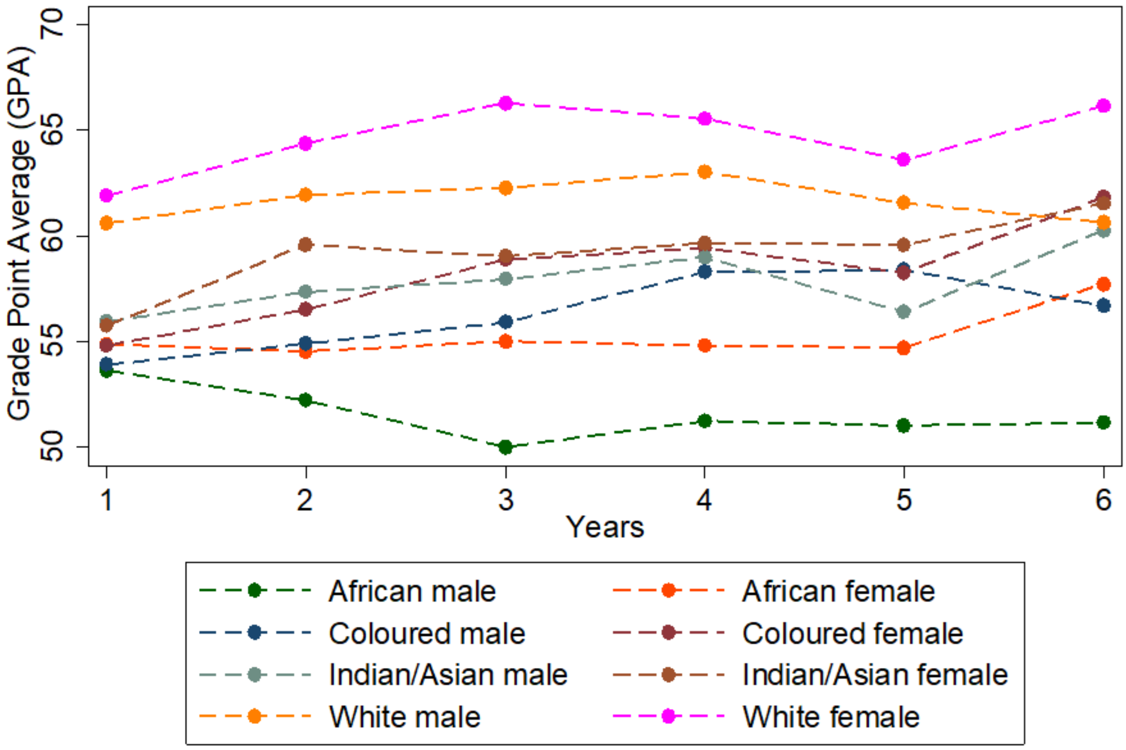

Figure 10 shows students' non-cumulative annual GPA by race and gender over a 6-year horizon.
Data: UCT IPD

Grove and Wasserman (2004) explain the 'freshman-slump' as a reallocation of time away from studying towards social interactions as students build new friendships and experience more of what student life has to offer. As students realise the importance of course performance for academic progression, students find ways to increase their focus, adapt to the system of study, and where possible, choose courses that suit their abilities better. In the US case, students experience the final-semester slump due to the allocation of time towards job applications and interviews, but this does not appear to be the case in aggregate in the South African context.

Following on Arcidiacono, Aucejo and Spenner (2012), GPAs were converted into programme ranks by creating a ranking variable based on the student's programme rank. This is preferred to course rank as the degree programmes at UCT are quite rigid and leave little room for individual course choice. This allowed for the removal of some differences in GPA that occurs between classes and programmes and presented the preferred variable as the basis of analysis.

Following on Arcidiacono *et al* (2012), GPAs are converted into programme ranks by creating a ranking variable based on the student's programme rank. Specifically,

$$Rank = 1 - \frac{(R-1)}{(N-1)} \tag{4.2}$$

where R equals the student's rank against their peers within the programme and N equals the total number of students within the programme.

Table 15 Median GPA by race for 3-year degrees

	African	Coloured	Indian	White
Year 1	59.8775	58.646	60.258	64.38
Year 2	57.79	58.955	59.55	65.33
Year 3	57.6	60.91	61	66.59
Yr3–Yr1	-2.2775	2.264	0.742	2.21

Table 15 shows the median GPA by race for students enrolled in 3-year degrees
Data: UCT IPD.

Table 15 shows the median GPA by race for students registered for 3-year degrees. In the South African context, GPA is based on a scale between 0 and 100, with average student grades typically between 55 and 65. Grades in excess of 70% are achieved by between 20% to 30% of students, depending on the degree programme and level of difficulty. The information in Table 9 aligns with the information presented in Figure 10, where annual average GPA does not exceed 70% for many students. From Table 9, all students except African students present an upward trend in performance over time. African students outperform Coloured students in the first year of study but do not maintain this momentum into subsequent years of study. One explanation for this observed difference in performance is African students' participation in the Academic Development Programme. African students comprise most students in these programmes in which students receive small, specialised lectures during the first year of study separate from mainstream students. These classes fall away in subsequent years and programme students stream into mainstream classes. This has the effect of dampening student performance for students originally enrolled in the Academic Development Programme (Smith, 2012). Coloured students make the greatest gains in terms of relative improvement based on the data in Table 9. Between first and final year of a 3-year degree, Coloured students' rank increases by 2.264 percentage points. African students show the opposite trend to Coloured students by moving down the distribution over time.

Table 16 shows the median GPA by race for students registered for 4-year degrees. A similar pattern emerges when GPAs are examined for students enrolled in 4-year degrees. The marginal gain in GPA is negative for African students, and for Indian students it is opposite in sign but of similar magnitude. Taken together, Tables 9 and 10 indicate that there are marginal gains in GPA for all students except African students. However, this might be misleading as GPAs are censored from above and courses may have different grading patterns. To overcome these two issues, students' individual year-by-year class ranking was created.

Table 16 Median GPA by race for 4-year degrees

	African	Coloured	Indian	White
Year 1	60.88	61	60.14	65.375
Year 2	59.19	61.61	61.8365	65.554
Year 3	57.63	62.06	61.04	65.68
Year 4	58.41	62.78	62.48	66.745
Yr4–Yr1	-2.47	1.78	2.34	1.37

Table 16 shows median GPA by race for students enrolled in 4-year degrees.
Data: UCT IPD

Table 17 shows students class rank unadjusted for mean class performance. In the left hand table, the median white student is at the 66th percentile and the median Coloured student is at the 45th percentile. This information is not visible in the mean GPA data given in Table 8. The pattern for African students is very interesting. African students end first year at the 50th percentile but fall to the 34th percentile by the end of their degree. This represents a decrease in class rank of approximately 15.7 percentage points, and they are the only race group to experience a reduction in class rank. GPA levels for African showed smaller losses relative to other students, indicating the using GPA alone as a measure of performance is insufficient. For those students who gain between first year and third year, the improvements in class rank are smaller than the improvements in mean GPA. However, these results do not take into account differential grading practices between courses.

Table 17 Median class rank by race, (unadjusted) for 3-year and 4-year degrees

Median class rank by race unadjusted for 3-year degrees

	African	Coloured	Indian	White
Year 1	0.5008	0.4583	0.5133	0.6609
Year 2	0.3931	0.4326	0.4557	0.6729
Year 3	0.3438	0.4639	0.4676	0.6932
Yr3-Yr1	-0.157	0.0056	-0.0457	0.0323

Median class rank by race adjusted, for 3-year degrees

	African	Coloured	Indian	White
Year 1	0.5328	0.4867	0.5322	0.6639
Year 2	0.4827	0.5133	0.4758	0.6686
Year 3	0.4331	0.5057	0.4967	0.6637
Yr3-Yr1	-0.0997	0.019	-0.0355	-0.0002

Table 17 shows the median class rank by race (unadjusted) for 3-year and 4-year degrees.
Data: UCT IPD

Table 17 also shows the median class rank adjusted for mean performance (right side table). The reduction in median class rank for African students is much smaller after the adjustment than without the adjustment. An interesting outcome is that of Indian/Asian and white students. Relative to Table 11 (left side table), these students also experience a reduction in class rank. Coloured students are the only group to show an improvement in class rank for those students registered for the 3-year degree.

Table 18 shows the median class rank for students by race enrolled in a 4-year degree. The initial performance of all students within the longer programme is representative of academically stronger students selecting into these programmes. The class rank of the median white student is now at the 69th percentile compared to the 53rd percentile for the median

African student. There is less of a difference in the performance of African and Coloured students in the longer programmes, but the aggregate class rank data shows that differential paths emerge through HE.

Table 18 Unadjusted median class rank by race for students enrolled in 4-year degrees

Median class rank by race unadjusted for 4-year degrees

	African	Coloured	Indian	White
Year 1	0.5323	0.5379	0.5105	0.6938
Year 2	0.4408	0.5352	0.5448	0.6796
Year 3	0.3454	0.5101	0.4692	0.662
Year 4	0.3851	0.5556	0.5448	0.7187
Yr4-Yr1	-0.1472	0.0177	0.0343	0.0249

Median class rank by race adjusted for 4-year degrees

	African	Coloured	Indian	White
Year 1	0.534	0.544	0.4674	0.6326
Year 2	0.4517	0.5571	0.4611	0.6137
Year 3	0.3731	0.5251	0.4322	0.6063
Year 4	0.3861	0.5012	0.4698	0.6262
Yr4-Yr1	-0.1479	-0.0428	0.0024	-0.0064

Table 18 shows the unadjusted median class rank by race for students enrolled in 4-year degrees.
Data: UCT IPD

The next step was to calculate class rank adjusted for selection into programmes. The mean performance by course was subtracted from the student's median grade, and class rank was recalculated on this adjusted basis (Arcidiacono, Aucejo & Spenner, 2012).

Table 18 (right side table) presents an interesting change in class rank for students registered for the 4-year degree. While African students still experience the decrease in class rank between first and fourth year, Coloured and white students also experience the reduction in rank. Indian/Asian students are the only group to show a marginal improvement in class rank, even though they remain below the 50th percentile throughout their studies. It is worth pointing out that academically stronger students enrol in 4-year degree programmes, given by the higher starting class rank in both the adjusted and unadjusted class ranks.

4.6 Academic Performance: Identifying Pathways

Students proceed along a variety of pathways in HE. These pathways are neither simple nor linear. Student trajectories through HE are heterogenous and no two students' journeys are necessarily the same. Unlike the US system where students enrol for a degree, in South Africa they are admitted to specific degree programmes and faculties rather than granted general admission to the institution. Most students remain enrolled in their admitting faculty. When students switch degrees, they typically switch within the same faculty. The percentage of students who switch across faculties remain small over time as the time cost of switching is potentially high. In instances where students do not meet the progression requirements for their enrolled programmes, universities will evaluate whether students should switch to a degree that is potentially a better match. This is often the case with commerce-based degrees where the 4-year programmes are constructed around a core set of courses that make up a 3-year degree. Should students in the 4-year degree not meet the progression requirements, a 3-year degree is recommended as the alternative. Most degrees offered in South Africa are 3-year degrees, but there are some 4-year degree programmes that lean towards professional-based disciplines like actuarial science and engineering.

As shown in Chapter 2, South Africa has a low completion or graduation rate in HE. There are a few factors that drive this outcome, but it will not be discussed in this chapter because there is not enough detailed data exploring the causes and driving factors of non-completion.[33]

When trying to disentangle different measures of student success, the following specific South African-based indicators stood out: Student progress unit, AYOS, and academic progress code. Each of these indicators are discussed in the following subsections to get a sense of the different pathways students experience.

4.6.1 Student progress unit

One measure of identifying student pathways is to look at the student progress unit (Dobson, Sharma, & Haydon, 1997; Gallagher & Conn, 1994; Linke, 1991). This is a ratio that

[33] This includes qualitative data from interviews and focus groups.

is a measure of student output to input. Specifically, it is the ratio of the number of credits a student pass per year based on the number of credits for which they registered. This ratio is indicative of performance at a particular stage but is not that useful as a continuous variable of performance over time. Students are required to register for a different number of courses each year that each have a different credit weighting depending on being a first-, second-, third- or fourth-year course. In addition, at the university under consideration, there is not a standard measure of credit equivalence between faculties. A first-year course in the Commerce faculty may carry an 18-credit weighting while a similar course in the Science faculty may carry a 16-credit weighting, indicating that expected time on task differs between faculties.[34] This makes it difficult to directly compare courses. It may even be the case that the same course offered in different faculties may carry different credit weightings depending on the student's home faculty, for example, mathematics or economics.

Initial investigations show that there are substantial differences in performance when measured according to the student progress unit or student/course success rate. African students pass 75% of all courses attempted in their first year of study compared to 89% for white students. The rates for Coloured and Indian students are 78% and 82%, respectively. These numbers compare favourably when students from this institution are measured against the general student population (see Figure 4 in section 2.4.1.1).

4.6.2 Academic year of study

At the start of each year a measure of progress is determined for every student. This measure is called AYOS. It is established by looking at the set of courses for which a student is enrolled and calculating backward how many years it would take the student to meet the

[34] 18 credits are equivalent to 180 notional hours spent on course content.

requirements for the award of the degree. For example, if a student is enrolled for a 3-year Bachelor of Commerce (BCom) degree with an intended specialisation in economics, and the student is enrolled for first-year Economics courses, then the student is coded AYOS-1.[35] This means they will take three more years to complete the requirements for a 3-year degree, including the current year of registration.

The measure of AYOS helps to explain how many students remain on track to complete their degrees in minimum time. AYOS is also an indicator of progress as it is typically the specialisation or major courses that are the cause of the delay in academic progression. If students fail their specialisation subjects in the first year, it means they are effectively equivalent to a new first year when they return for the second year of study as the time to completion remains at its maximum. Specialisation courses usually have very detailed entry and progression requirements, failing which students must repeat the required entry courses to proceed to the next level of study. This repetition of courses extends the time to completion, creating bottlenecks in the system.

Figure 11 shows the AYOS indicator by degree duration for different race groups. It is evident from the figure there are stark differences in meeting the progression requirements between African and White students. While 84.16% of White students progress to the next year of study as AYOS-2 students, only 74.74% of African students are on track in year 2. Of the 74.74% African students who are on track, only 60% proceed to AYOS-3 in their third year. That is, only 44% of African students are on track to complete the degree in minimum time.

Figure 12 shows an even more dramatic picture of performance emerges for students who complete 4-year degrees. The overall proportion between years 1 and 3 for those who are on

[35] UCT does not offer majors. Instead, UCT refers to majors as specialisations.

track to complete the degree in minimum time are higher for the 4-year degree compared to the 3-year degree. However, once students get to their fourth and final year of study, the true differences in performance for on-track students become evident. Here only 28.61% of African students are on track compared to 66.3% of White students. Six out of every 10 White students will complete the degree in minimum time while fewer than three out of every 10 African students will complete their degree in the shortest amount of time. This is clear evidence that even for the most successful students in HE, the paths to completion are substantially different.

Figure 11 AYOS: 3-year degree

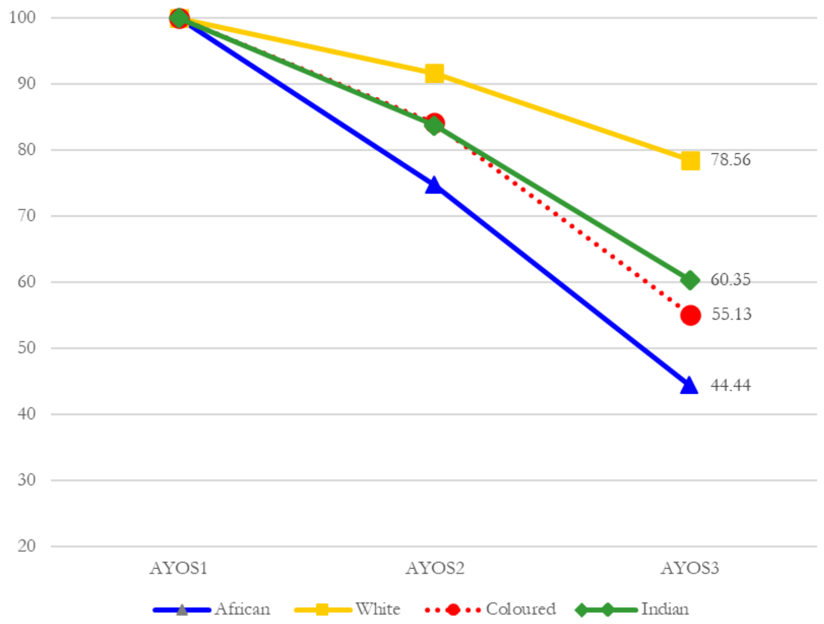

Figure 11 shows the AYOS or the proportion of students on track to complete in minimum time in a 3-year degree. Data: UCT IPD

Figure 12 AYOS: 4-year degree

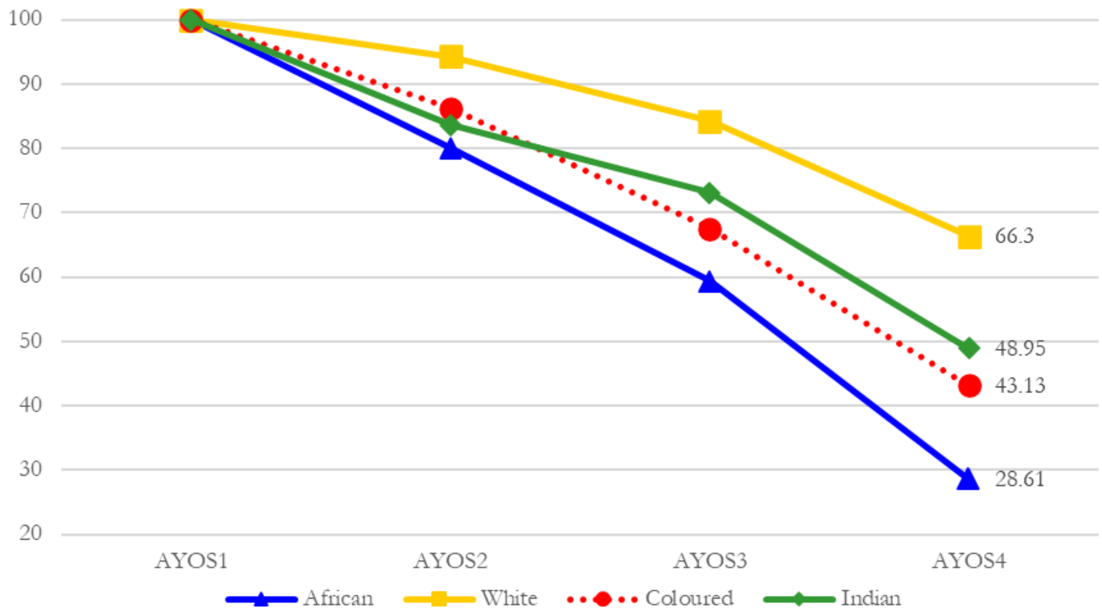

Figure 12 shows the AYOS or the proportion of students on track to complete in minimum time in a 4-year degree. Data: UCT IPD

4.6.3 Academic progression by code

Student progress through the system can be examined by focusing on the academic progression code awarded to each student at the end of the academic year. The university assigns each student an academic progression code based on their performance relative to the rules within each degree programme. This measure of success is based on students either receiving an academic progression code of 'academically eligible to continue' or qualifying for the award of the degree.[36]

Figure 13 shows the proportion of all first-time entering students undergraduate student success rates after 5 years.[37] A key point to note about the 5 year versus 6 year debate is that all programmes in Engineering and half of Commerce are 4-year degrees. The university has

[36] Academically eligible to continue is given the code *CONT*.

[37] The standard approach was to evaluate student success at n + 2; however, the university has done the analysis on time t = 5, not taking into account any differences for 3- or 4-year degree durations.

also expanded its offering of extended or augmented degrees, which allows students accepted into the academic development programme to complete the same courses but over a longer period of time. These options are available to previously disadvantaged students only to allow greater support of students to assimilate into the HE environment, ultimately allowing for slower progression during the first and second years of study. The aim is to provide superior academic support to students in smaller classrooms with expert lecturers in the hope that the more individualised attention in smaller classes will lead to improved academic success in later years.

As expected, white students are the best performers over a 5-year period, and 81% of all white students have graduated with a first degree after 5 years with only 4% continuing their studies. Compared to the average, white students significantly outperform all other race groups when evaluated using this measure. In contrast, African students have a 30 percentage-point lower graduation rate after 5 years. Of greater concern is the exclusion rate that is six times larger for African students compared to white students. At 31%, the exclusion rate for African students is double the university average over the period, pointing to a markedly different experience in HE for African students compared to all other race groups. An interesting observation from the data represented in this format is that African students have the lowest dropout rate when in good academic standing. While many may argue that this is evidence showing greater inclusion of African students at a historically white institution, it does not provide any information on how students finance their studies, a key determinant for students continuing through to graduation and not taking into account the high correlation between race and socio-economic background. With a growing student body of African students, the percentage of dropped out in good standing hides the actual number of students in this bracket leading to a biased representation of the data.

This data also highlights the significantly different completion rates by race. The maximum potential completion rate for white students is 85%, compared to 63% for African students specifically, and 70% for all Black students.[38] It is the first descriptive statistic to highlight the very different experience of HE for students by race.

This section shows that students experience heterogenous pathways through HE. Irrespective of the indicator used to measure academic performance, white students significantly outperform African students. Indian and Coloured students are located near the mean for some student indictors and are not often the worst performing group in HE.

Figure 13 Student progression status after 5 years

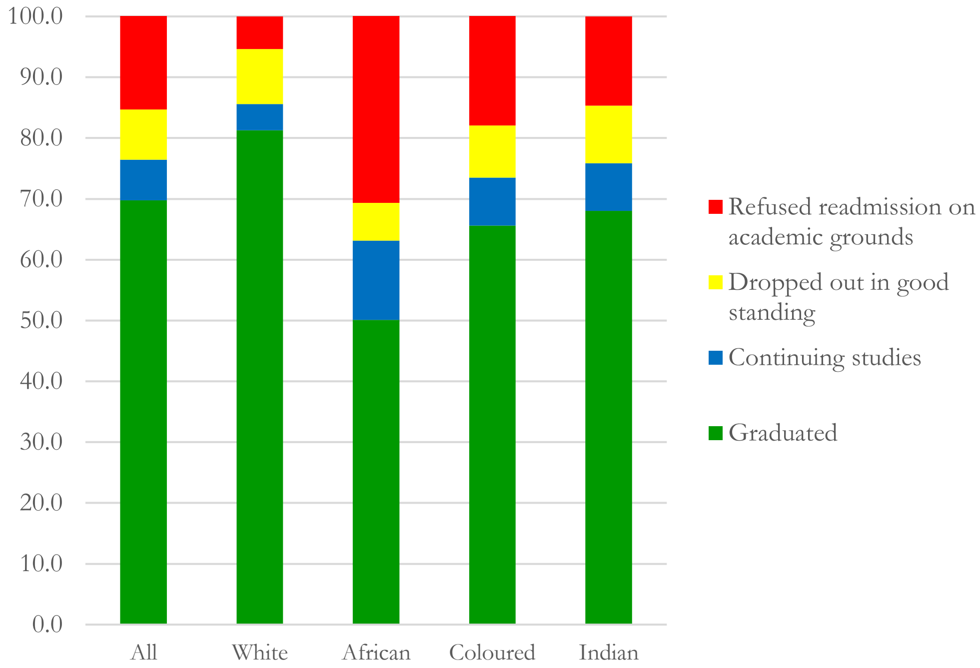

Figure 13 shows student progression status by race 5 years after enrolment.
Data: UCT IPD

[38] Here the term Black is used to capture African, Coloured and Indian/Asian students.

4.7 Estimation Strategy

The primary objective of this part of the study was understand the factors that influence performance in HE by using methods that try to control for the endogeneity of performance and the fact that there are differences between students who drop out and those who persist through to graduation.

An important limitation of longitudinal data is missingness. Data may be missing at random or missing completely at random if the missingness depends only on the observed responses. Where the data is missing at random or missing completely at random, the missing data process is ignored, and consistent estimates of population values may be obtained using maximum likelihood and using all the available data (Little & Rubin, 2002). This missingness is treated as ignorable or noninformative. However, it is of greater concern to econometricians when data is not missing at random (NMAR). This type of missingness is defined by the informativeness of the missing data. Little (1995) states that when the missing data depends on 1) the missing values themselves or 2) individual values of a latent variable such as a random coefficient that appears in the growth model, the missingness is not ignorable. The first case is referred to as outcome-dependent missingness (Little, 1995). Student dropout, which is the opposite of retention due to a deterioration in performance is categorised as outcome-dependent missingness. The second case is random-coefficient-dependent missingness, which is when the probability of dropout could be predicted by the student's performance trajectory through the intercept and slopes within the model. Therefore, NMAR requires explicit modelling of the missingness process together with the response variable. In this data, dropout was notably large at almost 25% of the student population, and may have caused biased results if not appropriately addressed. Put differently, retention (persistence) was not 100% due to dropout.

In the presence of data that is NMAR, any estimates derived from ordinary least squares estimations are affected by selection bias. Specifically, due to students who dropout at various points during their studies, not every student who starts their degree will complete their degree.[39] This non-random dropout is commonly classified NMAR because the dropout generating mechanism implies a missingness that is not random, which reduces the original sample of potential graduates, who were the target of this study.

In order to estimate the econometric models, we must first define the parameters of analysis. Students may decide to remain enrolled or dropout. This is the participation decision. For those who choose to remain enrolled, we can observe their GPA. In turn, the observed GPA can be used to calculate programme rank. Importantly, GPA will only be observed for those who choose to remain enrolled, that is, we only observe GPA for those retained in the system. Therefore the participation decision references retention and the outcome decision references rank, derived from GPA.

4.7.1 Heckman two-step approach

The Heckman selection model was deemed most appropriate to use in the presence of data that is NMAR because it was likely that unobservable factors may affect both the outcome of interest and the probability of selection (Miranda & Rabe-Hesketh, 2006).

4.7.1.1 Identification of the retention decision

The Heckman selection model is a two-equation model where the outcome regression equation is given by

[39] For the purposes of this study, stopout is ignored as the percentage of students who stopout is small.

$$y_i = x_i\beta + \mu_1 \tag{4.3}$$

where $y_i$represents the student's programme rank. The selection model is given by

$$z_i^* = w_i\gamma + \mu_2 \tag{4.4}$$

With $$z_i = \begin{cases} 1 \; if \; z_i^* > 0 \\ 0 \; if \; z_i^* \leq 0 \end{cases}$$

and it is assumed that the following holds

$$\mu_2 \sim N(0,1)$$

$$\mu_1 \sim N(0, \sigma^2)$$

$$corr(\mu_1, \mu_2) = \rho \tag{4.5}$$

where y_i denotes the dependent variable; x_i denotes the independent variables; and β denotes the parameters to be estimated. μ_1 is assumed to be a normally distributed error term with a mean of zero and a standard deviation of σ. z_i denotes observable characteristics including overlapping with x_i; γ denotes the parameters to be estimated; and μ_2 is a distributed error term with mean zero and a standard deviation equal to 1. ρ is the correlation between the two errors terms to be estimated. The parameter $\lambda = \sigma\rho$ is the estimated selection coefficient and is known as the inverse Mills ratio. If the coefficient on the inverse Mills ratio is statistically significant, then selection bias is present.

In the context of this chapter, the dependent variable is the change in programme rank between the first and last year. The independent variables are divided into three main categories of personal characteristics, socio-economic background and academic characteristics.

The proportion of courses failed in the first year served as the exclusion restriction in this analysis. This variable, known as an instrument in the economics literature, should ideally influence the probability of treatment without affecting the outcome variable. In reality, it remains difficult to measure and test that it is not correlated to the error term. Therefore, the choice of a valid instrument depends on economic reasoning and intuition. The choice of proportion of courses failed in the first year as the exclusion restriction assumes this variable would be correlated with attrition but is unlikely to impact on GPA. It is because proportion of courses failed in the first year is unlikely to impact programme rank, it meets the exclusion criteria.

In the main estimation of this analysis, it was assumed that a regression model can be used to explain changes between the first and last (graduating) year programme rank using the unadjusted programme rank score as the main dependent variable. The results for the selection equation are shown in Table 19 in section 4.8.1.

4.7.1.2 Identification of the outcome equation

The change in programme rank was estimated via the following equation:

$$y_{it} = \alpha + \beta_1 Z_i + \beta_2 K_i + \beta_3 X_i + \beta_4 IMR_i + \varepsilon_{it} \tag{4.6}$$

where y_{it} represents an academic outcome of student i in time t, Z_i is their personal characteristics, K_{it} is socio-economic characteristics, X_{it} represents factors such as academic characteristics, IMR_i is the Inverse Mills Ratio and ε_{it} represents unobserved determinants of academic outcomes. If the IMR is negative and significant then negative selection is present. In this instance, the coefficients would be biased downward without the correction. If the IMR is positive, then positive selection is present. The coefficients would be biased upward if no correction was made.

The analysis proceeds with the estimation of the change in rank between the first and last years. In each of the estimation, race and gender variables are interacted to estimate any differences between the race and gender groups.

4.8 Results

The results for the analysis are presented in this section. First, the results for the selection equation are shown. The probability that a student remains enrolled is estimated as the main participation equation. To shed light on the participation decision, Figure 14 shows the percent of students who remain enrolled at UCT after each year. The biggest dropout occurs at the end of the first year or the start of the second year. The figure shows that just less than 85% remain enrolled in year 2, which equates to dropout of slightly more than 15% at the end of the first year. The next steeped drop is between year 2 and 3, where approximately 4% of students leave. The retention rate is 80.62% after 3 years.

Figure 14 Retention rate over time

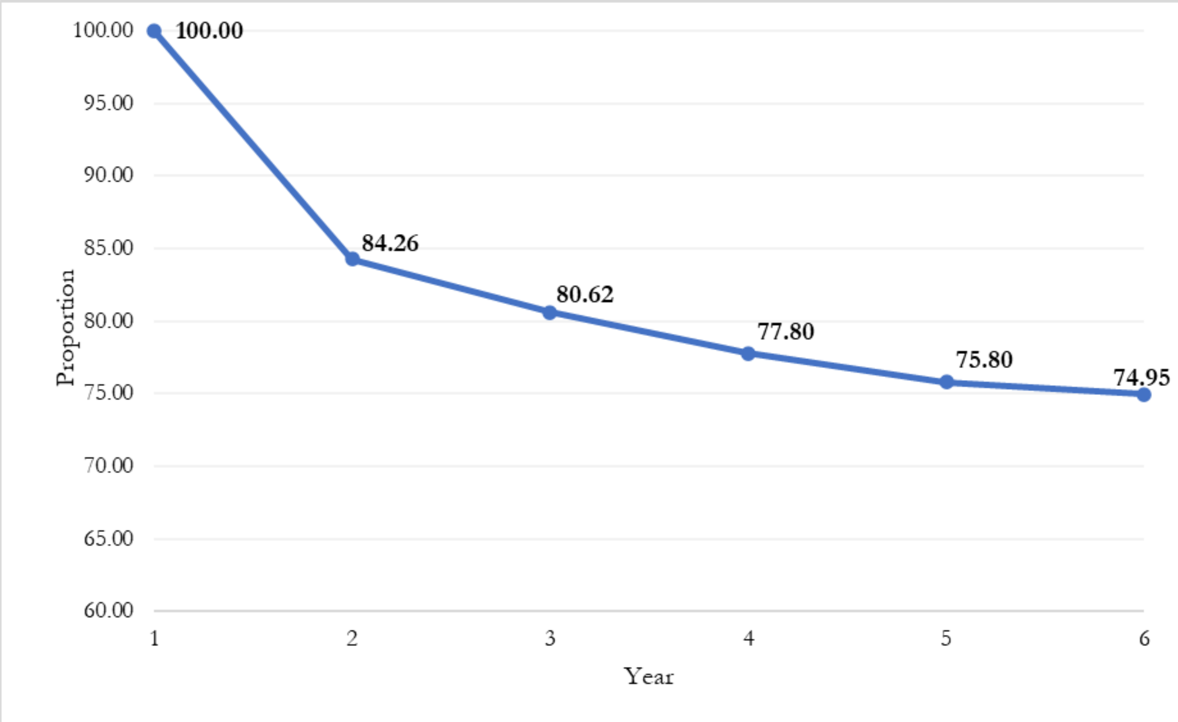

Figure 14 shows the retention rate, or the percent of students who remain enrolled over a six-year period.
Data: UCT IPD

4.8.1 Estimation of retention

Table 19 shows the results for the retention equation, that is, the probability that a student remains enrolled at the university. The estimations suggest that the likelihood of students remaining at university are influenced by race and gender, entrance score, if the student is enrolled in an academic programme and degree switching. African male is statistically significant across all the specification, with a consistently negative coefficient. The size of the coefficient shrinks substantially between specifications (1) and (6), as more indicators are added to the estimations. The other interesting personal characteristic is White male, which is also statistically significant across the first five specifications. The sign of the coefficient changes from positive to negative once all the independent variables are included. For socio-economic characteristics, which largely reflects students' backgrounds, entrance score is statistically significant in specifications (3) – (6). Initially, entrance score was positive, indicating that students with higher entrance scores are more likely to remain enrolled. In the full specification the sign of the entrance score coefficient becomes negative, indicating that students with higher entrance scores are less likely to remain enrolled at the university. Students enrolled in extended academic programmes, which are programmes in support of redress and equity, have negative signs where the variable is included. These students are less likely to remain enrolled at the university. Students enrolled in extended programmes are students who are African, Coloured or Indian. No White students were allowed to register for extended programmes during the period under consideration.

Table 19 Probit estimation of student retention

	Probit estimation of student retention (probability that a student will remain enrolled)					
	(1)	(2)	(3)	(4)	(5)	(6)
Personal characteristics						
White male	0.131***	0.128***	0.124**	0.167**	0.177**	-0.031***
	(0.044)	(0.046)	(0.052)	(0.068)	(0.073)	(0.007)
African female	-0.089	-0.0139	-0.083	0.134*	0.138**	-0.036**
	(0.095)	(0.080)	(0.099)	(0.080)	(0.084)	(0.016)
African male	-0.573***	-0.494***	-0.576***	-0.209**	-0.201**	-0.109***
	(0.097)	(0.082)	(0.098)	(0.094)	(0.096)	(0.013)
Coloured female	0.051	0.097	0.065	0.195**	0.192*	-0.003
	(0.094)	(0.089)	(0.100)	(0.098)	(0.099)	(0.011)
Coloured male	-0.097	-0.047	-0.092	0.015	0.027	-0.051***
	(0.078)	(0.064)	(0.077)	(0.061)	(0.061)	(0.013)
Indian fermale	0.080	0.109	0.028	0.065	0.072	-0.003
	(0.105)	(0.104)	(0.112)	(0.122)	(0.124)	(0.014)
Indian male	0.043	0.066	0.01	0.073	(0.095)	-0.066***
	(0.081)	(0.077)	(0.090)	(0.090)	(0.084)	(0.016)
Socio-economic characteristics						
Entrance Score			0.011***	0.023***	0.024***	-0.004***
			(0.002)	(0.004)	(0.003)	(0.001)
School: DET				-0.177**	-0.172**	-0.022
				(0.083)	(0.081)	(0.023)
School: HoD				-0.027	-0.023	-0.016
				(0.124)	(0.127)	(0.025)
School: HoR				0.042	0.038	-0.006
				(0.072)	(0.071)	(0.016)
Private School				0.03	0.0334	0.021***
				(0.057)	(0.058)	(0.0088)
Affluence				0.191*	0.194*	0.011
				(0.103)	(0.103)	(0.015)
Academic characteristics						
3 year degree					0.083	-0.059***
					(0.079)	(0.022)
Financial Aid						-0.002
						(0.001)
Humanities						0.049**
						(0.023)
Science						0.087***
						(0.025)
Engineering						0.028
						(0.0201)
English HL						0.0415***
						(0.009)
Residence						-0.019***
						(0.006)
Extended programme		-0.241**				-0.052**
		(0.116)				(0.023)
Degree Switching			-0.593***			-0.056**
			(0.181)			(0.0256)
Proportion of courses failed	-0.0289***	-0.028***	-0.028***	-0.027***	-0.027***	0.001***
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Observations	6949	6949	6949	6949	6949	6949
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses ***p<0.01, ** p<0.05, *p<0.10. Data source: UCT IPD 2006, 2007 and 2008

4.8.2 Estimation of rank

Table 20 presents the coefficients for the Heckman selection model where the dependent variable is the change in unadjusted rank. The proportion of courses failed in the first year of registration was the exclusion restriction in the selection equation (the probability that the student will remain enrolled). This variable controlled for the observation that students who fail many courses in their first year are less likely to remain enrolled, but if they do manage to complete their studies, they will show an improvement in GPA and rank. While this is a weak instrument, when encountering data limitations, trade-offs must be made.

The result of the likelihood ratio tests across all specifications suggested that the assumption of uncorrelated errors between the rank and selection equations could be rejected. This implied that it was appropriate to correct for selection bias, noting reservations about the specification of the model. Rho represents the estimated correlation coefficient between the error terms of the rank and selection equations. A positive sign of rho indicates that the unobservables are positively correlated with one another.

Table 20 Heckman selection estimates of change in unadjusted rank

	Dependent Variable: Change in unadjusted rank					
	(1)	(2)	(3)	(4)	(5)	(6)
Personal characteristics						
White male	-0.026***	-0.026***	-0.027***	-0.024***	-0.029***	-0.031***
	(0.006)	(0.013)	(0.006)	(0.007)	(0.007)	(0.006)
African female	-0.080***	-0.069***	-0.087***	-0.085***	-0.088***	-0.036**
	(0.018)	(0.013)	(0.018)	(0.024)	(0.023)	(0.015)
African male	-0.145***	-0.133***	-0.155***	-0.150***	-0.154***	-0.108***
	(0.014)	(0.011)	(0.015)	(0.019)	(0.018)	(0.012)
Coloured female	-0.004	0.002	-0.007	-0.011	-0.011	-0.003
	0(.012)	(0.009)	(0.013)	(0.013)	(0.013)	(0.011)
Coloured male	-0.041***	-0.034***	-0.048***	-0.050***	-0.057***	-0.051***
	(0.013)	(0.012)	(0.012)	(0.011)	(0.011)	(0.012)
Indian fermale	-0.012	-0.007	-0.012	-0.005	-0.010	-0.003
	(0.015)	(0.015)	(0.014)	(0.016)	(0.016)	(0.013)
Indian male	-0.068***	-0.064***	-0.067***	-0.062***	-0.072***	-0.065***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.018)	(0.016)
Socio-economic characteristics						
Entrance Score			-0.002***	-0.004***	-0.004***	-0.004***
			(0.000)	(0.001)	(0.001)	(0.000)
School: DET				-0.014	-0.015	-0.021
				(0.026)	(0.025)	(0.023)
School: HoD				-0.014	-0.012	-0.015
				(0.022)	(0.023)	(0.025)
School: HoR				-0.008	-0.007	-0.005
				(0.019)	(0.018)	(0.016)
Private School				0.020***	0.018**	0.021***
				(0.006)	(0.007)	(0.007)
Affluence				-0.002	-0.004	-0.010
				(0.015)	(0.015)	(0.015)
Academic characteristics						
3 year degree					-0.033*	-0.058***
					(0.019)	(0.021)
Financial Aid						-0.001
						(0.010)
Humanities						0.049**
						(0.022)
Science						0.086***
						(0.024)
Engineering						0.0284
						(0.020)
English HL						0.041**
						(0.009)
Residence						-0.018**
						(0.005)
Extended programme		-0.042				-0.052**
		(0.035)				(0.022)
Degree Switching			-0.065**			-0.056**
			(0.029)			(0.025)
Observations	6949	6949	6949	6949	6949	6949
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses ***p<0.01, ** p<0.05, *p<0.10. Data source: UCT IPD 2006, 2007 and 2008

The r-squared for each of the specifications in Table 20 are very low, and range from 3% for column (1) to 10% for column (6).

Starting with the first model, the reported marginal effects showed that African males experience the largest decline in rank relative to White females when all else is equal. In addition, African students are the only racial group to experience a statistically significant decline in relative rank compared to the base group of White female students. Across the specifications, the results show a dominant trend by males to be more likely to experience declines in relative rank, compared to White females. Male Indian students also experience a statistically significant decline in relative rank but the effect is relatively small at 0.06 of a standard deviation. The effects for Coloured female students were not statistically significant but was significant for Coloured male students across all estimations, which was not obvious when looking at the raw data presented in the previous sections. However, this initial specification only controls for race and gender, the key variables of interest in this analysis. Additional controls were added to the remaining models to improve specifications.

The next specification added a dummy variable for the extended programmes. These programmes are the key initiatives at the university to address equity by providing smaller, more specialised classes to African, Coloured and Indian students. Surprisingly, the coefficient on this variable was statistically insignificant. The third specification (3) added the entrance score based on the standardised Grade 12 examination grades to the model. The size of the coefficient was statistically significant and small, but it was the negative sign that was of interest. Intuitively, the negative sign can be interpreted as the higher the Grade 12 examination scores, the less likely a student is to experience a change in class rank. This is because students with higher Grade 12 scores tend to find themselves higher up the ranking system, leaving little room to improve rank over time. The coefficient was significant at the 10% significance level. In addition to adding the entrance score, this specification also added the switching variable.

Students change degree programmes at a high rate in the data, averaging just over 60% for the full and graduating samples. With such a high rate of degree switching taking place, it was important to interrogate if this action, which is voluntary in most occurrences, has an impact on grades. The results showed that degree switching has a statistically significant but negative effect on improving rank. This is unsurprising. Astorne-Figari and Speer (2019) find that even though changing majors is common in the US, it did not lead to grade improvements.

The addition of controls for high schooling yielded one of the most interesting results for the analysis. Adding dummy variables for the different types of high schools attended by students showed no statistically significant effects for the set of public schools attended. The only schooling variable that was significant was the indicator variable for private schooling, which was highly significant across models (4) – (6). This means that the private school variable captured something about students who attended these schools. It may also have captured an affluence or income effect as typically students who attend private schools come from wealthier income backgrounds. It may have also captured information about quality of education at these types of schools.

The final specification added dummy variables controlling for the duration of degree programmes. This full specification shed light on the key variables that appear to drive improvements in performance given through the change in the ranking variable. Students who attended private schools and whose home language is English show statistically significant improvements in rank over time. Males, students in the Engineering faculty, residential status, entrance score, and degree switching coefficients were all statistically significant and negative in their impact on improvement of rank. An important observation in Model 6 was that coefficient on degree switching changed to statistically significant in this final specification. This indicated that once a more complete picture of the determinants of performance was established, an improved presentation of the factors underpinning performance was available.

The findings in this chapter are supported by some of the findings in Chapter 3, namely that males are less incentivised to improve performance by aiming for high grades. Instead, males significantly underperform in their senior years of study relative to females. This outcome is also supported by the findings in Chapter 5.

The gender difference noted in Table 20 is striking as it supports other gender-driven findings in this study. It therefore merited additional investigation to explore the correlates of degree switching.

4.8.2.1 Degree switching correlates

The strongest predictors of academic achievement are race, gender and school-leaving exam scores. In previous estimations it was noted that degree switching showed as statistically significant when a more complete model specification was established. As the proportion of students who change degrees is high at over 60% for the sample, it was worth investigating the correlates of degree switching.

Table 21 presents the results for the correlates of degree switching. At an aggregate level it appears that there are different drivers of degree switching between males and females. This is shown through the differing factors that are significant in each of the equations in Table 15.

Surprisingly, none of the personal factors were significant. The African race variable has been significant in most specifications. Of the schooling background variables, House of Representatives or individuals who attended previously classified House of Representative schools are more likely to change degrees compared to individuals who attended House of Assemblies schools (the omitted category). House of Representatives was the name of schools previously attended by Coloured students. It is possible that students from these schools enrolled in programmes that did not match their expectations or interests, and therefore, they are more likely to change degrees at some point during their academic careers. Unfortunately, the data in this study was limited and the exact timing of degree changes could not be observed.

Table 21 Major switching by gender

	Dependent Variable: Degree Switching	
	Males	Females
Personal characteristics		
African	-0.076	0.018
	(0.117)	(0.098)
Coloured	-0.007	0.033
	(0.089)	(0.081)
Indian	-0.158	0.158
	(0.091)	(0.101)
School characteristics		
Entrance Score	0.003	0.169***
	(0.004)	(0.003)
School: DET	-0.085	-0.225**
	(0.117)	(0.095)
School: HoD	-0.094	-0.086
	(0.162)	(0.140)
School: HoR	0.227*	0.186**
	(0.130)	(0.087)
Private School	-0.096*	0.050
	(0.057)	(0.073)
Academic characteristics		
3 year degree	-0.230	0.004
	(0.335)	(0.313)
Financial Aid	-0.409***	-0.330***
	(0.085)	(0.091)
Humanities	-0.079	-0.162
	(0.399)	(0.375)
Science	-0.184	0.015
	(0.414)	(0.377)
Engineering	-0.256	-0.124
	(0.411)	(0.353)
Extended programme	-0.077	0.026
	(0.326)	(0.316)
Residence	0.405***	0.478***
	(0.078)	(0.096)
English HL	0.100	-0.079
	(0.108)	(0.082)
Change in unadjusted rank	0.458***	0.404**
	(0.152)	(0.166)
Major dummies	Yes	Yes
Time dummies	Yes	Yes
Observations	3531	3418

Note: Standard errors in parentheses ***p<0.01, ** p<0.05, *p<0.10. The sample includes all students who graduated. The dependent variable is a dummy for degree switching. Data source: UCT IPD 2006, 2007 and 2008

4.8.3 Study limitations

A number of limitations should be acknowledged for the analysis in this chapter. One of the key limitations of the analysis in this chapter was the lack of data pertaining to students' home background, including family income and parental education. For the cohorts analysed in this study, UCT did not ask students to voluntarily disclose this type of information because admission decisions were not based on income. It is only since 2010 that applicants were asked to submit this type of sensitive information with the view to provide greater insight into the first-time entering student cohort. This type of omitted variable bias was an important limitation that should be acknowledged. The implementation of the Heckman selection model helped correct for some of the selection biases present but may not have captured all the necessary information.

It may also be argued that focusing this study on one HEI might have made the findings less generalisable and the results less pertinent to other HE environments. Limiting the study to one institution removed the variability of grading patterns observable across HEIs in South Africa, which makes it more difficult to directly compare across all institutions. Instead, by focusing on UCT we were able to ensure that the results were not driven by differences in grading patterns across institutions. This provided more stable results.

The ideal analysis for data in HE was selection bias correction within panel data methods. The timed nature of academic decisions made it attractive to exploit the longitudinal variation in performance over time using panel data methods. In addition, student pathways through HE are not homogenous and panel data methods could better account for these varying observed pathways. In addition, developing countries are likely to suffer from high HE dropout, and the resulting selection bias in the data had to be corrected.

4.9 Conclusion

This chapter presents a way to consider and evaluate the DML incentive policy in the context of an African economy. While many colleges and universities have policies that recognize good student performance, very little is known about the impact these policies have on student outcomes. Only a few papers in the economics literature have examined them and, to the best of the author's knowledge, the international literature in both developing and developed economies is thin. It is difficult to locate the results of this study in the context of the broader literature on higher education systems around the world.

This chapter evaluates the impact of academic recognition policies, specifically the DML on student outcomes. Using a regression discontinuity approach, it shows that the DML as an academic excellence incentive policy has large negative (rather than the intended positive) effects over the short and long-run on the academic performance of South African students. The effects of the policy on exit outcomes such as graduation and dropout appear to be within the expected range but also show the unanticipated negative effect for treated students relative to the control group.

Two main specifications were run. The first specification was a baseline specification where no baseline or pre-determined covariates were included. This approach is recommended by Cattaneo, et al (2021) as canonical RD designs do not necessitate the inclusion of covariates. For this specification, the exit outcome specification delivers results with the expected magnitude of coefficients but with unexpected signs. The size of the coefficients for the GPA analysis are significantly larger than those for exit outcomes, yielding more unstable results for student GPAs.

However, Cattaneo et al (2021) strongly recommends the reporting of treatment effects with and without covariates as the treatment effects are expected to be stable across covariate inclusion. The results for exit outcomes with covariate adjustment are robust to the inclusion of all pre-determined covariates. The negative impact of the policy on treated students is still observed in these estimations. While the results for the GPA analysis show of the coefficients shrinking, the size of the estimates remain of concern. These results suggest that the impact of the DML policy on future student GPA should be an area of further research.

These results seem particularly counterintuitive relative to the findings of international empirical studies which show that DML reward policies have significant and positive impacts over the long run. However, they are in line with Bénabou and Tirole's (2002) theoretical expectations regarding extrinsic motivation in a situation of asymmetric information between an agent and principal. A few reasons are suggested for the difference in findings between this study and that of international investigations.

Regression discontinuity estimates local average treatment effects around a threshold or cutoff point. The treated group is compared to the control groups, where treatment and control groups are most similar around the cutoff. Importantly, Local average treatment effects provide evidence on whether a policy may be expanded or reduced at the margin, however, they do not inform whether or not a policy should exist.

The findings of this study may be of use to academic administrators. The results suggest that the award of the DML does not achieve its stated goals for the students involved. It may make more sense for administrators to re-imagine the policy at the margin, perhaps changing the cutoff slightly, or conducting a formal RCT, and then re-evaluating the impact on academic and labour market outcomes. Without knowing longer-run impacts of the policy, it would be informative for academic administrators to first understand post-graduate and labour market effects before considering a full scrapping or large changes to the policy, which current estimates cannot inform as they do not provide average treatment effects. Future research may benefit from collaboration in investigating this type of policy across institutions in South

Africa, and this could include variation in implementation across institutions to provide further insights into the effects of university policies. By focusing first on expanding the knowledge on postgraduate education and labour market outcomes, the additional information will provide more certain pathways for future policy approaches.

The findings of the chapter do not accord with existing literature on the topic. For example, Seaver and Quarton (1973) found positive impacts of the DML policy on the treated students. Similarly, Wright (2018) also finds positive impacts on student GPA.

The negative impact of the DML persist in this analysis when examining the results with and without covariates. This raises the possibility that additional information, obtained through surveys of students near the cutoff, could provide the data needed to fully investigate the underlying causal mechanisms at play. Additional qualitative information would help inform the quantitative outcomes observed in the education sector.

There are a few limitations to note with respect to this analysis. First, the study presents local average treatment effects, which cannot be extrapolated to the entire population. On the basis of these one cannot say whether the policy should be deleted in its entirety. Researchers may only recommend the removal of the policy if average treatment effects are established, which this study is unable to do.

Second, it would be interesting and valuable to academic administrators to know the impact of the timing of the DML award on future academic outcomes. Shifting the award to the second year of study, and not offering it in the first year may drive changes in behaviour for undergraduate students which are not observed when the DML is awarded after first year. Put differently, the award of the DML closer to degree completion may result in different behavioural outcomes. Third, leading on from the second limitation, it would be useful to see if the DML incentive policy has any impact on postgraduate study and labour market outcomes.

This would help to show the full long-term impact of the policy on outcomes beyond higher education.

While the results presented in this chapter are largely discouraging, they do present an opportunity for university administrators to create research-driven policy agendas, which will enable the university to revise, rethink and update the student incentive structure.

Chapter 5: The Determinants of Academic Outcomes: A Competing Risks Approach

5.1 Introduction

A HE qualification leads to significant gains in the labour market. Individuals with tertiary level qualifications are significantly more likely to be formally employed than individuals who only completed secondary schooling. In addition, the average income of tertiary educated workers is up to three times that of individuals with only a matric qualification (Branson, Leibbrandt & Zuze, 2009). The return on post-secondary education varies significantly around the world, but this return is especially high for South Africa (Branson, Leibbrandt & Zuze, 2009; Keswell & Poswell, 2004; OECD, 2014). In addition to the high labour market return that HE graduates can expect, these individuals typically also face lower unemployment rates. The average unemployment rate of 7% over time for tertiary graduates is significantly lower than the 32% for individuals who completed secondary schooling but did not continued to tertiary education (Statistics South Africa, 2019).[40] It also compares well with the national average unemployment rate of 25% over the last 10 years.

Enrolling in tertiary education does not guarantee an individual will successfully complete their studies. South Africa still struggles to address the inequity in education attainment. This is one of the deep-rooted outcomes of the apartheid regime, which contribute to driving divergent economic paths for individuals of different race groups. The country also has one of the lowest HE completion rates in the world, last measured at 15% nationally (Letseka & Maile,

[40] Using StatsSA data over the period 2010-2015 from the Quarterly Labour Force Surveys (QFLS)

2008).[41] The population share of tertiary graduates at 6% of the adult population is significantly lower than that of other middle-income or developed economies (OECD, 2014). In the context of a poor secondary schooling system and the high cost of tertiary education, students bear significant financial risks when they drop out before completion. Therefore, the opportunity cost of not completing is high. Forgone income and the cost of tuition bear significantly on students, and there is little support from tertiary and financial institutions available in South Africa. The South African government has made a substantial proportion of all expenditure on HE available as financial support to students, but an extremely stringent means test for eligibility excludes most students from accessing these funds. Students who drop out before completion are left with sizeable student loans or debts. The costs, both direct and indirect, imposed on individuals and society when students do not complete tertiary education are high. Student debt remains even though an individual has not completed their education, and academic institutions lose government subsidies they would have received had the student graduated. The South African National Treasury reported that the high dropout rate cost taxpayers about R4.5 billion over the period 2000–2005 (Letseka & Maile, 2008).

Another consideration is that students' academic career paths may vary significantly. Some students will complete degrees in the minimum specified time, and others may take a few years longer to complete the same degree. About 20% of students complete degrees on time in South Africa (Council on Higher Education, 2018). Some students take a break during their studies and return to complete their degrees at a later stage. This is commonly known as stopout. It is particularly difficult to measure as there is not fully established national database on student activities. Other students change degrees completely and possibly start from scratch; time to completion vary depending on how transferable subjects are between degrees. Some

[41] Measured in 2000.

students drop out and do not return to their studies, incurring costs not only to themselves but to society at large.

Understanding when dropout is likely to occur in South Africa is important given concerns about the quality of the education system. This will also help policy makers tailor interventionist policies much better in the education sector, potentially leading to improved outcomes for education retention, progression and attainment. It is therefore important to identify the factors that influence students' decision to drop out or persist with their studies. From an institution's perspective, this is important in terms of resource allocation as the funds available to support institutions from the government are not increasing fast enough to sufficiently cover a growing body of students requiring financial assistance.[42] From the students' perspective, the cost of acquiring a tertiary education is high and better information will help students make more informed decisions. Improving the performance of both students and the HE sector should be fundamental to a society like South Africa where pathways out of poverty are largely facilitated through educational attainment.

The key contribution of this chapter is studying the determinants of dropout and graduation in HE in an African country to understand the unique issues at play in this context. The literature emanating from developing countries is scarce relative to developed countries. The analysis further evaluated by gender and race. In addition, attention was given to some factors whose influence may vary over the duration of studies, such as race effects and the switching of degrees. Focusing on one HEI allowed for the control of effects experienced by students facing the same organisational context, level of academic support, and other services while

[42] Real growth in expenditure is low or negative.

eliminating the heterogeneity that could be introduced by including multiple institutions (see Christie, Munro & Fisher, 2004; Patrick, 2001).

The aim of this section of the study was two-fold. Firstly, this work will help university administrators identify factors that play a role in predicting student success and dropout. This is especially important when the schooling system varies significantly in quality between the public and private spheres of both the primary and secondary education sector (Spaull, 2013a). Moreover, as universities are adapting admission criteria to ensure greater access and equity in student enrolments, understanding which factors contribute to success is critical. Many view a university education as the gateway to some of the most lucrative and well-paying jobs in the labour market; student potential to succeed in HE is significant in terms of equity and access. This section also contributes to the research on performance in HE by examining the variation in performance over time. This will enable university administrators to identify key bottlenecks in student progression and allow for improved targeting in terms of academic interventions. Lastly, it will help identify the types of interventions that may be necessary at different levels of study, and how these vary by faculty or discipline.

Time to dropout or degree completion can be determined by several factors. This study focused on several aspects. Firstly, the determinants of dropout and graduation were examined, and the results were broken down by race and gender to identify any differences. Section 5.2 presents the literature on the determinants of dropout, degree completion and other relevant research. Section 5.3 discusses the methodology, and section 5.4 describes the data used in this study. Empirical results are discussed in section 5.5. The final section considers conclusions and the implications of the results of this study.

5.2 Literature Review

The literature review below largely focuses on dropout or student attrition. The review is structured around groups of determinants that influence the decision to drop out. The first section highlights the key contributions to the theoretical framework described in Section 5.3by Tinto (1975), Bean (1980) and the college choice nexus model (St. John et al, 2000). The remaining groups are clustered around student characteristics such as background and ability, HE experience, and labour market performance.

There is a substantial literature on the factors that determine university attrition or dropout. Tinto's (1975), Bean's (1980) models and the college choice nexus model (St. John et al., 2000) are the most comprehensive theoretical models of student attrition. Emerging from the psychology literature, Tinto's student integration model theorises that students assimilate and interact with their environment on two key levels. The first interaction is an academic one where the students interact with faculty and staff. The second is the social component of interaction with peers and participation in extracurricular activities. One of the first longitudinal studies of dropout behaviour, Tinto posits that students who integrate both academically and socially into the university community are more likely to feel a sense of belonging within the institution, and therefore, more likely to graduate. Thus the interaction between students and the institution ultimately impacts a student's decision to persist or not. Although it acknowledges the importance of external factors, such as job market opportunities and home background, a major shortcoming of this model is the exclusion of these external factors from the institutional environment, which Tinto assumes to be given and constant. Bean's student attrition model builds on and is slightly more comprehensive than Tinto's model in that it includes external factors, such as employment opportunities, as well as behavioural factors as predictors of persistence. The college choice nexus model (St. John et al., 2000) suggests there is a strong relationship between a student's college choice and persistence at college. The model posits a three-stage process: Socio-economic and academic ability factors affect whether a student goes to college; next, the student compares the costs and benefits of attending a given college; and once enrolled, the student’s college experiences and academic performance affect whether they drop out or continue (St. John et al., 2000). St. John et al. (2000) finds that financial aid positively impacts student retention as it decreases the cost of attending college. Poor grades and other negative college experiences impact negatively on persistence, making it more likely for students to drop out (St. John et al., 2000).

The dropout literature focusing on student characteristics is extensive. Early research by Tinto (1975) and Bean (1980), and later research by Tinto (1993) show that student characteristics play an important role in explaining not only dropout, but also persistence. Student characteristics or personal characteristics are shown to be important predictors of dropout. Research covering gender effects show less robust results. Montmarquette, Mahseredjian, & Houle (2001), McNabb, Pal, & Sloane, (2002) and Smith and Naylor (2001) find that females are less likely to drop out of university. Goldin, Katz and Kuziemko (2006) find that rising age of first marriage, rising expectations of future employment opportunities and improved labour force participation partly explains the better performance and completion rates for females.

Another exogenous factor related to dropout is race. Studies in the US and UK contexts refer to student performance or dropout of minority students (Alon, 2007; Light and Strayer, 2000; Vignoles and Powdthavee, 2009). Minority students are less likely to graduate, shown to take longer to graduate, and more likely to dropout prior to completion (Bowen and Bok, 1998). Alon and Tienda (2005) take the analysis a step further by investigating a “mismatch” hypothesis, which predicts a lower graduation rate for minority students who attend selective institutions. They do not find support for the “mismatch” hypothesis, instead they find strong evidence for improved performances by minority students who attend selective institutions, based on higher quality teaching and learning environments, and higher quality peers.

Ability is another prominent factor that should be considered. This is largely proxied through high school attended and the grade achieved on school-exit exams or university entrance exams. Smith and Naylor (2001) and Stratton, O'Toole & Wetzel (2008) show that students with better high school marks are less likely to drop out of university. However, other studies have shown that students with higher high school grades are more likely to dropout (DesJardins et al, 1999). This relates the ideas of high school grades and prior education performance, expectations and HE performance. This links to Tinto (1975) and Bean's (1980) idea of academic integration and updating beliefs based on academic performance. Stinebrickner and Stinebrickner (2012) find that at enrolment, students tend to underestimate the possibility of a bad grade performance but update their beliefs over time by taking into account new information. The revealing of information, some of it through HE academic performance, plays a role in the dropout decision. Horn (1998) finds that students who dropout are more likely to be older, have children, or work fulltime.

The experience of HE once enrolled should also be taken into account. The completion of all first-year subjects is an important determinant for year two enrolment, drop out and subsequent degree completion. Stinebrickner and Stinebrickner (2014) model the influence of poor grade performance on dropout. The link between Becker's (1964) human capital model and Tinto (1975) and Bean (1980) should not be overlooked here. In line with Tinto (1975) and Bean (1980), they find that a poor early university performance reduces the enjoyability of remaining in school and lowers the post-school financial return that may be expected.

Light (1996) finds that local wage rates and unemployment rates are statistically significant in determining academic success and persistence of studies. Financial aid is an important determinant for both graduation and dropout (Ishitani & DesJardins, 2002; Sampaio, 2012; Stratton, O'Toole & Wetzel, 2008).

Another prominent factor influencing drop out is that of financial aid, which the literature highlights come in many forms, including scholarships, fee exemptions, food stamps and housing (Aina et al, 2022). Early research on financial aid focused on the traditional definition of aid – that of aid given to low-income or need-based students whose family income was below certain thresholds. Alon (2007) assesses the effectiveness of financial aid in the persistence of minority students in the United States. Using measures for aid eligibility and aid amount, he found a negative relationship between need-based eligibility and the probability of graduation, but a positive relationship between aid amount and the probability of graduation, conditional on eligibility. A key issue relating to financial aid arose from this strand of literature – the puzzling negative impact of financial aid on persistence, and the endogenous access to financial aid (Aina et al, 2022). Typically, students who are eligible for aid are non-randomly selected as they tend to be low-income students. Non-random selection confounds the output of an OLS regression due to endogeneity. Stratton, O'Toole & Wetzel (2008) analyse two types of aid including student loans and scholarships, finding that students who receive scholarships are more likely to drop out relative to students who receive loans. However, they fail to take into the account the non-random selection of financial aid students. Alon (2007), using an IV strategy, shows that there is a positive relationship between grant aid and persistence. Dynarski (2004, 2008) and Scott-Clayton (2011) examine the causal effects of financial aid in the form of merit-based aid in the US, finding that this type of aid increased enrolment and to a lesser extent increased degree completion.[43]

[43] For a detailed discussion on the role of financial aid see Deming and Dynarski (2010).

Ishitani and DesJardins (2002) find that dropout rates vary with both the value and timing of financial aid. Stratton, O'Toole & Wetzel (2008) focuses on the determinants of dropout and stopout. Stopout is defined as the temporary absence from college or university enrolment. While stopout is more short-term in nature, dropout is deemed to be a long-term phenomenon. Employing a multinomial logit, they find significant differences between stopout and dropout. The type of financial aid received is an important determinant of both phenomena. Financial aid is significantly more important for dropout than stopout as financial aid typically equals loans that must be repaid. Sampaio (2012) studied students in the Brazilian HE system. Implementing a variation of a survival data model, the author finds that better performance in the entrance exam is correlated with a positive increase in the probability of dropout for earlier quantiles but significantly decreases the probability of dropout for higher quantiles. However, once degree major fixed effects are included to control for differences in degree types, entrance scores become positively correlated with the probability of dropout at all quantiles of the distribution. The author also finds that students in majors with low entrance requirements are more likely to switch majors.

The empirical literature can be broken down into two broad approaches. The first approach is the static type models of OLS-based analysis. Montmarquette, Mahseredjian, and Houle (2001) finds that class size of compulsory first-year courses affect a student's probability of dropout in a non-linear manner. Using teacher experience, the authors find that class sizes in excess of 87 students increased the probability of dropout and that smaller classes have a positive impact on persistence. A major shortcoming of this static literature is that students are observed at two points in time only, usually at initial enrolment and after one term (semester) or one year. Other studies suffer from small sample sizes and too short time periods (Stratton, O'Toole & Wetzel, 2008). This does not conform to our theoretical background based on Becker (1964), Tinto (1975) and Bean (1980) that includes some measures for academic assimilation. There is a second body of literature that uses survival analysis; it is small but growing. Scott and Kennedy (2005) laid the foundation for discrete-time event history analysis. Motivated to understand pathway effects of educational attainment, the authors set up the competing risks model to evaluate dropout and degree attainment. Murray (2014) is one of the first South African authors to investigate the factors affecting graduation and dropout using a competing risks framework and focuses on credit points to graduation as the dependent variable of the analysis. The author argues that this approach circumvents the stopout phenomena discussed above. Using a sample of students from the University of KwaZulu-Natal in South Africa, the author finds that residence and financial aid status are important determinants for both dropout and graduation. The receipt of financial aid and the provision of residence-based accommodation significantly assists students to graduate. The author also finds significant differences by race and gender. White females graduate at faster rates relative to African males as white females repeat fewer courses.

While the analysis in this study did not include foreign students, empirical evidence regarding students' nationalities is mixed. Arias and Dehon (2011) studied dropout and degree completion using discrete-time methods for competing risks survival analysis. They find that foreign students are more likely to experience consecutive enrolments without completing degree requirements. Foreign students have a more difficult time integrating into academic environments because the live away from home without the necessary support. Having a mother with a HE qualification contributes significantly to reducing the risk of dropping out while significantly increasing the chance of graduation. Not surprisingly, having a strong mathematical background significantly reduces the risk of dropping out during the first few years of study. In addition, a strong mathematics background significantly increases the chance of graduating in minimum time.

Tinto (1975) and Murray (2014) clearly distinguish between voluntary and involuntary withdrawal. Voluntary withdrawal is dropout of the student's own volition. Reasons for this may include poor early performance before any major exams are written, external factors such as family or financial issues, or realising they registered for an unsuitable or incompatible degree programme. Involuntary withdrawal focuses on the discontinuation of academic studies because of academic exclusion where students have not met the minimum progression requirements. This can also include failing the same course numerous times. It is important to note these distinctions as the quality of education in developing economies suggests poorer academic preparation for HE, and thereby, potentially leads to higher levels of dropout driven by poor academic performance resulting in involuntary academic exclusion (Clerici, Giraldo, & Meggiolaro, 2014).

5.3 Theoretical background

A dominant model within the economic literature is that of Becker's (1964) human capital model. This model assumes the decision to invest in education is the result of comparisons between expected costs and benefits. Individuals will complete a level of education, such as secondary school or university-level education, if the expected net present value of lifetime earnings is positive at the time of decision (Becker, 1964).

Before enrolling at university, students do not have perfect information about programmes of study, nor do they have updated information about the difficulty of different options, or the effort level required to pass subjects or courses. In addition, students may not know if their chosen professions will be interesting and intellectually rewarding once they enter the labour market. In essence, limited information may prevent students from correctly evaluating expected costs and benefits of the decision to invest in education.

Altonji (1993) and Stinebrickner and Stinebrickner (2003) expand on Becker's human capital model by introducing a dynamic element into decision making to take uncertainty into account. What this does is to more accurately reflect true decision making by students by introducing sequential decision making and allowing students to update information over time. From a university student's perspective, this means that they may update their information and preferences once new information is acquired.

Following Aina, et al (2021), if in every time period t ($t = 0,1,2,\ldots,x$) student i enrols at university, this may be written as

$$U\left(NPV_t^i, B_{NMt}^i\right) > C_{NM}(e_t^i) \tag{5.1}$$

where NPV is the expected net present value of meeting the requirements for the award of an undergraduate degree, B_{NM} is the expected non-monetary benefits, and C_{NM} , which is a function of effort e, is the expected non-monetary costs of studying. Then, expected monetary benefits and monetary costs (direct and indirect), determines the NPV of completing the university degree:

$$NPV_t^i = \sum_{j=x+1}^{L} \frac{Y_{Dt}^j}{(1+r)^j} - \sum_{j=1}^{x} \frac{C_{Mt}^j}{(1+r)^j} - \sum_{j=1}^{L} \frac{Y_{Nt}^j}{(1+r)^j}, \tag{5.2}$$

where Y_D represents annual earnings of a university graduate, C_M represents direct monetary costs of completing the degree, Y_N represents the annual earnings of a high school graduate, L is assumed to be the retirement age, and r represents the discount rate. The university student should also consider Y_N as representative of foregone earnings while studying at university. All values, including monetary benefits and costs may change during the time in which the student is enrolled at university.

Students may revise and update the education decision throughout the period of post-secondary study. In addition, the expected monetary benefits and costs associated with university education may change, resulting in a change in the NPV above. Closer to degree completion students may acquire more information about the non-monetary benefits associated with their degrees, updated information on labour market conditions and additional information about non-monetary costs (effort), which depends on their ability. Given this updated and additional acquisition of information, students may adapt their behaviour in accordance with the type of information acquired.

If expected monetary benefits and costs remain the same, a student may choose to

i. remain enrolled if the non-monetary benefits exceed the costs (effort) of study, $B_{NM} > C_{NM}$,

ii. drop out in non-monetary benefits are less than non-monetary costs, $B_{NM} < C_{NM}$.

Individual beliefs and information are continuously updated throughout enrolment. Through this updating of information, if it becomes sub-optimal to remain enrolled, irrespective of year of study, students may choose to dropout. Here costs exceed benefits. Alternatively, it may be assumed that students still observe the value of enrolment in university, or more specifically, still believe that benefits exceed costs, they will remain enrolled. However, time to graduation may be extended, especially in situations where students remain enrolled but change degrees. This then becomes a dynamic model where dropout, graduation in minimum time, or extended time enrolled before graduation are options.

Aina et al (2021) links this form of the human capital model to Tinto's model of dropout. Tinto's (1975) model of dropout is one of the earlier models to integrate individual level cost-benefit decision making into a formal model of dropout. Tinto asserts that students may integrate into the academic and social spheres of a HEI to feel a sense of belonging. Academic integration includes academic performance among other things. Social integration focuses on how well individuals assimilate to that HEI's environment (Tinto, 1993). Individuals set goals about education that are constantly updated to inform about their willingness to persist. To understand this process of persistence and dropout, information such as high school background and individual attributes should be known to the institution, as this informs the student's decisions around costs and benefits, and therefore, dropout or persistence (retention). Bean's (1980) model is similar to that of Tinto in that it draws on the individuals background information and their subsequent interaction and commitment to the institution once enrolled. Both Tinto (1975) and Bean (1980) asset that students' background information is highly informative vis-a-vis their decision to remain enrolled and therefore persist.

The combination of human capital theory with Tinto (1975) and Bean (1980) allows for an explicit link between the three different but similar theoretical frameworks. Specifically, an individual will choose to remain enrolled (not dropout) if the NPV is positive at each point in time a decision is likely to be made, but that decisions are made at an individual level, taking into account a students' background, experience and external factors, both academic and social, in HE.

5.4 Methodology

Survival analysis, also known as event history or hazard modelling, is often used to estimate the timing of events or longitudinal outcomes. Originally used in medical research by biostatisticians (Cox, 1972), survival analysis has been extended to social science research, including economics, political studies and education. Initially when applied to the economics of education studies, individual outcomes (for example dropout or graduation) for students were investigated using 'single-risk' models. Survival models are constructed to measure the probability of transition between two outcomes (Austin, Lee, & Fine ,2016). In most instances there are only two possible outcomes, for example, employed and unemployed, or smoking and non-smoking.

In certain situations, more than two outcomes are possible. Standard single-risk models do not take into account the potential interdependence between competing outcomes (for example, dropout, stopout and graduation).[44] A method known as competing risks is preferred when there are more than two possible outcomes (Singer & Willett, 1993). Many outcomes are mutually exclusive (non-overlapping) and should cover all states or events.

In terms of this analysis, the outcomes were defined as dropout that is voluntary, dropout that is involuntary (academic exclusion), and graduation. Although single-risk models may have been used, it would have resulted in misspecification due to correlation between events (DesJardins, Ahlburg & McCall, 2002). The appropriate model to implement with three competing outcomes or risks was the competing risks model, which evaluates the hazard simultaneously for events (Allison, 1984; Scott & Kennedy, 2005). One key advantage of this method was that we were able to investigate both *whether* individuals are likely to drop out and *when* they are most likely to do so. Some types of models also allow for the relative risk of different groups over time to event to be established, which is another advantage over more static methods.

Data using survival analysis techniques are separated into discrete and continuous-time data. When time is measured in discrete units, the use of discrete-time methods is more appropriate (Allison, 1984). Discrete-time data usually include education-type data as students are observed at most twice per year (at the end of each semester), but usually only once per year. As one does not observe the exact timing of dropout or other academic experiences (other

[44] Stopout is defined as the non-continuous enrolment of students.

than graduation) it was best to use competing risks discrete-time methods because it was not necessary to know the exact timing of events, only whether they have occurred.

A common issue arising in survival analysis is that of censoring (Cleves et al., 2010). Data is censored when the outcome of interest is not observed. Data may be right or left censored. Left censored data occurs when the starting point for individuals in the sample is not known. Right censored data is observed when the outcome or events of interest have not yet occurred for some individuals in the sample. In the case of this study, individuals were considered right censored because they are still in the system and have not dropped out nor graduated from the university. Survival analysis methods take into account this type of data censorship, thus yielding unbiased results.

The key to understanding competing outcomes is to understand the nature of the distribution of outcomes, and in particular, the correlation between possible outcomes in this framework (Cleves et al., 2010). If the two potential outcomes are uncorrelated, standard survival analysis may be applied to the problem as future events are treated as censored data observations. However, dropping out from university is not uncorrelated with successfully completing academic studies as only one of the two events can take place first. As the outcomes of interest were correlated, even imperfectly, it remained appropriate to implement the competing risks framework.

5.4.1 Cause-specific hazard probabilities

With standard survival analysis hazard functions, cumulative hazard functions and survivor functions are usually calculated (Cleves et al., 2010). These are appropriate in situations where only two potential outcomes are possible. Where competing events arise, it is appropriate to calculate the cause-specific hazard function and the cumulative incidence function (CIF) instead of the survivor function (Cleves et al., 2010). The cause-specific hazard is the risk of failure from a specified cause given that no failure from any cause has occurred yet. The cause-specific hazard function is given by

$$h_i(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t, failure\ from\ cause\ i\ | T \geq t)}{\Delta t} \tag{5.3}$$

where T is the time to first failure. The overall hazard rate given by $h(t) = \sum_i h_i(t)$ is the total risk of any event occurring, regardless of the independence of causes (Cleves et al., 2010). Once failure has occurred the failure is from cause *i* with probability $\frac{h_i(t)}{h(t)}$.

5.4.2 Cumulative incidence functions

Within the competing risks framework, it is appropriate to calculate a CIF instead of the standard survivor function. The CIF is given by

$$CIF_i(t) = P(T \leq t \text{ and failure from cause } i) \tag{5.4}$$

Specifically, the CIF at time *t* for cause *i* is the probability of failing from cause *i* before time *t*. The CIF is thus a modified version of the failure function, which is one minus the survivor function. The relationship between $CIF_i(t)$ and the cause-specific hazards is given more formally as

$$CIF_i(t) = \int_0^t h_i(x) S(x) dx \tag{5.5}$$

$$= \int_0^t h_i(x) \exp\left\{- \sum_{j=1}^k H_j(x)\right\} dx \tag{5.4}$$

$$= \int_0^t h_i(x) \exp\left\{- \sum_{j=1}^k \int_0^x h_j(u) du\right\} dx \tag{5.6}$$

$S(x)$ denotes the overall survivor function, interpreted as the probability of being failure-free from any cause up to time x. Equation (5.4) identifies $H_j(x)$ as the cause-specific cumulative hazard function for cause *j*, which is the integral from 0 to x of the cause-specific hazard for cause *j*. The overall cumulative hazard or the accumulated risk of failure from any cause is given by $H_x = \sum_j H_j(x)$. The relationship between the overall survival and overall cumulative hazard, given by $S(x) = exp\{-H(x)\}$, is maintained in the competing risks environment (Cleves et al., 2010).

The Fine-Gray model of regression on the CIF is very popular as it allows for the calculation of a sub-distribution hazard. When implemented, a sub-distribution hazard ratio (SHR) is calculated for each covariate. Interpretation of SHRs are not the same as typical regression models. Instead, a SHR = 1 may be interpreted as there is no association between the covariate and corresponding CIF. A SHR greater that 1 implies that an increase of the covariate value is associated with an increased risk, and a SHR less than 1 implies that an increase in the covariate value is associated with a decreased risk. The further away the SHR is from 1, the larger the estimated associated risk on the CIF.

5.5 Data and Descriptive Statistics

The data for this part of the study came from the UCT IPD. Detailed data on undergraduate students who first enrolled at UCT between 2006 and 2008 was obtained for analysis. All applicants to the university are required to submit basic demographic information, secondary schooling details, degree choices, and any requests for financial aid or housing. Applicants are not required to disclose family background information or any other personal information. Thus information relating to applicants' parents, such as income, education and occupational background, are not routinely requested or available. The final sample included every student who enrolled for a three or four year degree and were first-time entering students, excluding students from the medical school. This is standard practice in South Africa when assessing students in the HE sector. The data provided detailed information for each student based on application, enrolment and progression, including GPA by year, field of study by faculty, degree programme and background information, including high school attended and Grade 12 examination scores. Unfortunately, if a student drops out of university, they are not trackable to any other institution or to the labour market. Data on personal characteristics, such as gender, residential status and financial aid status, are also known.

The covariates used in the empirical analysis were classified into three groups. The first group was related to demographic characteristics such as race and gender. The race variable was comprised of a set of four binary variables., namely African, Coloured, White and Indian. White was used as the reference group.[45] Gender was a binary variable with the reference group set to female. The second group of covariates was related to the educational and socio-economic background of the student and included Grade 12 examination score, home province and high school type. High school type was divided into schools based on their classification under the apartheid system. Schools were identified as Department of Education and Training schools (previously meant to educate African students), House of Representative schools (previously intended to educate Coloured students), House of Delegates schools (previously intended to educate Indian students), and private and Model C schools, which were combined under the umbrella of House of Assemblies and other schools (intended for the education of white students).[46] Dummy variables were assigned for each of these four types of schools, and House of Assemblies schools set as the omitted category. This categorisation was preferred to the identification of schools based on the current classification of quintiles as too few schools were identified in quintiles 2 and 3 in the data, leading to inferior results when classified using the quintile system. Due to the absence of information on household income, an affluence variable was constructed based on the type of school attended. Affluence was a binary variable

[45] The African/Black category is referred to as African in this chapter.

[46] Model C schools are formerly White schools established under apartheid that had the legal autonomy to decide who enrolled while operating on a semi-private basis. These schools admitted White scholars exclusively.

that equalled 1 if the student attended a private school and 0 for any other type of public school attended. This method is commonly applied in the health literature in South Africa. The third set of covariates related to the student's HE experience. This included faculty of first enrolment, degree type, first-year GPA, financial aid status, residence status, and number of courses failed in the first-year.

Following Ishitani & DesJardins (2002) who show that first year GPA is non-linearly related to dropout in the US, the student's first-year GPA was used to create a set of dummies to examine its effect on dropout and graduation These dummies coincided with the mapping of GPA to class of pass as it was easier to measure the relative risks of dropout and graduation with dummy variables than by using first-year GPA as a continuous variable. It was also hypothesised that GPA is non-linearly related to the probability of dropout, and therefore, a dummy variable specification allowed for better testing of this scenario (Ishitani & DesJardins, 2002). Students whose GPA was between 0% and 49% were used as the reference category. This also allows us to test Bean's (1980) assertion that students learn about themselves during their time of study. First year GPA is one such learning opportunity, as it allows students to better understand what they're capable of in the HE sector.

To indicate a student's home faculty, dummy variables were constructed for each of the faculties included in the study. Due to the nature of application and entry to the Health Sciences faculty, these students were excluded from the study. The reference category was the Commerce faculty. Residence status and financial aid status were also included in the analysis. Institutional data only captures these variables for the first year, thus the impact of financial aid over the course of a student's academic career could not be assessed. Similarly, residence status is available for the first year only. This limited the interpretation of the impact of staying in a university residence as the information for each year the student was actively enrolled was not available.

5.5.1 Descriptive statistics

Table 22 presents a basic summary of the statistics for the sample. The data consists of 6 788 observations of South African first-time entering students enrolled at the university between 2006 and 2008. The overall sample was 51% male. More than 70% of the cohort speak English as a home language, and significantly fewer African students (13%) have English as a first language. The racial breakdown reveals that White students made up 46% of the sample, followed by Africans (25%), Coloureds (19%) and Indians (9%). More than half the students registered for a three-year degree, and the majority of students registered in the Commerce and Humanities faculties. The Science and Engineering faculties make up a smaller proportion of the sample as they accept fewer students each year than the Commerce and Humanities faculties. Approximately 14% of students received financial aid, and 26% of the sample live in a university residence. These descriptive statistics compare favourably with the university as a whole.

Table 22 Descriptive statistics for the full sample

		Full Sample	Whole University
Gender	*(Male)*	0.51	0.50
Race	*African*	0.25	0.27
	Coloured	0.19	0.17
	Indian	0.09	0.08
	White	0.46	0.46
Matric score		81.5	81.3
3-year degree		0.58	0.60
Faculty	*Commerce*	0.32	0.31
	Humanities	0.31	0.32
	Engineering	0.19	0.18
	Science	0.16	0.15
English Home Language		0.71	0.69
Financial Aid		0.14	0.15
Residence		0.26	0.29
2006		0.35	0.32
2007		0.32	0.32
2008		0.34	0.36

N = 6788

Data source: UCT IPD.
Author's own calculations

Table 22 presents the descriptive statistics for the pooled first-time entering students cohorts from 2006 to 2008. Data: UCT IPD

Table 23 presents the data for the main career paths that students follow. About 20% of students complete a 3-year degree on time, and slightly fewer students complete it on time for the 4-year degree. This compares well to other countries (see Arias & Dehon, 2011). The third most common path for students to follow after first year is for students to take an extra year to complete a 3-year degree. This is typically done by students who stay enrolled for the same degree and/or those who were initially registered for a 4-year degree but switch to a 3-year option with many common courses. About half of all students who start a 3-year degree are academically excluded and do not graduate with a degree from this university. The next pathway is that of voluntary and involuntary dropout.

Table 23 Main career paths of students

	Main career paths	
1.	Degree on time (3 year degree)	19.65%
2.	Degree on time (4 year degree)	16.70%
3.	3-year degree + 1	14.53%
4.	All drop-out year 2	8.33%
5.	4-year degree + 1	7.50%
6.	All drop-out year 1	6.95%
7.	3-year degree + 2	6.30%
8.	All drop-out year 3	3.50%
9.	4-year degree + 2	2.80%
10.	Other	13.82%

Table 23 presents the most common career paths observed in the data. All dropout captures both voluntary and involuntary dropout.
Data: UCT IPD

Table 24 presents the possible academic outcomes by race. Just over 60% of all African students graduate with a degree, implying that slightly less than 40% drop out before completion. In contrast, more than 80% of white students graduate with a degree, and less than 15% exit the system before completion. The gap between the two groups is large. It is interesting to note that while African students are the majority of the academically excluded students, white students are the majority who drop out on a voluntary basis. For those students who voluntarily exit, most do so during the second year of study. However, most students who exit on an involuntary basis or who are academically excluded due to poor performance, do so at the end of the first year of study. The timing of the types of dropout is noted to happen at different times. Approximately 80% of students who are academically excluded exit the system by the end of the third year of study. The timing of exit is especially important when considering the cost of study and the high debt burden faced by students in HE. When analysed by race, the pattern of exclusion does not resemble that of the average institution.

Table 24 Dropout and completion rates by race

	African	Coloured	Indian	White
Voluntary drop-out	6.5%	8.5%	9.7%	10.1%
Involuntary drop-out	30.8%	17.4%	15.2%	5.2%
Graduation	62.6%	73.8%	74.5%	84.5%

Table 24 shows dropout and completion rates by race.

Data: UCT IPD

5.5.2 Hazard probabilities

The previous section provided useful information about the sample. In order to understand *when* exit or graduation might occur, the survival analysis methodology was applied to the data set. The first step was to determine when students are likely to experience one of the three events of interest. The sample hazard function was then calculated. For each year the set of students who are at risk was defined. Once this was established, the proportion of students who leave university or graduate given that they have survived all the previous years was calculated. In determining when students are likely to experience the events of interest, the first step was to determine who is at risk. This information is presented in Table 25.

Table 25 shows the discrete-time hazard for the three outcomes under consideration, defined as the conditional probability that the event occurs at time *t*, given that the event has not yet occurred. Table 25 shows that 6 788 students are at risk during the first year because they are enrolled at the university. Of these students, 472 exit from university, either voluntarily or involuntarily. As students cannot graduate with a degree in less than three years, the hazard rate for graduation in years 1 and 2 is zero. In the first year, the hazard rate for voluntary dropout, $h_i(t)$, is 1.1, calculated as 75/6788. In year 2 there are 6 316 students at risk as they did not experience any of the outcomes in year 1. At this point in time, the risk of voluntary exit increases while the risk of academic exclusion falls from a first-year high. The trend in academic exclusion continues falling from first year through to fourth year, but rise quite significantly in year 5 and then hovers around the 3% rate. Voluntary exit displays a different trend in that it reaches a peak in year 2, and then gradually declines; though it rises again to a peak in year 6 and remains fairly high in year 7. The hazard probability for graduation is 27% in year 3, approximately doubling in year 4 to 54.48%, and then steadily increases through to year 7. The difference between the number of outcomes in year 7 (123) and the number of individuals in the risk set (131) yields the number of censored observations in the data.

Table 25 Discrete-time hazard probabilities for the full sample

		Frequencies			Hazard Probabilities		
Year	Population	Voluntary Drop-out	Involuntary Dropout	Graduation	Voluntary Drop-out	Involuntary Dropout	Graduation
1	6788	75	397	0	1.10	5.85	0.00
2	6316	326	216	0	4.80	3.42	0.00
3	5774	49	194	1601	0.85	3.36	27.73
4	3930	71	107	2141	1.81	2.72	54.48
5	1611	42	89	949	2.61	5.52	58.91
6	531	40	16	351	7.53	3.01	66.10
7	131	9	4	102	6.87	3.05	82.26

Table 25 shows the discrete-time hazard probabilities for the full sample of N = 6 788.
Data: UCT IPD

Figure 15 shows the hazard function of graduation by race. The graph clearly reveals the differences in the graduation outcomes by race. It is encouraging that the patterns are consistent for all groups over time. For all students, the probability of graduation peaks at approximately 5.5 years, indicating that on average all students take at least one year longer to complete their degrees. Figure 15 does not indicate whether students are registered for 3- or 4-year degrees, but indicates the probability of graduating peaks after five years of enrolment for the sample. White students experience much higher hazards of graduating at each time period compared to all other race groups, especially African students. The hazards are fairly similar for Coloured and Indian students, but this holds only for the graduation outcome.

Figure 15 Graduation hazard function by race

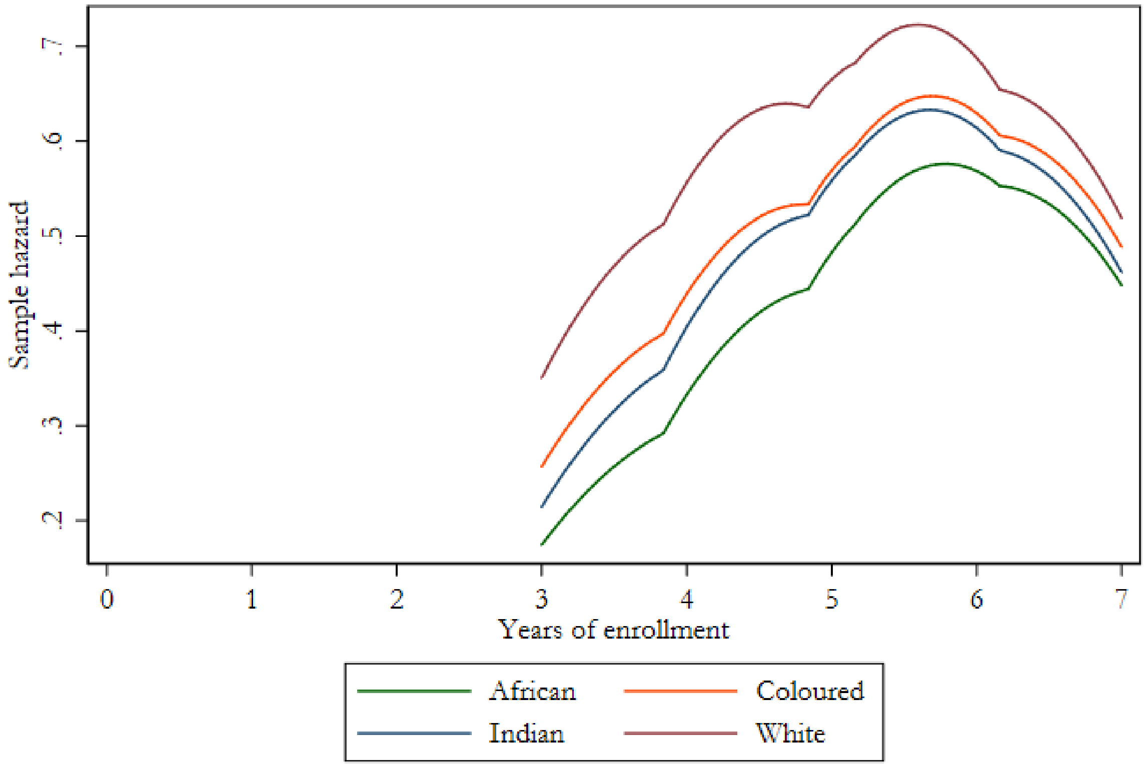

Figure 15 shows the graduation hazard function by race.
Data: UCT IPD

Figure 16 shows that females have a higher probability of obtaining a degree everywhere along the graduation hazard distribution. It appears that the gap narrows after five years of enrolment. A similar pattern emerges in Figure 17 when combing the race and gender information.

Figure 16 Graduation hazard function by gender

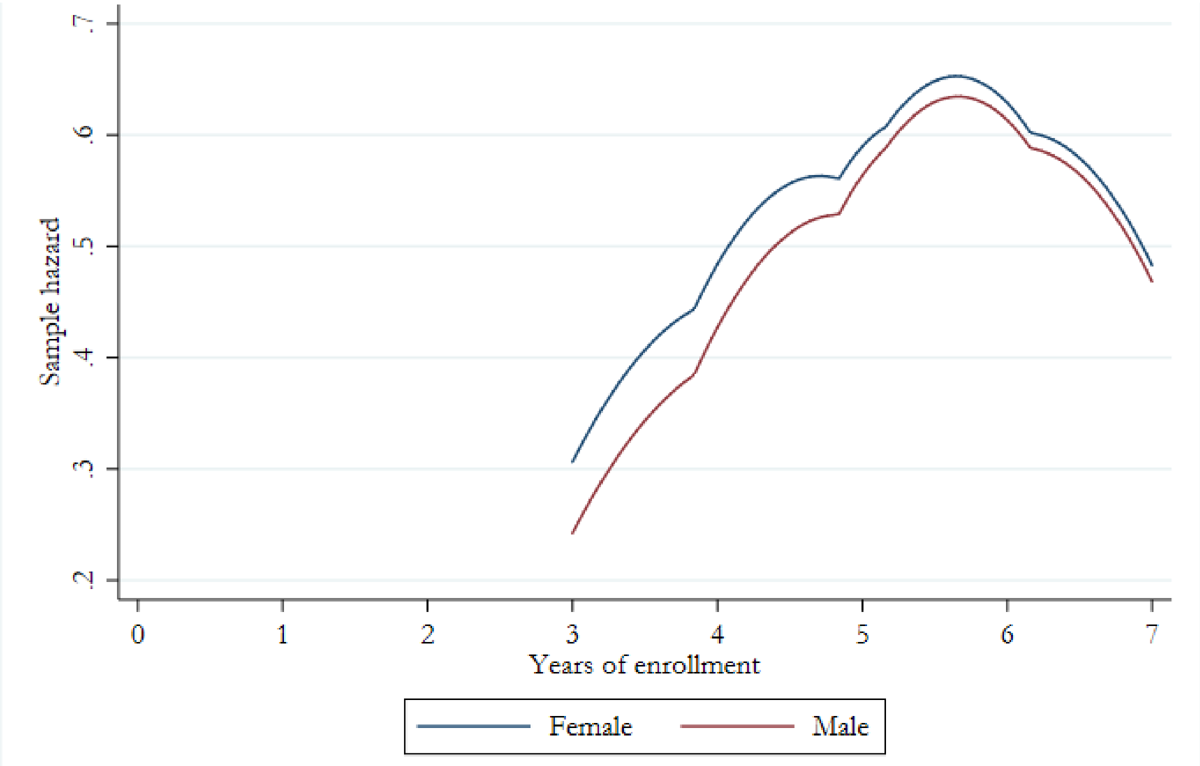

Figure 16 shows the graduation hazard by gender.
Data: UCT IPD.

Following on from the race and gender breakdowns presented in Chapter 4, race and gender breakdowns are presented across the three academic outcomes in this chapter too. The first academic outcome to show the race breakdown by gender is that of graduation. Figure 17 shows the graduation hazard function across the eight race and gender groups in this analysis. While the patterns displayed are expected, and similar in nature to those explored in Chapter 4, the extent of differences in the graduation outcome between White females and African males is surprising. The average time to graduation for White females is 3.7 years while the average time to graduation for African males is 5 years. Similarly, the median time to graduation for White females and African males is 4 years and 5 years respectively.

Figure 17 Graduation hazard function by race and gender

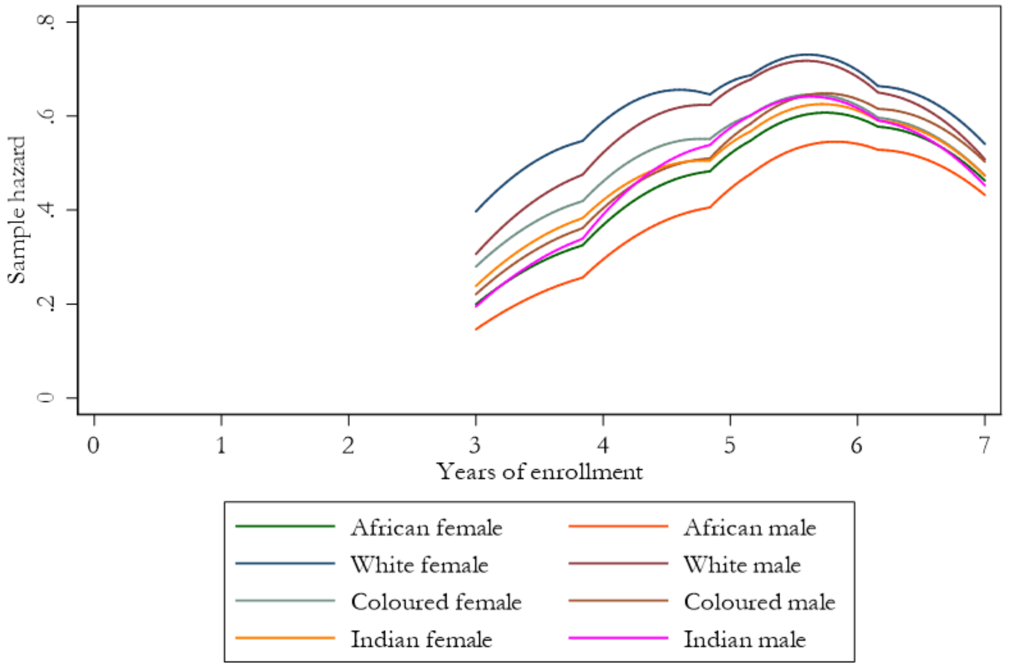

Figure 17 shows the graduation hazard by race and gender.
Data: UCT IPD.

Figure 18 shows that the hazard function for involuntary dropout reveals the extensive differences in outcomes between races. African students face significantly higher probabilities of dropout everywhere along the hazard function. White students face the lowest academic exclusion probabilities, but the peak of white exclusion is 5.5 years compared to approximately 3 years for African students. Therefore, African students are likely to be excluded earlier than white students, indicating that white students experience longer retention in HE. Overall, the trend for white students is very flat and is very different to that of other race groups. While Coloured students have marginally higher probabilities of graduating (see Figure 14), they also have higher probabilities of academic exclusion compared to Indian and African students.

Figure 18 Involuntary dropout hazard function by race

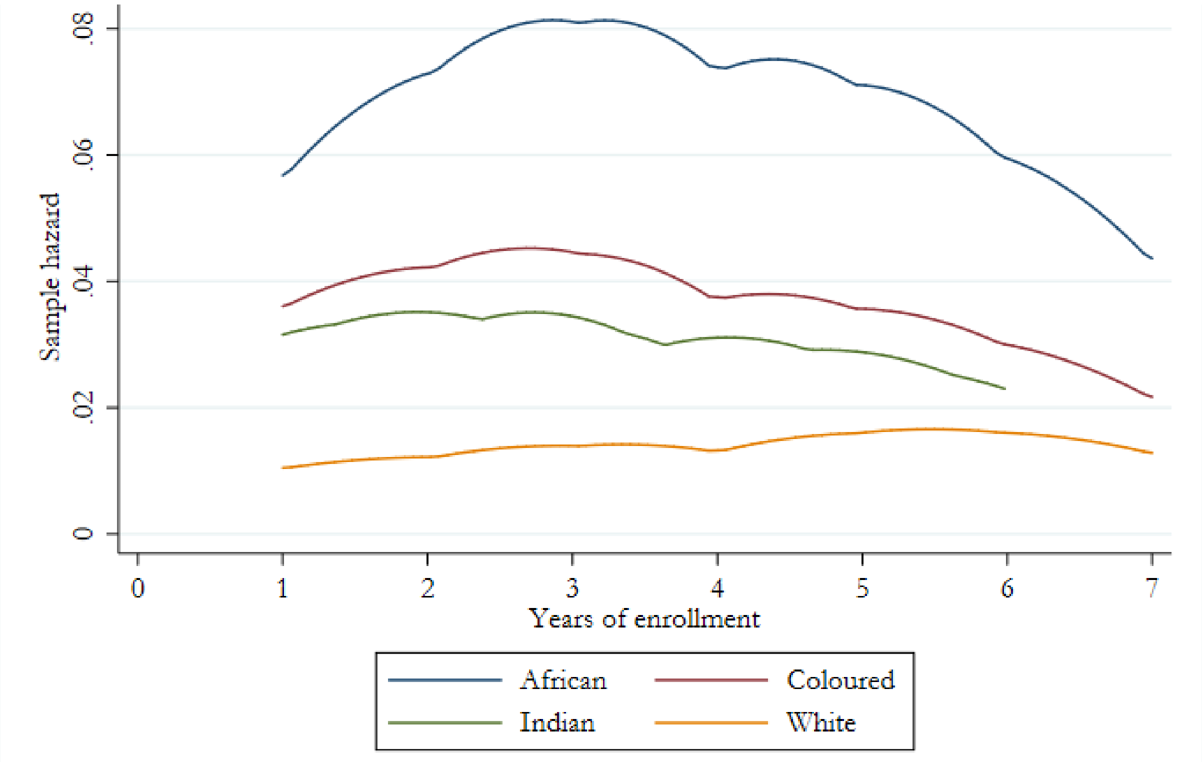

Figure 18 shows the hazard function for involuntary dropout by race.
Data: UCT IPD

The profile of academic exclusion in Figure 19 shows that the academic experiences of males and females are different. Males are significantly more likely to be academically excluded than females. The peak of the female distribution occurs earlier in the academic career compared to that of males. The peak for academic exclusion for males is reached in year 4 of enrolment, indicating that male students linger longer in the system and then miss progression targets after some time of enrolment. Compared to other countries where these types of studies have been done, academic exclusion rates are much lower at the institution in this analysis.

Figure 19 Involuntary dropout hazard function by gender

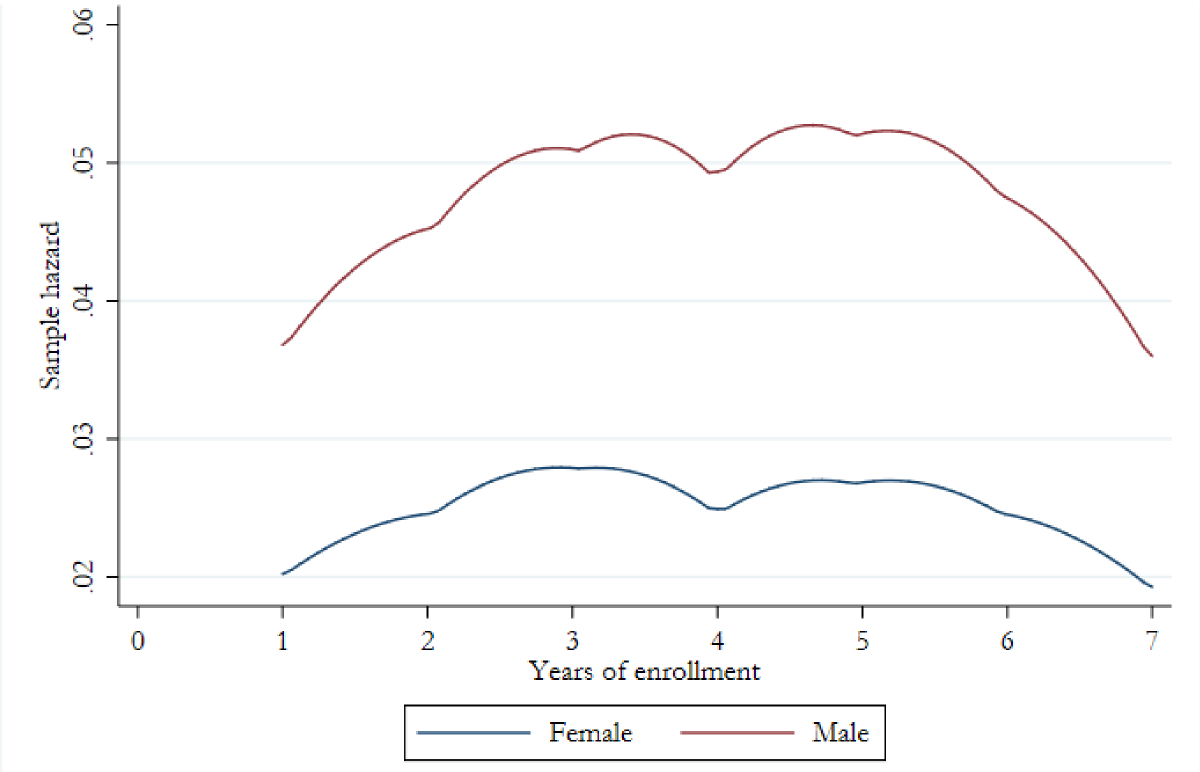

Figure 19 shows the hazard function for involuntary dropout by gender.
Data: UCT IPD.

Figures 18 and 19 show the racial and gender breakdowns separately, providing valuable insight into the extent of involuntary dropout for the two groups. Further breaking down the analysis into the race groups by gender, Figure 20 sheds additional light on the extent of differences between the gender and race groups. The involuntary dropout hazard for White females is the lowest of all groups, and consistently close to zero over the 7-year period for which we observe students. In contrast, African males have the highest involuntary dropout hazard at every time point observed in the data., Like the patterns observed in Chapter 4, after African males Coloured males have higher hazards of involuntary dropout in years 2 and 3. Thereafter, hazards decrease for all groups, but no consistent pattern emerges for any single group.

Figure 20 Involuntary dropout hazard function by race and gender

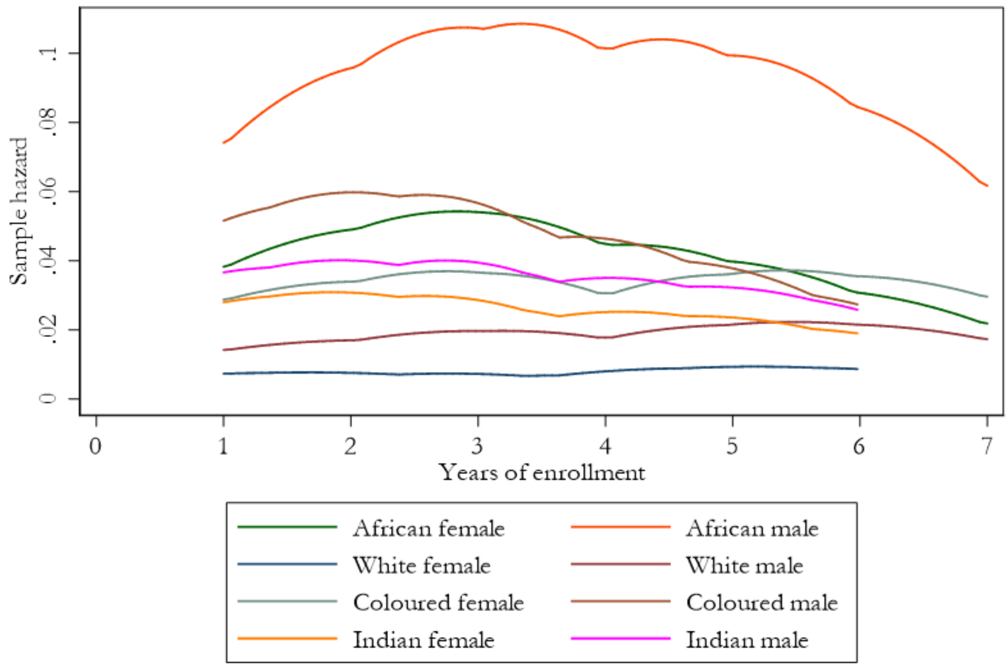

Figure 20 shows the involuntary dropout hazard function by race and gender.
Data: UCT IPD

To add completeness to the discussion on hazard rates for each of the three academic outcomes, Figure 21 presents the voluntary dropout hazard function by race and gender. The figure shows that the hazard rate with respect to voluntary dropout is different in nature compared to involuntary dropout. Differences in voluntary dropout are also observed between earlier and later years of enrolment. While no individual group dominates across the entire period over which enrolment is observed, White females show higher hazards of voluntary dropout in earlier years compared to African males, a reversal of the trend, albeit over a shorter period, when compared to involuntary dropout displayed in Figure 20.

Figure 21 Voluntary dropout hazard function by race and gender

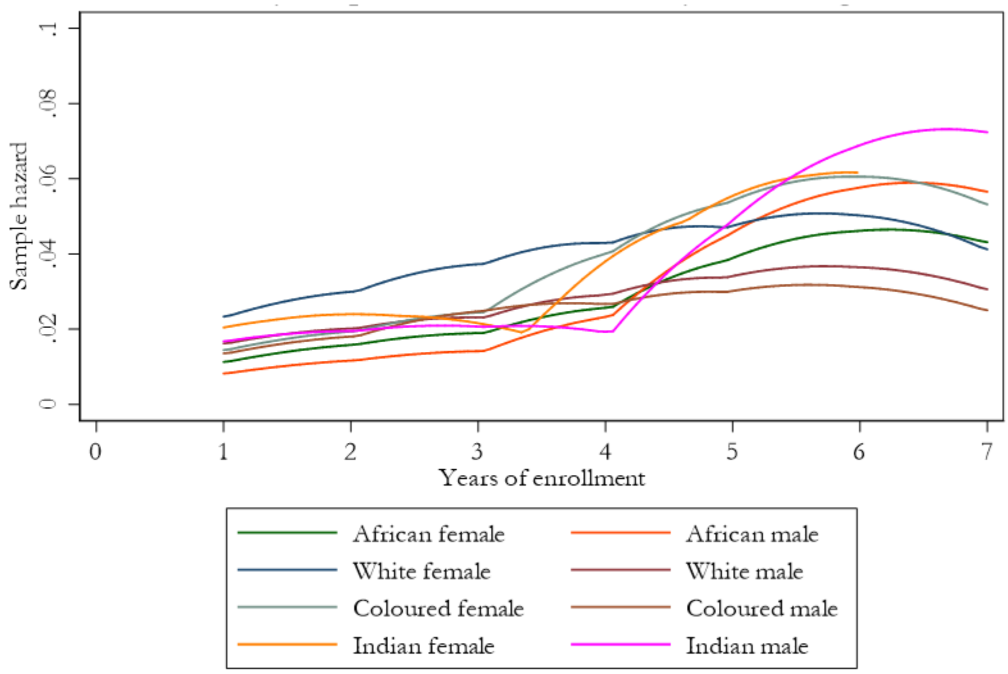

Figure 21 shows the voluntary dropout hazard function by race and gender.
Data: UCT IPD

The final hazard function presented in this descriptive section is the impact of academic standing of the first year of study on graduation hazards. Figure 22 shows the hazard functions based on the three main academic standings of good academic standing, concession to continue and academically excluded but readmitted. Students who meet the full progression requirements for a particular year and programme are coded as good academic standing. Students who marginally fail to meet the full progression requirements but have passed the most important courses for their degree programmes to date are usually coded as concession to continue. These students are treated in the same way as those in good academic standing. These academic outcomes indicate that students may proceed to year 2 of studies via different paths. Students who do not meet the minimum progression requirements for their degree programmes are coded as academically excluded. The hazard function patterns show an interesting occurrence: Students who proceeded to year 2 in good academic standing do not have the highest hazard rates for degree completion in the sample, but students who proceeded to year

2 with a concession to continue have the highest hazard rates for completion. Students who are academically excluded at the end of first year but are readmitted via appeal take longer to graduate compared to other students. One reason for this is that readmitted students are often advised to change degree programmes, implying they need to start with a completely new set of courses and a zero base of completed courses.

Figure 22 Graduation hazard function by academic standing in year 1

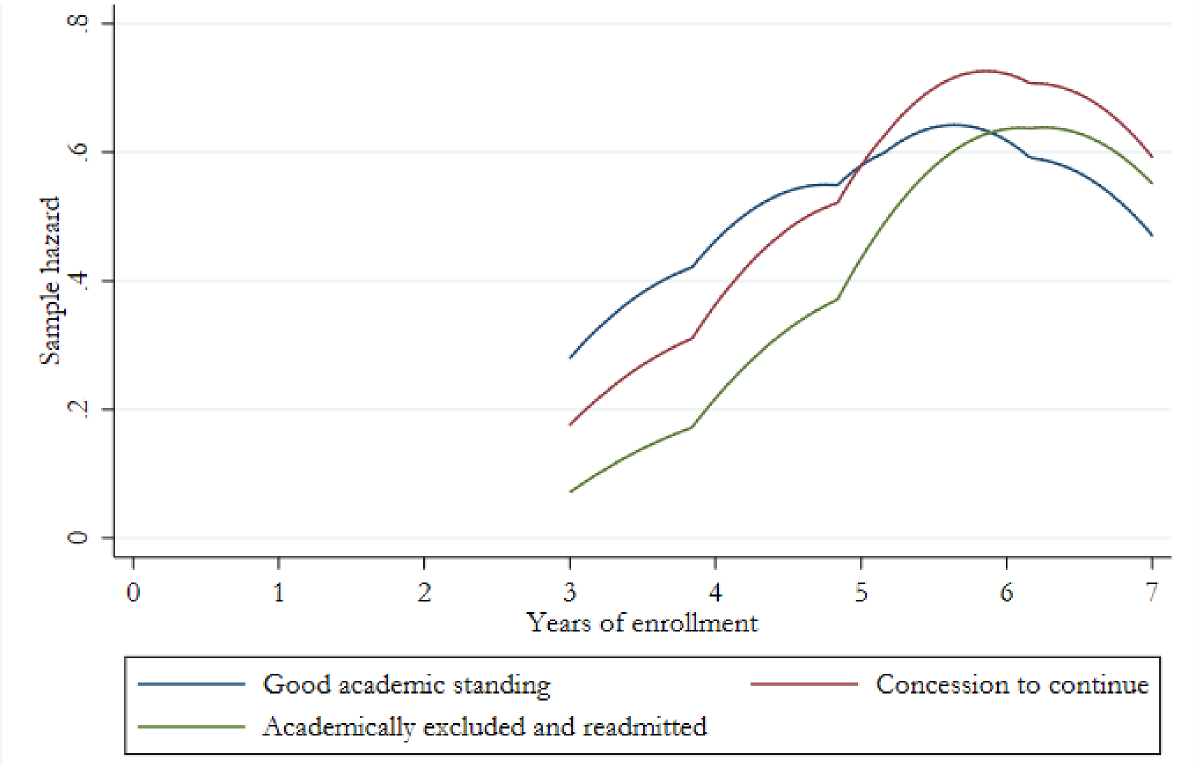

Figure 22 shows the graduation hazard function by academic standing in year 1.
Data: UCT IPD

The probability of graduating in the third and fourth year of enrolment is highest for students who have good academic standing at the end of year 1. However, students with concessions to continue experience a large jump in the hazard rate of graduation between years 5 and 6. An evaluation of the confidence intervals also suggested significant differences of years of enrolment to graduation between the different outcomes in year 1. Students in good academic standing take on average 4.2 years to complete, compared to 4.7 and 5.4 years for students with concessions to continue and readmitted students, respectively.

5.6 Results

5.6.1 Baseline profile of risk over time

To evaluate the effect of time as the sole predictor of academic outcomes, time dummies were introduced to the analysis. The time dummies were defined as 1 when the outcome for individual *i* was observed in year *t* and 0 otherwise, yielding six time dummies defined for all periods. Running the model with time dummies only represented the baseline profile of risk over time, referred to as the baseline model. In this baseline model, the dummy variables for years 6 and 7 were set as the omitted category.

Table 26 Coefficients and SHR for the baseline

	Years of registration						
	1	2	3	4	5	6	7
Sub-hazard ratios (SHR)							
Voluntary dropout	2.160	17.957	0.225	0.239	0.378	1.057	0.789
Involuntary dropout	23.982	3.570	0.596	0.206	0.471	0.232	0.202
Graduation	-----	-----	3.560	1.689	0.857	0.732	0.696
Coefficients							
Voluntary dropout	0.770	2.888	-1.490	-1.431	-0.974	0.055	-0.237
Involuntary dropout	3.177	1.272	-0.518	-1.582	-0.754	-1.459	-1.602
Graduation	-----	-----	1.270	0.524	-0.155	-0.312	-0.363

Note: All coefficients are significant at a 99% confidence interval, except for years 6 and 7 voluntary dropout, which are not statistically significant.
Source: Author's own calculations

Table 26 illustrates the basic shape of the hazard functions. It is important to note that the coefficients for graduation in years 1 and 2 are not displayed as no student may graduate ahead of time. Econometrically, these estimated coefficients were unstable due to large standard errors (Hosmer, 2000). For the sample as a whole the probability of voluntary dropout peaks in year 2, and thereafter it declines dramatically. For involuntary dropout, the shape is somewhat more undefined. It peaks in year 2, falls through year 4 and then increases again. Thus academic or involuntary exclusion can be seen as more volatile over time relative to voluntary dropout. It may be that the progression criteria on which involuntary exit is based is poorly formulated across the institution. Finally, graduation peaks in year 3 and gradually declines in subsequent years.

5.6.2 Full model and assumptions

The main results for the analysis are presented in Table 27. Each set of columns in Tables 27-34 display exponentiated coefficients, or estimated β, which is the SHR for the covariate concerned as well as the standard errors. The SHR may be interpreted as the instantaneous probability of failure from cause *i* at time *t* given no failure before time *t* or failure from another cause before time *t*. Put differently, this hazard may be thought of as that which generates failure events of cause *i* but does not remove subjects from the sample when competing events occur. Column 2 displays the SHR results for voluntary exit when involuntary academic exclusion and graduation are treated as censored in the data. SHRs may be interpreted as a rise in the incidence of the outcome variable for an increase in the independent variable when the SHR value is greater than one. If the SHR value is between 0 and 1, the interpretation becomes a lower incidence of the outcome variable for a given value of the independent variable.

Just like standard survival analysis, any competing risks survival analysis must be checked to ensure that the proportionality or proportional subhazards assumption holds. The proportionality assumption states that the effect of an explanatory variable is proportionate over time; the shape of the hazard function does not change over time. Put differently, the relative contribution of a given explanatory variable is the same in the first year and subsequent years. However, studies show that some variables do exhibit characteristics that are not time-dependent (Arias & Dehon, 2011). To test for the proportionality assumption, each covariate was interacted with time according to the time assumption preferred. Where this time-adjusted variable was statistically significant, it implied that the proportionality assumption was violated. To overcome the violation of the proportionality assumption, covariates interacted with time were included in the specification of the main model.

All results presented in Table 27 were adjusted for the violation of the proportionality assumption. After running initial estimations, some independent variables were found to exhibit non-proportional hazards over time. These included the two race variables of African and Coloured and the changing (switching) of degrees at the end of year 1. The remainder of the independent variables did not violate the proportionality assumption and their effects were proportional over time.

The matric score hazard ratio was consistently close to and above 1, and not statistically significant across all race groups. For the sample as a whole, the higher the matric score the lower the likelihood of either voluntary or involuntary dropout. However, there is no significant difference in terms of matric score for those who graduate and those who do not. As expected, individuals with higher entry scores are slower to voluntarily exit the system. This is similar to findings by Murray (2014), who finds a comparable pattern of exit. In this context, South African HE is similar to that of developed economies where voluntary dropout is not driven solely by prior academic achievement.

Arising from the health literature in South Africa, an affluence variable was introduced into the analysis. This variable was represented as 1 for individuals who attended private or quintile 5 schools, and 0 otherwise. As the school quintile system is largely based on the socio-economic wealth of the area immediately surrounding the school, it was expected that it captured some of the socio-economic effects of households located in these areas. Across all specifications the affluence variable was statistically insignificant. This could be due to collinearity with some other variable captured by the old schooling classification, or in the omitted variables, such as family background, including parental educational background.

Table 27 SHR estimates by academic outcome for the full sample

	Dependent Variable					
	Voluntary Dropout		Involuntary Dropout		Graduation	
	SHR	Std. Error	SHR	Std. Error	SHR	Std. Error
Male	0.846**	0.067	1.412***	0.089	0.830***	0.020
Indian	0.862	0.122	1.827***	0.254	0.783***	0.033
Affluence	0.823	0.178	0.933	0.102	1.082	0.071
Matric Score	0.985**	0.005	0.990***	0.003	1.004	0.002
African*Yr1	0.300**	0.167	8.011***	1.565	------------	
African*Yr2	0.492***	0.099	4.789***	0.825	------------	
African*Yr3	0.522**	0.154	2.372***	0.358	1.101	0.114
African*Yr4	0.496**	0.147	0.991	0.166	1.082	0.061
African*Yr5	0.139***	0.081	0.791	0.154	0.868**	0.497
African*Yr6	1.247	0.359	0.383***	0.130	0.683***	0.046
African*Yr7	1.595	0.682	0.294**	0.168	0.720***	0.077
Coloured*Yr1	0.473	0.215	16.113***	3.059	------------	
Coloured*Yr2	0.787	0.138	4.694	0.917	------------	
Coloured*Yr3	0.296***	0.105	1.955***	0.321	1.629***	0.129
Coloured*Yr4	0.418***	0.109	0.454***	0.119	1.035	0.039
Coloured*Yr5	0.860	0.227	0.600**	0.154	0.721***	0.036
Coloured*Yr6	0.987	0.372	0.404*	0.220	0.505***	0.035
Coloured*Yr7	0.484	0.417	0.576	0.463	0.538***	0.065
School: DET	0.399**	0.119	1.488***	0.171	0.843**	0.063
School: HoD	1.126	0.255	1.502**	0.285	0.873*	0.067
School: HoR	1.111	0.200	1.046	0.124	1.023	0.053
3 year degree	1.041	0.111	0.782***	0.073	1.406**	0.043
Financial Aid	0.814	0.123	1.133*	0.080	0.966	0.039
Humanities	1.430**	0.164	0.868	0.097	0.980	0.034
Science	0.910	0.142	1.366***	0.147	0.938	0.040
Engineering	1.034	0.131	1.183**	0.098	0.963	0.030
Extended programme	0.776*	0.108	1.328***	0.099	0.773***	0.032
Residence	1.192*	0.124	1.068	0.079	0.946*	0.031
English HL	0.876	0.108	0.733	0.079	1.189***	0.048
Failed 1	1.184	0.122	0.786	0.077	0.986	0.031
Western Cape	0.934	0.093	0.983	0.080	1.032	0.031
GPA50-59	0.746***	0.078	0.377***	0.030	2.023***	0.097
GPA60-69	0.508***	0.061	0.150***	0.018	2.737***	0.131
GPA70-74	0.301***	0.070	0.046***	0.016	3.187***	0.170
GPA75+	0.309***	0.084	0.023	0.013	3.158***	0.180
Switch2*yr2	13.720***	1.637	0.531***	0.072	------------	
Switch2*yr3	0.516**	0.164	1.489***	0.184	1.194**	0.104
Switch2*yr4	0.602*	0.158	0.917	0.147	0.951	0.039
Switch2*yr5	0.557*	0.190	0.866	0.153	0.757***	0.033
Switch2*yr6	1.345	0.450	0.525	0.256	0.893	0.090
Switch2*yr7	1.025	0.676	0.429	0.356	0.596***	0.065
Time dummies	Yes		Yes		Yes	

Note: ***p<0.01, ** p<0.05, *p<0.10.
Data source: UCT IPD 2006, 2007 and 2008, author's own calculations

The results support previous research for the general educational outcomes by gender (Branson and Leibbrandt, 2013). Across all estimations, gender was found to be significant. Males are significantly less likely to voluntarily exit or graduate from HE and significantly more likely to be academically excluded. This result was further supported in the breakdown by race and gender.

African, Coloured and Indian students are significantly more likely to be excluded earlier during their studies. This possibly reflects the struggle to adjust to the level of academic studies because of past educational exposure and experience.

Students registered in the Humanities faculty voluntarily drop out at faster rates (or sooner) than students from any other faculty. This could be driven by many students who enter the Humanities faculty doing their second rather than first choice of degree, leading to students being undecided about a future within the Humanities faculty. Similarly, not all out-of-faculty disciplines are available to students within the Humanities faculty, leading to students being less decisive about a degree in the Humanities. Science and Engineering students are significantly more likely to be academically excluded than Commerce students. The STEM disciplines within the Science and Engineering faculties are known to be technically challenging, leading to higher exclusion rates on academic grounds. There are two perspectives indicated here. Firstly, the Commerce faculty is more likely to attract students with higher matric grades due to its popularity among entering students. Secondly, there is great concern in South Africa about the quality of high school physics and mathematics as the lack of exposure to proper teaching techniques and subject matter greatly influences students' chances of succeeding in these fields post-secondary school. The interesting finding here is that this effect does not extend across all three outcome-states. For example, students in the Humanities are more likely to voluntarily exit than Commerce students, but there are no significant differences for the other potential academic outcomes.

A surprising outcome of the analysis is that of financial aid. Financial aid was positively correlated with involuntary dropout, but the relationship between financial aid and voluntary dropout was statistically insignificant. The relationship was also statistically insignificant between financial aid and graduation, as is evident in Table 27. Receiving financial aid increases the hazard ratio of involuntary dropout, albeit at the 90% confidence level. Therefore, financial aid students are more likely to experience continued enrolments without getting a degree. Students who receive financial aid enter with lower matric scores and progress more slowly through the system, failing more courses along the way.

In contrast, students who live in university residence voluntarily drop out at much faster rates than students who either live on their own or with their families. Students in residence are also less likely to graduate, indicating that the lack of on-site family support or being away from one's family or natural support structures significantly impairs students' ability to graduate. The higher the student's GPA in the first year, the less likely they are to exit the system, and if they do, they do so more slowly. This effect was not found to be linear across the GPA spectrum, as the size of the coefficients for the GPA75+ variable was closer to the one for voluntary exit and graduation. Some courses or programmes require students to meet certain progression requirements, and part of this effect is possibly captured in the GPA effect.

Home language has been shown to be an important indicator of success in HE (see Smith, 2012). Students whose home language is English are more likely graduate relative to any other home language group. This is an expected result as the language of instruction matching students' home language makes the learning of new content much easier. An interesting result is that of students on extended programmes. These programmes have been implemented across a few faculties as a means of improving support for and enhancing outcomes for students from previously disadvantaged backgrounds. Extended programme students are less likely to voluntarily dropout, significantly more likely to be academically excluded (not meeting academic progression rules), and significantly less likely to graduate over time. The results also indicate a consistent schooling effect present. Students who attended previously Department of Education and Training schools face significantly worse academic outcomes, but the effects are larger in magnitude compared to extended programme students. These students are significantly less likely to drop out voluntarily, significantly more likely to be academically excluded, and significantly less likely to graduate over time. Thus the effect of poor schooling persists throughout their HE career for students who attended this type of high school. Similar to the effect for students on the extended programme, students from Department of Education and Training schools are most likely to experience significantly worse outcomes.

Gender and racial differences were evident in the data, and these results are discussed in the following subsections.

5.6.3 Racial breakdown

Table 28 illustrates that there are very different processes underlying voluntary dropout by race. Very few variables were significant across separate specifications by race, indicating that the process of voluntary dropout varies by race. The only variable that was significant across all race groups was switching degrees at the start of year 2. In all instances students who switch were significantly more likely to exit the system voluntarily. White females were significantly more likely to exit HE voluntarily than White males.

An important observation from the results for the voluntary dropout specification is that the standard set of covariates do not tend to explain the phenomenon of voluntary dropout very well, indicating that further investigation is required to fully identify those factors associated with voluntary dropout, and more specifically, those factors that may have a causal relationship with respect to voluntary dropout or exit from HE. The analysis in this chapter is unable to make claims on causality.

Table 28 SHR estimates by race for voluntary dropout

	Dependent Variable: Voluntary dropout							
	African		White		Coloured		Indian	
	SHR	Std. Error	SHR	Std. Error	SHR	Std. Error	SHR	Std. Error
Male	1.201	0.249	0.794 **	0.082	0.918	0.181	0.791	0.213
Affluence	1.010	0.295	1.330	0.739	0.902	0.394	0.274 ***	0.133
Matric Score	0.991	0.013	0.983 **	0.007	0.973 **	0.011	1.004	0.016
School: DET	0.383 ***	0.132	3.441	2.669	0.871	0.701	1.801 ***	2.131
School: HoD	3.370 ***	1.230	5.001 ***	5.881	0.615	0.477	1.352	0.520
School: HoR	0.363	0.278	1.081 ***	5.881	1.556 *	0.353	1.524	0.936
3 year degree	1.606	0.484	0.903	0.120	0.826	0.241	0.841	0.303
Financial Aid	0.829	0.211	0.645	0.214	0.598 *	0.173	1.386	0.448
Humanities	1.220	0.351	1.228	0.177	1.922 **	0.638	1.654	0.674
Science	0.494 *	0.183	1.092	0.202	0.741	0.362	1.382	0.790
Engineering	0.947	0.326	1.073	0.172	1.119	0.390	1.172	0.415
Extended programme	0.762	0.186	1.316	0.397	0.946	0.242	0.863	0.390
Residence	1.699 **	0.408	1.168	0.150	1.322	0.496	1.109	0.421
English HL	0.611	0.215	0.886	0.160	0.934	0.263	0.979	0.568
Failed 1	0.973	0.238	1.274 *	0.164	0.859	0.264	1.512	0.642
Western Cape	0.833	0.243	1.095	0.129	0.648	0.239	0.709	0.277
GPA50-59	1.171	0.287	0.693 ***	0.092	0.778	0.208	0.439 *	0.192
GPA60-69	1.090	0.292	0.430 ***	0.064	0.589	0.186	0.364 **	0.166
GPA70-74	1.070 ***	3.100	0.311 ***	0.084	0.231	0.241	0.3901	0.234
GPA75+	6.880 ***	2.120	0.341 ***	0.105	0.759	0.499	5.571 ***	2.301
Switch2*yr2	11.867 ***	2.854	19.603 ***	2.738	12.878 ***	3.093	11.591 ***	4.166
Switch2*yr3	1.493	0.736	0.320 *	0.189	9.651 ***	2.051	0.891	0.745
Switch2*yr4	0.608	0.364	0.403 **	0.184	1.238	0.569	0.520	0.542
Switch2*yr5	0.260	0.264	0.342	0.242	0.677	0.392	1.097	0.665
Switch2*yr6	5.706 ***	2.017	1.877	1.233	0.850	0.877	7.861 ***	7.371
Switch2*yr7	5.825 *	6.215	1.861 ***	7.421	1.680	1.573	1.427	1.233
2007	1.138	0.282	1.228 *	0.145	0.590 **	0.143	0.986	0.359
2008	1.157	0.319	0.884	0.103	0.912	0.218	1.024	0.362
Time dummies	Yes		Yes		Yes		Yes	

Note: ***p<0.01, ** p<0.05, *p<0.10.
Data source: UCT IPD 2006, 2007 and 2008, author's own calculations

The racial breakdown for involuntary exit or academic exclusion presents interesting findings. Coloured males face the lowest incidence of involuntary exclusion compared to other race groups. The matric score effect is similar across all race groups, which is a surprising outcome. On the whole, the higher the matric score the less likely students are to be excluded. However, one would have expected to see larger differences by race groups. Previous research (Bokana & Tewari, 2014) shows that the matric score is significant in determining performance and outcomes. In this case, it appears that the effects are more muted across the racial distribution. White students in the Humanities faculty are significantly more likely to be academically excluded while the opposite is true for African students.

On the whole, African students face significantly worse outcomes in terms of academic exclusion compared to any other race group, shown by the results in Table 28. The results suggest that this process is not driven by performance in the matric exam but by some other mechanism not clear from the limited data at hand. This also confirms the information in the hazard functions presented earlier in the chapter.

Table 29 SHR estimates by race for involuntary dropout

	Dependent Variable: Involuntary dropout							
	African		White		Coloured		Indian	
	SHR	Std. Error	SHR	Std. Error	SHR	Std. Error	SHR	Std. Error
Male	1.521 ***	0.151	1.789 **	0.338	1.457 ***	0.204	1.626 **	0.366
Affluence	0.934	0.128	0.520	0.233	1.065	0.478	0.789	0.393
Matric Score	0.985 ***	0.005	0.980 *	0.011	0.980 ***	0.007	0.973 **	0.011
School: DET	1.660 ***	0.241	9.121 ***	5.371	0.991	0.745	0.681	0.719
School: HoD	1.258	0.495	1.610	1.180	0.662	0.328	1.305	0.341
School: HoR	1.769 *	0.520	2.076 **	0.679	0.881	0.146	1.134	0.546
3 year degree	0.891	0.137	0.583 **	0.130	0.598 **	0.128	1.062	0.411
Financial Aid	1.137	0.116	2.736 ***	0.873	1.054	0.199	0.992	0.417
Humanities	0.613 **	0.116	1.632 *	0.438	1.094	0.277	0.720	0.336
Science	1.761 ***	0.286	1.150	0.368	2.353 ***	0.617	1.209	0.564
Engineering	1.518 ***	0.203	1.162	0.253	2.060	0.409	1.362	0.379
Extended programme	1.610 ***	0.161	1.379	0.737	1.201	0.200	1.322	0.381
Residence	2.137 ***	0.217	1.411	0.326	1.256	0.363	1.114	0.309
English HL	0.672	0.176	0.750	0.199	0.886	0.176	0.741	0.356
Failed 1	0.913	0.119	0.524	0.142	0.890	0.217	0.343	0.177
Western Cape	0.969	0.124	1.446 *	0.320	0.748	0.204	0.621 *	0.174
GPA50-59	0.339 ***	0.036	0.179 ***	0.035	0.200 ***	0.034	0.1720 ***	0.048
GPA60-69	0.190 ***	0.029	0.062 ***	0.015	0.064 ***	0.018	0.083 ***	0.040
GPA70-74	0.087 ***	0.034	0.008 ***	0.008	0.024 ***	0.024	0.000	0.000
GPA75+	0.057 ***	0.036	3.371 ***	6.671	1.361 ***	3.211	0.054	0.052
Switch2*yr2	1.062	0.156	0.142 ***	0.085	0.790	0.208	0.604	0.264
Switch2*yr3	1.448 **	0.225	0.633	0.296	2.195 ***	0.521	1.618	0.761
Switch2*yr4	0.548 ***	0.118	0.126 ***	0.091	0.447 **	0.177	0.461	0.333
Switch2*yr5	0.511 ***	0.110	0.360	0.234	0.372	0.133	0.326 **	0.180
Switch2*yr6	0.073 ***	0.054	1.351 ***	6.641	0.693	0.449	0.000	0.000
Switch2*yr7	9.071 ***	3.861	4.121 ***	3.261	0.397	0.392	0.000	0.000
2007	1.320 **	0.146	0.813	0.152	1.031	0.169	1.082	0.284
2008	1.361 **	0.176	0.816	0.162	1.036	0.183	0.873	0.254
Time dummies	Yes		Yes		Yes		Yes	

Note: ***p<0.01, ** p<0.05, *p<0.10.
Data source: UCT 2006, 2007 and 2008, auhtor's own calculations

Focusing on graduation outcomes, similar patterns were observed. Many more variables were significant in the determination of graduation outcomes for African students compared to other race groups. African students in the Science and Engineering faculties were significantly less likely to graduate than African students in the Commerce faculty. This was also true for financial aid and residence status.

Table 30 SHR estimates by race for graduation

	Dependent Variable: Graduation							
	African		White		Coloured		Indian	
	SHR	Std. Error	SHR	Std. Error	SHR	Std. Error	SHR	Std. Error
Male	0.739 ***	0.042	0.849 ***	0.026	0.801 ***	0.046	0.808 ***	0.064
Affluence	1.020	0.083	0.929	0.133	1.050	0.185	1.495	0.429
Matric Score	1.003	0.003	1.000	0.002	1.009 ***	0.003	0.993	0.006
School: DET	0.866 *	0.074	0.553 **	0.168	1.103	0.433	0.921	0.149
School: HoD	1.178	0.160	0.743	0.646	1.035	0.183	0.887	0.107
School: HoR	1.006	0.179	1.059	0.232	0.915	0.061	1.106	0.165
3 year degree	1.292 ***	0.100	1.465 ***	0.058	1.608 ***	0.112	1.355 **	0.164
Financial Aid	0.888 *	0.059	0.931	0.092	1.080	0.079	0.868	0.123
Humanities	1.258 ***	0.110	0.957	0.041	0.906	0.071	0.850	0.119
Science	0.780 ***	0.074	1.073	0.058	0.855	0.084	0.777	0.144
Engineering	0.810 ***	0.064	1.016	0.039	0.747 ***	0.063	0.916	0.093
Extended programme	0.699 ***	0.048	0.852	0.098	0.731 ***	0.053	0.574 ***	0.089
Residence	0.689 ***	0.051	0.975	0.041	0.940	0.111	0.830	0.100
English HL	1.300 **	0.123	1.098 *	0.060	1.157 *	0.102	1.228	0.232
Failed 1	1.138 *	0.082	0.970	0.041	1.014	0.069	0.991	0.096
Western Cape	1.068	0.089	0.955	0.037	1.400 ***	0.133	1.249 **	0.123
GPA50-59	2.271 ***	0.210	2.278 ***	0.169	2.852 ***	0.280	3.325 ***	0.510
GPA60-69	3.006 ***	0.285	3.205 ***	0.230	3.762 ***	0.378	4.969 ***	0.849
GPA70-74	4.135 ***	0.483	3.565 ***	0.279	4.227 ***	0.533	6.079 ***	1.149
GPA75+	5.617 ***	0.772	3.366 ***	0.275	4.583 ***	0.626	6.471 ***	1.243
Switch2*yr2	----	----	----	----	----	----	----	----
Switch2*yr3	0.513 **	0.136	2.234 ***	0.133	1.524 **	0.274	1.755 *	0.572
Switch2*yr4	1.285 **	0.138	0.911 **	0.042	1.051	0.115	1.458 **	0.257
Switch2*yr5	0.991	0.088	0.562 ***	0.034	0.849 *	0.079	0.887	0.123
Switch2*yr6	0.934	0.138	0.524 ***	0.111	0.590 **	0.125	1.523	0.579
Switch2*yr7	0.779	0.167	0.415 ***	0.054	0.505 ***	0.089	0.326 ***	0.052
2007	0.891	0.062	0.961	0.033	1.032	0.066	1.033	0.099
2008	0.775 ***	0.055	0.939 *	0.034	0.807 ***	0.055	0.805 **	0.084
Time dummies	Yes		Yes		Yes		Yes	

Note: ***p<0.01, ** p<0.05, *p<0.10.
Data source: UCT 2006, 2007 and 2008, author's own calculations

All students enrolled for 3-year degrees were significantly more likely to graduate. Within the racial breakdown, the gender variable was statistically significant for all race groups. Male students, irrespective of race, were significantly less likely to graduate from the university, controlling for faculty, matric score and degree length. This is not surprising as it conforms to previous findings about gender and also confirms the information in Figure 15 (Murray, 2014).

5.6.4 Gender breakdown

An interesting finding in the gender breakdown is the type of high school attended. For females, while all except one of the school variables were statistically significant for the type of high school attended, the effect of high school on the incidence of voluntary dropout was only significant at the 90% level. For males the type of high school attended is much more important. Males who attended Department of Education and Training or House of Delegates schools were significantly more likely to be academically excluded and significantly less likely to graduate compared to males who attended Model C or private schools.

Table 31 SHR estimates for males

	Dependent Variable					
	Voluntary Dropout		Involuntary Dropout		Graduation	
	SHR	Std. Error	SHR	Std. Error	SHR	Std. Error
Indian	0.794	0.173	1.709 ***	0.309	0.817 ***	0.047
Affluence	0.923	0.271	1.009	0.153	0.983	0.098
Matric Score	0.990	0.009	0.991 **	0.004	1.001	0.002
African*Yr1	0.120 **	0.126	8.004 ***	2.054	------------	
African*Yr2	0.576	0.219	4.264***	0.969	------------	
African*Yr3	0.741	0.313	2.375 ***	0.499	1.006	0.189
African*Yr4	0.560	0.289	1.034	0.226	1.169 *	0.102
African*Yr5	0.230 **	0.167	0.893	0.224	0.884	0.078
African*Yr6	0.160	0.643	0.573	0.198	0.755 ***	0.078
African*Yr7	3.153 **	1.652	0.340	0.235	0.690 **	0.122
Coloured*Yr1	0.570	0.323	13.990 ***	3.345	------------	
Coloured*Yr2	0.743	0.225	4.156 ***	1.123	------------	
Coloured*Yr3	0.266 *	0.191	2.535 ***	0.537	1.339 *	0.226
Coloured*Yr4	0.500 *	0.195	0.419 **	0.145	1.194 ***	0.065
Coloured*Yr5	1.031	0.356	0.521 *	0.186	0.758 ***	0.055
Coloured*Yr6	0.354	0.366	0.557	0.371	0.529 ***	0.047
Coloured*Yr7	2.881 ***	1.441	2.301 ***	1.031	0.671 **	0.108
School: DET	0.341 **	0.142	1.890 ***	0.308	0.644 ***	0.077
School: HoD	1.192	0.453	1.636 *	0.438	0.776 **	0.096
School: HoR	1.173	0.319	1.159	0.195	0.918	0.079
3 year degree	1.083	0.173	0.701 ***	0.087	1.533 ***	0.073
Financial Aid	1.185	0.230	1.079	0.097	0.883 **	0.055
Humanities	1.286	0.233	1.153	0.172	0.860 ***	0.048
Science	0.803	0.194	1.635 ***	0.230	0.902	0.057
Engineering	0.923	0.161	1.336 **	0.141	0.934 *	0.038
Extended programme	0.728	0.162	1.196 **	0.109	0.798 ***	0.049
Residence	1.334 *	0.226	1.027	0.095	0.957	0.047
English HL	0.834	0.170	0.806	0.121	1.217 ***	0.075
Failed 1	1.281	0.214	0.737 **	0.089	0.993	0.044
Western Cape	1.064	0.177	0.982	0.105	1.010	0.043
GPA50-59	0.611 ***	0.102	0.413 ***	0.042	2.217 ***	0.153
GPA60-69	0.439 ***	0.088	0.182 ***	0.026	3.004 ***	0.211
GPA70-74	0.243 ***	0.092	0.056 ***	0.023	3.576 ***	0.280
GPA75+	0.148 ***	0.079	0.023 ***	0.016	3.814 ***	0.301
Switch2*yr2	11.586 ***	2.276	0.693 **	0.114	-------------	
Switch2*yr3	0.742	0.329	1.325	0.231	1.212	0.203
Switch2*yr4	0.190 **	0.134	1.194	0.224	0.981	0.062
Switch2*yr5	0.563	0.256	0.868	0.195	0.763 ***	0.049
Switch2*yr6	1.163	0.564	0.555	0.296	0.894	0.097
Switch2*yr7	1.379	1.250	4.851 ***	2.310	0.628 ***	0.102
Time dummies	Yes		Yes		Yes	

Note: ***p<0.01, ** p<0.05, *p<0.10.

Data source: UCT 2006, 2007 and 2008, author's own calculations

Once again, a surprising outcome is that of financial aid. For females, financial aid was associated with a lower incidence of voluntary drop out and higher incidences of academic exclusion. In contrast, the role of financial aid on males was significantly different. Males in receipt of financial aid were significantly less likely to graduate from university.

Table 32 SHR estimates for females

	Dependent Variable					
	Voluntary Dropout		Involuntary Dropout		Graduation	
	SHR	Std. Error	SHR	Std. Error	SHR	Std. Error
Indian	1.000	0.191	1.965 ***	0.449	0.763 ***	0.046
Affluence	0.773	0.254	0.843	0.141	1.117	0.093
Matric Score	0.980 ***	0.006	0.989 *	0.006	1.006 ***	0.002
African*Yr1	0.615	0.417	7.998 ***	2.542	------------	
African*Yr2	0.473 ***	0.112	5.141 ***	1.411	------------	
African*Yr3	0.417 **	0.177	2.191 ***	0.499	1.186	0.142
African*Yr4	0.502 *	0.181	0.781	0.227	1.038	0.073
African*Yr5	0.087 **	0.084	0.568 *	0.186	0.855	0.064
African*Yr6	1.119	0.467	0.073 **	0.089	0.645 ***	0.056
African*Yr7	0.468	0.488	0.238	0.253	0.754 **	0.099
Coloured*Yr1	0.363	0.278	19.919 ***	5.883	------------	
Coloured*Yr2	0.826	0.183	5.446 ***	1.661	------------	
Coloured*Yr3	0.325 ***	0.135	1.618 *	0.434	1.712 ***	0.139
Coloured*Yr4	0.384 ***	0.135	0.517	0.212	0.952	0.047
Coloured*Yr5	0.730	0.296	0.616	0.232	0.709 ***	0.048
Coloured*Yr6	1.354	0.596	0.239	0.211	0.486 ***	0.049
Coloured*Yr7	0.950	0.659	1.862	1.393	0.440 ***	0.076
School: DET	0.428 *	0.200	1.107	0.194	1.078	0.097
School: HoD	0.97	0.285	1.265	0.342	0.957	0.081
School: HoR	1.123	0.285	0.932	0.151	1.072	0.067
3 year degree	1.017	0.151	0.876	0.133	1.317 ***	0.053
Financial Aid	0.571 **	0.139	1.226 *	0.142	1.044	0.054
Humanities	1.556 ***	0.237	0.631 ***	0.107	1.070	0.046
Science	1.053	0.217	1.089	0.184	0.962	0.057
Engineering	1.212	0.232	0.970	0.140	1.002	0.051
Extended programme	0.792	0.142	1.661 ***	0.200	0.751 ***	0.041
Residence	1.141	0.160	1.149	0.150	0.946	0.042
English HL	0.906	0.147	0.629 ***	0.102	1.155 ***	0.059
Failed 1	1.126	0.154	0.860	0.142	0.987	0.042
Western Cape	0.863	0.111	0.995	0.124	1.066	0.045
GPA50-59	0.855	0.121	0.325 ***	0.044	1.853 ***	0.122
GPA60-69	0.572 ***	0.089	0.101 ***	0.023	2.499 ***	0.161
GPA70-74	0.360 ***	0.107	0.028 ***	0.021	2.863 ***	0.209
GPA75+	0.497 **	0.153	0.023 ***	0.023	2.603 ***	0.216
Switch2*yr2	15.429 ***	2.453	0.354 ***	0.086	-------------	
Switch2*yr3	0.398 **	0.180	1.717 ***	0.283	1.201 *	0.113
Switch2*yr4	0.907	0.263	0.539 *	0.179	0.930	0.049
Switch2*yr5	0.587	0.300	0.850	0.247	0.751 ***	0.045
Switch2*yr6	1.521	0.712	0.636	0.721	0.880	0.145
Switch2*yr7	1.018	0.759	0.517	0.461	0.594 ***	0.081
Time dummies	Yes		Yes		Yes	

Note: ***p<0.01, ** p<0.05, *p<0.10.

Data source: UCT 2006, 2007 and 2008, author's own calculations

The final outcome of interest is that of faculty of choice. Females are more likely to voluntarily drop out and less likely to be academically excluded from the Humanities faculty. For males the effects are rather different. Males are more likely to be academically excluded from the Science and Engineering faculty. One reason for this could be that males make up the majority of students in the Science, Technology, Engineering and Maths degrees. These disciplines are also more technical, requiring more technical skills, which it appears more females enter the HE system with. Males in the Engineering and Humanities faculties were significantly less likely to graduate. The extended programme effects for both genders were similar, indicating that something other than gender drives performance in the extended programmes.

5.6.5 Degree-duration breakdown

A final sub-analysis in this section is that of degree duration. The analysis is therefore broken down into 3-year and 4-year degree durations. The presentation of information in this subsection allows for additional information by degree duration, as presented in previous chapters.

Tables 33 and 34 show the results for the competing risks estimations for students enrolled in 3-year and 4-year programmes respectively. The tables show the same patterns as the previous tables. Interestingly, the graduation specifications are most influenced by university-related covariates, such as GPA in the first year and degree switching, which was already shown to be an important indicator in the Chapter 4.

Table 33 SHR estimates for 3-year degree duration

	Dependent Variable					
	Voluntary Dropout		Involuntary Dropout		Graduation	
	SHR	Std. Error	SHR	Std. Error	SHR	Std. Error
Male	0.830*	0.083	1.408***	0.123	0.870***	0.027
Indian	0.884	0.178	2.113***	0.431	0.771***	0.050
Affluence	1.213	0.342	0.947	0.142	0.931	0.078
Matric Score	0.980***	0.006	0.993	0.005	1.005**	0.002
African*Yr1	0.214**	0.166	11.314***	2.934	----------	
African*Yr2	0.417***	0.094	6.057***	1.442	----------	
African*Yr3	0.466**	0.148	2.481***	0.489	1.184*	0.116
African*Yr4	0.432**	0.152	1.120	0.239	1.033	0.072
African*Yr5	0.129***	0.095	1.001	0.277	0.865*	0.067
African*Yr6	1.002	0.428	0.495	0.213	0.719***	0.071
African*Yr7	1.450	0.797	0.009***	0.000	0.7645*	0.118
Coloured*Yr1	0.433	0.238	19.167***	5.497	----------	
Coloured*Yr2	0.754	0.155	5.194***	1.457	----------	
Coloured*Yr3	0.281***	0.108	1.760**	0.394	1.538***	0.102
Coloured*Yr4	0.369***	0.118	0.410**	0.147	0.990	0.048
Coloured*Yr5	0.967	0.304	0.842	0.271	0.683***	0.049
Coloured*Yr6	0.524	0.331	0.820	0.420	0.521***	0.047
Coloured*Yr7	0.588	0.513	0.678	0.503	0.555***	0.108
School: DET	0.606	0.120	1.638	0.268	0.749***	0.074
School: HoD	0.535*	0.256	1.656	0.409	0.951	0.095
School: HoR	1.240	0.201	1.049	0.171	1.053	0.062
Financial Aid	0.793	0.141	1.069	0.107	0.975	0.047
Humanities	1.267*	0.176	0.807*	0.105	1.061	0.043
Science	0.794	0.136	1.191	0.138	1.031	0.045
Engineering	0.878	0.253	0.953	0.213	1.136**	0.072
Extended programme	0.794	0.125	1.540***	0.156	0.772***	0.036
Residence	1.020	0.141	1.207*	0.129	0.947	0.042
English HL	0.688***	0.098	0.765*	0.107	1.218***	0.061
Failed 1	1.184	0.154	0.812	0.115	0.986	0.031
Western Cape	0.862	0.113	1.042	0.116	1.035	0.041
GPA50-59	0.683***	0.087	0.368***	0.043	2.070***	0.126
GPA60-69	0.533***	0.077	0.178***	0.028	2.685***	0.164
GPA70-74	0.297***	0.091	0.045***	0.023	3.243***	0.229
GPA75+	0.353***	0.119	0.035***	0.025	3.035***	0.231
Switch2*yr2	12.784***	1.911	0.562***	0.104	----------	
Switch2*yr3	0.485**	0.173	1.595***	0.256	1.181**	0.088
Switch2*yr4	0.458**	0.164	1.319	0.242	0.851***	0.040
Switch2*yr5	0.439*	0.206	1.122	0.261	0.694***	0.037
Switch2*yr6	1.924	0.863	1.026	0.511	0.718**	0.100
Switch2*yr7	1.241	0.861	0.956	0.716	0.588***	0.080
Time dummies	Yes		Yes		Yes	

Note: ***p<0.01, ** p<0.05, *p<0.10.

Table 34 SHR estimates for 4-year degree duration

	Dependent Variable					
	Voluntary Dropout		Involuntary Dropout		Graduation	
	SHR	Std. Error	SHR	Std. Error	SHR	Std. Error
Male	0.841	0.113	1.392***	0.129	0.836***	0.029
Indian	0.762	0.168	1.557	0.295	0.798***	0.043
Affluence	0.495**	0.157	0.852	0.137	1.307***	0.118
Matric Score	0.993	0.009	0.985***	0.005	1.004	0.003
African*Yr1	0.535	0.446	5.246***	1.525	-----------	
African*Yr2	0.732	0.370	3.484***	0.860	-----------	
African*Yr3	0.365	0.386	2.269***	0.545	0.635	0.178
African*Yr4	0.636	0.351	0.807	0.220	1.189**	0.104
African*Yr5	0.172*	0.170	0.527**	0.151	0.864*	0.068
African*Yr6	1.825	0.752	0.246**	0.146	0.613***	0.055
African*Yr7	2.098	1.514	0.401*	0.216	0.654***	0.093
Coloured*Yr1	0.504	0.403	12.888***	2.846	-----------	
Coloured*Yr2	0.906	0.295	4.227***	1.149	-----------	
Coloured*Yr3	0.269	0.278	3.047***	0.753	1.090	0.321
Coloured*Yr4	0.478	0.220	0.591	0.227	1.187***	0.068
Coloured*Yr5	0.583	0.307	0.372**	0.166	0.774***	0.049
Coloured*Yr6	1.653	0.727	0.001***	0.000	0.529***	0.047
Coloured*Yr7	0.001***	0.000	0.002***	0.000	0.553***	0.070
School: DET	0.226***	0.128	1.204	0.186	1.021	0.095
School: HoD	2.366***	0.773	1.301	0.382	0.762**	0.081
School: HoR	1.023	0.325	1.164	0.202	0.966	0.076
Financial Aid	0.875	0.250	1.2601**	0.126	0.962	0.062
Humanities	1.925***	0.383	1.059	0.224	0.869**	0.057
Engineering	1.148	0.175	1.281**	0.127	0.893***	0.035
Extended programme	0.667	0.217	1.060	0.121	0.887	0.072
Residence	1.446**	0.242	0.978	0.103	0.923*	0.043
English HL	1.458	0.366	0.684**	0.113	1.113*	0.065
Failed 1	1.307	0.234	0.812**	0.115	0.992	0.044
Western Cape	1.037	0.166	0.868	0.104	1.035	0.041
GPA50-59	0.763	0.146	0.364**	0.043	1.9067***	0.136
GPA60-69	0.434***	0.104	0.106***	0.021	2.689***	0.190
GPA70-74	0.276***	0.101	0.043***	0.022	2.990***	0.231
GPA75+	0.263***	0.122	0.012***	0.012	3.361***	0.271
Switch2*yr2	15.665***	3.252	0.496***	0.101	-----------	
Switch2*yr3	0.820	0.624	1.384	0.286	0.999	0.288
Switch2*yr4	0.954	0.364	0.431**	0.159	1.170**	0.090
Switch2*yr5	0.772	0.367	0.578*	0.178	0.935	0.060
Switch2*yr6	0.925	0.510	0.002***	0.000	1.116	0.135
Switch2*yr7	0.003***	0.000	0.003***	0.000	0.623**	0.138
Time dummies	Yes		Yes		Yes	

Note: ***p<0.01, ** p<0.05, *p<0.10.

The results in Tables 33 and 34 are supportive of the claims in Chapter 4 that students who enrol in 3-year and 4-year degrees are characteristically different. Table 33 and 34 show that the covariates for each of the three exit outcomes are different for the two groups.

5.7 Conclusion

In this chapter detailed evidence was provided on the determinants of academic outcomes using discrete-time methods for competing risks survival analysis. Understanding this subject matter is important because of the implications for both students and HEIs. The high and rising costs of HE paired with increasing importance of HE qualifications in the labour market cannot be discounted in the current economic environment.

By observing students over an extended period of time, an informative sense of the determinants of dropout and graduation was gauged. The results of the analysis provide guidance on the key determinants affecting student retention, namely race, gender, and performance in first year studies. The lack of evidence that financial aid status is correlated with either voluntary dropout or graduation after controlling for academic and socio-economic background factors is worth noting.

A key outcome of this chapter is that an understanding of the temporal element of education paths is important for understanding different academic outcomes students may experience in a HE setting. The results indicate that race and gender are consistently important determinants across all three potential academic outcomes. Higher entrance scores given by students' performance on the final Grade 12 exam are associated with lower probabilities of either type of dropout. However, matric scores do not significantly influence graduation outcomes, indicating that some part of the school effect is eliminated over time. Students' residence status is another important factor determining academic outcomes. Residential status is associated with higher rates of voluntary and involuntary dropout and lower rates of graduation. Lastly, students' academic performance in first year, given by their GPA for the first year, is an important indicator of future academic outcomes. A GPA of at least 50% is associated with lower rates of voluntary and involuntary dropout and higher rates of graduation, but the effect varies across the grade distribution as one would expect. The outcomes for individuals registered on academic programmes are of concern too. Students on academic programmes are more likely to be involuntarily excluded and less likely to graduate or voluntarily exit HE than mainstream students. This is cause for concern as these programmes are an initiative intended to address transformation and equity in HE and attract significant resources from within and outside the university.

The one outcome that was expected to be significant across all three academic outcomes was financial aid. The results indicate that financial aid is not correlated with involuntary exit. There is also no statistical difference in graduation rates between financial aid recipients and students on other forms of funding. Many previous studies worldwide find that financial aid is one of the most important factors preventing dropout and improving success or graduation rates (Arias & Dehon, 2011).

In order to reach more informed and reliable conclusions, universities must improve the depth of data they collect from potential applicants. A lack of detailed information about individuals enrolled in HE, such as parental background and socio-economic status, prevents a more detailed analysis of the determinants of HE academic outcomes. Previous studies, both local and international show that parental education and family background are important determinants of academic performance (Buchmann & DiPrete, 2006). As family background is shown to be important and a significant determinant of academic outcomes (Lassibille, 2011), this information must be routinely collected from students upon application.

Further investigation into student outcomes should also entail modelling the hierarchical nature of HE data. Multilevel or hierarchical modelling split along faculty of enrolment and degree of study would be an excellent starting point. Moreover, an investigation that takes into account the multilevel nature of HE data and the existence of competing risks are likely to yield detailed and valuable analysis to the academic administrator and policymaker.

Chapter 6: Conclusion

Educational attainment is highly consequential in South Africa. McMahon (2009) presents a thorough review of the private and social benefits of higher education. At a macro level, the overall level of educational attainment is important for a country's economic development and growth. At an individual level, completion of the highest possible levels of education plays a role in labour market outcomes, earnings, marriage prospects, improvements in access to opportunity and the promotion of social mobility.

Throughout the world, and especially in the United States, the importance of higher education has been recognised as one of the leading factors contributing to the country's economic growth and development. Especially in the US context, the strength of the economy is premised on the high rate of higher education completion among the population.

However, over time, while the level of education attainment has remained at the forefront of education policy, disparities in educational outcomes by race, gender and socioeconomic status are increasingly being investigated. For those who care about the economic prospects of this country, and its social fabric, evaluating the successes and failures in higher education is significantly important.

This thesis investigates the correlates of higher education outcomes in South Africa, using the University of Cape Town as a case study. It contributes to the economics of education literature on incentives, performance and outcomes in HE by contributing three papers on the South African experience. The first contribution explores the use of incentives as a motivator for student achievement. The second contribution examines the pathways through higher education. The third contribution investigates how students exit from higher education in South Africa and analyses the competing nature of academic outcomes.

6.1 Findings

In the first substantive chapter, the thesis finds that the DML as an academic incentive policy has a large negative impact on academic performance, both in the short run and in the long run. The short run is measured by the student's annual non-cumulative GPA. The cumulative GPA, over the entire course of a student's study in HE is used to measure long run performance. Over both these measures, the DML is found to have a negative impact. However, the stability and precision of the results have been questioned. This may be interpreted as the DML not being an adequate incentive for good students to maintain or improve their academic performance and that the DML does not reinforce academic achievement. These results appear to be counterintuitive but are supportive of Bénabou and Tirole's (2002) theoretical expectations regarding coasting, where good students become over-confident in subsequent periods, exerting less effort than is expected on tasks. The second finding from the first substantive chapter is the impact of the DML policy on exit outcomes. We believe that the estimates for exit outcomes are estimated with more precision, resulting in more stability of the results. However, the results are still opposite to expectation, indicating that students treated with the DML are more likely to dropout. The treated students are also less likely to graduate compared to the control group.

The second contribution by this study examined the pathways through HE by investigating student performance over time. It also introduced a ranking variable approach to student achievement. The data provided supportive evidence that there are differing pathways through HE based on race and gender in South Africa. The main finding was that race, gender and performance on final school-leaving examination are important determinants of academic achievement. Surprisingly, high school type attended was not statistically significant, but the descriptive nature of the analysis means we cannot draw causal inferences from the results.

The results show that gender plays a big role in explaining part of the differences in outcomes related to student achievement. The gender effects were particularly notable as the underlying process of programme rank improvements for males appears to be different to females. Males were shown to experience a greater decline in rank between the first and graduating year relative to females. The fall in programme rank is largest for African males. Female students outperform male students across the distribution of GPA, a finding that is consistent with the growing international literature (Van Broekhuizen & Spaull, 2017). A similar relationship was noted for degree switching where the correlates for degree switching appear to differ by gender. An expected finding was that of entrance scores. Students who enter HE with higher entrance scores perform relatively well in terms of programme rank, and show little improvement as they tend to remain at the top of the distribution throughout their studies.

The third and final finding in this study presented detailed evidence on the determinants of academic outcomes using discrete-time methods for competing risks survival analysis. While graduation is the preferred route of exit, voluntary and involuntary exit before completion remains prominent for a significant number of students. Interestingly, and contrary to other international studies, this study did not find support for financial aid status contributing to either voluntary dropout or graduation, even after controlling for academic and socio-economic background factors. Students on extended academic programmes were shown to be more likely to be involuntarily excluded and less likely to graduate or voluntarily exit HE than mainstream students. While the results in Chapters 4 and 5 are not causal, they present a basic understanding of the timing of exit that students may face. For example, the results show that African students experience involuntary dropout differently to other students. This is cause for concern as these programmes are initiatives intended to address transformation and equity in HE, attracting significant resources from within and outside the university. This gives rise to concern about policies that aim to retain students in the system. This may include a focus on student advising initiatives and how best to implement them. This study also found support for higher entrance scores, as given by students' performance on the final Grade 12 exam, being associated with lower probabilities of either voluntary or involuntary dropout. However, matric scores were not found to significantly correlate with graduation outcomes, indicating that some part of the school effect is eliminated over time. Another important factor in the South African context, that of students' residence status, was shown to be associated with higher rates of voluntary and involuntary dropout as well as lower rates of graduation. The study also found that students' academic performance in first year, as given by their GPA, is an important indicator of future academic outcomes. A GPA of at least 50% is associated with lower rates of voluntary and involuntary dropout and higher rates of graduation, but the effect varies across the grade distribution as one would expect. The final finding of the competing risks analysis is the supportive evidence of students learning from their past academic performance. In all estimations, the coefficients for the first year GPA dummy variables are statistically significant, with the dummy variables for higher GPAs showing larger influences on the SHR for graduation.

6.2 Contributions

The economics of higher education is a growing area of research, especially in developing countries where natural experiments to evaluate policies are harder to come by. This thesis contributed to the HE literature by examining pathways through HE and incentives.

The body of HE literature in developed countries is well established. Countries like the US lead this research area as natural experiments around policies in HE is fairly abundant. However, the same cannot be said for many developing countries.

The work presented in this thesis provides evidence of the impact of recognition policies in a developing country context. The impact of recognition policies on exit outcomes is found to be counter-intuitive with respect to theoretical expectations. However, the results for the effects of the DML policy on GPA are less certain as the estimates are large relative to expectations.

This thesis also contributes to the literature on pathways through education by providing an attempt to map and understand the pathways through HE in a South African context. The results provide evidence for HE experiences to differ by race and gender in the SA context, and that the experience of African males in HE is significantly worse compared to the strongest group of White females.

Overall, this thesis makes a modest contribution to the literature on the impacts of recognition policies on students. In addition, the thesis shows that in the South African context, HE has yet to overcome many of the issues that plague the sector, including racial and gender disparities in educational performance and outcomes.

6.3 External validity

A key concern with a case study approach to economic issues is that of external validity. It is important that readers and users of research can understand the approaches adopted to ensure that concerns about internal and external validity have been taken into account. Internal validity questions may be settled when evaluating the techniques used to assess a particular intervention or policy. Quasi-experimental methods such as regression discontinuity designs have been shown to meet or even exceed the minimum criteria for internal validity (Lee & Lemieux, 2009). Identification strategies that meet the baseline criteria for RDD permits researchers to assert causality about policy interventions or natural experiments encountered in different settings via comparison with valid counterfactuals. Proceeding from this, researchers question the generalisation and extrapolation of results beyond the study samples in what is commonly known as external validity. Claims about external validity arise when considering how an intervention may be replicated or scaled up and continue to hold in a different context.

Policymakers face the problem of knowing if a study's findings, conducted elsewhere, would continue to hold in their specific context. This is especially pertinent when RDD is conducted on small and very niche populations where results derived generally provide feedback on local average treatment effects. This means that findings are limited to the specific populations that have been studied and may not extrapolate to larger populations were natural experiments could give rise to average treatment effects. In an ideal world, studies are on populations where treatment effects cover a sufficiently large group such that local average treatment effects converge with average treatment effects. A study by Oreopoulos (2010) evaluating the effects of a change to the minimum school-leaving age in the UK is one rare study where results cover a general population.

The case study analyses in this thesis do not meet the criteria for external validity. The nature of the data analysed also does not support extrapolation of results towards external validity. However, the case study results present robust descriptive evidence of pathways and outcomes in HE for South Africa, and should contribute to the creation of knowledge of the functioning of the HE sector in South Africa.

6.4 Limitations and avenues for further research

South Africa is a very diverse country, colloquially known as the rainbow nation, comprised of many small parts. By this description, the contributions of this thesis should be considered as a small part of a large puzzle – HE in South Africa and Africa.

The use of administrative data at a historically White institution implies that the results are not generalisable to a larger population as the underlying sample population is not nationally representative. The case study approach to this thesis presents small chunks of information and highlights opportunities for future research. While the underlying research methodologies conform to standard practice within the discipline, case studies do not allow for generalisation to a representative population at a national level.

In chapter 3 with current data limitations in this thesis, the analysis does not allow for the underlying causal mechanisms to be clearly identified. At best, at least one or two mechanisms may be discounted, but without additional data, especially survey and qualitative data, many studies may not be able to establish the necessary rigour to make causal claims. While at least one potential explanation of choosing programmes that are easier has been ruled out, further data and research is required to shed light on the drivers of the negative outcomes.

Similarly, in any analysis of educational outcomes, such as chapters 4 and 5, more holistic information about students provides more opportunities to fully explore student experiences in HE. The results in chapters 4 and 5 respectively did not shed enough light on the role of financial aid. As new research and econometric methods are developed, the funding and financial side of HE will become important avenues of research. This type of research, such as the role of financial aid, and the introduction of merit aid in South Africa, and its impact on academic outcomes, will be important an important future research contribution.

Another limitation of the study was the unavailability of data that could link students to further study or labour market outcomes. In acknowledging that this would have been an excellent contribution given the labour market conditions in South Africa, it does represent an opportunity for future research. The results obtained in this paper is a motivator to enhance data collection not only at UCT but the HE sector in general.

As the South African government seeks to continue spending a large proportion of its education budget on universities, it is imperative that universities improve their data collection within the sector. For example, detailed surveys of student experiences and interactions across the entire South African HE sector would provide an abundance of opportunities for more generalised research, especially for samples that would be national representative. The establishment of longitudinal surveys in the sector would also enhance the research capacity of the sector. The research provided in this thesis could serve to inform future research in South Africa by providing a descriptive pool of evidence through case study analysis.

References

Aina, C., Baici, E., Casalone, G. & Pastore, F. 2021. The determinants of university dropout: A review of the socio-economic literature. *Socio-Economic Planning Sciences*: 101102.

Allison, P. 1984. *Event history analysis: regression for longitudinal event data*. Newbury Park: Sage Publications.

Alon, S., 2007. The influence of financial aid in leveling group differences in graduating from elite institutions. *Economics of Education Review*, *26*(3):296-311.

Alon, S. & Tienda, M. 2005. Assessing the “mismatch” hypothesis: Differences in college graduation rates by institutional selectivity. *Sociology of education*, *78*(4):294-315.

Altonji, J.G., 1993. The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics*, *11*(1, Part 1):48-83.

Arcidiacono, P., Aucejo, E.M. & Spenner, K. 2012. What happens after enrollment? An analysis of the time path of racial differences in GPA and major choice. *IZA Journal of Labor Economics*. 1(1):5.

Arias O.E. & Dehon, C. 2011. *The road to success: analysing dropout and degree completion at university* (Working paper 2011–025). E-cares.

Astorne-Figari, C. & Speer, J.D. 2019. Are changes of major major changes? The roles of grades, gender, and preferences in college major switching. *Economics of Education Review*. 70:75–93.

Aucejo, E. 2013. Explaining cross-racial differences in the educational gender gap.

Austin, P.C., Lee, D.S. & Fine, J.P. 2016. Introduction to the analysis of survival data in the presence of competing risks. *Circulation*, *133*(6):601-609.

Bean, J.P. 1980. Dropouts and turnover: the synthesis and test of a causal model of student attrition. *Research in Higher Education*. 12(2):155–197.

Bello, A.L. & Valientes, R.M. 2008. Grade inflation: fact or myth? *Philippine Review of Economics*. 45(1):93–108.

Bénabou, R. & Tirole, J. 2000. *Self-confidence and social interactions* (No. w7585). National Bureau of Economic Research.

Bénabou, R. & Tirole, J. 2003. Intrinsic and extrinsic motivation. *The review of economic studies*, *70*(3):489-520.

Bettinger, E.P. & Baker, R.B. 2014. The effects of student coaching: an evaluation of a randomized experiment in student advising. *Educational Evaluation and Policy Analysis*. 36(1):3–19.

Betts, J.R. & Morell, D. 1999. The determinants of undergraduate grade point average: the relative importance of family background, high school resources, and peer group effects. *Journal of Human Resources*. 43(2):268–293.

Bokana, K.G. & Tewari, D.D. 2014. Determinants of student success at a South African university: An econometric analysis. *The Anthropologist*, *17*(1):259-277.

Bowen, W. & Bok, D. 1998. *The shape of the river*. Princeton, NJ: Princeton University Press.

Bowen, W.G., Chingos, M.M. & McPherson, M.S. 2009. *Crossing the finish line: completing college at America's public universities*. Vol. 52. Princeton, NJ: Princeton University Press.

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.

Branson, N. & Leibbrandt, M. 2013. *Educational attainment and labour market outcomes in South Africa, 1994–2010* (Working paper 1022). OECD Economics Department. OECD Publishing. Available: http://dx.doi.org/10.1787/5k4c0vvbvv0q-en

Branson, N., Leibbrandt, M. & Zuze, T.L. 2009. *The demand for tertiary education in South Africa*. South African Labour and Development Research Unit.

Buchmann, C. & DiPrete, T.A. 2006. The growing female advantage in college completion: The role of family background and academic achievement. *American sociological review*, *71*(4):515-541.

Calonico, S., Cattaneo, M.D. & Titiunik, R. 2014. Robust data-driven inference in the regression-discontinuity design. *Stata Journal*. 14:909–946.

Casey, M.D., Cline, J., Ost, B. & Qureshi, J.A. 2018. Academic probation, student performance, and strategic course-taking. *Economic Inquiry*. 56(3):1646–1677.

Cattaneo, M.D., Jansson, M. & Ma, X. 2018. Manipulation testing based on density discontinuity. *The Stata Journal*. 18(1):234–261.

Cattaneo, M.D., Keele, L. and Titiunik, R., 2021. Covariate Adjustment in regression discontinuity designs. *arXiv preprint arXiv:2110.08410*.

Cattaneo, M.D., Titiunik, R. and Vazquez-Bare, G., 2019. Power calculations for regression-discontinuity designs. *The Stata Journal*, *19*(1), pp.210-245.

Caviglia-Harris, J. & Maier, K. 2020. It's not all in their heads: The differing role of cognitive factors and non-cognitive traits in undergraduate success. *Education Economics*, *28*(3):245-262.

Christie, H., Munro, M. & Fisher, T. 2004. Leaving university early: Exploring the differences between continuing and non-continuing students. *Studies in Higher education*, *29*(5):617-636.

Clerici, R., Giraldo, A. & Meggiolaro, S. 2014. The determinants of academic outcomes in a competing risks approach: evidence from Italy. *Studies in Higher Education*. 40(9):1535–1549.

Cleves, M., Gould, W., Gutierrez, R. & Marchenko, Y. 2010. *An introduction to survival analysis using Stata*. 3rd ed. College Station, TX: Stata Press

Council on Higher Education. 2012. *Vital stats. Public higher education 2010*. Pretoria: CHE.

Council on Higher Education. 2015. *Vital stats. Public higher education 2013*. Pretoria: CHE.

Council on Higher Education. 2018. *Vital stats. Public higher education 2016*. Pretoria: CHE.

Cox, D.R. 1972. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological), 34*(2): 187-202.

Crandall, V. & McGhee, P. 1968. Expectancy of reinforcement and academic competence. *Journal of Personality*. 36:635–648.

Deming, D. & Dynarski, S. 2010. College aid. In *Targeting investments in children: Fighting poverty when resources are limited*: 283-302. University of Chicago Press.

Department of Basic Education. 2015. https://www.education.gov.za/Resources/Reports.aspx

Department of Higher Education and Training. 2018. *Statistics on post-school education and training in South Africa: 2016*. Pretoria

Department of Higher Education and Training. 2019. *2000 to 2016 first time entering undergraduate cohort studies for public higher education institutions*. Pretoria

DesJardins, S.L., Ahlburg, D.A. & McCall, B.P. 2002. Simulating the longitudinal effects of changes in financial aid on student departure from college. *Journal of Human Resources*. 37(3):653–679.

Dobson, I., Sharma, R. & Haydon, A. 1997. *Commencing undergraduates in Australian universities: enrolment and performance trends 1993–1995. Survey of applicants for undergraduate higher education courses (ACN 008 502 930)*. Canberra: Australian Vice-Chancellors' Committee.

Dynarski, S. 2000. Hope for whom? Financial aid for the middle class and its impact on college attendance. *National Tax Journal*, *53*(3):629-661.

Dynarski, S. 2002. Race, income, and the impact of merit aid. *Who should we help*: 73-92.

Dynarski, S. 2004. The new merit aid. In *College choices: The economics of where to go, when to go, and how to pay for it* (63-100). University of Chicago Press.

Dynarski, S. & Scott-Clayton, J. 2013. Financial aid policy: Lessons from research.

Fisher, G. & Scott, I. 2011. *Background paper 3: the role of higher education in closing the skills gap in South Africa* (Closing the skills and technology gap in South Africa). The World Bank.

Fletcher, J.M. & Tokmouline, M. 2010. The effects of academic probation on college success: lending students a hand or kicking them while they are down?

Fryer Jr, R.G. & Levitt, S.D. 2004. Understanding the black-white test score gap in the first two years of school. *Review of economics and statistics, 86*(2):447-464.

Gallagher, M. & Conn, W. 1994. *Diversity and performance of Australian universities* (No. 22). Canberra: DEET.

Gallop, C.J. & Bastien, N. 2016. Supporting Success: Aboriginal Students in Higher Education. *Canadian Journal of Higher Education*, *46*(2):206-224.

Gelman, A. & Imbens, G. 2019. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*. 37(3):447–456.

Goldin, C., Katz, L.F. & Kuziemko, I. 2006. The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic perspectives, 20*(4):133-156.

Goldrick-Rab, S.N., Harris, D., Kelchen, R. & Benson, J. 2012. *Need-based financial aid and college persistence experimental evidence from Wisconsin*. Available: http://dx.doi.org/10.2139/ssrn.1887826

Grove, W.A. & Wasserman, T. 2004. The life-cycle pattern of collegiate GPA: longitudinal cohort analysis and grade inflation. *The Journal of Economic Education*. 35(2):162–174.

Horn, L. 1998. Stopouts or stayouts? Undergraduates who leave college in their first year (NCES 1999-087), U.S. Department of Education, Washington, DC. Aid on student departure from college. *The Journal of Human Resources*. 37(3):653–679.

Hosmer, D.W., 2000. Lemeshow S. Applied logistic regression.

Imbens, G.W. & Lemieux, T. 2008. Regression discontinuity designs: a guide to practice. *Journal of Econometrics*. 142(2):615–635.

Ishitani, T. & DesJardins, S.L. 2002. A longitudinal investigation of dropout from college in the United States. *Journal of College Student Retention: Research, Theory and Practice*. 4(2):173–201.

Jewell, R.T., McPherson, M.A. & Tieslau, M.A. 2013. Whose fault is it? Assigning blame for grade inflation in higher education. *Applied Economics*. 45(9):1185–1200.

Keswell, M. 2004. February. Education and racial inequality in post apartheid South Africa. In *DPRU/FES Second Annual Conference on Labour Markets*.

Keswell, M. & Poswell, L. 2004. Returns to education in South Africa: A retrospective sensitivity analysis of the available evidence. *South African Journal of Economics*. 72(4):834–860.

Khuluvhe, M. & Netshifhefhe, E. 2021. Funding and Expenditure Trends in Post-School Education and Training.

Kohn, A. 2002. The dangerous myth of grade inflation. *The Chronicle of Higher Education*. 49(11):B7.

Lee, D.S. & Lemieux, T. 2009. *Regression discontinuity designs in economics* (No. w14723).

Leeds, D.M. & DesJardins, S.L. 2015. The effect of merit aid on enrollment: a regression discontinuity analysis of Iowa's National Scholars Award. *Research in Higher Education*. 56(5):471–495.

Letseka, M. & Maile, S. 2008. *High university drop-out rates: a threat to South Africa's future*. HSRC Policy Brief. Pretoria, South Africa: Human Sciences Research Council.

Lindo, J.M., Sanders, N.J. & Oreopoulos, P. 2010. Ability, gender, and performance standards: evidence from academic probation. *American Economic Journal: Applied Economics*. 2(2):95–117.

Light, A. 1996. Hazard model estimates of the decision to reenroll in school. *Labour Economics*. 2:381–406.

Linke, R.D. 1991. *Performance indicators in higher education: report of a trial evaluation study commissioned by the Commonwealth Department of Employment Education and Training* (Vol. 1, Report and Recommendations). Canberra: AGPS.

Little, R.J. 1995. Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association*. 90(431):1112–1121.

Little, R.J. & Rubin, D.B. Eds. 2002. Bayes and multiple imputation. In *Statistical analysis with missing data*. 2nd ed. Wiley. 200–220.

Loury, L. & Garman, D. 1995. College selectivity and earnings. *Journal of Labor Economics*. 13(2):289–308.

Martin, N.D., Spenner, K.I. & Mustillo, S.A. 2017. A test of leading explanations for the college racial-ethnic achievement gap: Evidence from a longitudinal case study. *Research in Higher Education*. *58*(6):617-645.

Massey, D.S. & Mooney, M. 2007. The effects of America’s three affirmative action programs on academic performance. *Social Problems*. 54(1):99–117.

McCrary, J. 2008. Manipulation of the running variable in the regression discontinuity design: a density test. *Journal of Econometrics*. 142(2):698–714.

McMahon, W.W., 2009. *Higher learning, greater good: The private and social benefits of higher education*. JHU Press.

McNabb, R., Pal, S. & Sloane, P. 2002. Gender differences in educational attainment: The case of university students in England and Wales. *Economica*, *69*(275):481-503.

Meggiolaro, S., Giraldo, A. & Clerici, R. 2017. A multilevel competing risks model for analysis of university students’ careers in Italy. *Studies in Higher Education*. 42(7):1259–1274.

Milligan, K., Moretti, E. & Oreopoulos, P. 2004. Does education improve citizenship? Evidence from the United States and the United Kingdom. *Journal of public Economics*, *88*(9-10):1667-1695.

Ministry of Education. 2001. *National Plan for Education*. Pretoria.

Miranda, A. & Rabe-Hesketh, S. 2006. Maximum likelihood estimation of endogenous switching and sample selection models for binary, ordinal, and count variables. *Stata Journal*. 6:285–308.

Montmarquette, C., Mahseredjian, S. & Houle, R. 2001. The determinants of university dropouts: a bivariate probability model with sample selection. *Economics of Education Review*. 20:475–484.

Murray, M. 2014. Factors affecting graduation and student dropout rates at the University of KwaZulu-Natal. *South African Journal of Science*. 110(11/12):Art. #2014–0008. Available: http://dx.doi.org/10.1590/sajs.2014/20140008

Nonyana, J.Z. & Njuho, P.M. 2018. Modelling the length of time spent in an unemployment state in South Africa. *South African Journal of Science*, *114*(11-12):1-7.

Oates, G.L.S.C. 2009. An empirical test of five prominent explanations for the black-white academic performance gap. *Social Psychology of Education*. 12(4):415–441.

OECD. 2014. *Education at a glance 2013: OECD indicators*. OECD Publishing. Available: http://dx.doi.org/10.1787/eag-2013-en

OECD. 2021. *Population with tertiary education (indicator)*. OECD Publishing. Available: http://dx.doi.org/10.1787/0b8f90e9-en (Accessed on 23 February 2021)

Oosthuizen, M. & Bhorat, H. 2005. *The post-apartheid South African labour market* (Development and Poverty Research Unit Working Paper 05/093). Development Policy Research Unit. University of Cape Town.

Oreopoulos, P. & Petronijevic, U., 2013. Making college worth it: A review of research on the returns to higher education.

Patrick, W.J., 2001. Estimating first-year student attrition rates: An application of multilevel modelling using categorical variables. *Research in Higher Education*, *42*(2):151-170.

Phillips, M. 2000. Understanding ethnic differences in academic achievement: empirical lessons from national data. In *Analytic issues in the assessment of student achievement*. US Department of Education. 103–132.

Phillips, M., Crouse, J. & Ralph, J. 1998. Does the black-white test score gap widen after children enter school. In *The black-white test score gap*. C. Jencks & M. Phillips, Eds. Brookings Institution Press. 229–272.

Quaye, S.J., Harper, S.R. & Pendakur, S.L. eds., 2019. *Student engagement in higher education: Theoretical perspectives and practical approaches for diverse populations*. Routledge.

Reardon, S.F. 2013. The widening income achievement gap. *Educational Leadership*. 70(8):10–16.

Rojstaczer, S. 2016. Grade inflation at American colleges and universities. Available: www.gradeinflation.com (accessed on 2 November 2018).

Roser, M. & Ortiz-Ospina, E. 2013. *Tertiary education*. Available: https://ourworldindata.org/tertiary-education

Rosovsky, H. & Hartley, M. 2002. *Evaluation and the academy: are we doing the right thing*. Cambridge, MA: American Academy of Arts and Sciences.

Sampaio, G. 2012. Three essays on the economics of education. PhD Thesis. University of Illinois.

Schochet, P.Z. 2008. *Technical methods report: Statistical power for regression discontinuity designs in education evaluations*. Mathematica Policy Research.

Schoer, V., Ntuli, M., Rankin, N., Sebastiao, C. & Hunt, K. 2010. A blurred signal? The usefulness of National Senior Certificate (NSC) Mathematics marks as predictors of academic performance at university level. *Perspectives in Education*. 28(2): 9–18.

Scott, M. & Kennedy, B. 2005. Pitfalls in pathways: some perspectives on competing risks event history analysis in education research. *Journal of Educational and Behavioral Statistics*. 30(4):413–442.

Scott-Clayton, J.,= 2011. On money and motivation a quasi-experimental analysis of financial incentives for college achievement. *Journal of Human resources*, *46*(3):614-646.

Seaver, W.B. & Quarton, R.J. 1973. Social reinforcement of excellence: Dean’s List and academic achievement. Available: https://eric.ed.gov/?id=ED079644

Singer, J.D. & Willett, J.B. 1993. It’s about time: using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational Statistics*. 18(2):155–195.

Smith, L.C. 2012. The effect of selected academic development programmes on the academic performance of academic development students at a South African university: an empirical study.

Smith, J. & Naylor, R. 2001. Determinants of degree performance in UK universities: a statistical analysis of the 1993 student cohort. *oxford Bulletin of Economics and Statistics*, *63*(1):29-60.

Sonner, B.S. 2000. A is for "adjunct": examining grade inflation in higher education. *Journal of Education for Business*. 76(1):5–8.

Spaull, N. 2013a. *South Africa's education crisis: the quality of education in South Africa 1994–2011*. Centre for Development and Enterprise.

Spaull, N. 2013b. Poverty & privilege: primary school inequality in South Africa. *International Journal of Educational Development*. 33(5):436–447.

Spaull, N. 2015a. Education quality in South Africa and sub-Saharan Africa: an economic approach. Doctoral dissertation. Stellenbosch University.

Spaull, N. 2015b. Schooling in South Africa: how low-quality education becomes a poverty trap. *South African Child Gauge*. 12:34–41.

Statistics South Africa. 2019. *Education series volume V: higher education and skills in South Africa* (Report 92-01-05).

Statistics South Africa, 2022. Mid-year population estimates, South Africa. Pretoria: Government of South Africa.

Stinebrickner, R. & Stinebrickner, T. R. 2003. Understanding educational outcomes of students from low-income families: Evidence from a liberal arts college with a full-tuition subsidy. *The Journal of Human Resources*. 38(3):591–617.

Stinebrickner, T. & Stinebrickner, R. 2012. Learning about academic ability and the college dropout decision. *Journal of Labor Economics*, *30*(4):707-748.

Stinebrickner, R. & Stinebrickner, T. 2014. Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model. *Journal of Labor Economics*, *32*(3:601-644.

St. John, E., Cabrera, A., Nora, A. & Asker, E. 2000. Economic influences on persistence reconsidered: how can finance research inform the reconceptualization of persistence models. In *Reworking the student departure puzzle*. J. Braxton, Ed. Nashville, TN: Vanderbilt University Press. 29–47.

Stratton, L., O'Toole, D. & Wetzel, J. 2008. A multinomial logit model of college stopout and dropout behavior. *Economics of Education Review*. 27:319–331.

Thistlethwaite, D.L. & Campbell, D.T. 1960. Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, *51*(6):309.

Tinto, V. 1975. Dropout from higher education: a theoretical synthesis of recent research. *Review of Educational Research*. 45(1):89–125.

Tinto, V. 1993. *Leaving college: rethinking the causes and cures of student attrition*. 2nd ed. Chicago, IL: University of Chicago Press.

Vandenberghe, V. & Robin, S. 2004. Evaluating the effectiveness of private education across countries: a comparison of methods. *Labour economics*, *11*(4):487-506.

Van Broekhuizen, H. & Spaull, N. 2017. *The 'Martha Effect': The compounding female advantage in South African higher education* (No. 14/2017).

Vignoles, A.F. and Powdthavee, N., 2009. The socioeconomic gap in university dropouts. *The BE journal of economic analysis & policy*, *9*(1).

Wright, N. 2018. *Perform better, or else: academic probation, public praise and students decision-making*. Andrew Young School of Policy Studies Research Paper Series.

Appendix A

A1 National Senior Certificate (NSC) Achievement Criteria

Table 35 Matric exam achievement criteria

Achievement	Achievement Description	Marks %
7	Outstanding achievement	80 – 100
6	Meritorious achievement	70 – 79
5	Substantial achievement	60 – 69
4	Adequate achievement	50 – 59
3	Moderate achievement	40 – 49
2	Elementary achievement	30 – 39
1	Not achieved	0 – 29

Table 35 shows the matric (grade 12) achievement criteria. 50% is an achievement of 4 on the achievement scale of 1-7.

A2 Minimum requirements for admission to Higher Education on the NSC syllabus

1) Admission to Bachelor's degree

Students must obtain a minimum of 30% (achievement level 2 – elementary achievement) in the language of teaching and learning, in addition to a minimum of 50% (achievement level of 4) in four chosen subjects from the designated subject list:

Table 36 NSC Designated subject list

Designated Subject List
Accounting
Agricultural Science
Business Studies
Consumer Studies
Dramatic Arts
Economics
Engineering Graphics and Design
Geography
History
Information Technology
Languages
Life Sciences
Mathematics
Mathematical Literacy
Music
Physical Science
Religion Studies
Visual Arts

2) Admission to Diploma

Students must obtain a minimum of 30% (achievement level 2 – elementary achievement) in the language of learning and teaching, in addition to a minimum of 40% (achievement level 3) in four chosen subjects from the designated subject list.

3) Admission to Higher Certificate

Students must obtain a minimum of 30% (achievement level 2 – elementary achievement) in the language of learning and teaching.

Additional requirements may be stipulated by institutions and programmes.

A3 Power analysis

Table 37 Power analysis: full sample

	BW(CJM)			BW(2)		
	Below c	**Above c**	**Total**	**Below c**	**Above c**	**Total**
	5.179	6.643		2	2	
Graduation	337	467	804	245	561	806
Involuntary dropout	404	695	1099	617	1827	2444
Voluntary dropout	282	343	625	108	158	266
Annual non-cumulative GPA Yr 2	268	272	540	456	404	860
Annual non-cumulative GPA Yr 3	519	617	1136	500	651	1151
Annual non-cumulative GPA Yr 4	419	503	922	472	633	1105
Cumulative GPA Yr 2	106	93	199	360	335	695
Cumulative GPA Yr 3	172	176	348	449	503	952
Cumulative GPA Yr 4	182	209	391	484	556	1040

Table 37 shows the power analysis for the full samples used in Chapter 4. The CJM bandwidth is shown on the left and the narrowest bandwidth of ±2 percentage points around the cutoff is shown on the right.

Table 38 Power analysis: 3-year degree

3-year degree	BW(CJM)			BW(2)		
	Below c	**Above c**	**Total**	**Below c**	**Above c**	**Total**
	5.179	6.643		2	2	
Graduation	389	492	881	269	503	772
Involuntary dropout	561	591	1152	675	1605	2280
Voluntary dropout	283	265	548	90	91	181
Annual non-cumulative GPA Yr 2	478	565	1043	381	417	798
Annual non-cumulative GPA Yr 3	711	930	1641	531	690	1221
Annual non-cumulative GPA Yr 4	408	420	828	347	389	736
Cumulative GPA Yr 2	493	509	1002	300	302	602
Cumulative GPA Yr 3	652	694	1346	419	445	864
Cumulative GPA Yr 4	713	630	1343	500	429	929

Table 38 shows the power analysis for the 3-year degree sub-sample used in Chapter 4. The CJM bandwidth is shown on the left and the narrowest bandwidth of ±2 percentage points around the cutoff is shown on the right.

Table 39 Power analysis: 4-year degree

4-year degree	BW(CJM)			BW(2)		
	Below c	**Above c**	**Total**	**Below c**	**Above c**	**Total**
	5.179	6.643		2	2	
Graduation	240	383	623	146	337	483
Involuntary dropout	40	31	71	7	6	13
Voluntary dropout	317	514	831	172	402	574
			0			0
Annual non-cumulative GPA Yr 2	449	375	824	548	339	887
Annual non-cumulative GPA Yr 3	391	431	822	334	293	627
Annual non-cumulative GPA Yr 4	425	558	983	590	829	1419
			0			0
Cumulative GPA Yr 2	418	352	770	499	318	817
Cumulative GPA Yr 3	447	448	895	448	369	817
Cumulative GPA Yr 4	468	504	972	516	491	1007

Table 39 shows the power analysis for the 3-year degree sub-sample used in Chapter 4. The CJM bandwidth is shown on the left and the narrowest bandwidth of ±2 percentage points around the cutoff is shown on the right.

A4 Polynomial estimates on exit outcomes

Table 40 Impact of DML on exit outcomes: second order polynomial estimations (no covariates)

	(1)	(2)	(3)	(4)	(5)
Panel A: Graduation	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Robust	-0.043***	-0.057***	-0.074***	-0.085***	-0.112***
	(0.014)	(0.014)	-0.016	(0.016)	(0.019)
Conventional	-0.007	-0.015	-0.024	-0.045***	-0.062***
	(0.012)	(0.014	(0.016)	(0.016)	(0.019)
Observations	2153	1989	1686	1264	858
Panel B: Voluntary Dropout					
Robust	0.020*	0.023*	0.034**	0.037**	0.052***
	(0.011)	(0.017)	(0.014)	(0.016)	(0.018)
Conventional	0.002	0.005	0.008	0.020	0.024
	(0.011)	(0.013)	(0.014)	(0.016)	(0.018)
Observations	2153	1989	1686	1264	858
Panel C: Involuntary Dropout					
Robust	0.044***	0.052***	0.055***	0.066***	0.061***
	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
Conventional	0.033***	0.033***	0.040***	0.047***	0.060***
	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
Observations	2153	1989	1686	1264	858

Impact of the Dean's Merit List Policy: Exit outcomes with second order polynomials. Estimates are presented for the full sample. Standard errors are clustered along the running variable. * implies p value<0.1, ** implies p < 0.05, and *** implies p < 0.01. This specification does not include covariates. The cutoff has been recentered on zero.

Table 40 shows the impact of the DML on exit outcomes with no covariates and second order polynomials.

Table 41 Impact of DML on exit outcomes: second order polynomial estimations (with covariates)

	(1)	(2)	(3)	(4)	(5)
Panel A: Graduation	**BW=CJM**	**BW = ± 4.8**	**BW = ± 4**	**BW = ± 3**	**BW = ± 2**
Robust	-0.036**	-0.047**	-0.063***	-0.067***	-0.087***
	(0.017)	(0.019)	(0.020)	(0.021)	(0.023)
Conventional	-0.001	-0.010	-0.018	-0.036**	-0.046**
	(0.013)	(0.014)	(0.015)	(0.016)	(0.018)
Observations	2153	1989	1686	1264	858
Panel B: Voluntary Dropout					
Robust	0.019	0.021	0.031*	0.037**	0.054**
	(0.015)	(0.017)	(0.017)	(0.019)	(0.021)
Conventional	0.002	0.003	0.007	0.018	0.024
	(0.011)	(0.013)	(0.014)	(0.015)	(0.016)
Observations	2153	1989	1686	1264	858
Panel C: Involuntary Dropout					
Robust	0.016**	0.025***	0.032***	0.030***	0.033***
	(0.007)	(0.007)	(0.008)	(0.009)	(0.008)
Conventional	0.004	0.007	0.010*	0.018***	0.023***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.008)
Observations	2153	1989	1686	1264	858

Impact of the Dean's Merit List Policy: Exit outcomes with second order polynomials. Covariate-adjusted estimates are presented for the full sample. Standard errors are clustered along the running variable. * implies p value<0.1, ** implies p < 0.05, and *** implies p < 0.01. The cutoff has been recentered on zero.

Table 41 shows the impact of the DML on exit outcomes with covariates and second order polynomials.