# A hybrid simulation analysis of graduation success at Stellenbosch University

by

Sha-abaan Slamang

Thesis presented in fulfilment of the requirements for the degree of
**Master of Commerce (Quantitative Management)**
in the Faculty of Economic and Management Sciences at Stellenbosch University

Supervisor: Dr Lieschen Venter

December 2023

# Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2023

# Abstract

Stellenbosch University aims to be one of the leading research institutions on the African continent. High school matriculants (both locally and internationally) can apply to the university with the hope of graduating through one of its ten faculties. The selection of students for enrolment into a degree programme in the Economic Management Sciences Faculty is based on academic merit.

For a student to progress to each academic year, they must obtain the minimum academic credits required to continue. The aim of this research is to design and implement a hybrid simulation model of student progression using agent-based and systems dynamics modelling. This model is applied to analyse graduation success in the Faculty.

Following the validation of the hybridised student progression model, four intervention scenarios are tested for three degree programmes to increase the number of minimum time graduates. These intervention scenarios focus on a decrease on the influence of the perception of degree difficulty on graduate success, an increase in admission requirements, an increase in student engagement, and a combination approach where all three intervention scenarios are implemented.

In the calibrated model for Mathematical Sciences students it results that they are better assisted through higher selection criteria while Management and Economic Sciences students are better assisted through higher engagement. Both the calibrated and uncalibrated models are analysed to control the bias of overfitting. In the non-calibrated model, all students appear to be better assisted through the decreased perception of degree difficulty. The model use is illustrated by means of this specific case study, but it is able to assist decision support in multiple contexts wherever graduation success is a metric of interest.

# Opsomming

Stellenbosch Universiteit beoog om een van die voorste navorsingsinstansies op die vasteland van Afrika te wees. Hoërskoolmatrikulante (plaaslik sowel as internasionaal) kan by die universiteit aansoek doen met die hoop om deur een van die tien fakulteite te gradueer. Die keuring van studente vir inskrywing vir 'n graadprogram in die Fakulteit Ekonomiese Bestuurswetenskappe is gebaseer op akademiese meriete.

Vir 'n student om na elke akademiese jaar te vorder, moet hulle die minimum akademiese krediete verwerf wat nodig is om voort te gaan. Die doel van hierdie navorsing is om 'n hibriede simulasiemodel van studentedeurvloei te ontwerp en te implementeer deur gebruik te maak van agentgebaseerde en stelseldinamiese modellering. Hierdie model word toegepas om graduering-sukses in die Fakulteit te ontleed.

Ná die validering van die hibriede studentedeurvloeimodel word vier intervensie scenarios op drie graadprogramme getoets om die aantal gegradueerdes binne die minimum tyd te verhoog. Hierdie intervensie scenario's fokus op 'n afname in die invloed van die persepsie van hoe uitdagend die graadprogram is, 'n toename in toelatingsvereistes, 'n toename in studentebetrokkenheid, en 'n kombinasiebenadering waar al drie intervensie scenarios geïmplementeer word.

In die gekalibreerde model vir Wiskundige Wetenskappestudente blyk dit dat hulle beter bygestaan word deur hoër keuringskriteria, terwyl Bestuurs- en Ekonomiese Wetenskappestudente beter bygestaan word deur hoër betrokkenheid. Beide die gekalibreerde en ongekalibreerde modelle word ontleed om te beheer vir die vooroordeel van data-oorpassing. In die ongekalibreerde model blyk dit dat alle studente beter gehelp word deur die verminderde persepsie van die moeilikheidsgraad van die kwalifikasies. Die modelgebruik word deur middel van hierdie spesifieke gevallestudie geïllustreer, maar is daartoe in staat om besluitnemingsondersteuning in verskeie kontekste te bied, waar gradueringsukses 'n maatstaf van belang is.

# Acknowledgements

The author wishes to acknowledge the following people for their various contributions towards the completion of this work:

- Dr Lieschen Venter for her guidance, insights, advice, always willing to accommodate my crazy hours, dyslexic thought patterns, and always going above and beyond to support me.
- My family for their endless support, late nights, and snacks.
- Loumarie Kistner for the many long hours, labour intensive data extracts, and willingness to assist.
- Dr Gert Young and Hamman Schoonwinkel for their hard work and persistence in providing data.
- SCITYLANA x KO for being my yellow duckies (so called rubber ducks) when I needed to debug my code.
- My colleagues at the Shoprite Group, and my manager in particular, for the time off and accepting increased work duties to complete my research.
- My fellow Masters students from the Postgraduate Lab and staff at the Department of Logistics for their pleasant company and many coffee breaks.
- Thank you Allah for blessing me.

# Contents

**List of Acronyms** **xi**

**1 Introduction** **1**

1.1 Problem description . . . 3

1.2 Objectives and scope . . . 4

**2 Literature review** **7**

2.1 Determining causality in complex systems . . . 7

2.2 Analyses of the South African basic education system . . . 8

2.3 Analyses of the South African higher education system . . . 10

2.4 Methodology . . . 11

2.4.1 System dynamics simulation modelling . . . 11

2.4.2 Agent-based simulation modelling . . . 12

2.5 Chapter summary . . . 13

**3 Determinants for timely graduation in higher education systems** **15**

3.1 Academic preparedness . . . 15

3.2 Socio-economic status and financial stability . . . 17

3.3 Family support and educational background . . . 19

3.4 Institutional factors and resources . . . 20

3.5 Student motivation and engagement . . . 21

3.6 Study habits and time management skills . . . 23

3.7 Social integration . . . 24

3.8 Choice of major and course load . . . 25

3.9 Determinant feedback in higher education systems . . . 26

3.10 Chapter summary . . . 28

**4 Data and assumptions** **29**

4.1 Data preparation for student progression . . . 29

4.2 Data preparation for systemic support of student progression . . . . . . . . . . . 32
4.3 Chapter summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 36

**5 Model formulation** **37**
5.1 Simulating student progression . . . . . . . . . . . . . . . . . . . . . . . . . . . 37
5.1.1 Agent attributes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 38
5.1.2 The state chart and state transitions . . . . . . . . . . . . . . . . . . . . . 39
5.1.3 Scripting and implementation . . . . . . . . . . . . . . . . . . . . . . . . 41
5.2 Simulating system support . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46
5.2.1 System formulation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46
5.3 Hybridisation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 47
5.4 Chapter summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 48

**6 Validation and results** **49**
6.1 Validation of student progression in equilibrium . . . . . . . . . . . . . . . . . . 49
6.1.1 Behavioural replication test . . . . . . . . . . . . . . . . . . . . . . . . . . 49
6.1.2 Structural validation test . . . . . . . . . . . . . . . . . . . . . . . . . . . 50
6.2 Validation of systemic support in equilibrium . . . . . . . . . . . . . . . . . . . 51
6.2.1 Weight calibration . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 51
6.2.2 Behavioural replication test . . . . . . . . . . . . . . . . . . . . . . . . . . 52
6.3 Validation of the dynamic hybrid student progression model . . . . . . . . . . . 54
6.3.1 Behavioural replication test . . . . . . . . . . . . . . . . . . . . . . . . . . 54
6.3.2 Extreme conditions tests . . . . . . . . . . . . . . . . . . . . . . . . . . . 54
6.4 Sensitivity analysis . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 55
6.5 Results and interventions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 56
6.5.1 The influence of decreased degree difficulty . . . . . . . . . . . . . . . . . 57
6.5.2 The influence of increased academic preparedness . . . . . . . . . . . . . . 58
6.5.3 The influence of increased student engagement . . . . . . . . . . . . . . . 59
6.5.4 The influence of a combined intervention . . . . . . . . . . . . . . . . . . 60
6.5.5 Summary of results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 61
6.6 Interventions at equilibrium . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 61
6.6.1 The influence of decreased degree difficulty . . . . . . . . . . . . . . . . . 62
6.6.2 The influence of increased academic preparedness . . . . . . . . . . . . . . 63
6.6.3 The influence of increased student engagement . . . . . . . . . . . . . . . 64
6.6.4 The influence of a combined intervention . . . . . . . . . . . . . . . . . . 65
6.6.5 Summary of results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 65

6.7 Chapter summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 66

**7 Conclusion** **67**

7.1 Summary of analysis . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 67

7.2 Summary of results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 68

7.3 Achievement of objectives . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 68

7.4 Future work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 68

**Appendix** **71**

**Bibliography** **71**

**A Sensitivity analysis** **77**

# List of Figures


3.1 A generic causal loop diagram for higher education graduation success. . . . . . . 26

3.2 The main feedback loops in the generic higher education system. . . . . . . . . . 27

5.1 The state chart for agents within the SPM. . . . . . . . . . . . . . . . . . . . . . . 41

5.2 SPM stock-flow diagramme . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46

6.1 The percentage of minimum time graduates in the student progression submodel. 50

6.2 The percentage of minimum time graduates in the systemic support submodel. . 53

6.3 The percentage of minimum time graduates in the hybridised student progression model. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 54

6.4 Extreme conditions testing for minimum and maximum goodness . . . . . . . . . 55

6.5 Systemic support strengths at the status quo. . . . . . . . . . . . . . . . . . . . . 56

6.6 Base case results for graduate success in the three degree programmes at the status quo. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 56

6.7 Systemic support strengths for a decreased degree difficulty intervention. . . . . . 57

6.8 Minimum time graduates for a decreased degree difficulty intervention. . . . . . . 57

6.9 Systemic support strengths for an increased academic preparedness intervention. 58

6.10 Minimum time graduates for an increased academic preparedness intervention. . 58

6.11 Systemic support strengths for an increased student engagement intervention. . . 59

6.12 Minimum time graduates for an increased student engagement intervention. . . . 59

6.13 Systemic support strengths for the combination intervention. . . . . . . . . . . . 60

6.14 Minimum time graduates for the combination intervention. . . . . . . . . . . . . 60

6.15 Systemic support strengths for a decreased degree difficulty intervention. . . . . . 62

6.16 Minimum time graduates for a decreased degree difficulty intervention. . . . . . . 62

6.17 Systemic support strengths for an increased academic preparedness intervention. 63

6.18 Minimum time graduates for an increased academic preparedness intervention. . 63

6.19 Systemic support strengths for an increased student engagement intervention. . . 64

6.20 Minimum time graduates for an increased student engagement intervention. . . . 64

6.21 Systemic support strengths for the combination intervention. . . . . . . . . . . . 65

6.22 Minimum time graduates for the combination intervention. . . . . . . . . . . . . . 65

A.1 Systemic support strengths for changes to the weight of engagement. . . . . . . . . 77
A.2 Systemic support strengths for changes to the weight of academic preparedness. . 78
A.3 Systemic support strengths for changes to the weight of degree difficulty. . . . . . 78

# List of Tables

1.1 An example of module requirement types at the EMS Faculty of SU. . . . . . . . 2

4.1 A sample of the raw data to populate the SPM for selected focal areas. . . . . . 29
4.2 A sample from the module dimension table. . . . . . . . . . . . . . . . . . . . . . . . . 30
4.3 A sample from the module combination table. . . . . . . . . . . . . . . . . . . . . . . 31
4.4 A sample from the prerequisite dimension table. . . . . . . . . . . . . . . . . . . . . 31
4.5 A sample of the input data to calculate the triangular distributions of pass probabilities. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32
4.6 A sample from the raw data transcription of the SFF. . . . . . . . . . . . . . . . 32
4.7 A sample from the captured student feedback data. . . . . . . . . . . . . . . . . . 33
4.8 A sample from the module dimension table for the SFF. . . . . . . . . . . . . . 33
4.9 A sample from the school naming dimension table. . . . . . . . . . . . . . . . . . 34
4.10 The initial values for the determinants to describe systemic support . . . . . . . 36

5.1 The different AnyLogic features used to generate agents and simulate agents within the SPM. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37
5.2 A summary of the AnyLogic features used to generate agents and simulate agents within the SPM. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 40
5.3 The hybridisation table function for the SPM. . . . . . . . . . . . . . . . . . . . . 47

6.1 A sample of the actual and simulated final achievement marks for BCom (Management Sciences) with a RMSE of 6.1%. . . . . . . . . . . . . . . . . . . . . . . . 51
6.2 The calibrated weights for the determinants of graduate success in the systemic support submodel. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 53
6.3 The average change in graduate success per interval change in the determinant weights for intervals of $0.1 \in [0.1]$. . . . . . . . . . . . . . . . . . . . . . . . . . . 55
6.4 The effect of each intervention strategy on the percentage of minimum time graduates for the weight calibrated SPM. . . . . . . . . . . . . . . . . . . . . . . . . . 61
6.5 The effect of each intervention strategy on the percentage of minimum time graduates for the weight uncalibrated SPM. . . . . . . . . . . . . . . . . . . . . . . . 66

# List of Acronyms

**ASCS:** Academic Self-Concept Scale

**ABS:** Agent-based simulation

**Bacc:** Bachelor of Accounting

**BCom:** Bachelor of Commerce

**LLB:** Bachelor of Law

**BCIA:** Black, Colour, India, and Asian

**BI:** Business Intelligence

**CLD:** Causal loop diagramme

**CLT:** Central Limit Theorem

**SU LaunchLab:** Centre for Student Entrepreneurship

**CECS:** College of Engineering and Computer Sciences

**CIRP:** Cooperative Institutional Research Programme

**C:** Corequisite module

**CAPS:** Curriculum Assessment Policy Statements

**DBE:** Department of Basic Education

**DHET:** Department of Higher Learning and Training

**ECD:** Early Childhood Development

**EMS:** Economic and Management Sciences

**EDP:** Extended Degree Programme

**FGCS:** First-Generation College Student

**GPA:** Grade point average

**HE:** Higher education

**HEMIS:** Higher Education Information Management System

**IE:** Industrial Engineering

**IG:** Information Governance

**IS:** Information Systems

**IPEDS:** Integrated Postsecondary Education Data System

**ILS:** Inventory of Learning Styles

**LASSI:** Learning and Study Skills Inventory

**LM:** Logistics Management

**MIT:** Massachusetts Institute of Technology

**MSLQ:** Motivated Strategies for Learning Questionnaire

**NA:** Not applicable

**NSC:** National Senior Certificate

**OLDA:** Ohio Longitudinal Data Archive

**OR:** Operations Research

**P:** Prerequisite module

**PP:** Prerequisite pass module

**PMT:** Prestatie-Motivatie Test

**PSO:** Private Student Organisations

**RMSE:** Root Mean Square Error

**SU:** Stellenbosch University

**SEM:** Structural Equation Modeling

**SFF:** Student Feedback Form

**SPM:** Student Progression Model

**SHSA:** Study Habits, Skills, and Attitudes

**SSHA:** Survey Of Study Habits and Attitude

**ETH Zurich:** Swiss Federal Institute of Technology in Zurich

**SD:** System Dynamics

**TE:** Transport Economics

**UCF:** University of Central Florida

CHAPTER 1

# Introduction

Stellenbosch University (SU) aims to be one of the leading research institutions on the African continent. SU has expanded its offering from the 1918 offering where students could enrol into one of four faculties, namely Arts, Science, Education and Agriculture. In 2023, SU has ten faculties spread across five campuses. These faculties are AgriSciences, Arts and Social Sciences, Economic and Management Sciences, Education, Engineering, Law, Medicine and Health Sciences, Military Science, Science, and Theology [59].

High school matriculants (both locally and internationally) can apply to SU with the hope of graduating through one of SU's ten faculties. Learners must meet the minimum admission criteria as set out by each faculty to apply. To be considered for selection for an undergraduate programme at the Faculty of Economic and Management Sciences (EMS), learners are required to meet the minimum standard of 50% in Language or 60% in their First Additional Language when obtaining their National Senior Certificate (NSC). In general for any degree programme within the Faculty, excluding the Extended Degree Programme (EDP), a minimum of 60% in Mathematics and 65% NSC average is required. Some degree programmes require a higher or lower Mathematics and NSC average [54].

Admission into a degree programme is subject to the number of students admitted in the Faculty which is limited by the enrolment capacity. The selection for a degree programme is based on academic merit. A prospective student's academic merit is considered in two categories. Firstly it is determined by whether they have met the minimum admission requirements as set out in the Faculty yearbook published annually. In the yearbook for the EMS Faculty an introduction of the minimum requirement for admission, selection and other degree specific information is published [54]. For the second selection category, a prospective student's selection mark is calculated. The selection mark is the average achievement mark of the prospective student's Grade 11 subjects excluding life orientation. If a prospective student completed more than six subjects, the six highest marks are used to compute the selection mark. If a prospective student completed fewer than six subjects, at least four subject marks are used to calculate the selection mark. Mathematics and either English or Afrikaans must be among the subjects used to determine a prospective student's selection mark.

After completion of the selection process, a prospective student receives feedback on their application. If they have not met all of the requirements they are notified that their application was unsuccessful. Selected degree and diploma programmes have a limit to the number of students that may be provisionally admitted and this limit is unique for each degree programme. General programmes such as BCom (Management Sciences) and BCom (Economic Sciences) have no restriction on the number of students admitted additional to the the general upper bound on all students. The Faculty notifies students of provisional admission by the end of June or July preceding registration for general programmes while feedback for programmes with additional restrictions on the number of admissions are communicated by the end of September of the year preceding registration.

For a student to progress to graduation they must obtain the minimum academic credits required to continue in the programme. The credit system employed at SU is the Higher Education Information Management System (HEMIS). An undergraduate student is required to pass at least 50% of their total credits during each year to continue in their programme. To pass a module and meet the HEMIS requirements, a student must achieve a final achievement mark of 50% per module. The final mark is calculated when a student completes a minimum of two main assessments for a semester module and a minimum of three main assessments for a year module [55]. A student is eligible for graduation when they have obtained the minimum number of credits required for their degree, subject to the student completing the minimum number of credits and modules for each year in their study.

A student may fail a module but still be able to enroll for succeeding modules. This depends on whether a module is a prerequisite pass module, a prerequisite module or a co-requisite module. A module is a prerequisite pass module when the student is required to pass it as a preceding module(s) with a final achievement mark of at least 50% to be able to enrol for a new module(s). A module is a prerequisite module when the student fails the module but still obtains a final mark of at least 40% which allows them to enrol for a new module(s). A module is a co-requisite module when a student must take it as a required module in the same academic year to enrol for a module(s) in the current year. In Table 1 an example of each type of module requirement is shown.

| **Type** | **Selected module** | **Required module(s)** |
|---|---|---|
| Prerequisite pass module | Economics 214 | Economics 114 & 144 |
| Prerequisite module | Operations Research 314 | Operations Research 244 |
| Corequisite module | Law of Taxation 411 | Mercantile Law 471 |

TABLE 1.1: *An example of module requirement types at the EMS Faculty of SU.*

An undergraduate student is a student who is enrolled into a bachelor's degree programme or a diploma. A bachelor's degree takes three to five years to obtain depending on the faculty and degree for which the student is enrolled. A postgraduate student is a student who decides to further their education by completion of a postgraduate diploma, honours degree, masters degree or doctorate [54].

At SU an undergraduate student may repeat a module if they did not achieve a final achievement mark of at least 50%. An alternative reason for repeating a module could be that the student wishes to improve their final mark to qualify for a postgraduate degree. When a student repeats a module their final achievement mark is updated to reflect the second attempt final achievement mark. If a student does not pass a module and their degree programme and another module choice does not require this module, they may simply disgard the module and select a different one [55].

During enrolment certain degree programmes require students to select a range of modules. These modules are known as electives. Certain degree options do not allow students to select electives. These degrees have a set structure which students must follow while others allow students to choose a minimum number of electives from a list of first-year modules. These electives assist students to better prepare for the specialist area they would like to focus on later in the degree programme. In their second year students select an focal area along with three other elective modules. These electives modules and focal area limits the variety of the electives in their third year.

Depending on the degree programme a student might be required to enroll for modules across multiple faculties. These programmes are often a combination between two degrees. For example, a combination between a Bachelor of Accounting (BAcc) and a Bachelor of Law (LLB) is combined into a BAccLLB programme. This combination allows the student to consider a career as either a professional attorney or accountant.

In 2022 SU had a total of 32 535 students enrolled into various degree programmes. Of the 32 535 students enrolled, 51.6% were White students, 23.3% were Black African students, 17.7% were Coloured students, 3.4% were Indian students, 0.4% were Asian students and 3.6% of students withheld information regarding their race [53]. During 2022, SU had the majority of its student body (*i.e.,* 56.4%) identify as female. The Management Sciences student body comprised 27.7% of the total student body.

The average module pass rate at SU has been above 85% since 2013. SU has shown the highest throughput rate of 84% compared to the average of 68% for other higher education and training instutions in South Africa [53]. In 2019, the EMS Faculty had an average module pass rate of 82.2% for first years, 82.7% for second years. and 83.7% for third years. This gives the EMS faculty an average throughput rate of 82.9% in 2019 [53].

South Africa's youth unemployment for young adults between the ages of 15 to 24 years old has risen to 61% with an increasing trend since 2019 [17]. At the end of 2022, youth unemployment rates were 7.8% for White adults, 36.8% were Black African adults, 26.5% were Coloured adults, and 13.7% were Indian adults [52]. Youth unemployment is not only a problem for a specific gender, the male and female youth unemployment rates were 32.8% and 36.4%, respectively [51].

Graduate unemployment is a contentious issue in South Africa. These issues include the quality of schooling, the schooling curricula, university access, university curricula, emigration, immigration restrictions, and the nature of the South African economic growth [67]. Graduate unemployment comprised 2.7% of total unemployment while matriculant unemployment comprised 38.2% and school leavers unemployment comprised 51.5%, respectively. The remainder is comprised of unemployed adults who hold some other tertiary qualification (7.6%) [50]. Van der Berg and Van Broekhuizen [67] explained that employers prefer to hire university graduates as they are better equipped for the labour market. The ManpowerGroup reports a labour shortage of about 34% in filled positions in the South African labour market. In their top ten skills shortages to fill these positions include engineers, chartered accountants, auditors, financial analysts, teachers and cybersecurity experts [29].

## 1.1 Problem description

The minimum degree span is the duration of a degree programme set out by the Department of Higher Learning and Training (DHET). If a student is unable to pass all the required modules within a given year it could lead to an increase in the time a student requires to complete a degree programme. In the EMS faculty the average undergraduate degree programme is three years with some exceptions requiring a longer duration of four years. This minimum duration of a degree programme is also linked to the number of credits that a student must obtain to complete the programme.

South Africa's history of racial segregation in combination with shortages of skilled individuals within the South African labour market necessitates an analysis as to how SU can increase their graduate throughput to relieve the skill shortages. SU has implemented several initiatives to address youth unemployment by for example, the establishment of the Centre for Student Entrepreneurship (SU LaunchLab) to provide graduates with opportunities to gain work experience and develop their skills.

The aim of this study is to construct a computer simulation model by means of a systems analysis approach to simulate the progression of students within the EMS Faculty of SU. Such a model can allow policy makers to understand the complex interaction between the factors that influence student progression at SU. Through this analysis systemic reasons as to why students are unsuccessful to complete their degree within the minimum degree span can become apparent. This analysis will assist in adapting institutional strategies and practices to better support graduation within the minimum degree duration.

To better understand student progression three research questions are analysed. The first research question is which systemic factors hinder graduation within the minimum-time? The second research question is which policies holistically contribute to achieving an increase in the number of minimum-time graduates within the EMS Faculty? A question arises as the possible presence of a counter-intuitive causality where degree difficulty might not be the most influential determinant of the number of minimum-time graduates for disparate sciences.

Three different degree programmes are analysed. Graduation success for the bachelor's degrees for Management Sciences, Economic Sciences, and Mathematical Sciences are compared. These three degree programmes were selected as they represent the three main streams where there are only compulsory modules and elective modules. These three degree programmes also contain modules from each department within the EMS Faculty where students were not limited to a specialisation area such as Law or Accounting.

## 1.2 Objectives and scope

The goal of this thesis is to construct a computer simulation model of the EMS Faculty of SU to analyse student progression and the systemic factors that hinder the successful graduation of the student body in the minimum-time. The following objectives are pursued:

**Objective I:** Perform a thorough literature study to

- summarise existing mathematical and simulation models of higher education systems;
- introduce the simulation paradigms employed in the model. The methods under consideration are agent-based simulation modelling, system dynamics modelling, and a hybrid approach; and
- determine a causal-loop structure for the generic higher learning system from significant determinants reported in literature.

**Objective II:** Collect and process institutional datasets to

- identify trends in this specific higher education system;
- quantify factors that influence the academic performance and progression of EMS students; and
- determine initial values and rates for student progression.

**Objective III:** Construct a computer simulation model using the processed data from Objective II. The model is to be constructed in the AnyLogic simulation package [3]. The aim of the model is to replicate the current student progression in the EMS to analyse different intervention strategies so as to increase the number of EMS students graduating in minimum-time.

**Objective IV:** Achieve validation of the model to determine if it represents the observed system adequately by means of the validation methods suggested by Forrester and Senge [20] and Shreckengost [48]. The parameters in the model are to be tested to determine the sensitivity of the model to different initial values.

**Objective V:** Illustrate the results of the simulation and the effects of different intervention strategies. Various intervention strategies are to be tested to determine which intervention(s) should be implemented to improve the student progression and their magnitude of effect over time.

**Objective VI:** Summarise the findings in a formal report document and include future work for potential further development.

CHAPTER 2

# Literature review

In this chapter the contributions of important authors on modelling causality and higher education systems are listed. The problem background is further informed by means of a summary of previous attempts to solve it. Lastly, a primer on the computer simulation paradigms employed in this study is also offered.

## 2.1 Determining causality in complex systems

Determining causality in complex systems is a challenging task but it is an important concept in many disciplines from philosophy to econometrics. There are a multitude of researchers focused on understanding and determining causality.

Causality (also called cause and effect) refers to the influence that a singular event, process, state or object can have on another where one of these are partially responsible for another and is also dependent on the other [27]. For example, a generalisation of causality is that smoking causes lung cancer when it is perhaps more accurate to state that there exists some relationship between smoking and lung cancer [23].

Coffman [11] discussed the challenges of establishing causality in nonlinear complex systems. Coffman emphasised that causality in complex systems is not straight-forward and cannot be understood through linear cause and effect relationships but should instead be analysed through interactions and feedback loops via multiple variables and factors. He argued that determining causality in nonlinear complex systems required a different approach that accommodated for the nonlinearity, interdependencies, and time delay. He suggested that the use of developmentalism over reductionism methodology was an improvement in both natural and social sciences.

Galea *et al.* [22] explored the role of causal thinking and complex system approaches in the field of epidemiology. They highlighted the limitations of traditional epidemiological approaches that often focus on individual risk factors and linear causal relationships. These traditional approaches fail to show the interconnectedness and the influences that each factor had on the outcomes. They proposed the adaption of system dynamic modelling for its ability to capture nonlinearity, feedback loops, emergence, and self-organisation. They also conceded, however, that causality was challenging to determine within complex systems.

Angrist and Pischke [2] provided a comprehensive guide to applied econometrics with a focus on causal inference to better equip researchers with the tools and methodologies to draw reliable causal conclusions from empirical data. They highlighted the importance of causal inference in economics and the challenges involved in establishing causality. They provided a detailed overview of a variety of approaches to determine causality such as instrumental variables, regression discontinuity designs, and difference-in-differences methods.

Determining causality in complex systems is an important and difficult, but not an impossible task. An argument for causality in a system might at least reasonably be defended through a thorough overview of system determinants and their relationships recorded in the literature. Khan *et al.* [30] found that through a systemic review of both meta-analytic and criteria-based method can be used to determine casual inference.

## 2.2 Analyses of the South African basic education system

Gustafsson and Taylor [24] explored the influence of improving school administration on basic education outcomes in South Africa. They analysed metrics describing Grade 12 learners from 1 005 schools in three South African provinces and conducted surveys to better understand the school conditions, leadership, teacher qualifications, and resources. They found that schools with better administration tended to have higher learner achievement scores and lower rates of grade repetition. The schools were measured by the quality of their communication channels, effective use of data, and adequate support for school leaders. Gustafsson and Taylor concluded that improving the quality of administration at a school would be a cost-effective strategy for improving its outcomes. These improvements should focus on building the capacity of administrators and providing them with the necessary resources and support to effectively manage their schools.

Spaull and Taylor [49] developed a measure for education success which included both a metric for its quantity and its quality. Educational quantity refers to the measure of access that learners had to education while educational quality refers to the measure of learning they achieved as a result of this access. Their measure combined multiple data sources into one unified data source. This data included measures for schooling, enrolment rates, matriculation rates, attrition rates, and learning outcomes sourced from household surveys, administrative records, and standardised tests for each of the eleven countries selected for their study.

They found that some of the countries with high enrolment rates had low levels of educational quality, suggesting that enrolment rates alone were not sufficient for measuring success. They also found that household income, urbanisation, and gender were factors that influenced educational access. They suggested that policymakers should take these factors into consideration when designing and implementing policies and interventions.

Venter and Passerini [72] constructed an agent-based simulation model of South Africa's early childhood development (ECD) system to analyse the factors that influenced a child's readiness for primary school. An agent in the model was characterised by their age, family type, health status and ECD enrolment. An agent existed in one of the three states: Achieving the required development standard, falling behind, or being at risk. Similar as for Venter's system dynamics simulation modelling of ECD systems [75], Venter and Passerini found that bimodality in cognitive development outcomes were perpetuated between the socio-economic systems for children in Quintiles 1 to 3 (*i.e.,* the quintiles of lower socio-economic status) and Quintiles 4 to 5 (*i.e.,* the quintiles of higher socio-economic status). They analysed three interventions strategies in an attempt to bridge the inequality between these two systems. The first intervention was an improvement in the health status of the children. The second intervention was to improve the family structure in which the children were raised where the ideal structure is a nuclear family which consists of two parents and the child(ren). The third intervention was to improve the quality of the ECD programmes so to improve the children's level of school readiness. They concluded that a continuous combination of all three of these intervention strategies yielded the most success in bridging the gap between the two systems but that more data were required to attribute a larger set of characteristics to the child agents for an improved analysis.

Venter [69] constructed a system dynamics simulation model of public primary schools in the Western Cape. She focussed on uncovering the factors that influenced the academic success of learners in Grades 1 to 7. These factors were measures for school resources, teacher effectiveness, engaged time, school readiness, academic performance, learner motivation, family support, and family poverty. She analysed the effect of these factors on learners' literacy and numeracy outcomes. She again found that bimodality was present between the different socio-economic systems for learners in Quintiles 1 to 3 and Quintiles 4 to 5. To mitigate this bimodality she explored three interventions. The first intervention was an improvement in the effectiveness of the time learners spent in the classroom. The second intervention was to improve the quality of the time learners spent outside of the classroom through their social and family interaction. The final intervention was to improve both the learners' school environment as well as the learners' social environment at home. She found that to improve literacy and numeracy performance, interventions to improve home circumstances had the greatest influence. Venter defined home interventions as regular social, psychological and academic support from parents and the community.

Perrie and Searle [40] expanded on Venter's systems dynamics simulation model of the primary school system [69]. They constructed an agent-based model of the primary school system for Grade 1 to 4 learners focusing on learner progression. Their analysis extended to finding a suitable data mining approach to select the datasets for the model. This step was required as no existing datasets were available that provided all the necessary data for their modelling approach. They used logistic regression to approximate the probabilities associated with the various states in the model. An agent in the model was characterised by their age, family type, health status, gender, their last completed school grade and the parents' level of education. An agent would fall into one of the six states: Exceeding the academic standard, achieving the standard, falling one year behind, falling two years behind, falling three years behind, or falling four or more years behind the standard. Their model is technically and operationally feasible but due the lack of available data, an accurate representation of the Quintile 1 to 5 systems could again not be achieved by means of the agent-based simulation paradigm.

Venter and Slamang [74] used Venter's simulation of the South African primary school system and adapted it for the high school system within the Western Cape. They constructed a system dynamics simulation model of public high school management within the Western Cape where they analysed the factors that influenced the progression of learners in Grades 7, 9, and 12. The factors they analysed where measures for school resources, teacher effectiveness, engaged time, school readiness, learner motivation, learner attrition, class size, social support, and family poverty. Similar to Venter's primary school model, they found that bimodality existed between the two socio-economic systems for learners in Quintiles 1 to 3 and Quintiles 4 to 5. In an attempt to minimise this inequality they explored four intervention strategies. The first intervention was an increase in the effectiveness of the time the learner spent in the classroom to improve their academic performance. The second intervention was an increase in the number of educators for schools in socio-economic Quintiles 1 to 3 in combination with the first intervention. The third intervention was an improvement in learners' social circumstances outside of the classroom. The final intervention focussed on changing the learners' social circumstances and decreasing the learner's family poverty. They found that no intervention strategy significantly improved performance of the Quintile 1 to 3 system but that a combination of interventions produced a slight improvement, especially when an intervention focussed on improving the metric for school resources. They note that improved data collection was required to increase confidence in the intervention results.

## 2.3 Analyses of the South African higher education system

Venter and Rantsane [73] attempted a system dynamics simulation of student progression within the BAcc programme at SU. The BAcc programme was selected as a first attempt simulation of progression from enrolment to graduation without complicated elective module selection. The model simulated progression and graduation rates of students by race and gender. The model depended on broad and unclear assumptions which may have produced misleading results but is a necessary first analysis of the complexity in simulation student progression in higher education systems.

Venter and Jansen van Rensburg [71] used this model as a stepping stone to improve their approach of simulating student progression. They adapted the model so that stocks represented the historic year for which a student was enrolled instead of representing the specific modules. A stock is an element that collects items that accumulate or drain. They included stock dimensions to represent the students' academic year, whether they repeated any year, and the number of times they repeated. The aim of the model was to increase the graduation rates of BCIA students enrolled in the EMS Faculty of SU. They considered three intervention strategies to improve the diversity of students enrolled as well as the graduation rates of each group of students. The first intervention sought to increase the attainment of more high performance BCIA students. The second intervention explored an increase in the success of first year BCIA first year students. The final intervention was to decrease repetition rates of BCIA students. They concluded that the first intervention strategy had the most success in improving the diversity of graduate cohorts within the BAcc programme, subject to their assumptions and model initialisation data.

Venter and Grobler [70] analysed the population of teacher students graduating to become teachers in the basic education system of South Africa. Their system dynamics model analysed the relationships between the various factors impacting the training of new teachers. Some of the factors that they explored were graduate emigration, teacher's salary, student motivation and policy review. Teacher students were grouped either as a student enrolled into a higher ranked institution or a student enrolled into a lower ranked institution. This distinction between a high- and low-ranking institution was based on the institution's internationally ranked academic performance.

Venter and Grobler simulated three intervention strategies to improve and maintain a high-quality teacher graduate. The first intervention focussed on improving the ranking of the institutions a student was enrolled at to increase the number of high-quality teachers. The second intervention was to increase the teacher starting salary to motivate students to graduate in minimum time and to retain existing educators. The last intervention analysed the influence of student motivation and their likelihood of attrition or change of career paths. They concluded that the most effective way to satisfy the demand for high-quality teachers without causing an over or under supply in the market was to improve the quality of teacher training programmes at the largest university in South Africa.

Bostwick *et al.* [6] explored the long-standing debate in higher education of whether offering modules by semesters or by quarters led to better student performance. They analysed data from the Integrated Postsecondary Education Data System (IPEDS) and the Ohio Longitudinal Data Archive (OLDA) for four-year American colleges and universities degrees/courses. They found that colleges and universities that converted from quarters to semesters decreased the probability of a student graduating on time. They concluded that longer semesters caused students to enrol for more courses per term. This affected the success of students due to an increase in workload as well as delaying the selection of a student's major.

Robledo [45] dissertated a dynamic enrolment simulation model for the University of Central Florida (UCF) for planning and decision-making. This model was used as a decision support tool to support the university management by using a hybrid simulation approach using system dynamics, agent-based simulation, and discrete event simulation modelling.

The model was constructed on two main levels. The first was a high level abstraction from a macro level making use of system dynamics to illustrate the overarching enrolment process of the UCF. The stocks represented the different academic years while the flows simulated the progression from one year to another and attrit of each year influenced by the attrition and success rates.

The second level was low level abstraction from a micro level, which was modelled through the agent-based and discrete event paradigms. This level simulated students who declared Industrial Engineering (IE) has their major at the start of their university career compared to those who had not declared their major yet. The model also considered the students enrolled in IE courses but have not declared IE as their major. The student agents' states were their academic year, major declaration (*i.e.,* IE, undeclared, other), and the probability of attrition or new major change.

To bridge the gap between the higher and lower level a medium transitory level was constructed. This level focussed on the IE department and creating a level between the systems model and the agent-based and discrete event model. The stocks represented the year in which the student was enrolled for with flows of College of Engineering and Computer Sciences (CECS) students. The combination of all levels result in the measurement of a retention level, a students to faculty ratio, and academic workload and growth rates.

Robledo analysed the student to faculty ratio to predict the increase of faculty members needed for 2015. He determined that an increase of 43% is required to remain in line with the recommended ratio of one educator (faculty member) for every thirty students.

## 2.4 Methodology

The simulation techniques considered for this thesis are systems dynamics and agent-based simulation. These paradigms are combined to construct a hybrid simulation model by which the research questions may be answered.

### 2.4.1 System dynamics simulation modelling

System dynamics (SD) is a simulation technique to analyse different outcomes at a macro perspective [69]. SD essentially simulates a collection of variables that interact with each other over time to form a unified whole. To simulate behaviour an SD approach makes use of computer simulation which imitates the behaviour of a system of the physical reality through automated computation [21].

Four structures are considered in SD. The first is a stock. A stock is an element that collects items that accumulate or drain. For example, a high school accumulates learners as they enrol for Grade 8 and drains learners as they matriculate Grade 12. The second element is a flow. A flow is the rate at which a stock changes over time. This change can be the accumulation or drainage. For example, the rate at which learners pass Grade 12 in 2019 was 39.4% [58]. The third element is a converter. A converter contains logic to convert input data and manipulate it to return an output value. For example, if the enrolment rate increases by 10% per year, then the flow is multiplied by the converter which is set to 10%. The final element is a connector. A connector passes the information between the converters, flows and stocks.

A SD model cannot function without feedback as a model without feedback is simply static. For example, information about a problem leads to action which leads to a result, but this is incomplete. In reality there are different complexities, reactions and influences before a decision is made. There are two types of feedback loops, namely positive and negative, also referred to as reinforcing or balancing. The feedback loops reference the behaviour of a system. A positive feedback loop increases the behaviour in the same direction to produce reinforcing behaviour. A negative feedback loop on the other hand references behaviour in the opposite direction which produces balancing or stabilising behaviour [21].

SD makes use of difference equations and delta time [77] where difference equations are discrete approximations to differential equations [47]. The differential equations define the relationship between two rates of change [77]. Delta time ($\triangle t$) is the period of calculation used for the simulation. This time period is usually small and incremental [77]. For example, if modelling the population of a university using SD, the population would be the students enrolled for university and the growth rate would be the enrolment of students less the graduates, expressed as

$$\text{population}(t) = \text{students}(t - \triangle t) + \text{net growth}(\triangle t). \tag{2.1}$$

SD is a useful tool to test how a large, complex system would behave for a specific intervention. If an individual's behaviour is modelled, an agent-based simulation modelling approach should be considered where individual behaviour is incorporated through local interactions (on a micro level). Macro-level (or system level) is then observed as a result of many local interactions [77]. SD, on the other hand, allows for the exploration of aggregated behaviour on a population level (which is the macro-level), instead of focussing on the interactions between individuals [77]. Other methods such as the Monte Carlo simulation and discrete event simulation use discrete and stochastic events for a top-down modelling approach, while SD is a continuous process for deterministic modelling.

### 2.4.2 Agent-based simulation modelling

Agent-based simulation (ABS) is a computer simulation approach comprised of autonomous individuals or objects (called agents) who interact with each other and their environment with a possible end-goal [79]. The interaction between the agents and environment changes over time depending on the strategy and decisions an agent may deploy. ABS seeks to understand the behaviour of the system and how the individuals affect this system. ABS follows the bottom-up modelling approach.

In ABS there are four elements required in a model [41]. The first element is the agent with a specific goal, characteristics, behaviour, and decision-making rules. Agents can represent individuals, organisations, or any other relevant entities in a system. For example, within a high school, the learners would be the agents. The learner's end goal would be to matriculate with specific behaviour that the agent must perform to achieve this goal.

The second element is the collection of rules of how agents interact with each other, referred to as the agent topology. Agents follow certain rules or decision-making processes that govern their behaviour. These rules are pre-defined based on prior knowledge, observations, or a theoretical framework. The rules of an agent can either be deterministic or stochastic, and they can change over time. For example, within a high school, the learners interact with peers or a possible secondary agent being the educators. Learners must then interact with educators at least five times per week to gain learning.

The third element is the environment in which the agent exists and operates. The environment represents a physical setting, resources, constraints, and other external factors that can influence agents' behaviours and interactions. The environment can either be dynamic or conservative, and changing over time as it would impact the decisions and outcomes of agents. For example, within the high school, the learner's environment could be the classroom and library. The environment of the learner impacts their behaviour or state.

The fourth element is the state of the agents which represents the agents' internal characteristics, attributes, and information. A state could be an age, location, beliefs, preferences, or any other relevant aspects that influence an agent's behaviour and decision-making. The state of an agent can change over time as the agent interacts with the environment and other agents. For example, within the high school, their state would be Grade 10, 11, 12, or matriculant where the state represents the progress of the learner through the system.

## 2.5 Chapter summary

Higher education systems are complex and large systems that are subject to both clear and latent causal influences. Causality is not easy to determine, but may be reasonably mapped within such a system as shown by the authors listed in this chapter. The modelling of education systems may be approached by means of a number of quantitative methods. Forecasting, machine learning, structural equation modelling, multi-level modelling, and simulation are all methods that have been successfully applied. This chapter provides a summary of specific attempts at simulation modelling in the context of basic and higher learning systems. The work summarised in this chapter serves as a sound foundation upon which the simulation model proposed for this thesis may be constructed.

CHAPTER 3

# Determinants for timely graduation in higher education systems

In this chapter determinants for timely graduation in a generic higher education system is collected and their causal influence is deduced from literature. Various factors can interact in complex ways, and their relative importance may differ for individual students. It's essential to consider the unique context of South African higher education and the specific institutions and populations when analysing the significance of these factors.

## 3.1 Academic preparedness

Adelman [1] explored the relationship between academic intensity, attendance patterns, and bachelor's degree graduation rates. Academic intensity is the degree to which high school learners engage in rigorous academic coursework. Adelman conducted a national longitudinal study from 1980 to 1993 where he followed high school learners from Grade 10 until the age of 30. The learners were given thirteen years to matriculate high school, enter higher education, obtain a four-year college qualification, and then obtain a bachelor's degree. He collected data through surveys of students, teachers, school administrators, test scores, and academic transcripts.

Adelman found that academic intensity and attendance patterns during high school were both significant determinants for college degree graduation. He found that students who enrolled for more challenging subjects in high school were more likely to graduate from college, irrespective of the student's ethnicity and socio-economic status. Students who attended college continuously were also more likely to graduate than those who took an extended break. He also found that students who attended higher education full-time were more likely to graduate than those who attended part-time.

Adelman highlighted the importance of early academic preparation for college success. He found that academic intensity was a better determinant for graduation than making use of standardised testing and he suggested that policymakers should focus on increasing access to rigorous coursework while still in high school and on supporting students to attend higher education full-time.

Tinto [60] explored some reasons as to why students attrit of higher education institutions by developing a theoretical framework. The framework used a so-called "interactionalist" theory to identify these factors as adequate academic preparation, sufficient financial aid, family background, and the social and academic environment of the higher education institution.

The theory emphasised two key influences as individual and institutional factors. The individual factors were students' academic preparation, motivation, academic performance, and goals, while institutional factors were the social environment (*i.e.,* peers, mentors, and involvement in campus activities), and academic environment (*i.e.,* quality of teaching, advising, and support services). Together these are the student's "academic integration" and "social integration". A student was more likely to continue with their degree programme if they felt a sense of belonging at the institution. And if the student was motivated and had a clear understanding of their academic goals and the steps needed to achieve them, they were less likely to attrit. Tinto suggested that institutions should provide clear and realistic academic expectations and provide support and guidance to students to assist them to achieve those goals.

Astin [5] aimed to answer the question of what really mattered in colleges in terms of student learning, development, and success, and what were the factors that influenced these. He analysed data collected from the Cooperative Institutional Research Programme (CIRP), which surveyed over 200 000 students at more than 600 colleges and universities in the United States of America.

Astin found four main factors that influenced student success within higher education as student involvement, the quality of interactions with faculty and peers, the degree of the academic challenge, and the importance of a supportive campus environment. Students who were involved in campus activities, such as clubs and organisations, were more likely to continue in college as it laid a foundation for a social integration and students who had positive relationships with their lecturers, supervisors and peers were likely to continue and achieve higher grades.

Kuh *et al.* [32] investigated the relationship between student engagement and academic success within colleges. They defined student engagement as the degree to which students were involved in academic and social activities on campus. They explored factors that may influence student success, such as pre-college academic preparation, demographic characteristics, and institutional characteristics and found that for students with robust pre-college academic preparation, their prior preparation did not significantly influence their academic success once they were immersed in the college experience, living on campus, enrolled in courses, and working part-time.

They found that student engagement was a strong determinant for first-year success. The students who attended college part-time and commuted to campus did not have much sense of community, as the classroom was their only interaction with other students. Thus, the authors suggested that educators should use the classroom to create communities of learning with increased engagement between students.

Kuh *et al.* [32] further found that students struggled with the engagement in one of four aspects. The first aspect concerned whether the coursework was academically challenging, and if it was, the student would likely persist in college. The second aspect revolved around the student's participation in active learning within the classroom, which involved answering and asking questions as well as engaging in discussions and debates. The third aspect was if the student collaborated with others. Students who collaborated with their peers outside of class cultivated and reinforced academic goals and motivation within their group. The final aspect hinged on the frequency of student-faculty interactions; increased interactions often led to students becoming more invested in and committed to their studies. The authors recommended that colleges and universities should focus on creating environments that fostered student engagement to improve academic success.

Warburton *et al.* [78] explored the academic preparation and post-secondary success of first-generation college students (FGCS) whose parents did not attend college. FGCSs were less likely to enroll in a four-year programme, and when they did they were less likely to continue until graduation. The authors' first key finding was that FGCSs were more likely to come from low-income families and would attend community colleges or other non-selective institutions. They are also less likely to complete a college degree than their peers who had parents who had attended college.

Their second finding was that academic preparation was a key factor in the success of FGCSs. Students who enrolled for more challenging high school subjects, such as advanced mathematics and science classes, were more likely to succeed in college and gain acceptance in their chosen college. The academic preparation in areas such as reading and writing were particularly important for FGCSs as they may lack the academic skills and background necessary for success in college. These students would require additional support.

Their third finding highlighted the importance of support services for FGCSs. Students who received additional supportive services, such as tutoring, academic advisement, and financial aid counseling, were more likely to complete a college degree than those who do not receive these services. The authors emphasised the need for policymakers and educators to provide better support and resources for FGCSs to help them succeed in college. This support included increased access to high-quality academic preparation programmes, financial aid, and improved quality of education at community colleges and other non-selective institutions.

## 3.2 Socio-economic status and financial stability

Pascarella and Terenzini [38] combined their knowledge and experiences of over three decades to focus on how college affected students. They focused on a wide range of outcomes such as student competency development, personal growth, socialisation and career development. These four aspects were core to their six research questions as to whether students changed during college, the nature of the changes unique to the college experience, whether different colleges influenced student change differently, whether different experiences influenced student change within an institution, whether the effects of college varied amongst different types of students, and the long-term effects of college.

Pascarella and Terenzini also examined the role of various institutional factors in shaping student outcomes, such as the influence of faculty, curriculum, and campus climate. Campus climate was defined as the likelihood that a student engaged at campus, the quality of the social and physical environment, and the norms, values, and expectations that govern interpersonal and institutional behaviour. One of their key findings was that college had a positive influence on students where, for example, they found that colleges had an influence on student learning and cognitive development. Students who attended college were more likely to develop critical thinking, problem-solving, and analytical skills, as well as increased understanding and knowledge about their chosen fields of study.

Pascarella and Terenzini stressed the importance of a good campus climate and how institutional factors can shape a student's outcomes. These institutional factors where the quality of education, the availability of student supportive services (*i.e.,* academic advisement, tutoring, career counseling, and mental health services), financial aid, the location of the institution (*i.e.,* proximity to housing, public transport, and jobs), and the size and type of institution (*e.g.,* large or small institution; public or private).

The authors also identified a number of challenges faced by higher education institutions. These challenges were the need to adapt to changing technology (*e.g.*, online learning and distance education), the cost of higher education, access to higher education remaining unequal with many students from low-income backgrounds facing significant financial barriers to attending college, demographics of students that attend college (students from minority backgrounds, and students with disabilities), and the effectiveness and accountability of the institution to performance and achieve academic success. They suggest that there is strong need for institutions to collaborate and provide innovative solutions to overcome these challenges and to commit to equity and access for all students.

Cabrera *et al.* [9] developed an integrated model for student retention. They incorporated both pre-college and college experiences, as well as the student's psychosocial and academic characteristics. They focused on seven key factors that influence student retention, such as pre-college characteristics (*e.g.,* high school scores), academic and social integration within the college, student satisfaction, goal setting, institutional fit, financial aid, and the student's external support and responsibilities (with regards to friends and family).

They made use of structural equation modeling (SEM) to test the validity of their model. Their model showed that pre-college characteristics, academic and social integration, satisfaction, and goal setting were all significant predictors of student persistence while institutional fit was a significant predictor of satisfaction, commitment, and persistence. This would suggest that that students who feel a sense of belonging and connection to their institution are more likely to persist. In addition, they found that financial aid had a direct effect on persistence, but also indirectly affected persistence as it impacted student satisfaction and goals. A student's external support and responsibilities had a negative impact on persistence. The authors suggested that institutions should focus on creating supportive and inclusive environments that foster academic and social integration, student satisfaction, and goal setting, while also providing adequate financial support to students.

DesJardins and Chen [15] investigated the influence of financial aid on student attrition rates and especially considered racial and ethnic differences. They found that students who received financial aid were at a lower risk for leaving the institution and specifically for students from low-income families. They also found that the influence of financial aid on student attrition varied between different races and ethnicities with African American and Hispanic students receiving more financial aid than White and Asian students.

DesJardins and Chen also analysed reasons for the disparity between racial and ethnic groups. They analysed the differences in financial need, academic preparation, and social support and found that financial need and academic preparation did not fully explain the disparity observed, but that social support (*i.e.,* family income and parental education) was the reason. DesJardins and Chen suggested that an increase in financial opportunities for students from low-income families and minority groups would reduce educational inequality and promote greater access to higher education.

Paulsen and St. John [39] analysed the relationship between social class, college costs, and student retention by investigating how financial factors influence students' decisions to attend college and if these decisions affect their likelihood of graduation. They found that social class was a significant factor in both college choice and student retention. Students from higher social classes were more likely to attend more expensive colleges, while those from lower social classes were less likely. They also found that college cost had a stronger effect on student retention irrespective of their social class. Students that attended more expensive colleges were more likely to attrit.

Paulsen and St. John found that once a student enrolled their financial need would still influence their retention and degree attainment. Students who had less financial support from their family would take longer to complete their degree or would attrit and this would be the same for students who worked part-time (*i.e.,* more than 20 hours per week). The authors suggested that institutions should address students' financial challenges by providing more financial support and resources to low-income students. They also suggested that additional academic and social support should be provided to help students navigate higher education.

Titus [64] examined the role of institutional financial support to influence graduation for students from low socio-economic backgrounds. Titus too found that student from low-income families were less likely to complete their degrees than students from higher-income families due to the financial context of the higher education. Students who attended institutions with higher levels of financial resources and lower levels of tuition would have higher graduation rates than those attended institutions with lower levels of resources and higher tuition rates.

Titus highlighted the need for financial aid for low-income students. He found that institutions relied heavily on merit-based financial scholarship to attract high-achieving students from low-income families eventually attrit. He suggested that institutions and policy makers should have merit-based and need-based financial aid to ensure that students from low-income families had access to the financial resources.

## 3.3 Family support and educational background

Ishitani [28] analysed factors that influence graduation for FGCSs. His multi-variate logistic regression model showed that FGCSs were more likely to attrit before obtaining a degree compared to their non-first-generation peers. He identified several factors that increased the risk of attrition as lower levels of academic preparation,and extra-curricular responsibilities (*e.g.*, social clubs, part-time jobs and family). He also found that students who had parents that attended higher education had more knowledge regarding financial aid and academic support.

Ishitani highlighted the importance of social and academic integration. He found that FGCSs who had strong interactions with faculty and peers, were involved in extracurricular activities and societies, and had a sense of belonging on campus were more likely to continue with their degree. He suggested the need-based grants and scholarships should be established to assist students as financial aid is an important factor for FGCSs to continue after the first two-years of college.

Engle and Tinto [19] analysed the challenges that FGCSs from low-income households faced during college. They analysed the data in terms of the type of institutions, the number of students in part-time employment, tuition, financial aid, and the programme type (whether certificate, associate degree, and bachelor's degree). They highlighted several key challenges faced by FGCSs as inadequate academic preparation, financial constraints, limited access to support services, and difficulty navigating the college system.

Engle and Tinto emphasised the importance of fostering a sense of belonging and community on campus for FGCS.s They suggested that programmes should promote social integration, peer support, student engagement, and motivation within campus life. They also suggested that the higher education institution can improve FGCS graduation rates by providing targeted supportive services, such as academic advisement, tutoring, and mentoring, and financial aid and scholarships for these students.

Choy [10] investigated the experiences of high school graduates and FGCSs. She focused on their access to higher education, their persistance through out the programme, and graduation.

She found that FGCSs were less likely to obtain a college degree, as well these students were more like to enroll for part-time or short duration certifications or programmes.

Choy noted that FGCSs were at disadvantage even after controlling for external factors, such as educational expectations, academic preparation, peer influence, support from family, and family income. She found that financial aid assisted in gaining access to higher education, but that students still faced financial barriers to complete their programme. She found that once FGCSs graduated, they were in the same position as other graduates within the labour market and with access to different employment opportunities. Choy recommended supportive programmes to encourage FGCSs and to prepare them prior to enrolment into college.

## 3.4 Institutional factors and resources

Tinto [62] examined the role of the lecture hall on student attrition and graduation in higher education to understand the experiences and interactions that would influence a student to persist in their degree. He expanded on the work of Endo and Harpel [18] and Tinto [61] by adapting the student persistence theory and linking the lecture hall and learning within his theoretical model. He analysed the students' pre-college attributes (*i.e.,* skills, prior schooling, family background), goals and commitment, institutional experiences (*i.e.,* lecturer halls, laboratories, social environment, academic system), and personal integration within the higher education.

Tinto emphasised the importance of student engagement, active participation, and meaningful interactions with lecturers, supervisors, mentors and peers in the learning process. When a student felt connected to peers and educators, they were more likely to develop a sense of belonging and commitment to completing their degree. Similiar to prior researchers [33, 37, 63], Tinto recommended that a sense of community within the lecture hall can contribute to higher levels academic success and that educational institutions should prioritise building supportive environments and providing educators with the necessary resources and training to promote student engagement and retention.

Hamrick *et al.* [25] analysed the relationship between institutional characteristics and resource allocation on graduation rates in higher education. Institutional characteristics refer to the type of institution, selection process, student diversity, and location of the institution, while resource allocation referred to instructional and student affairs expenditures. They found that several institutional characteristics such as selectivity, student-faculty ratio, and student demographics lead to higher graduation rates. Institutions that had a higher percentage of full-time faculty and a diverse student body were more likely to have higher graduation rates.

Hamrick *et al.* also explored the effects of financial resources, instructional resources, and academic supportive resources on graduation rates. They found that financial resources (*e.g.,* donation size and per-student headcount spending) had a positive influence on graduation rates. Higher library expenditures, technology investments, and academic support also had a strong association with higher student graduation rates. They suggested that institutions should focus on redistribution and balancing of financial and instructional resources to enhance student success in higher education.

Schultz [46] presented an economist's view on resources and resource allocations for higher education institutions. His first view point was student sovereignty and their self-interest. He identified that students were not fully aware of the cost to produce a graduate and from the point view of the economy, the allocation of resources to higher education would be sufficient to support them. The second view point was recognising the social benefits of higher education.

Through attending higher education there would be an increase in the mobility of skilled labour force that would have better job opportunities than those who did not attend higher education. The third view point was academic entrepreneurship. Schultz focused on the potential value of students and faculty and he emphasised that financial resources such as government funding and private contributions are important in supporting the infrastructure, faculty, and research activities of universities. He highlighted the importance of human capital in establishing the quality and effectiveness of higher education.

Liefner [34] examined the relationship between funding, resource allocation, and performance in higher education as well the how resource allocation influenced academic performance. He compared six different universities by their source of funding, the theoretical budgeting methods used, and the factors that would influence the success of the universities. He compared the Swiss Federal Institute of Technology in Zurich (ETH Zurich), University of Basel, University of Twente, University of Bristol, Massachusetts Institute of Technology (MIT), and University of Texas. He focused on the financial resources that would influence teaching support, research, infrastructure development, and student supportive services as these factors would overall influence the quality and effectiveness of higher education.

Liefner identified factors that would influence the long-term success of the universities as faculty qualification, students ability, university culture, and resource allocation. He found that although each of the universities had different budgets and resources, they were able to attract qualified educators that elevated the reputation and success of the institution. He suggested that if university administration used performance-based budget allocation, they should keep in mind the long-term success of the institution which is influenced by qualifications and abilities of the faculty staff.

## 3.5 Student motivation and engagement

Robbins *et al.* [43] conducted a meta-analysis into skills and the relationship between psychosocial, study skills and various college outcomes such as college GPA, student retention, and graduation rates. They identified twelve possible factors from literature as achievement motivation, academic goal setting, institutional committee strength, social support, social involvement, academic self-efficacy, general self-concept, academic skills, financial support, size of institutions, institutional selectivity, and environmental influences.

They identified that achievement motivation, academic self-efficacy, time management skills, goal orientation, and social integration were the strongest determinants for student success. They also explored additional factors such as gender, race, the number of academic years, and institutional characteristics and found that the initial factors were still valid across different factors. Robbins *et al.* suggested the that institutions develop interventions and support programmes to enhance student success by focusing on improving achievements motivation, academic self-efficacy, time management skills, and social integration.

Astin [4] analysed the role of student involvement in the development and success of students in higher education. He described student involvement as the quantity and quality of the physical and psychological energy that students invest in the college experience. These include academic work, participation in extracurricular activities, and interaction with institutional personnel. Austin emphasised the importance of student involvement in higher education and how it contributes to student learning, personal development, and academic achievement. He proposed a developmental theory that focused on the dynamic relationship between students and their college environment and how their experiences influenced their growth. He called this theory the "student involvement theory".

He listed factors that impact student involvement such as resources, student place of residence, degree programme, academic involvement, athletic involvement, student-faculty interaction, and their involvement in student government. He found that the level and quality of student involvement have a significant impact on their personal and intellectual development. He suggested that higher education institutions should provide opportunities for student engagement by establishing campus communities and interactions between students and faculty.

Busato *et al.* [8] calculated the correlation between intellectual ability, learning style, personality, and achievement motivation on the academic success of psychology students in higher education. Their studied sampled 409 first year psychology students at University of Amsterdam over three years. They analysed data collected during test week, the students' first examination at university level, the number of study points a student accumulated, the Inventory of Learning Styles (ILS) survey, the Prestatie-Motivatie Test (PMT), and various intellectual ability tests. The intellectual ability tests consisted of testing the students' abilities in verbal analogies, vocabulary, number series, number speed, embedded figures and drawing conslusions. The study considered the Big Five personality factors of openness to experiment, extraversion, agreeableness, conscientiousness and neuroticism.

Busato *et al.* found that students' intellectual abilities, achievement motivation, and conscientiousness personality had a positive influence on their academic success, while no learning style had direct influence on their academic success. The ILS survey was found to be useful on an individual level to assist in identifying a students strengths and weaknesses but not on an institutional level. A student's fear of failure in combination with their personality type and academic achievement could either have positive or negative impact on their academic success.

Burger and Naude [7] analysed the predictors of academic success in the entry and integration stages of students' academic careers. They focussed on three variables in their study as the type and quality of school the student attended, the student's academic self-concept, and the student's integration into higher education. Along with these variables they explored the extent to which a student's Grade 12 performance would influence their academic success. A student's academic self-concept is the student evaluation of their own academic abilities [7].

The study sampled 164 students in the Humanities faculties in South African higher education institutions. The students were categorised according to their academic years where first and second year students were categorised into the entry stage, and third and fourth year students were catgorised into the integration stage. The students' academic records, Grade 12 academic performance, and application scores were analysed along with the Academic Self-Concept Scale (ASCS).

Burger and Naudé conducted a multiple regression analysis to determine variation in students' academic success and their integration into HE. They found an approximate 31% variance for students in the entry stage and an approximate 27% for students in the integration stage. This variance could be explained when considering all the variables on academic success. When considering each variable individually, a student's academic self-concept was the best predictor for academic success with 19.4% variance for student in the entry stage and 14.2% for students in the integration stage.

## 3.6 Study habits and time management skills

Van der Zanden *et al.* [68] reviewed the domains and predictors of first year success and which predictors are associated with each domain using existing literature. They constructed a conceptual framework of three domains for academic success, a student's academic achievement, their critical thinking skills, and their social-emotional well-being. Academic achievement was an indicator of the quality and efficiency of the educational programme they were enrolled for. A student's critical thinking skills described higher-order thinking and scientific reasoning. A student's social-emotional well-being was their transition into HE and adulthood, and if they were able to cope with the increase in interdependent and responsibility.

Van der Zanden *et al.* identified a student's demographic, their academic preparation, motivation and study skills, their self-evaluations, their social support and integration, the institutional and organisational variables, and out of class stressors as significant determinants. The most common determinants were related to all three domains are the students' previous academic achievements, intrinsic motivation, study skills, and relationships with parents and peers. Van der Zanden *et al.* also identified specific determinants that linked to each domain. In the academic achievement domain two educational psychological determinants were identified academic self-concept and effort, in the critical thinking domain two learning environment determinants were identified as inquiry based learning and challenging teaching, and in the social-emotional well-being domain two psycho-social determinants were identified as self-concept and sense of belonging.

Crede and Kuncel [12] investigated the role of study habits, skills, and attitudes (SHSA) in predicting academic performance among college students. They reviewed wide range of past studies on the relationship between SHSA and academic performance, with the aim to conduct an in-depth analysis of the factors that impact student success. They examined various factors and grouped them according to different surveys, they were Survey of Study habits and attitude (SSHA) and Learning and Study Skills Inventory (LASSI). SSHA focused on factors such as delayed avoidance (or procrastination), work methods, study habits, teacher approval, educational acceptance, study orientation, and study attitude. While LASSI focused on student's attitude, motivation, time management, anxiety, concentration, information process, selecting main ideas, study aids, self-testing, and test strategies.

They found SHSA best explained the three reasons for student success. The first was that a student's study skills, habits and attitude would influenced their academic performance as without these skills students would struggle to pass. This indicated that a student's study habits, skills, and attitudes were robust predictors of academic performance. The second reason is that a student's SHSA scores explained the variance between a students college GPA and their pre-college preparation test. A student could do well in pre-college preparation test but not college as they lack SHSA. The third reason was that a student's personality traits partially influenced their study attitudes and study habits. They suggested that educational institutions provide students with resources and support to enhance their study habits, skills, and attitudes.

Kitsantas *et al.* [31] analysed the role of a student's self-regulation and cognitive abilities in predicting their academic success in college. They focused on the influence a student's self-regulation had in terms of the student's time management skills, task valuation, self-efficacy, and their test anxiety. They collected data about the student's demographics, prior abilities, and academic performance and each student answered the Motivated Strategies for Learning Questionnaire (MSLQ) about their motivation, self-regulation, and cognitive abilities.

They found that when predicting the second semester academic performance, a student's time management and self-efficacy had a significant influence on their academic performance while the other factors such as task valuation, test anxiety, and meta-cognitive abilities did not have a significant influence on the student's performance. They also found that the student's gender played a significant role in their second year of higher education where females with higher pre-college preparation and GPA achieved higher levels of academic success. They suggested that lecturers, supervisors, and adminstrators should focus on a student's time management skills as students with lacking time management skills were more likely to perform poorly or attrit in the following year.

## 3.7 Social integration

Rienties *et al.* [42] analysed factors that influenced the academic performance of international students with specific focus on their ethnicity, academic integration, and social integration. They conducted a quantitative analysis using survey data of international students studying at universities in the Netherlands. From these surveys they found that several factors contributed to the academic performance of international students. Firstly, academic performance was positively influenced by students' academic integration were the institution's reputation place the largest influence on the student's retention. Secondly, "Western" international students had better academic success as their motivation and learning attitude were higher than domestic students as these students made a conscious decision to succeed in their studies. Thirdly, academic and social adjustment had a compensatory relationship on academic performance. They suggested that institutions should focus on promoting academic and social integration as it could improve the academic performance of international students.

Hausmann *et al.* [26] investigated the relationship between students' sense of belonging and their intentions to persist in college. They focused specifically on African American and White first-year students in their investigation. They identified the factors that generated a sense of belonging for students. They explored factors such as student background (*i.e.,* race, gender, financial support and pre-college preparation), social and academic integration, social or community support (*i.e.,* peers and parents), institutional commitment, and intentions to persist. These factors are similar to the others identified by other researchers [18, 38, 61, 62].

They found that students who had regular interactions with peers, lecturers, and mentors as well as those who had supportive communities had a greater sense of belonging. It was important for these interactions to start at the beginning of the academic year as later in the academic year this sense of the belonging would decline. They also found that the academic integration and student's background did not impact their sense of belonging as students would still be able to establish a supportive environment. Hasumann *et al.* tested an intervention strategy were students would receive a gift or mail. They found that student who received them felt valued and part of the university community but compared to those we did not, there was only a slight difference in their sense of belonging. They suggested that institutions should generate inclusive and supportive campus environments that promote a sense of belonging for all students, particularly for African American students. Institutions should provide resources and support networks to enhance the students sense of belonging and increase their intentions to persist within college.

Noyens *et al.* [35] investigated the relationship between students' academic motivation and social integration during their first year of higher education. To understand the relationship, they conducted a longitudinal study of first-year students in Belgium where they would answer question about social integration (*i.e.,* difficulty interacting with other), motivation, intrinsic motivation, identified regulation (*i.e.,* personally important), introjected regulation (*i.e.,* the desire to be perceived as smart), external regulation (*i.e.,* the opinions of others, and motivation).

They found bidirectional relationships between academic motivation and social integration. Students who had positive interactions with peers, mentors and lecturers had a better quality and quantity of motivation. Student who were classified under introjected regulation and external regulation had a negative relationship with social integration, suggested that not all students study for joy, passion or interest. Students who experienced higher levels of social integration during their first year of college were more likely to develop and sustain their academic motivation over time. They suggest that universities and educators should create supportive and inclusive environments that facilitate both academic and social engagement.

## 3.8 Choice of major and course load

Szafran [56] investigated the relationship between academic course load and academic success to determine whether a lighter academic load lead to better academic performance and graduation rates. He sampled data from a large public university in the United States and analysed students' GPA, retention, number of credits taken and completed, course difficulty, gender, race, pre-college preparation tests, size of high school cohort, and part-time employment.

He found that students with a stronger academic background enrolled for more college-credit hours while students with weaker academic background enrolled for more development hours, which resulted in a similar total credit load. He also found that students would not unenrol from a course if they were heading towards failing but instead persist even if they had some difficulty. Szafran suspects that it would be due to financial aid or campus residency requirements that the students should not attrit. He also found students who enrolled for more hours obtained a higher GPA while students who enrolled for a more difficult course obtained a lower GPA. He suggested that lecturers and administrators should provide guidance and support to help students make informed decisions about their course loads based on their individual circumstances, interests and abilities.

Zou *et al.* [80] explored the relationship between university graduates' GPA and their salary. Their aim was to provide supporting evidence on the importance of graduate's GPA in determining their potential earnings in the future. They sampled students from the business and economic department at a Chinese university. The students were assumed to be homogeneous as their majors were closely related. In their regression model analysed at factors such as the individual and employment characteristics (*i.e.,* gender, degree level, major, unemployment period after graduation), family background (*i.e.,* socio-economic status, parents occupation, parents education), GPA, and if the student obtained any non-academic awards.

They found that there was a positive relationship between GPA and salaries. When a students' GPA increased by 1 point their starting salary would increase on average by 29.6%. When comparing the graduates starting salary with their current salary, graduates with a higher GPA earned a slightly higher salary than those with a lower GPA. In addition, they found that when there was a drive for higher education policies to churn out graduates, the labour market would then ignore the graduates' academic performance in the requirement process. They also found that male graduates had a higher starting salary compared to female graduates. This gender disparity highlights the presence of other factors, such as gender-based discrimination. They found that postgraduate degree, nonacademic awards, and parental occupation had no significant impact in the long run on salary while the other family background factors had a slight impact on the graduates on their current salary.

Dale and Krueger [13] investigated the economic benefits of attending a more selective college to estimate the impact that this would have on students' earnings after graduation. They explored factors of race, gender, high school GPA, high school ranking, type of higher education (*i.e.,* public or private), the number of application to higher education (*i.e.,* applied to three other higher education institutions other than the one you are currently attending.)

They found that students who attended more selective colleges earned a higher salary compared to those who attended less selective colleges. Dale and Krueger stated the reason that a student who attended a more selective college may have a higher unobserved ability (cognitive and/or non cognitive) and could be the reasoning for the higher salary. They found the students from a disadvantaged background that attended a selective college may raise their families income as they would benefited from attending the a more elite college.

## 3.9 Determinant feedback in higher education systems

In this chapter, various literature about graduate success, student performance, student retention, and student success inform the development of the casual loop diagramme shown in Figure 3.1. These factors were identified from the literature described in Sections 3.1 to 3.8.

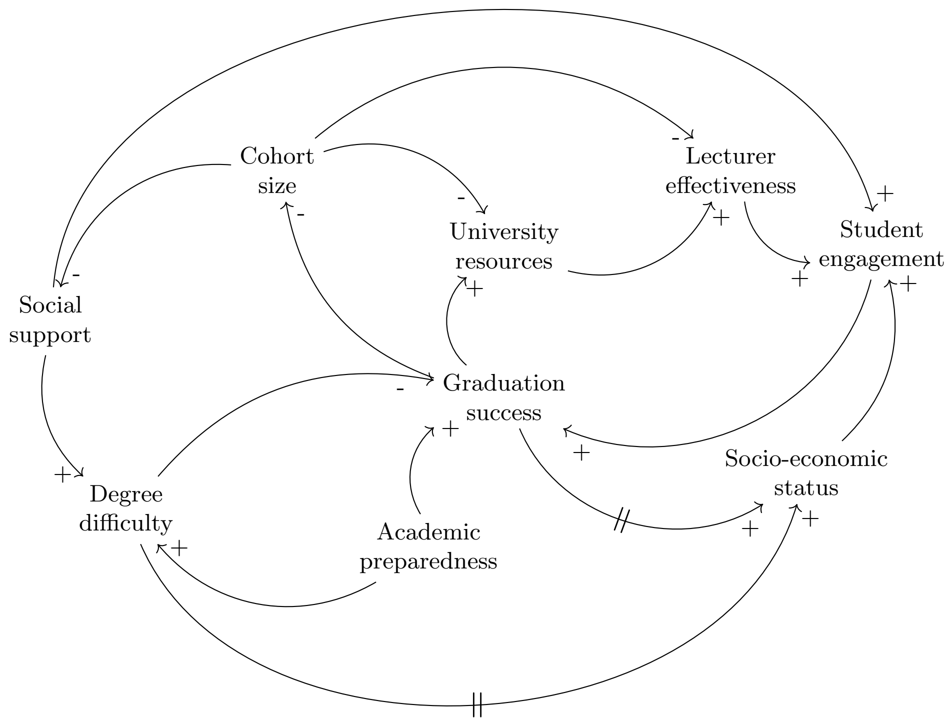



FIGURE 3.1: *The causal loop diagramme for graduate success in the higher education system determined from literature.*

Through an in-depth systematic literature review, nine determinants are assumed to have an impact on student progression within higher education. These are academic preparedness, social support, socio-economic status, university resources, student engagement, lecturer effectiveness, cohort size, qualification difficulty, and graduate success.

Academic preparedness is the student's final matriculation results as the foundation for their university career. These results could be the South African National Senior Certificate (CAPS), Cambridge International curriculum or any other international curriculum that gains the student entrance into higher education. The social integration (also called social support) is the environment, peers, and sense of community in which the student exists. Student engagement is the influence the students' engagement and motivation had on their academic work. If a student has a supportive environment their engagements improve.

A student's socio-economic status is the material support they receive from their family, the resources available to them while at home, and the ability of the student to focus on their studies without the need for part-time employment to support themselves. The degree programme that a student is enrolled for influences their success and career prospects as the more challenging and specialised a degree programme might be, the better the chance that the student would have in obtaining employment within the labour market (*i.e.,* degree difficulty).

The success of a student is their progression towards graduation. Without graduate success the student would remain within the system and is at risk of eventual attrition without graduation. As students graduate their cohort size decreases as their are fewer students repeating modules. A large cohort size negatively influences lecturers in terms of workload and their effectiveness to serve the cohort (*i.e.,* lecturer effectiveness) decreases. To improve lecturer effectiveness, sufficient resources (*i.e.,* the metric for university resources) and training are required. These resources may be obtained through subsidy and donations from past alumni or based on the institution's reputation.

Figure 3.1 contains the causal relationships between each of the eight endogenous factors and one exogenous factor. The system is subject to a number of major and minor feedback loops. The main loops influencing graduate success subject to the determinants within the policy maker's ability to manipulate are two reinforcing loops and one balancing loop as depicted in Figure 3.2.

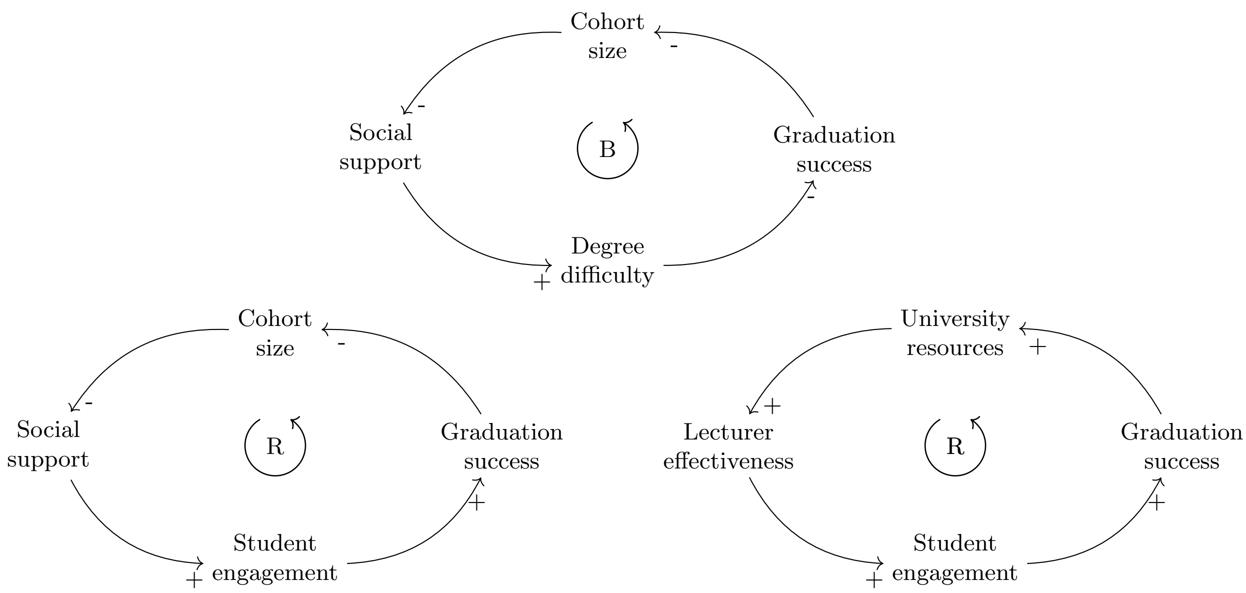



FIGURE 3.2: *The main loops that influence graduate success are subject to two reinforcing loops and one balancing loop.*

The first reinforcing loop is achieved as a decrease in cohort size leads to a decrease in social support. When there are few peers present students are able to engage more with the lecturer than with these peers, leading to a lack of social interconnectedness. These engaged interactions with lecturers motivate students to prepare and revise the lectured module content more, thus leading to an increase in student engagement. This drive in student motivation results in a growth in graduate success.

The second reinforcing loop is achieved through an increase university resources. This availability enables lecturers to be more effective within the lecture halls and tutorial rooms. An effective lecturer has a positive influence on student motivation and student retention resulting in an increase in student engagement. This drive in student motivation results in growth in graduate success.

The balancing loop influencing degree difficulty is achieved as a decrease in the cohort size results in students finding the degree programme more challenging, increasing the difficult of the degree programme, which ultimately leads to a decrease in graduation success.

## 3.10 Chapter summary

The first question when analysing complex non-linear systems is which elements should be included and how can these measures be determined. In this chapter a selection of determinants were sourced from literature and metrics for their quantification were suggested. Nine determinants were isolated from literature. The second question is how are the elements related and which causal feedback loops exist. Later in this chapter the causal relationships between the determinants were assumed from literature and mapped together in a causal loop diagramme. From the causal loop diagramme likely policy levers were suggested by identifying the reinforcing and balancing feedback structures within the system.

CHAPTER 4

# Data and assumptions

In this chapter the data describing graduation success in the EMS Faculty of SU are introduced. Data preprocessing is described and the method by which metrics are quantified and normalised as goodness scores, is presented

## 4.1 Data preparation for student progression

Within Stellenbosch University most of the data regarding student progression is stored at the Division for Information Governance (IG), within the Centre for Business Intelligence (BI). The Centre for BI provides information management, standardised reporting, and analytics to support institutional decision-making [66].

The data to populate the SPM was received from the BI unit in `.xlsx` format. This data contained records of the students' home language, their enrolled degree programme, their focal area, the faculty in which their degree is homed, their selected modules, number of credits, module academic year, final achievement mark, number of module attempts and final outcome of the module (*i.e.*, failed, passed, not assessed, insufficiently assessed, passed with distinction, discontinued, or not having met the requirements for examination). Table 4.1 contains a sample of this data.

| **Enrolment year** | **Student** | **Language** | **Programme** | **Focal area** | **Module** | **Credits** | **Module year** | **Final mark** | **Subject remark** | **Attempt** |
|---|---|---|---|---|---|---|---|---|---|---|
| 2017 | WA875 | Eng | BCom | LM | Eco 114 | 12 | 1 | 67 | Pass | 0 |
| 2017 | WA875 | Eng | BCom | LM | IS 112 | 6 | 1 | 97 | Pass with distinction | 0 |
| 2018 | WA875 | Eng | BCom | LM | TE 244 | 16 | 2 | 48 | Failed | 0 |
| 2019 | WA875 | Eng | BCom | LM | TE 244 | 16 | 2 | 79 | Pass with distinction | 1 |
| 2019 | WA875 | Eng | BCom | LM | LM 344 | 12 | 3 | 93 | Pass with distinction | 0 |
| 2019 | WA875 | Eng | BCom | LM | LM 354 | 12 | 3 | - | Insufficient assessment | 0 |

Table 4.1: *A sample of the raw data received to populate the SPM for selected focal areas of Logistics Management (LM) and selected modules in Economics (Eco), Information Systems (IS), Transport Economics (TE), and LM.*

The content and nature of modules have been subject to change over a number of years at SU.

Modules that spanned over two semesters could have been shortened to one semester or discontinued (*e.g.,* Quantitative Management), or split into multiple modules over two academic years (*e.g.,* Data Science, and Business Analytics). Some modules changed their administrative home department or anchor degree programmes and as a result their module naming and code was ammended (*e.g.,* Logistics Management changed to Supply Chain and Logistics Management).

To model these changes, a module dimension table was created. This dimension table matches the former module naming with the lastest module naming. The dimension table was compiled by comparing each of the academic yearbooks to determine if the content of the module remained the same, changed over time, was discontinued, or changed naming. A sample portion of this dimension table is listed in Table 4.2. A similar approach was conducted for degree programmes where naming was changed over time. The use of the dimension tables ensures that no observations are lost when a module changed name or was replaced by a new module. For this study, only undergraduate degree programmes that fall in the EMS faculty was considered when preparing the input data.

| **Old Module** | **Current Module** | **Notes** |
|---|---|---|
| Academic Literacy For EMS 111 | Academic Literacy For EMS 111 | No changed |
| Actuarial Science 242 | - | Discontinued |
| Actuarial Science 274 | Actuarial Science 211 | Replaced |
| Auditing 288 | Auditing 288 | Unchanged |
| Business Analytics 244 | Business Analytics 244 | Unchanged |
| Information Systems 188 | Information Systems 114 | Split 114 and 144 |
| Logistics Management 324 | Logistic and Supply Chain Management 324 | Naming change |
| Logistics Management 344 | Logistic and Supply Chain Management 344 | Naming change |
| Operations Research 214 | Operations Research 214 | Unchanged |
| Quantitative Management 214 | - | Discontinued |
| Statistical Methods 176 | Statistics and Data Science 188 | Replaced |
| Statistics 186 | Statistics and Data Science 188 | Replaced |

TABLE 4.2: *A sample from the module dimension table.*

At SU a student enrolls for a degree programme. This degree programme is either a specialised programme (*e.g.,* Bachelor of Accounting, or Data Science, or Food Science) or a generalised programme (*e.g.,* BCom (Management Sciences) or Bachelor of Engineering). Dependent on the programme students either select various elective modules or they have a fixed set of compulsory modules. Elective modules can generate a multitude of different module combinations based on the modules students have completed in their prior academic year and given that they have met all the prerequisites.

Table 4.3 lists a sample of the array table used to map the possible module combinations based on the recommended modules, prerequisites, and the assumption that the student would not seperate a pair of modules (*e.g.,* the student would select both Business Analytics 214 and 244 and not only select Business Analytics 214 in their second year.) Each module is assigned a value of either 0 or 1 to indicate whether or not a student is enrolled for the module. Each record in the table corresponds to a specific academic year, degree, and focus combination. For example, in BCom (Management Sciences) with a focal area in Human Resource Management there are only compulsory modules in the first year, 37 valid module combinations in the second year, and only six module combinations in the third year. In the second year, a student can select a variety of recommended modules and a total number of 56 combinations could be generated. Of those combinations 19 would not allow the student to have sufficient credits in their third year based on the recommended modules and module prerequisites. These 19 combinations were excluded from the module combination table.

| **Degree** | **Focus** | **Academic year** | **First year option** | **Second year option** | **Third year option** | **Module 1** | **Module 2** | **Module 3** | **...** | **Module *n*** |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | ... | 0 |
| 1 | 1 | 2 | 1 | 1 | 0 | 0 | 0 | 1 | ... | 0 |
| 1 | 1 | 3 | 1 | 1 | 1 | 0 | 0 | 0 | ... | 1 |
| 2 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | ... | 1 |
| 2 | 1 | 1 | 2 | 0 | 0 | 1 | 0 | 1 | ... | 0 |
| 2 | 1 | 1 | 3 | 0 | 0 | 0 | 1 | 0 | ... | 1 |

Table 4.3: *A sample from the module combination table.*

Table 4.4 lists a sample from the dimension table of the prerequisite modules. In this table, the module of interest, along with the module(s) required, and the requirement type is listed. A requirement type "P" means that a student must have a final mark of least 40% before they can enroll for the module. The requirement type "PP" means that the student must have passed the module before they can enrol for another module while a requirement type of "C" means that the student must be enrolled for the module at the same time or at least completed it in the prior academic year. For this study only requirement types of "P" and "PP" are considered.

| **Module** | **Prerequisite type** | **Required module** |
|---|---|---|
| Auditing 288 | P | Financial Accounting 178 |
| Financial Accounting 278 | PP | Financial Accounting 178 |
| Agricultural Economics 242 | P | Economics 114 |
| Investment Management 254 | P | Statistics and Data Science 188 |
| Business Analytics 214 | PP | Statistics and Data Science 188 |
| Logistics and Supply Chain Management 314 | PP | Logistics and Supply Chain Management 214 |
| Operations Research 322 | PP | Probability Theory and Stats 144 |
| Actuarial Science 241 | PP | Mathematics 114 |
| Actuarial Science 241 | PP | Mathematics 144 |
| Actuarial Science 241 | PP | Probability Theory and Stats 144 |
| Actuarial Science 241 | PP | Actuarial Science 112 |
| Actuarial Science 241 | PP | Actuarial Science 211 |

Table 4.4: *A sample from the prerequisite dimension table.*

The probability of a student passing a module is calculated according to a triangular distribution for each student for each year. This distribution is based on the lowest achieved mark, highest achieved mark, and average mark per module. A triangular distribution was selected due to the inconsistent number of enrolments per module over the analysis period and therefore standard methodology was applied. In Table 4.5 a sample of the input data by which the triangular distributions were generated, is shown.

| Module | Average achievement mark | Minimum achievement mark | Maximum achievement mark | Enrolments |
|---|---|---|---|---|
| Actuarial Science 142 | 53 | 21 | 80 | 985 |
| Business Analytics 214 | 72 | 29 | 94 | 150 |
| Business Analytics 318 | 53 | 2 | 72 | 43 |
| Business Management 113 | 57 | 16 | 91 | 15 323 |
| Computer Science 354 | 71 | 62 | 77 | 35 |
| Economics 114 | 61 | 19 | 97 | 15 648 |
| Human Capital Metrics 344 | 58 | 40 | 76 | 13 |
| Management Accounting 278 | 58 | 19 | 88 | 4 587 |
| Mathematical Statistics 344 | 60 | 29 | 90 | 783 |
| Operations Research 322 | 58 | 32 | 82 | 117 |

TABLE 4.5: *A sample of the input data to calculate the triangular distributions of pass probabilities.*

## 4.2 Data preparation for systemic support of student progression

At SU as a module nears completion, students are invited to complete a voluntary survey, called the student feedback form (SFF). In this survey students are asked a variety of questions about their demographics, academic habits, and experience of the module and lecturer(s). The SFF includes between 30 and 40 questions depending on the department and faculty. In the EMS faculty the survey consists mainly of closed-ended questions where the student selects a response to correspond to a range or statement. In the raw data, the answers are assigned to a numerical value of between one and six as shown in Table 4.6.

| **Question** | **Score of 1** | **Score of 2** | **Score of 3** | **Score of 4** | **Score of 5** | **Score of 6** |
|---|---|---|---|---|---|---|
| My matriculation average falls in the interval: | 50-59 | 60-69 | 70-79 | 80-89 | 90+ | NA |
| Module diffculty, relative to other modules you followed this year, was: | Very easy | Easy | Medium | Difficult | Very difficult | NA |
| The lecturer's explainations were clear: | Disagree strongly | Disagree | Neutral | Agree | Agree strongly | NA |

TABLE 4.6: *A sample from the raw data transcription of the SFF.*

The SFF responses were captured in `.xlsx` format A sample of this data is listed in Table 4.7. Participation to complete the SFF is voluntary, and as such the data might be biased by responses received possibly from only the top achieving students in the module or by disgruntled students. Due to the anonymity of the SFF it is difficult to determine the nature of the respondent.

The first step is wrangle the data describing systemic support submodel is cleaning and pre-processing the data. For example, in Table 4.7 respondents "22 465" and "22 468" did not fully complete the SFF and only answered selected questions, so producing output that is not a true reflection of feedback. Respondent "22 510" on the other hand had impossible values captured, so that computation of the average lecturer scores are skewed. Therefore all responses where the

| **Response** | **Module** | **Q1.1** | **Q2.2** | **Q2.6** | **Q3.1** | **Q3.14** |
|---|---|---|---|---|---|---|
| 22 455 | EE_SMI09 | 2 | 4 | 50 | 3 | 60 |
| 22 464 | EE_SMI09 | 4 | 4 | 100 | 5 | 100 |
| 22 465 | EE_SMI09 | | | 15 | | |
| 22 468 | EE_SMI09 | 6 | 5 | | 6 | 80 |
| 22 476 | EE_SMI09 | 3 | 2 | 75 | 4 | 80 |
| 22 477 | EE_SMI09 | 4 | 4 | 65 | 5 | 65 |
| 22 510 | EE_SMI09 | -999 | 4 | -999 | 5 | -999 |

TABLE 4.7: *A sample from the captured student feedback data.*

student did not complete the SFF entirely, and had impossible data values were deleted from the final dataset.

Because module and lecturer names in the SFFs were depersonalised, a customised dimension table was created from the feedback form sent to the department to store the module names. For example, the module name "EE SMI09" was matched to the department's naming conventions, which include both English and Afrikaans naming conventions, class groups, and a variety of spellings over the years. Each module was also aggregated to its degree programme, and a value of 1 was assigned if the module was presented in the degree programme, or a value 0 if it was not. A sample from the module dimension table can be seen in Table 4.8.

| **SFF naming** | **Department naming** | **Final nnaming** | **Bachelor of Accounting** | **BCom (Mathematical Sciences)** | **BCom (Management Sciences)** | **BCom (Economic Sciences)** |
|---|---|---|---|---|---|---|
| W_ BEL399_ 16446 | Bestuursreken-ingkunde 278 (Eng) | Management Accounting 278 | 1 | 0 | 0 | 0 |
| VZ_ ACT112_ 13757 | Bestuursreken-ingkunde 278: Eng | Management Accounting 278 | 1 | 0 | 0 | 0 |
| 18EE_ VAN07 | Bestuursreken-ingkunde 278: Engels | Management Accounting 278 | 1 | 0 | 0 | 0 |
| 19EE_ VAN06 | Business management 113 | Business Management 113 | 1 | 1 | 1 | 1 |
| SFB_ 30782_ 19EE_ JAC01 | Business Management 113: Group 1 | Business Management 113 | 1 | 1 | 1 | 1 |
| 18EE_ KRU04 | Business Management 113: Group 2 | Business Management 113 | 1 | 1 | 1 | 1 |
| 19EE_ SOL04 | Entrepreneurship and Innovation Management 244 | Entrepreneurship And Innovation Management 244 | 0 | 0 | 1 | 1 |
| 18EE_ DEJ01 | Entrepreneurship en Innovation Management 244 | Entrepreneurship And Innovation Management 244 | 0 | 0 | 1 | 1 |

TABLE 4.8: *A sample from the module dimension table for the SFF.*

Additional data from the Centre for BI was used to determine the number of lecturers, number of students enrolled per module, the names of students' secondary school, and their residential suburb (*i.e.,* campus resident, campus private accommodation or home town). Similar to the SFF module naming there were also differences in the module naming between the SFF and

Centre for BI so that a similar dimension table was created as shown in Table 4.8.

In the data received for students' secondary school and residential classification, a number of the schools where not assigned a schooling quintile. A schooling quintile refers to the categorisation of a school into one of five quintiles based on its financial resource allocations. Quintile 1 is the "poorest" quintile and Quintile 5 is the "least poor" quintile. Quintiles 1 to 3 schools charge no school fees, while Quintiles 4 and 5 schools require payment of school fees [14]. A list of schools for each province was obtained from Department of Basic Education (DBE) [16]. A school dimension table was created similar to Table 4.8 as the name of the school on the Centre for BI's database and the DBE differed in terms of language (English, Afrikaans, or isiXhosa), punctuation, or abbreviations. In Table 4.9 a sample of the schooling dimension table is shown.

| **SU Naming** | **DBE Naming** |
|---|---|
| Abbotts College Claremont | Abbott'S College (Claremont) |
| Beaulieu Kollege | Beaulieu College |
| Good Hope Gemeenskaps Kollege | Good Hope Community School |
| Holy Rosary Convent | Holy Rosary Convent School |
| Inspired Schools (Pty) Ltd T/A Reddam House Constantia | Reddam House |
| Michaelhouse | Michaelhouse Secondary School |
| Northern Academy Sec School | Northern Academy |
| Orient Islamic Sek Skool | Orient Islamic School |
| Plettenberg Bay Christian School | Plettenberg Bay Christian |
| Southcity Christen Kollege | Southcity Christian Schools |
| St Andrew'S School For Girls | St Andrews School For Girls |
| St Benedict School (Kzn) | St Benedict School |
| Three Rivers Christian Academy | Three Rivers Christian Acdemy |

Table 4.9: *A sample from the school naming dimension table.*

The SFF responses were used compute the initial values for five out of the nine determinants in the casual loop diagramme from Section 3.9. These deteminants are academic preparedness, student engagement, degree difficulty, university resources, and lecturer effectiveness. The other four determinants were computed using data obtained from the Centre for BI. A goodness score is used to normalise variables of disparate units. The goodness scores are calculated as the total number of responses as a fraction of the highest observed value. The goodness scores normalise the values obtained from the student feedback forms and the BI dataset to return normalised values. Scores between the upper and lower bounds are interpolated linearly.

A student's academic preparedness is measured as the student's final matriculation results. In the SFF students indicate the range within which their average matriculation result fell. The higher the range the more prepared a student is assumed to be for their university career. As this measure is already a fractional score of [0,1], no goodness score calculation is required.

Student engagement is measured of the students' score for engagement and motivation towards their academic work. Students were asked about their class attendance, the number of hours they spent outside of class time, and their interest in the content at the start and end of the modules. These questions provided insight into their behaviour and feeling towards the module. A goodness score of 0 represents a fully disinterested and unmotivated student while a goodness score of 1 represents a fully enthusiastic and involved student.

A score for university resources is quantified as a measure of the quality of the facilities, training and learning infrastructure available. At SU, lecturers partake in regular training programmes and have departmental support groups. Students were asked about the learning resources provided to them as well as physical teaching and learning environments (*i.e.,* lecture hall size, lighting, and computer labs). A goodness score of 0 represents a total lack of resources while a goodness score of 1 represents a surplus wealth of resources.

Degree difficulty is quantified by how challenging the student found the modules. The metric is quantified through measures for the pace of the modules, and the difficulty and workload of the module relative to other modules. A goodness score of 0 represents an overwhelmingly challenging module while a goodness score of 1 represents a module that the student did not find challenging in the least.

Lecturer effectiveness is a measure of how the student perceived the lecturers' teaching style and motivation. In the SFF, eleven questions were asked about the lecturer. These questions were about the lecturer's enthusiasm, teaching methods, communication, and whether the lecturer was stimulating and fair towards the students. In the SFF the lecturer effectiveness score is calculated as a score based on the feedback using the average mark on a continuum from one to five and seen as a general impression score of the lecturer. A goodness score of 0 represents an unsatisfactory performing lecturer as seen by students while a goodness score of 1 represents an outstanding performing lecturer as seen by students.

Cohort size is defined as the average enrolment per module as a percentage of the international best standard of thirty students to one lecturer. This ratio is dependent on the various factors such as type of institution, the level of education (*i.e.,* undergraduate or graduate), the discipline being taught, the availability of support staff, and the educational objectives [65]. A goodness score of 0 represents a high student to lecturer ratio while a goodness score of 1 represents a student to lecturer ratio close to the international best standard.

A student's social support is defined as the student's social integration into the university. A student who stays within Stellenbosch area would have greater social integration as they are closer to campus where they could more easier make use and benefit from the social club, resident groups, private student organisations (PSO), mentorship programme while a student who lived further from campus who have the similar access to these benefits but would have commute to SU to uptake the benefit. A goodness score of 0 represents a student with little or no social support while a goodness score of 1 represents a student with an abundance of social support.

A student's socio-economic status is defined as the financial support available to them. A student from a lower socio-economic quintile high school is assumed to have less financial support from their family while a student from a higher socio-economic quintile is assumed to have more support from their family. A goodness score of 0 represents a student from Quintile 1 school while a goodness score of 1 represents a student from a Quintile 5 school.

The graduate success measure is the measure of the current system strength. This strength is obtained from the student progression submodel as the percentage of students who graduated in minimum time. A score of 0 represents that there were no graduates in minimum time while a score of 1 represents all students graduated in minimum time.

The final goodness score only considers second and third year modules. First year students were excluded to control for the homogeneity of the system. First year students enroll for generalised modules and only after first year would students begin to specialise in their specific degree programmes, for the three programmes selected for analysis in this study. In Table 4.10 the inital values for the systemic support submodel can be shown per degree programme.

| Determinant | BCom (Economic Sciences) | BCom (Management Sciences) | BCom (Mathematical Sciences) |
|---|---|---|---|
| Academic preparedness | 0.58 | 0.58 | 0.64 |
| Cohort size | 0.59 | 0.59 | 0.65 |
| Degree difficulty | 0.38 | 0.38 | 0.36 |
| Graduate success | 0.22 | 0.34 | 0.18 |
| Lecturer effectiveness | 0.79 | 0.79 | 0.77 |
| Socio-economic status | 0.97 | 0.98 | 0.98 |
| Social support | 0.72 | 0.69 | 0.73 |
| Student engagement | 0.54 | 0.54 | 0.52 |
| University resources | 0.78 | 0.78 | 0.75 |

TABLE 4.10: *The initial values for the determinants to describe systemic support withing the EMS Faculty of SU.*

## 4.3 Chapter summary

In this chapter the various datasets used to populate the initial values and input data to quantify student progression, graduate success, and systemic support in SU EMS were described. The data used to measure student progression were received from the Centre for BI which contained records of the students' home language, degree programme, focal area, final achievement mark among other information. The data used to measure systemic support were a combination of the data received from the Centre for BI and the student feedback forms. A goodness score was calculated for each of the determinants to normalise the metrics of disparate units.

CHAPTER 5

# Model formulation

In this chapter the Student Progression Model is introduced. This hybrid model consists of the student progression submodel using an agent-based simulation modelling approach and a systemic support submodel using a system dynamics simulation modelling approach. The submodels are brought together to form the hybridised whole for analysis of graduate success. This chapter describes the logic, algorithms, and equations of the simulation model at the heart of this study.

## 5.1 Simulating student progression

The Student Progression Model (SPM) makes use of a hybrid paradigm to combine agent-based and system dynamics simulation modelling and the model is constructed on the AnyLogic workbench. AnyLogic [3] is a multi-objective simulation tool which supports the design, implementation, and intregation of multiple simulation paradigms within a single model.

The underlying scripting language of AnyLogic is Java Script. The main environment contains the agents and their environment. An agent can exist only within one state at any given time. The agents transition between states according to specific conditions. These conditions are criteria where each condition causes an agent to either remain in their current state or to transition to another state.

When the SPM is initialised agents with specific attributes are generated within the appropriate state. Various objects, functions, and variables are used to characterise and initiate the agents. These features are described in Table 5.1.

| **Name of Feature** | **Description** | **AnyLogic icon** |
|---|---|---|
| Object class | The object class represents the agent structure. | |
| Parameter | Parameters are used to assign attributes to agents. | |
| Function | Functions are used to perform calculations, which would return an value or or issue a command. | |
| Event | Events are used to execute an action at a specific time during model runtime. | |
| Variable | Variables are used as temporary storage for expressions that change over time. | |

TABLE 5.1: *The different AnyLogic features used to generate agents and simulate agents within the SPM.*

### 5.1.1 Agent attributes

A student is modelled as an agent and is assigned attributes at model initialisation. Each agent is assigned a degree programme that is the degree qualification in which they are progressing towards graduation. For example, a student may be enrolled to obtain a BCom degree in Mathematical Sciences.

For each degree programme, each agent is also assigned a focus which is the focal area and major the student has declared. Along with the focal area, the agent is assigned an option, sub-option, and sub-sub-option parameter which is the specific combination of modules they can select to obtain the credits for their degree. The sub-option and sub-sub-option parameters together form the unique number used to identify the specific combination of modules the student has selected for each of the required academic years, respectively. For example, a student in the BCom (Mathematical Sciences) programme might have Operations Research (OR) as their focus along with their specific combination of modules selected throughout their three years of study to achieve this focus. The option parameter would contain the unique value of the combination of elective and compulsory modules the student selected for first year, the sub-option parameter would be the combination of modules selected for second year based on the electives selected for the prior academic year, and the required parameter would the combination of modules selected in third year based on the modules combination selected in the prior academic years. Each student is initialised with their degree, focus, option, sub-option, and sub-sub-option at the start of the simulation through the initialisation event that is triggered only once.

Additionally, each student is assigned an academic and historic year which updates dynamically as the student progresses during the simulation period. Of all of the modules that a student is enrolled for at any time during the simulation, the academic year records the lowest year of study in which the module is presented. For example, OR 214 is presented in the second year of study and OR 314 is presented in the third year of study. If a student failed OR 214 previously and is currently enrolled for OR 214 and OR 314, their academic year is equal to a value of 2. The student's historic year is the number of simulation years they have spent thus far towards graduation. For example, if a student is registering to study for the third time, their historic year is equal to a value of 3. Another example would be if a student is in their fourth year of enrolment for a three-year degree programme so that their historic year is equal to a value of 4.

Each agent is assigned four more attributes to record their selected modules, realised modules, passed modules, and historically passed modules, respectively. Each of these attributes are arrayed parameters and variables. The module parameter lists all of the modules the student has selected for their current academic year based on the value of their option parameter. The module parameter therefore lists all of the compulsory modules the student must pass before they progress to a higher academic year.

The realised modules parameter lists a combination of the modules the student has selected for their current academic year and those expected for their historic year. For example, if a student failed OR 214, they are unable to enrol for OR 326 as one of the requisite modules. They would need to enrol for OR 214. The student would be able to enrol for other third year modules such as Project Management 314 and Mathematics 314 as OR 214 is not a requisite module. The realised modules parameter accommodates for this requirement and only the possible modules given the student's academic year, historic year, and the modules they have passed.

The passed modules parameter is associated with the student's current academic year and lists all the modules they passed during the year. If, for example, the student failed OR 214 and then enrolls for a few third year modules (as listed in the realised modules parameter) and successfully passed all of these modules, their passed module variable would reflect this. The historical passed modules variable contains the final achievement mark of the student for the duration of their degree programme. These variables update once for every simulation year by means of an event.

The student's HEMIS score is also calculated for each simulation year. This score determines whether the student will transition to the next academic year state or remain within the same year. The HEMIS score is the sum of all the credits for the passed modules in the student's current academic year (as a variable representing the total HEMIS) as a fraction of the sum of all the credits for the required modules in the current academic year (as a variable representing the required HEMIS). If a student has a value of 1 they would progress to the next academic year or else they would remain within the same academic year.

Each agent is characterised by eleven parameters, three variables, and three events. These characteristics are summarised in Table 5.2. An array is a single-dimensional collection of elements in AnyLogic, while a hyperarray is a multidimensional array that allows for more complex data storage and representation.

### 5.1.2 The state chart and state transitions

In the student state chart the necessary parameters, potential states, and underlying logic to enable a student to transition between different states are defined. Every student agent is intialised into the state representing their academic year which could be any one of the states representing first year, second year, third year, graduation, or attrition. A student cannot be initialised into multiple states and exist only within one state at a time, so that the student's current state represents their current academic year. A student's current state is determined by their list of required modules for each year of their degree programme. If, for example, a student failed only one of their modules in their first year, they would remain in the state containing first year students until they pass this module. The student is able to enrol for the following year's modules if the module they failed is not a prerequisite pass module. The different states are contained in the state chart shown in Figure 5.1.

For a student to transition to a state representing a higher year, they must pass all the required modules for an academic year as described in the Faculty yearbook [54]. A student's state is determined by the lowest year of all of the modules for which they are enrolled. It is therefore possible to bypass states. For example, a student who passes all of their lagging first year modules along with all of their second year modules would transition to the state representing their third year without ever entering the state representing their second year. The different paths by which a student may transition between the various states are determined by their last academic year completed. These conditions are represented by $C_1$ to $C_{10}$ in Figure 5.1. The student does not physically move from module to module and year to year, but rather the change happens through parameter and state changes during the model runtime.

Once the student has reached the graduation state, the output measure can be analysed to determine the time students spent within the system. It is assumed that a student is enrolled in a specific degree and they would either complete the degree or attrit. The student would therefore not change degree programmes.

Figure 5.1 maps the paths for a student to transition between states by the transition conditions. After a student is enrolled, their academic year is set to their first and they transition from the state `Enrolment` to the state `First year` through condition $C_1$. From the `First year` state, if a student successfully passed all their required modules they progress to `Second year` through condition $C_2$. Similarly if the student passed all their required modules in their `Second year`,

| Name of feature | Feature type | AnyLogic icon |
|---|---|---|
| Degree | Hyperarray parameter | Degree |
| Focus | Hyperarray parameter | Focus |
| Option | Parameter | Option |
| Sub-option | Parameter | Sub_Option |
| Sub-sub-option | Parameter | Sub_Sub_Option |
| Academic year | Dynamic parameter | Academic_Year |
| Historic year | Dynamic parameter | HistoricYear |
| Modules | Dynamic hyperarray parameter | Modules [..] |
| Realised modules | Hyperarray parameter | Realised_Req_Modules [..] |
| Passed modules | Hyperarray parameter | Passed_Modules [..] |
| Historical passed modules | Hyperarray parameter | Historical_Passed_Modules [..] |
| Required HEMIS | Variable | Required_HEMIS |
| Total HEMIS | Variable | Total_HEMIS |
| HEMIS score | Variable | HEMIS_Score |
| Initialisation | Event | Initialisation |
| Update historic year | Event | Update_HistoricYear |
| Update modules | Event | Update_Modules |

TABLE 5.2: *A summary of the AnyLogic features used to generate agents and simulate agents within the SPM.*

they progress to the `Third year` through condition $C_3$. After the student has successfully passed all of their `Third year` modules they progress to the `Graduation` state through condition $C_4$.

If a student failed a module they remain in current state of `First year` or `Second year` or `Third year`, depending on the lowest year ranked module they are still to pass. If the student is currently in `First year` state and their academic year is a value of 2 and they have successfully passed the modules they were trailing from `First year` as well as their `Second year` they progress to `Third year` through condition $C_5$. If a student is unable to pass all the required modules or if none of their `First year` modules were a pre-requisite for a `Second year` they follow the normal progress from `First year` to `Second year` through condition $C_2$. If the student successfully passed all the required modules to obtain their degree at the end of simulated year they progress to the `Graduation` state through condition $C_6$. The same rule applies for a `Second year`. If they had second and third year modules for which they were enrolled and

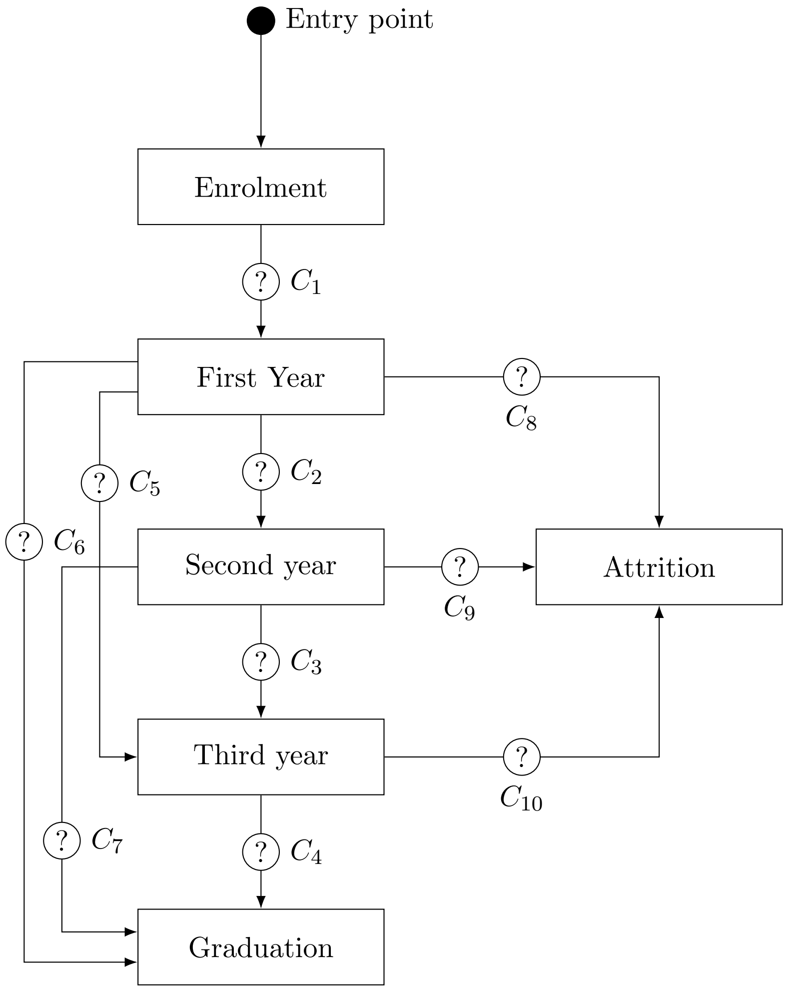



FIGURE 5.1: *The state chart for agents within the SPM.*

successfully passed all the required modules to obtain their degree at the end of the simulated year they progress to `Graduate` state through condition $C_7$.

If a student struggles to complete their year and has not met the minimum HEMIS credits of 50% to register for their next academic year they change states according to conditions ($C_8$), ($C_9$), or ($C_{10}$) depending on their academic year. The student would only be enrolled for a maximum of six years and after six years they would no longer meet the minimum HEMIS credits to continue with their degree programme.

### 5.1.3 Scripting and implementation

Let $S \in [s_1, \ldots, s_i, \ldots, s_n]$ be the total population of enrolled students. Let $\mathbf{D}_i = [d_{1i}, d_{2i}, d_{3i}, d_{4i}, d_{5i}]$ be their choice of path to graduation described by degree $d_{1i}$ with focal area $d_{2i}$ and unique combination of modules in the first year $d_{3i}$, second year $d_{4i}$, and third year $d_{5i}$, respectively that student $i$ is assigned and let $\mathcal{D}_i$ be this array from the import data file where the elements are valued by means of a uniform distribution over all possible options. Let $C_j$ be the total HEMIS credits awarded for each module $j \in [1, 211]$ and $Z_i$ be the percentage of total HEMIS credits achieved by student $i$. Let $A_i \in [1, 2, 3]$ be the academic year for student $i$ which corresponds to the student agent's state in the state chart of the simulation model. Let $H_i \in [1, 2, 3, 4, 5, 6]$

be the historic year for student $i$. Let $\mathbf{E}_i = [e_{1i}, \ldots, e_{ji}, \ldots, e_{211i}]$ be the array of modules that student $i$ will enrol for as they progress towards graduation for their chosen degree programme where

$$e_{ji} = \begin{cases} 1 & \text{if student } i \text{ will enrol for module } j, \\ 0 & \text{otherwise } \forall\ i, j. \end{cases}$$

Let $\mathbf{U}_i = [u_{1i}, \ldots, u_{ji}, \ldots, u_{211i}]$ be the array of unlocked modules student $i$ is allowed to enrol for as a result of module interdependency where

$$u_{ji} = \begin{cases} 1 & \text{if module } j \text{ is unlocked for student } i, \\ 0 & \text{otherwise } \forall\ i, j. \end{cases}$$

Let $\mathbf{P}_i = [p_{1i}, \ldots, p_{ji}, \ldots, p_{211i}]$ be the array of modules student $i$ has successfully passed towards graduation for their chosen degree programme where

$$p_{ji} = \begin{cases} 1 & \text{if student } i \text{ has passed module } j, \\ 0 & \text{otherwise } \forall\ i, j. \end{cases}$$

Let $\mathbf{R}_j = [r_{1j}, \ldots, r_{kj}, \ldots, r_{211j}]$ be the array of prerequisite modules a student must pass before they can enrol for module $j$ where

$$r_{kj} = \begin{cases} 1 & \text{if module } k \text{ is a prerequisite module for module } j, \\ 0 & \text{otherwise } \forall\ j, k. \end{cases}$$

Let $\mathbf{M}_i = [m_{1i}, \ldots, m_{ji}, \ldots, m_{211i}]$ be the array of final marks student $i$ has achieved for module $j$ where $m_{ij} \sim \text{Triangle}(a, b, c)$ if the student completed module $j$ and $m_{ij} = 0$ otherwise $\forall\ i, j$.

Finally let $t$ be the length of the simulation period in years.

In Algorithm 1 each student is initialised with each of the modules they must complete to graduate with their chosen degree. This initialisation depends on their choice of degree, focal area, and the module combination in their respective academic year. Initially, the student has access to all required modules irrespective of their interdependency with pre-requisite modules. For example, a student would be populated with all their required and electives modules they have chosen for their degree programme. A student enrolled for BCom (Mathematical Sciences) with a focal area in Operations Research in their third year, would have an array with active elements where modules are in Operations Research and depending on their selected elective modules, have active elements where modules are in Business Analytics, Financial Mathematics, Project Management, or Mathematical Statistics.

Algorithm 2 determines which students are falling behind and failed a module during the current year and the interdependency constraints of pre-requisite modules are set. The required modules and unlocked modules for each student are determined based on their expected succeeding academic year. Algorithm 1 is called in line 5 to populate all of the modules the student may enrol for during the succeeding academic year. If the student failed any module during the current academic year, this module remains open and unlocked for the student to repeat (as indicated in lines 7 to 9). If any new module for the expected next academic year is lacking some pre-requisite module due to the student's current failed module set, then that module becomes locked and inaccessable within the student's unlocked modules set (as indicated in lines 13 to 16).

**Algorithm 1** Initialise the modules for which each student may enrol

```
function SetModules(A_i, \mathbf{D}_i, \mathbf{E}_i, \mathbf{U}_i)
    for j = 1 to |\mathbf{E}_i| do
        if A_i = 1 then
            if e_{ji} \in d_{3i} then
                e_{ji} \leftarrow 1
                u_{ji} \leftarrow 1
        else if A_i = 2 then
            if e_{ij} \in d_{4i} then
                e_{ji} \leftarrow 1
                u_{ji} \leftarrow 1
        else if A_i = 3 then
            if e_{ji} \in d_{5i} then
                e_{ji} \leftarrow 1
                u_{ji} \leftarrow 1
        else
            e_{ji} \leftarrow 0
            u_{ji} \leftarrow 0
    return \mathbf{E}_i, \mathbf{U}_i
```

**Algorithm 2** Update the list of modules for which each student may enrol

```
function UpdateUnlockedModules(A_i, Z_j, \mathbf{E}_i, \mathbf{U}_i, \mathbf{P}_i, \mathbf{D}_j, \mathbf{M}_i)
    A'_i \leftarrow A_i
    if A_i = 3 then
        A'_i \leftarrow 3
    (\mathbf{E}'_i, \mathbf{U}'_i) \leftarrow SetModules(A'_i, \mathbf{D}_i, \mathbf{E}_i, \mathbf{U}_i)
    for j = 1 to |\mathbf{E}_i| do
        if e_{ji} \neq p_{ji} then
            e_{ji} \leftarrow e_{ji}
            u_{ji} \leftarrow u_{ji}
        else
            e_{ji} \leftarrow e'_{ji}
            u_{ji} \leftarrow u'_{ji}
    for j = 1 to |\mathbf{E}_i| do
        for k = 1 to |\mathbf{D}_j| do
            if m_{ji} \geqslant 40 then
                u_{ji} \leftarrow 0
```

In Algorithm 3 the HEMIS credits that a student has obtained by passing the modules for which they are enrolled is calculated as a fraction of the HEMIS credits the student would have obtained if they had passed all of the modules for which they are enrolled. Therefore, if the student passed all of their modules, their fraction $Z_i$ would be an upperbound value of 1, while if they had passed none of their modules, their fraction $Z_i$ would be a lowerbound value of 0.

**Algorithm 3** Calculate the fraction of HEMIS credits that a student has obtained

**function** CALCULATEHEMIS($C_j, \mathbf{E}_i, \mathbf{M}_i, \mathbf{P}_i$)
    **for** $j = 1$ to $|\mathbf{E}_i|$ **do**
        **if** $e_{ij} = 1 \;\&\; m_{ij} \geqslant 50$ **then**
            $p_{ij} \leftarrow 1$
    $Z_i \leftarrow \dfrac{\sum c_j p_{ij}}{\sum c_j e_{ij}}$
**return** $Z_i$

Algorithm 4 begins by determining the number of years that a student's academic year differs from their historic year. For example, if a student registered to progress towards their degree three times but is still to pass a second year module, this difference is a value of one. The student is in academic year two and historic year three. If this difference is a value of one, then the modules the student is able to enrol for during their next academic year (*i.e.*, academic year three) are unlocked so that the student may progress even while trailing modules from the previous academic year (*i.e.*, academic year two). This is achieved with line 3 in Algorithm 4.

However, if this value is greater than a value of one then, depending on the student's academic year, different modules are unlocked. If the academic year is third year, the student is near graduation and all remaining modules still outstanding towards obtaining their degree are unlocked by means of the procedure call in line 6. If their academic year is second year, the modules that the student may enrol for during their next academic year is again unlocked so that the student may progress even while trailing modules from the previous academic year (where line 8 is as line 3). If their academic year is first year, the student is at risk to not have enough HEMIS credits to complete their degree. The severity of this risk is calculated by determining their HEMIS fraction for their second year modules by means of Algorithm 3.

If the student's HEMIS fraction $Z_i$ is less than half, then only the modules the student is able to enrol for to complete their first and second years are unlocked as it is unlikely that a student trailing so many modules would have access to any modules from third year. If the student's HEMIS fraction $Z_i$ is more than half (but not equal to the total), then the modules the student is able to enrol for during their first, second and third years are unlocked by the procedure calls in lines 14 and 15. If a student's HEMIS fraction $Z_i$ is a value of 1, they have passed all of their second year modules so that the modules they are able to take in their third year is unlocked along with the modules they are still trailing a module or two from from their first year.

Algorithm 5 describes the workflow that is executed for each student agent during each iteration for the duration of the simulation period. If a student failed to secure enough HEMIS credits as determined in line 12, the student attritions from the system and this is modelled by their academic year value being reset to a value of 0. This corresponds to the student agent moving to the `Failed` state in the state chart as seen in Figure 5.1.

**Algorithm 4** Compare academic and historic years to unlock modules

**procedure** UpdateModulesByHEMIS($A_i, H_i$)
  **if** $|H_i - A_i| = 1$ **then**
    UpdateUnlockedModules($A_i + 1, \mathbf{E}_i, \mathbf{U}_i, \mathbf{P}_i, \mathbf{D}_j, \mathbf{M}_i$)
  **else if** $|H_i - A_i| > 1$ **then**
    **if** $A_i = 3$ **then**
      SetModules($A_i, \mathbf{D}_i, \mathbf{E}_i, \mathbf{U}_i$)
    **else if** $A_i = 2$ **then**
      UpdateUnlockedModules($A_i + 1, C_j, \mathbf{E}_i, \mathbf{U}_i, \mathbf{P}_i, \mathbf{D}_j, \mathbf{M}_i$)
    **else if** $A_i = 1$ **then**
      $Z_i \leftarrow$ CalculateHEMIS($C_j, \mathbf{E}_i, \mathbf{M}_i, \mathbf{P}_i$)
      **if** $Z_i < 0.5$ **then**
        UpdateUnlockedModules($A_i + 1, \mathbf{E}_i, \mathbf{U}_i, \mathbf{P}_i, \mathbf{D}_j, \mathbf{M}_i$)
      **else if** $Z_i > 0.5$ **then**
        UpdateUnlockedModules($A_i + 1, \mathbf{E}_i, \mathbf{U}_i, \mathbf{P}_i, \mathbf{D}_j, \mathbf{M}_i$)
        UpdateUnlockedModules($A_i + 2, \mathbf{E}_i, \mathbf{U}_i, \mathbf{P}_i, \mathbf{D}_j, \mathbf{M}_i$)
      **else if** $Z_i = 1$ **then**
        UpdateUnlockedModules($A_i + 2, \mathbf{E}_i, \mathbf{U}_i, \mathbf{P}_i, \mathbf{D}_j, \mathbf{M}_i$)

**Algorithm 5** Simulate student progression

**procedure** Main
  **for** $i = 1$ to $|S|$ **do**
    $A_i, H_i \leftarrow 1$
    $\mathbf{E}_i, \mathbf{M}_i, \mathbf{P}_i.Z_i \leftarrow 0$
    **while** $t \leqslant 3$ **do**
      SetModules($A_i, \mathbf{D}_i, \mathbf{E}_i, \mathbf{U}_i$)
      UpdateModulesByHEMIS($A_i, H_i$)
      **for** $j = 1$ to $|\mathbf{E}_i|$ **do**
        $m_{ji} \sim \text{Triangle}(a, b, c)$
        **if** $m_{ji} \geqslant 50$ **then**
          $p_{ji} \leftarrow 1$
      $Z_i \leftarrow$ CalculateHEMIS($C_j, \mathbf{E}_i, \mathbf{M}_i, \mathbf{P}_i$)
      **if** $Z_i = 1$ **then**
        $A_i \leftarrow A_i + 1$
      **else**
        $A_i \leftarrow 0$
      $H_i \leftarrow H_i + 1$
      $t \leftarrow t + 1$

## 5.2 Simulating system support

In Chapter 3 nine determinants were isolated from literature and their the causal relationships were established. In this section the mechanics of the causal relationship is depicted in Figure 5.2 through a stock-flow digramme. The system support model makes use of SD and the mathematical formulation is given in Section 5.2.1.

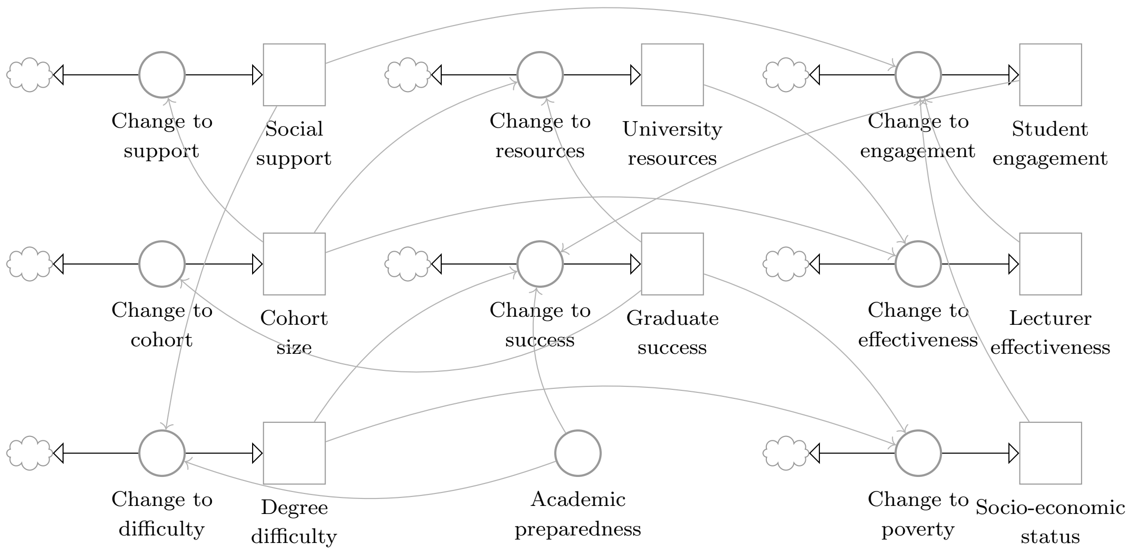



FIGURE 5.2: *The stock-flow diagramme for the system component of the SPM.*

### 5.2.1 System formulation

Let $F^S$ be the goodness score for social support, $F^C$ be the goodness score for cohort size, $F^R$ be the goodness score for university resources, $F^L$ be the score for lecturer effectiveness, $F^E$ be the s goodness score for student engagement, $F^P$ be the goodness score for socio-economic status, $F^Q$ be the goodness score for degree difficulty, and $F^G$ be the goodness score for graduate success, $F^A$ be the goodness score for academic preparedness. The changing values in each variable during the simulation is described by (5.1) to (5.8). The SPM system is driven by averaged linear difference equations to describe the relationships between variables. The calculation of averages requires the parameterisation of the associated weights ($w_S^S$ to $w_A^G$) to ensure the integrity of the causal strength between the variables is modelled correctly.

OECD [36] reported that young higher education graduates have good earning potential compared to those who did not have a post-secondary or non-tertiary education. After about eight to twelve years after graduation a graduate's median annual earnings reaches a comfortable level (at the halfway point of their life-time earnings) [57]. To control for the time before graduates reach this point and are able to change the SES of their community or their progeny, an information delay of an average of ten years is included in (5.6) [21, 44].

$$\frac{dF^S}{dt} = \frac{w_S^S F^P + w_C^S F^C}{w_S^S + w_C^S} - F^S, \tag{5.1}$$

$$\frac{dF^C}{dt} = \frac{w_C^C F^C + w_G^C F^G}{w_C^C + w_G^C} - F^C, \tag{5.2}$$

$$\frac{dF^R}{dt} = \frac{w_R^R F^R + w_C^R F^C + w_G^R F^G}{w_R^R + w_C^R + w_G^R} - F^R, \tag{5.3}$$

$$\frac{dF^L}{dt} = \frac{w_L^L F^L + w_C^L F^C + w_R^L F^R}{w_L^L + w_C^L + w_R^L} - F^L, \tag{5.4}$$

$$\frac{dF^E}{dt} = \frac{w_E^E F^E + w_S^E F^S + w_L^E F^L + w_P^E F^P}{w_E^E + w_S^E + w_L^E + w_P^E} - F^E, \tag{5.5}$$

$$\frac{dF^P}{dt} = \frac{1}{10}\left(\frac{w_P^P F^P + w_G^P F^G + w_Q^P F^Q}{w_P^P + w_G^P + w_Q^P} - F^P\right), \tag{5.6}$$

$$\frac{dF^Q}{dt} = \frac{w_Q^Q F^Q + w_S^Q F^S + w_A^Q F^A}{w_R^Q + w_Q^Q + w_S^Q + w_A^Q} - F^Q, \text{ and} \tag{5.7}$$

$$\frac{dF^G}{dt} = \frac{w_G^G F^G + w_Q^G F^Q + w_E^G F^E + w_A^G F^A}{w_G^G + w_Q^G + w_E^G + w_A^G} - F^G. \tag{5.8}$$

## 5.3 Hybridisation

The SPM amalgamates the student progression submodel of Section 5.1 and systemic support submodel of Section 5.2. In the student progression submodel the number of minimum time graduates in the final outcome of the students who has achieved the `Graduate` state while in the system support model the minimum time graduates are seen in graduate success stock ($F^G$).

The hybridisation of the two submodels is located at the calculation of $m_{ji}$, the final achievement mark for a module. The final achievement mark $m_{ji}$ is amplified and dampened as the system strength ($F^G$) dynamically improves or worsens. A table function was used to return a magnitude by which to either amplify or dampen $m_{ji}$. In Table 5.3 the table function is described, where the function interpolates linearly to return the multiplier value. Let $\mathbf{L}_i = [l_{1i}, \ldots, l_{ji}, \ldots, l_{211i}]$ be the multiplier returned from the table function where $F^G$ is the lookup value. Algorithm 5 is modified in line 9 to include the hybridisation and the hybrid student progression model is described by Algorithm 6.

| $F^G$ **for BCom (Economic Sciences)** | $F^G$ **for BCom (Management Sciences)** | $F^G$ **for BCom (Mathematical Sciences)** | **Multiplier** ($l_i$) |
|---:|---:|---:|---:|
| 0 | 0 | 0 | 0 |
| 0.219 | 0.344 | 0.183 | 1 |
| 1 | 1 | 1 | $\frac{1}{m_{ji}}$ |

Table 5.3: *The hybridisation table function for the SPM.*

**Algorithm 6** Simulate student progression with systemic influence

```
procedure Main
    for i = 1 to |S| do
        A_i, H_i ← 1
        E_i, M_i, P_i.Z_i ← 0
        while t ⩽ 3 do
            SetModules(A_i, D_i, E_i, U_i)
            UpdateModulesByHEMIS(A_i, H_i)
            for j = 1 to |E_i| do
                m_ji ~ l_ji Triangle(a, b, c)
                if m_ji ⩾ 50 then
                    p_ji ← 1
            Z_i ← CalculateHEMIS(C_j, E_i, M_i, P_i)
            if Z_i = 1 then
                A_i ← A_i + 1
            else
                A_i ← 0
            H_i ← H_i + 1
            t ← t + 1
```

## 5.4 Chapter summary

The methodology for the student progression submodel, the systemic support submodel, and the hybridised student progression simulation model is described. The different AnyLogic elements and objects used to generated SPM are explained. For the student progression submodel the student state chart, necessary parameters, potential states, and underlying logic to enable a student to transition between different states are defined. The systemic support submodel's stock-flow diagramme depicts the casual relationship identified in Chapter 3. The stock-flow diagramme is driven by partial differential equations and the hybridisation between the two submodels is presented.

CHAPTER 6

# Validation and results

In this chapter a case for the model's validity is made and model sensitivity is tested. Behavioural replication and structural integrity tests are applied to the student progression and systemic support submodels to illustrate their validity, and the same tests are applied to the hybridised SPM. Sensitivity analysis provides a ranking of the variables most likely to produce significant intervention results. These interventions are performed and insights of recommended policies for improved graduate success are obtained.

## 6.1 Validation of student progression in equilibrium

Model validation confirms whether the structure and behaviour of the model produces a trustworthy replication of the observed behaviour. The modeller must ensure that every element of the real world system is represented within the simulation. During behavioural validation testing, the modeller determines whether the model is designed correctly, while during structural validation testing, the modeller determines whether the design is implemented correctly. As a first test, the independent submodel is validated by soley testing the student progression submodel with the assumption that the systemic support submodel remains at a static equilibrium.

### 6.1.1 Behavioural replication test

The behavioural replication test compares the simulated number of students in the SPM with the observed data as well as the percentage of students that graduated within minimum time. Within the SPM a triangular distribution is used to determine the final achievement marks per student per module. This stochasticity necessitates the use of the Central Limit Theorem (CLT) to determine the required number of simulation runs for the desired confidence per year as calculated in (6.1) where

$$n = \left(\frac{Z\sigma}{V}\right)^2, \tag{6.1}$$

Here $n$ is the required number of simulation runs, $Z$ is the Z-score for the desired confidence level of 95.5%, $\sigma$ is the standard deviation of the number of minimum time graduates, $V$ is the desired margin of error of 5%. After calculating an $n$ value for each year and degree programme, the value ranged from 5 to 10 simulation runs. The maximum of 10 simulation runs was therefore selected for each year and degree programme.

After 10 simulation runs the average percentage of minimum time graduates of these runs were compared to the average observed values per year. The simulated values for BCom (Management Sciences), BCom (Mathematical Sciences), and BCom (Economic Sciences) are acceptably similar to the observed data as shown in Figure 6.1. The root mean square error (RMSE) is used to measure the fit of the data compared. A smaller RMSE value is preferred as it indicates a better estimation of the data and here RMSE values of 5.8%, 5.6%, and 5.8% are obtained for each degree programme, respectively.

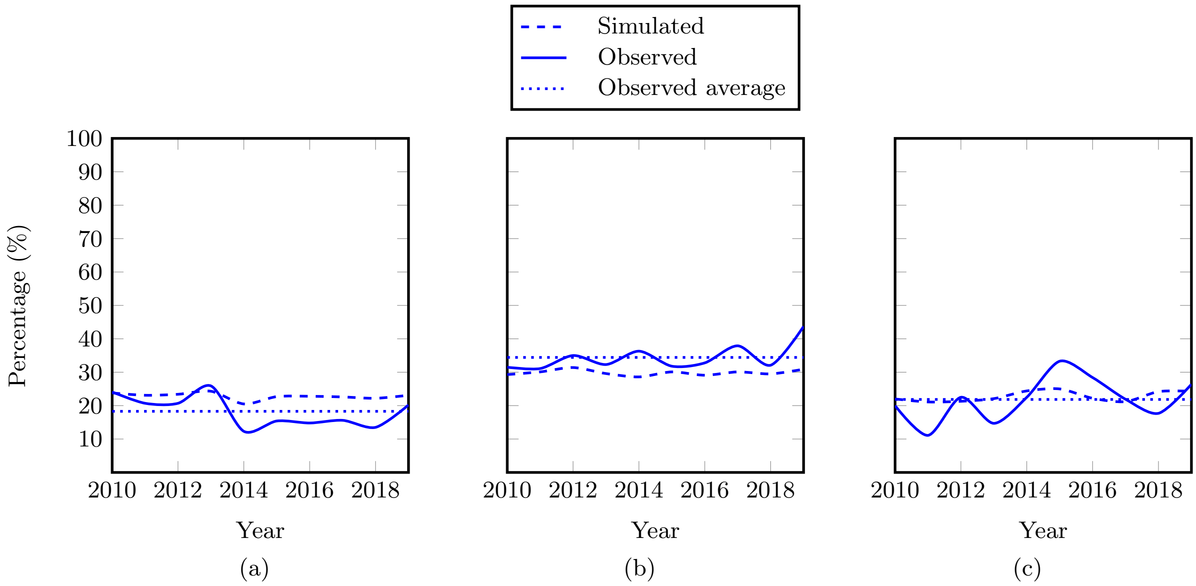



FIGURE 6.1: *The percentage of students obtaining BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively within minimum time in the student progression submodel.*

### 6.1.2 Structural validation test

A model parameter test was conducted to verify that the parameters represent real world observations. In the SPM the validity of the triangular distribution to approximate the average final achievement marks per module was tested. The averaged values were used to decrease the influence of outlier observations.

A further structural validation test was conducted to identify whether there existed errors in the underlying logic of the model. To test the structure of the model the simulated average module final achievement mark was compared to the observed final achievement marks. The simulated values at module level were acceptably similar to the actual data observed as seen in Table 6.1. This would indicate that there were no structural errors for the BCom (Management Sciences) students and so, by induction it may be assumed that the model is valid for all three degree programmes in this analysis too.

| Module | Observed average final mark | Simulated average final mark |
|---|---|---|
| Agricultural Economics 234 | 62 | 61 |
| Business Analytics 214 | 72 | 68 |
| Computer Science 354 | 71 | 65 |
| Economics 114 | 61 | 66 |
| Entrepreneurship And Innovation Management 214 | 57 | 63 |
| Financial Accounting 188 | 59 | 66 |
| Financial Planning 314 | 65 | 64 |
| Industrial Psychology 114 | 61 | 65 |
| Information Systems 112 | 75 | 71 |
| Introduction to Transport and Logistics Systems 144 | 56 | 62 |
| Introduction to Intercultural Communication 344 | 87 | 61 |
| Investment Management 324 | 61 | 65 |
| Logistics and Supply Chain Management 314 | 65 | 65 |
| Logistics and Supply Chain Management 324 | 59 | 64 |
| Marketing Management 214 | 55 | 61 |
| Operations Research 244 | 63 | 65 |
| Philosophy 144 | 72 | 65 |
| Political Sciences 114 | 70 | 67 |
| Public and Development Management 114 | 62 | 61 |
| Statistics 348 | 66 | 64 |
| Transport Economics 214 | 58 | 59 |

Table 6.1: *A sample of the actual and simulated final achievement marks for BCom (Management Sciences) with a RMSE of 6.1%.*

## 6.2 Validation of systemic support in equilibrium

Behavioural replication and structural validity tests are again performed on the second submodel for systemic support. Before these tests could be conducted, a weight calibration experiment was required to determine the best fit weights for each determinant in the system.

### 6.2.1 Weight calibration

In Section 5.2 the stock-flow diagramme for the dynamic higher education system is formalised and the set of partial differential equations is listed in (5.1) to (5.8). This system describes a generic higher education system as formulated in Chapter 3. The question now arises as to the specific causal weighting for the three degree programmes selected for analysis in SU EMS Faculty.

The AnyLogic workbench has an automatic calibration experiment by which the weights for each programme can be determined. The automatic calibration has native optimisation engines which use the Genetic or OptQuest engine. The Genetic engine is an optimisation approach based on the evolutionary algorithm which aims to generate a diversity number of possible solutions during each iteration until the best feasible solution is reached. The OptQuest engine is developed by OptTek Systems Inc which uses a general "black-box" global optimisation algorithm to identify the best possible outcomes with unparalleled efficiency using a variety of algorithms [3]. The calibration experiment determines the best possible values to minimise the difference between the simulation output and the observed output. The RMSE is again used to identify the best possible solution [3]. The user sets the lower and upperbound for the calibration variables, and in the systemic support submodel, the weights are normalised to better analyse their relative importance on the main output variable.

Both optimisation engines were used and resulted in the same RMSE values. In Table 6.2 the resultant weights after calibration are shown. The larger the weight, the more influence and inertia the associated determinant exerts. For example, in all three programmes specific to SU EMS, the students' academic preparedness was found to have little causal influence on their graduate success or their perception of degree difficulty. Social support was found to have little causal influence on students' engagement and motivation but a student is influenced by their lecturers and professors.

Where the super- and subscript symbols are the same, the weight represents the inertia in a variable. If this weight is nearer to a value of 1, then the inertia is great in that the variable resists change away from its current value. However, if this value is nearer to a value of 0, then the inertia is small in that the variable is easily changed away from its current value by the other determinants acting in on it. For example, social support has significantly low inertia in the system for BCom (Mathematical Sciences), but high inertia in the other two systems.

The calibration shows the different magnitudes of causal inference specific to each system modelled. For example, in the system for BCom (Mathematical Sciences) student engagement is primarily a function of lecturer effectiveness, while student engagement for BCom (Economic Sciences) is a function of both lecturer effectiveness and social support.

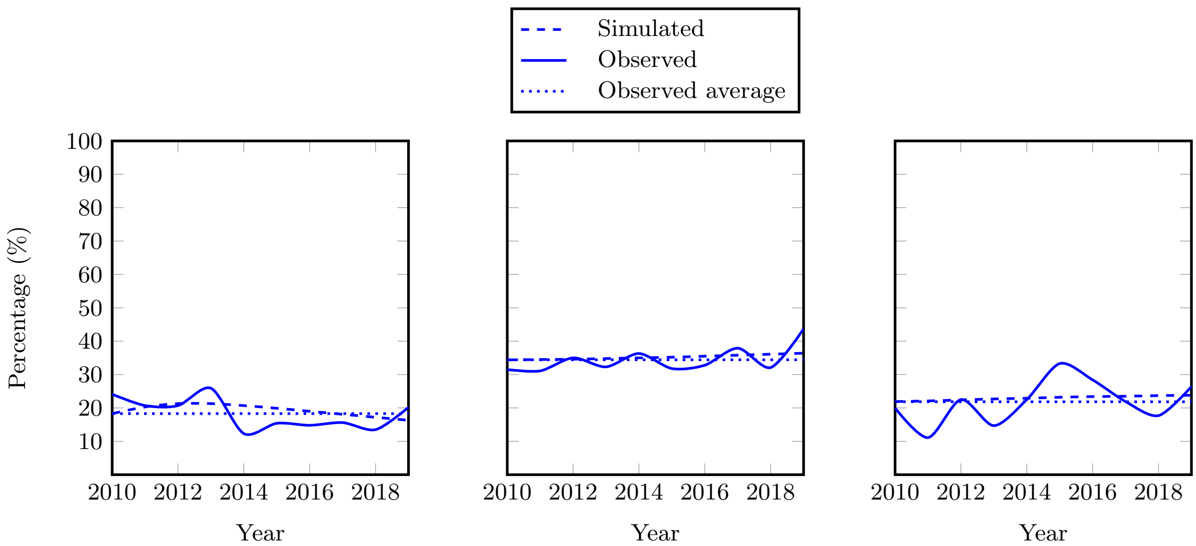
Table 4.10 shows that the initial value for graduate success is the lowest of all initial values determined for the systemic support submodel for all degree programmes, respectively. As such the calibration attempts to maintain the initial value as it is the target variable against which all weights are determined. The easiest way to maintain this value is by placing full weight on the inertia variable. It is therefore reasonable to assume that the determinants of graduate success as mapped in Figure 3.1 might lack some latent variable which better explains the initial value of graduate success, or that the bias in the determination of the initial values for self-reported determinants is unreasonably optimistic.

### 6.2.2 Behavioural replication test

The behavioural replication test compares the simulated percentage of students that graduated within minimum time to the observed percentage of students that graduated within minimum time. The submodel is calibrated by means of the weights reported in Section 6.2.1. The simulated values for BCom (Mathematical Sciences), BCom (Management Sciences), and BCom (Economic Sciences) are acceptably similar to the observed data as shown in Figure 6.2, and the RMSE values of 4.5%, 3.5%, and 6.0% are obtained for each degree programme, respectively.

| Weight | BCom(Mathematical Sciences) | BCom(Management Sciences) | BCom(Economic Sciences) |
|---|---|---|---|
| $w_S^C$ | 0.963 | 0.909 | 0.169 |
| $w_S^S$ | 0.001 | 0.004 | 0.352 |
| $w_R^G$ | 0.378 | 0.555 | 0.001 |
| $w_R^C$ | 0.582 | 0.440 | 0.318 |
| $w_R^R$ | 0.265 | 0.096 | 0.234 |
| $w_E^S$ | 1.000 | 0.001 | 0.001 |
| $w_E^L$ | 0.001 | 0.001 | 0.983 |
| $w_E^E$ | 0.001 | 1.000 | 0.508 |
| $w_E^P$ | 0.001 | 0.001 | 0.001 |
| $w_C^C$ | 0.001 | 0.298 | 0.569 |
| $w_C^G$ | 1.000 | 0.737 | 0.838 |
| $w_G^A$ | 0.001 | 0.001 | 0.001 |
| $w_G^Q$ | 0.001 | 0.012 | 0.003 |
| $w_G^G$ | 1.000 | 1.000 | 0.784 |
| $w_G^E$ | 0.001 | 0.001 | 0.002 |
| $w_L^C$ | 0.575 | 0.427 | 0.001 |
| $w_L^R$ | 0.001 | 0.791 | 0.999 |
| $w_L^L$ | 0.001 | 0.726 | 0.811 |
| $w_Q^S$ | 1.000 | 0.001 | 0.953 |
| $w_Q^A$ | 0.001 | 0.149 | 0.001 |
| $w_Q^Q$ | 0.001 | 0.618 | 0.650 |
| $w_P^G$ | 0.846 | 0.001 | 0.847 |
| $w_P^Q$ | 0.604 | 0.929 | 0.791 |
| $w_P^P$ | 0.052 | 0.001 | 0.766 |

TABLE 6.2: *The calibrated weights for the determinants of graduate success in the systemic support submodel.*




(a) (b) (c)

FIGURE 6.2: *The percentage of students obtaining BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively within minimum time in the systemic support submodel.*

## 6.3 Validation of the dynamic hybrid student progression model

Behavioural replication and structural validity tests are again performed but this time on the hybrid SPM with dynamic interaction between the two submodels.

### 6.3.1 Behavioural replication test

The behavioural replication test compares the simulated percentage of students that graduated within minimum time with the observed percentage of students that graduated within minimum time. The SPM makes use of the hybridised model logic as described in Section 5.3. The simulated values for BCom (Management Sciences), BCom (Mathematical Sciences), and BCom (Economic Sciences) are acceptably similar to the observed data as seen in Figure 6.3 and the RMSE values of 6.0%, 5.1%, and 6.0% are obtained for each degree programme, respectively

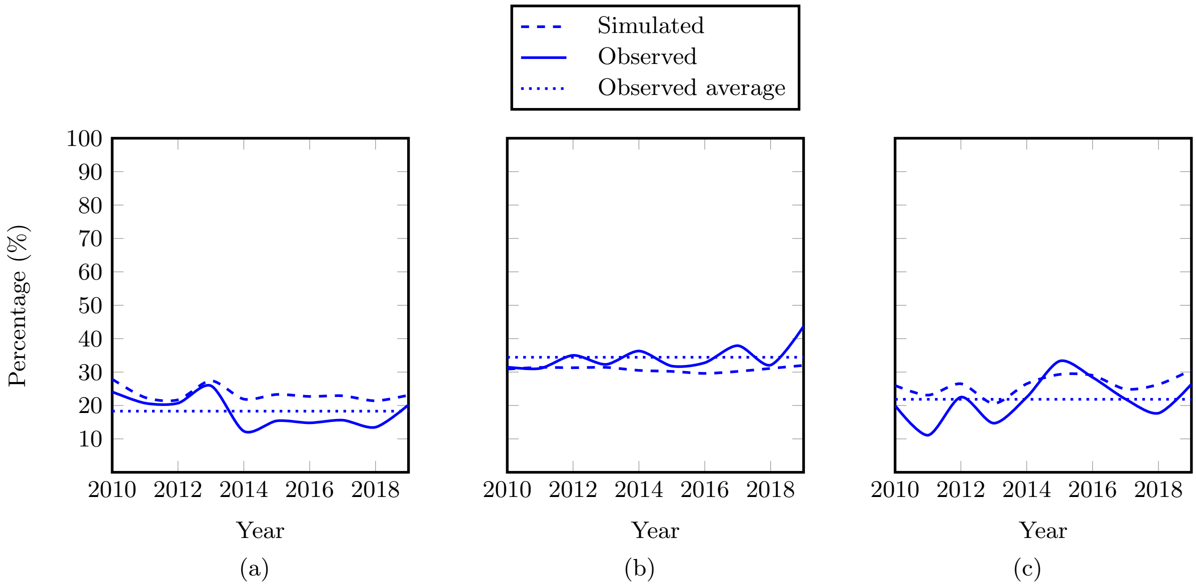



FIGURE 6.3: *The percentage of students obtaining BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively within minimum time in the hybrid student progression model.*

### 6.3.2 Extreme conditions tests

The model is tested under extreme conditions to determine whether expected behaviour under these are replicated. This test exposes any structural faults within the model at the margins of the input dataset [20].

The first test conducted was to initialise all values in the stock-flow diagramme equal to a value of one. This represents a perfectly functional support system where all students should pass within minimum time. In the second extreme conditions test, the inital stock values where initialised to a value of zero. This represents a completely dysfunctional support system where no students would pass within minimum time or pass at all. The behaviour exhibited is as expected and shown in Figure 6.4.

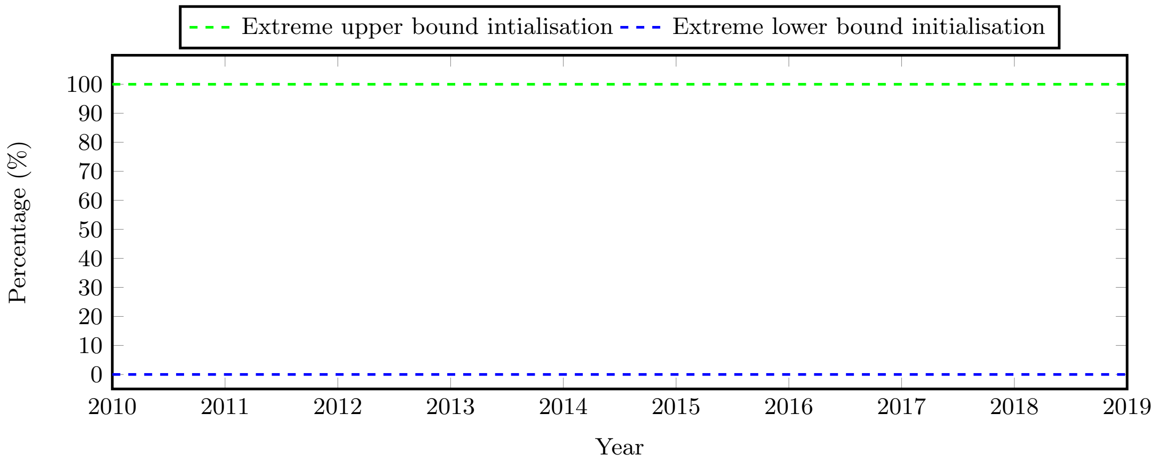



FIGURE 6.4: *Graduation success for extreme conditions where all determinants are at their maximum goodness, and where all determinants are at the minimum goodness, respectively*

## 6.4 Sensitivity analysis

The calibrated weights listed in Table 6.2 include a strong inertia to maintain the initial value for graduate success. This bias for the inertia is therefore stronger than any other determinant which might be analysed for interventions towards improvement in minimum time graduates. The bias could be as a result of a context specific latent variable or unreasonably optimistic initial values on self-reported determinants, thus leading to potential overfitting during the calibration process.

Direct determinants of graduate success are however found with adequate confidence from literature. It is therefore worthy of the analysis to determine which of these would improve graduate success the most for possible policy recommendation. The sensitivity of the graduate success output measure ($F^G$) was analysed through testing the changes to the weights $w_G^E$, $w_G^A$, and $w_G^Q$ in 0.1 increments on the interval [0.1,1]. Table 6.3 shows the sensitivity for each degree programme. The different programmes responded differently to changes in these weights. This behaviour is expected given the differences in the calibrated weights in Table 6.2. A detailed display of the sensitivity is shown in Appendix A.

| **Programme** | $w_G^A$ | $w_G^E$ | $w_G^Q$ |
|---|---|---|---|
| BCom(Mathematical Sciences) | 4.50% | 2.66% | 2.94% |
| BCom(Management Sciences) | 2.17% | 1.78% | 1.73% |
| BCom(Economic Sciences) | 2.76% | 3.41% | 3.39% |

TABLE 6.3: *The average change in graduate success per interval change in the determinant weights for intervals of 0.1 ∈ [0.1].*

The intervention strategies for analysis focus on the determinants that directly influence the graduate success measure. These interventions are decreased degree difficulty, increased academic preparedness, and increased student engagement.

## 6.5 Results and interventions

Figures 6.5 and 6.6 show the base case output for graduate success for the three degree programmes if systemic support remains at the initialised status quo for the simulation period of ten years. The percentage of minimum time graduates for each degree programme over this period is 18%, 34%, and 22% respectively. Three intervention scenarios are analysed for each degree programme in an attempt to increase this metric.

Given the calibrated weights in Table 6.2 and the initialised values in Table 4.10, the determinants draw down to their more realistic base case values to better explain the behaviour observed in $F^G$. $F^P$ is subject delay to such a measure that it approximates the behaviour of an exogenous variable. $F^A$ is modelled as an exogenous variable and as such the behaviour of $F^A$ remains constant. For BCom (Management Sciences), the goodness score for $F^E$ draws up due to the inertia placed on the $w^E_G$ in the calibration process.

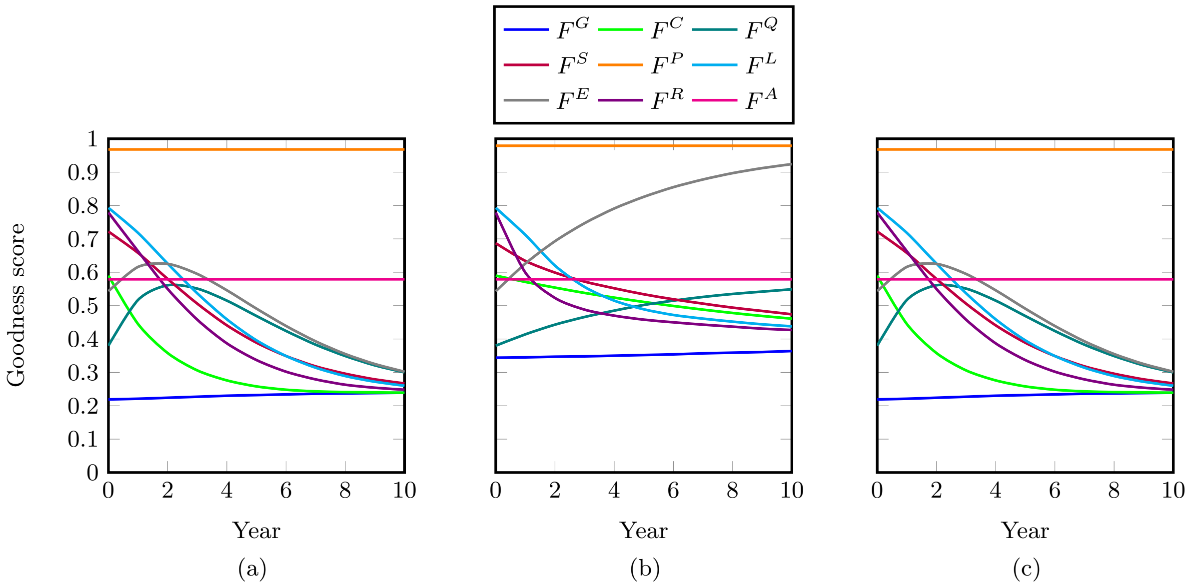



FIGURE 6.5: *Systemic support strengths for BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively for the ten year simulation period at the status quo.*

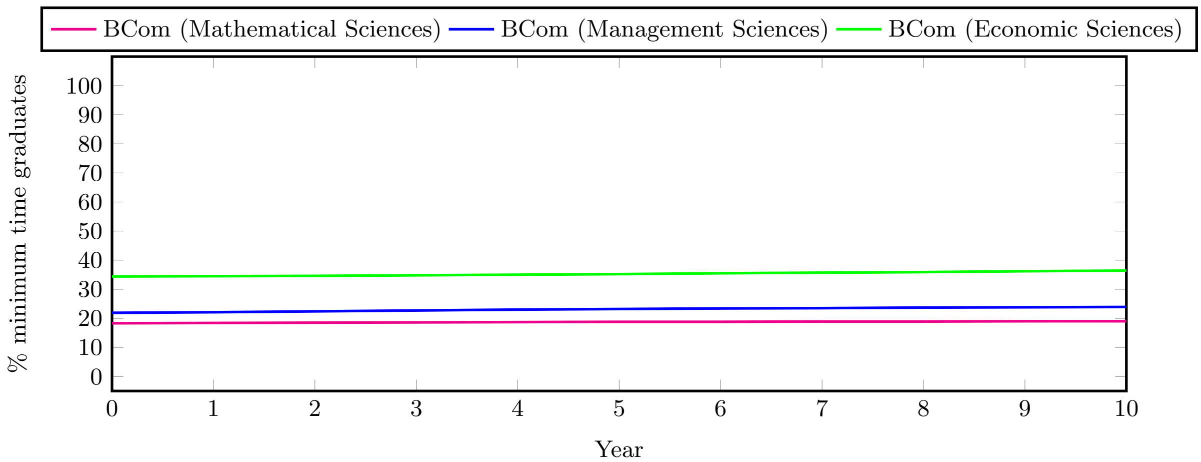



FIGURE 6.6: *Base case results for graduate success in the three degree programmes at the status quo.*

### 6.5.1 The influence of decreased degree difficulty

As a first intervention the influence of decreased degree difficulty on minimum pass time is analysed. In practice this intervention could manifest as the instilling of confidence in students as it is a predominantly subjective perception by the student that could be improved by mentorship, inspiration, improved marketing and expectation management. Figures 6.7 and 6.8 shows the change in graduate success for each degree programme when the initial value for $F^Q$ is set at the maximum goodness score of 1 at the start of the simulation period. The weight $w_G^Q$ is set to a value of 1 so that the impact of $F^Q$ on $F^G$ is comparable to the inertia of $F^G$.

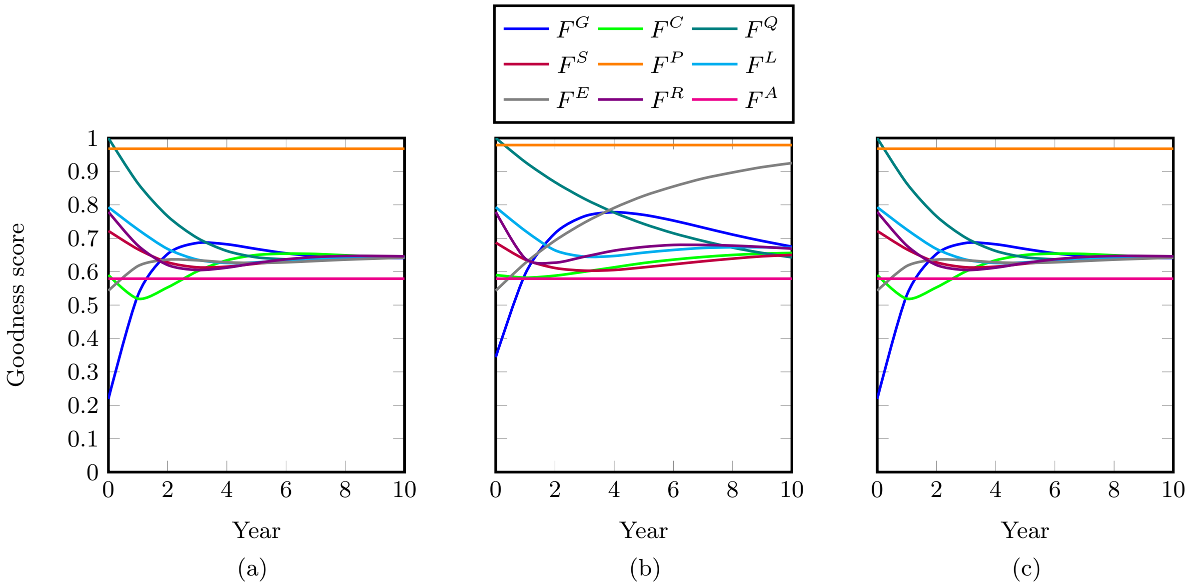



FIGURE 6.7: *Systemic support strengths for BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively for the ten year simulation period for the decreased degree difficulty intervention.*

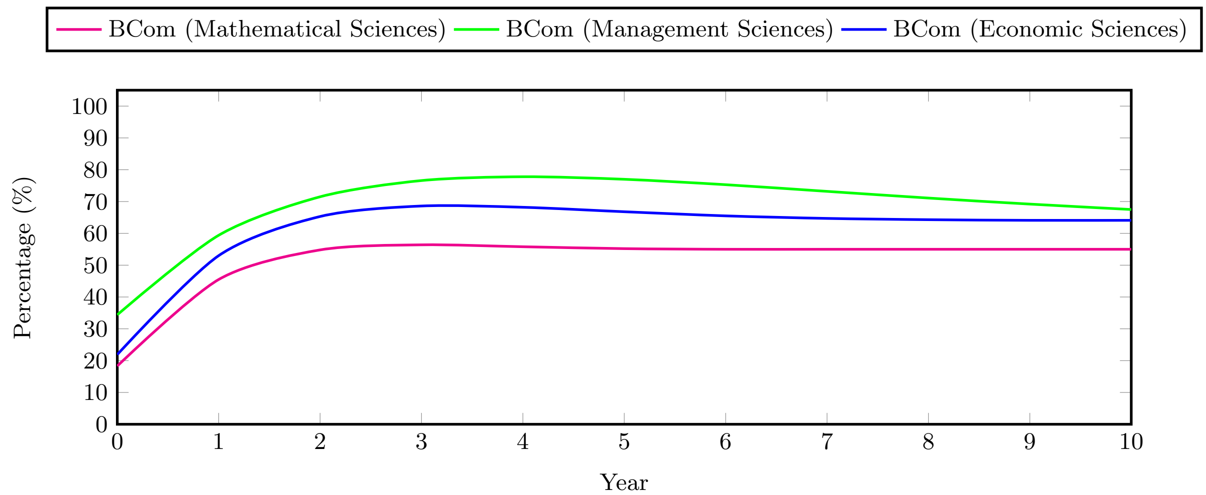



FIGURE 6.8: *The percentage of minimum time graduates per degree programme for the simulation period for the decreased degree difficulty intervention.*

### 6.5.2 The influence of increased academic preparedness

As a second intervention the influence of increased academic preparedness on minimum pass time is analysed. In practice this intervention could manifest as increased admission requirements for all degree programmes. Figures 6.9 and 6.10 shows the change in graduate success for each degree programme when the initial value for $F^A$ is set at the maximum goodness score of 1 at the start of the simulation period. The weight $w_G^A$ is set to a value of 1 so that the impact of $F^A$ on $F^G$ is comparable to the inertia of $F^G$.

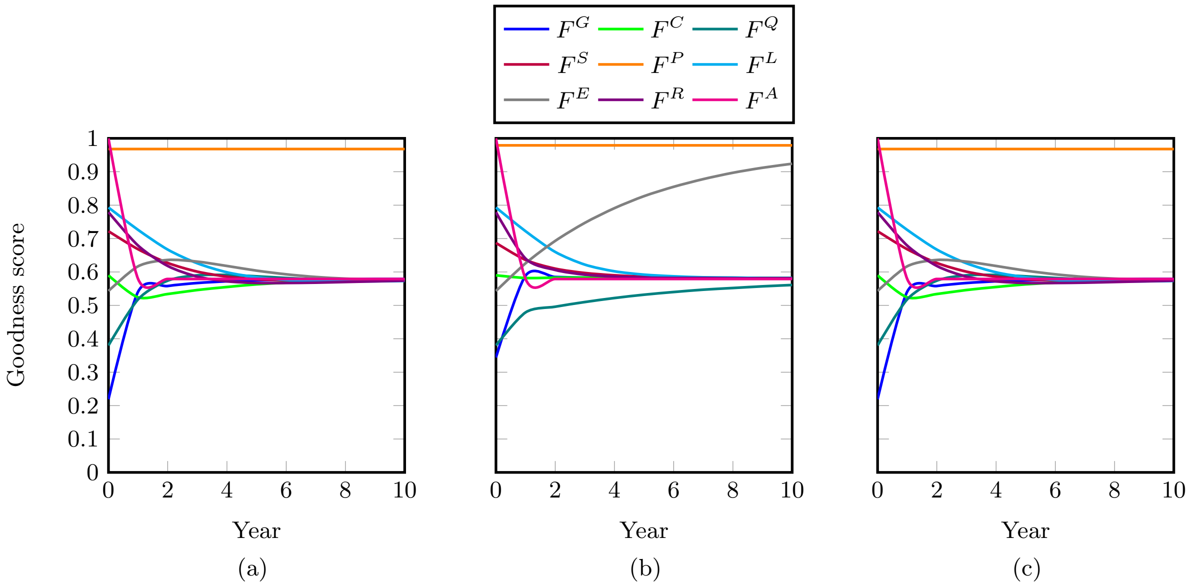



FIGURE 6.9: *Systemic support strengths for BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively for the ten year simulation period for the increased academic preparedness intervention.*

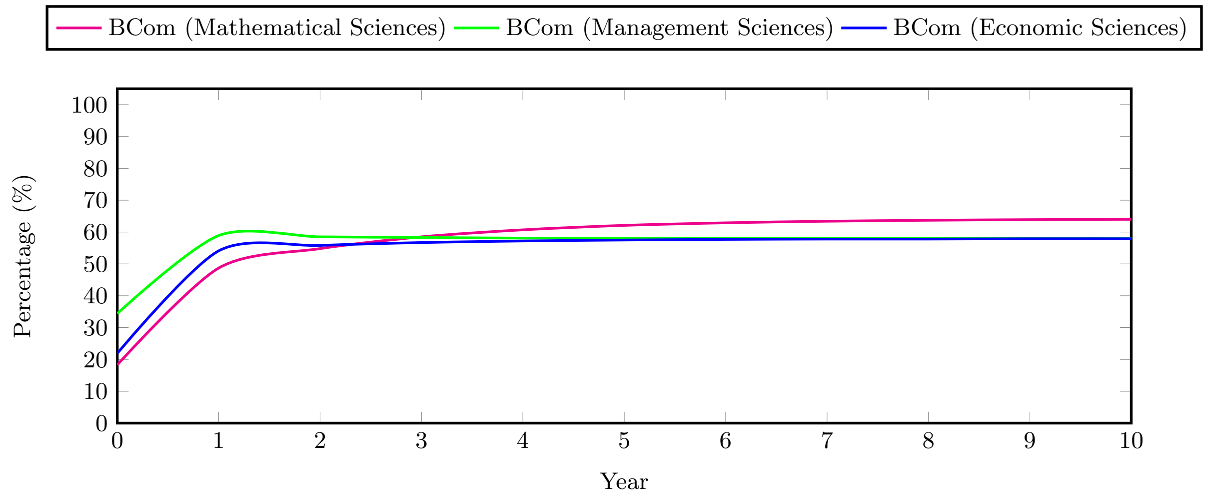



FIGURE 6.10: *The percentage of minimum time graduates per degree programme for the simulation period for the increased academic preparedness intervention.*

### 6.5.3 The influence of increased student engagement

As a third intervention the influence of increased student engagement on minimum pass time is analysed. In practice this intervention could include compulsory tutorial classes, a flipped classrooms teaching approach, or mentor driven study groups. Figures 6.11 and 6.12 shows the change in graduate success for each degree programme when the initial value for $F^E$ is set at the maximum goodness score of 1 at the start of the simulation period. The weight $w_G^E$ is set to a value of 1 so that the impact of $F^E$ on $F^G$ is comparable to the inertia of $F^G$.

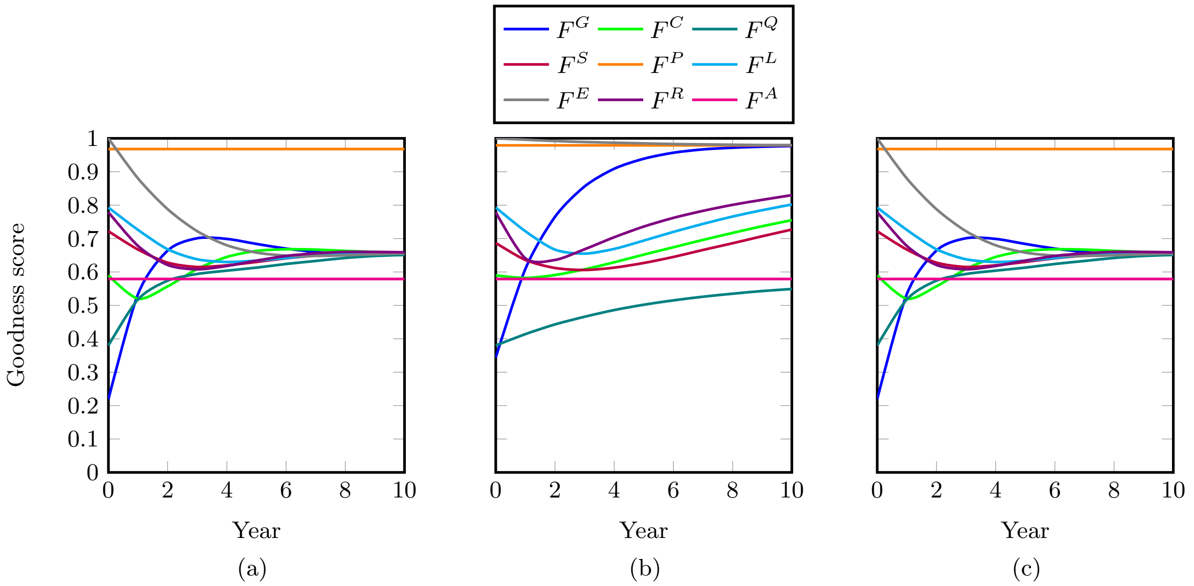



FIGURE 6.11: *Systemic support strengths for BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively for the ten year simulation period for the increased student engagement intervention.*

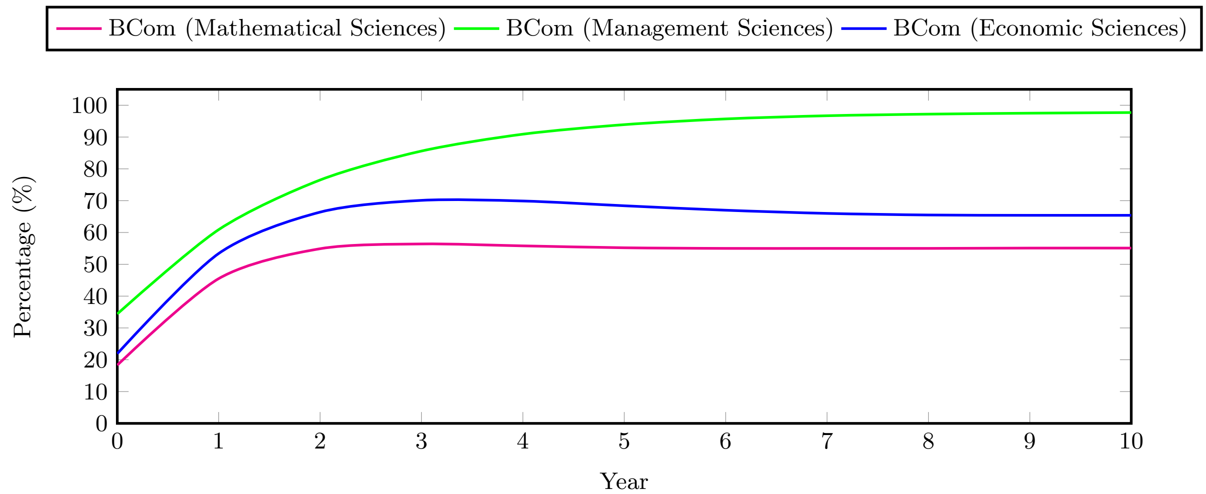



FIGURE 6.12: *The percentage of minimum time graduates per degree programme for the simulation period for the increased student engagement intervention.*

### 6.5.4 The influence of a combined intervention

As a fourth intervention the influence of all three interventions combined on minimum pass time is analysed. Figures 6.13 and 6.14 shows the change in graduate success for each degree programme when the initial value for $F^A$, $F^Q$, and $F^E$ is set at the maximum goodness score of 1 at the start of the simulation period. The weights $w_G^E$, $w_G^A$, and $w_G^Q$ are set to a value of 1 so that the impact of these on $F^G$ is comparable to the inertia of $F^G$.

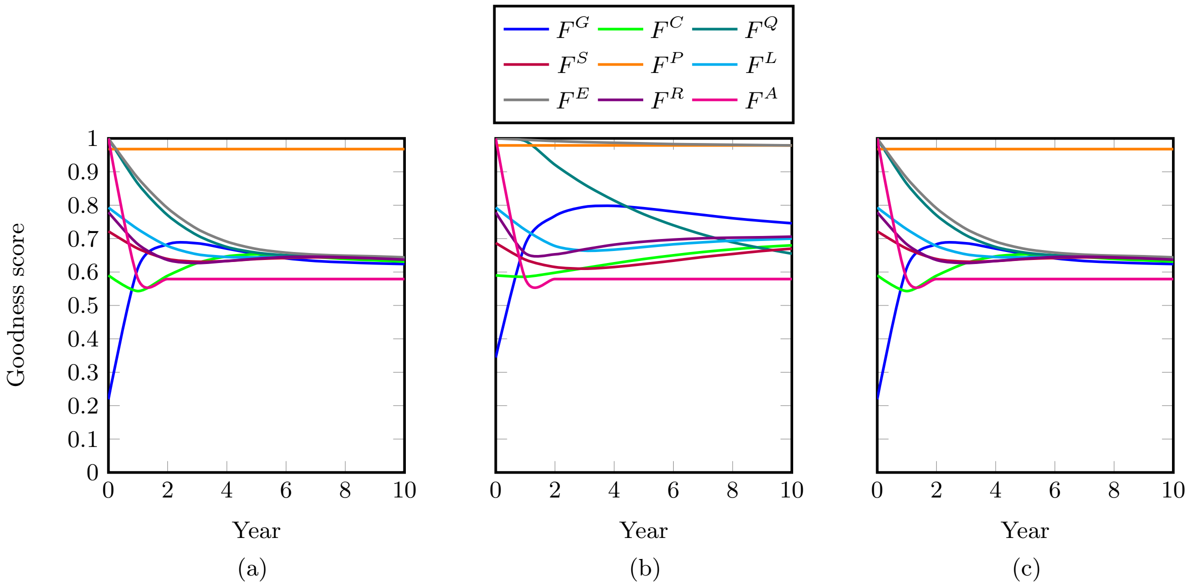



FIGURE 6.13: *Systemic support strengths for BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively for the ten year simulation period for the combination intervention.*

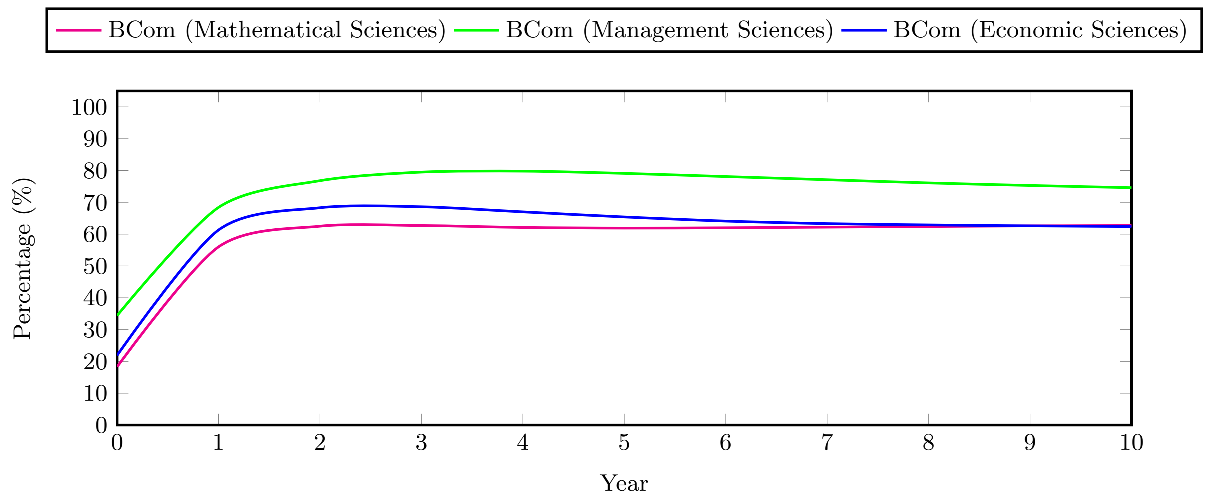



FIGURE 6.14: *The percentage of minimum time graduates per degree programme for the simulation period for the combination intervention.*

### 6.5.5 Summary of results

Mathematical Sciences students appear to be better supported through higher selection criteria while Management and Economic Sciences students appear to be better supported through higher engagement. A summary of the each of the intervention strategy is shown in Table 6.4.

Venter and Vosloo [76] found that a combination approach to intervention strategies is recommended. This allows for the interventions to not only focus on a specific area but rather multiple areas that would result in a better solution. Although a combination approach is best it is not always affordable or practical to implement. They also found that intervention strategies should not only be once off but rather periodic.

In Table 6.4 the large increase in the graduation success is expected from such a large increase in the determinant's initial values. Although degree difficulty and student engagement are self-reported at a goodness of about 0.4 and 0.6, respectively, these determinants realistically have a goodness score of closer to about 0.3. Setting these determinants to an initial value three times that during the interventions will expectedly cause a significant improvement in the resultant graduate success. It is unlikely to ever achieve perfect student engagement and academic preparedness so that the resultant graduation success when applying these interventions may be considered a practical upperbound. Therefore additional interventions on other determinants should also be analysed.

| Intervention | BCom(Mathematical Sciences) | BCom(Management Sciences) | BCom(Economic Sciences) |
|---|---|---|---|
| Base case | 18.3% | 34.4% | 21.9% |
| Decreased degree difficulty | 56.4% | 76.6% | 68.6% |
| Increased academic preparedness | 58.5% | 58.3% | 56.7% |
| Increased student engagement | 56.4% | 85.6% | 70.1% |
| Combination | 62.7% | 79.5% | 68.6% |

TABLE 6.4: *The effect of each intervention strategy on the percentage of minimum time graduates after three years from first convergence to equilibrium.*

## 6.6 Interventions at equilibrium

The three intervention scenarios analysed for the case study of SU are also analysed if the system were in equilibrium and no calibration of the weights where applied, assuming all determinants have the same weighting. When all the weights are set to be equal and given the initialisation values in Table 4.10, the systemic support submodel would produce a higher minimum graduation rate than the observed value. Unlike Section 6.5, the graduate success here draws up to better reflect the output of a system reasonable for the reported strength. The analysis of the uncalibrated system allows for better insight into the specific case study as there are no weight accommodations attempting to reconcile the biased self-reported metrics with the observed metrics. In this way, potential errors of overfitting are avoided.

### 6.6.1 The influence of decreased degree difficulty

As a first intervention the influence of decreased degree difficulty on minimum pass time is again analysed. Figures 6.15 and 6.16 shows the change in graduate success for each degree programme when the initial value for $F^Q$ is set at the maximum goodness score of 1 at the start of the simulation period.

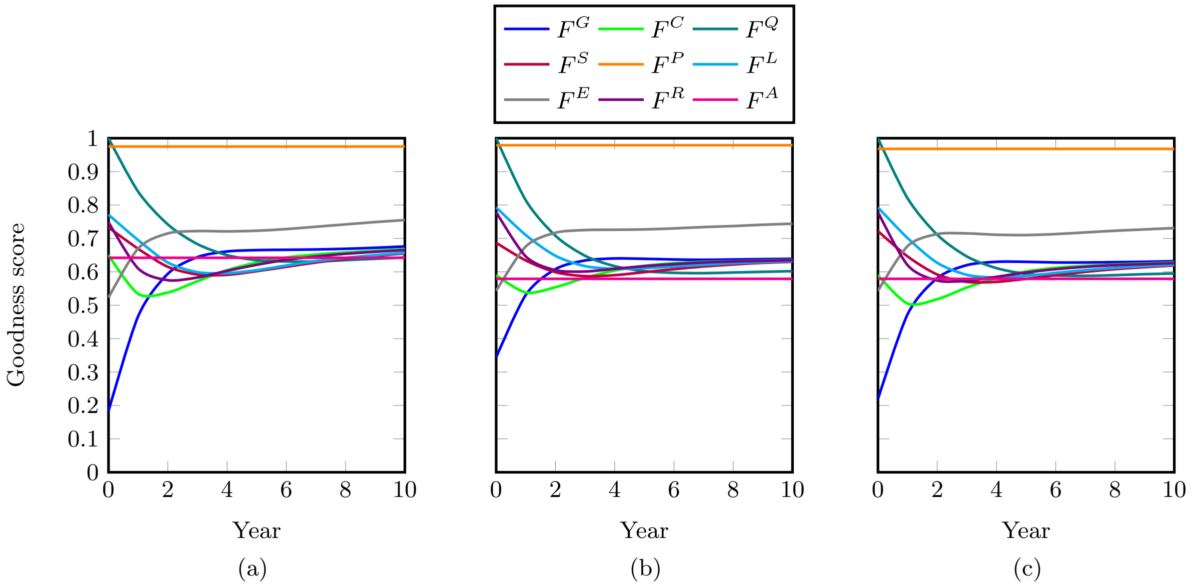



FIGURE 6.15: *Systemic support strengths for BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively for the ten year simulation period for the decreased degree difficulty intervention for systemic support without weight calibration.*

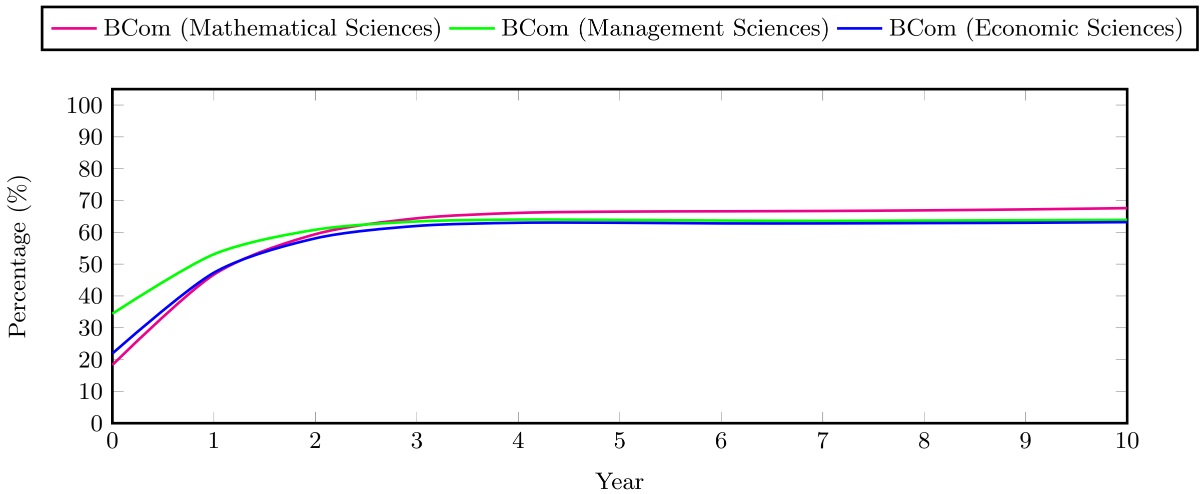



FIGURE 6.16: *The percentage of minimum time graduates per degree programme for the simulation period for the decreased degree difficulty intervention for systemic support without weight calibration.*

### 6.6.2 The influence of increased academic preparedness

As a second intervention the influence of increased academic preparedness on minimum pass time is again analysed. Figures 6.17 and 6.18 shows the change in graduate success for each degree programme when the initial value for $F^A$ is set at the maximum goodness score of 1 at the start of the simulation period.

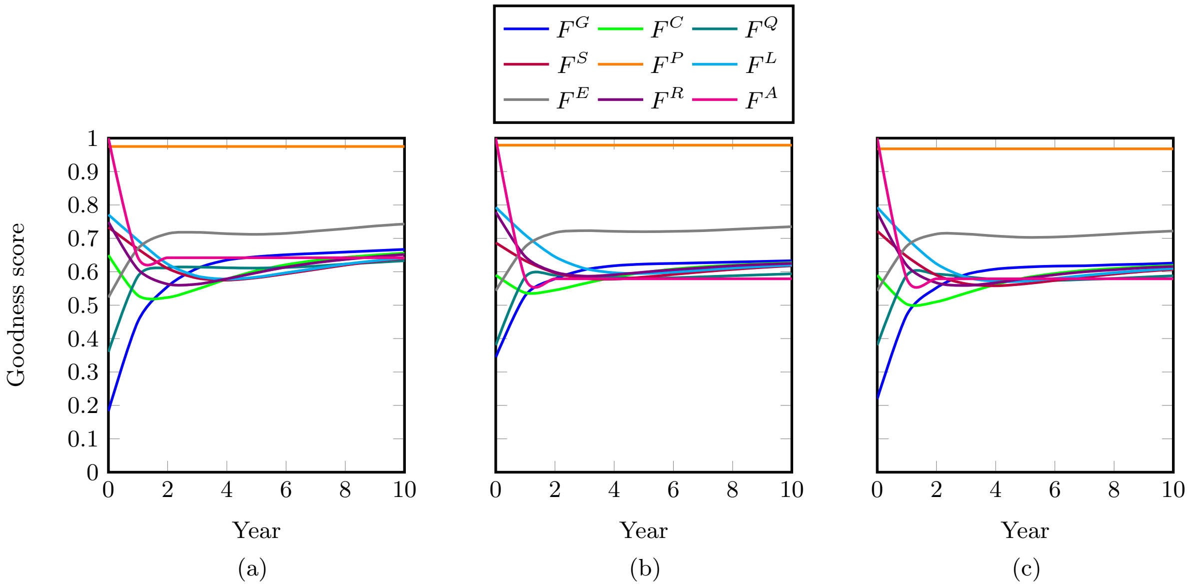



FIGURE 6.17: *Systemic support strengths for BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively for the ten year simulation period for the increased academic preparedness intervention for systemic support without weight calibration.*

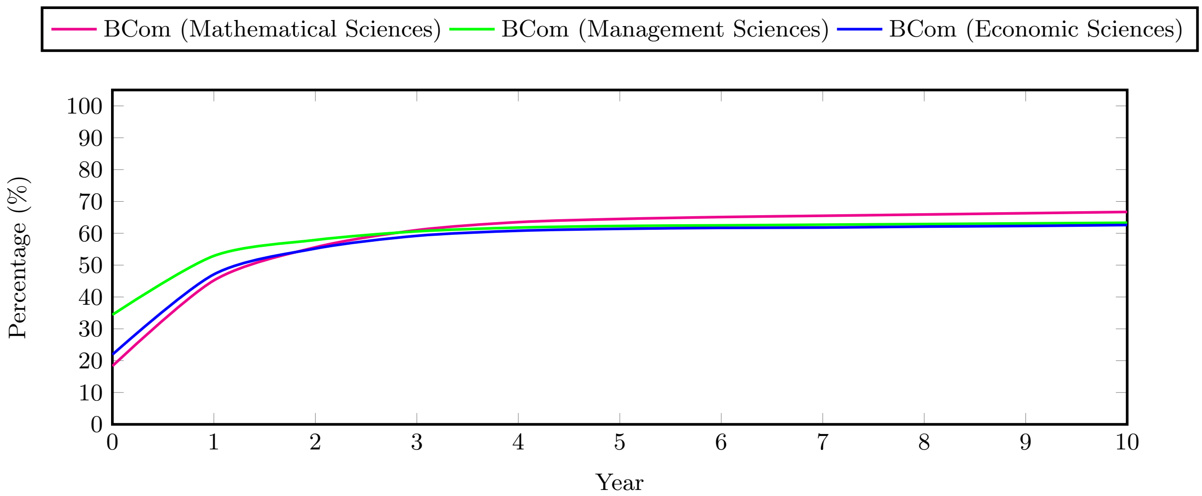



FIGURE 6.18: *The percentage of minimum time graduates per degree programme for the simulation period for the increased academic preparedness intervention for systemic support without weight calibration.*

### 6.6.3 The influence of increased student engagement

As a third intervention the influence of increased student engagement on minimum pass time is again analysed. Figures 6.19 and 6.20 shows the change in graduate success for each degree programme when the initial value for $F^E$ is set at the maximum goodness score of 1 at the start of the simulation period. The weight $w_G^E$ is set to a value of 1 so that the impact of $F^E$ on $F^G$ is comparable to the inertia of $F^G$.

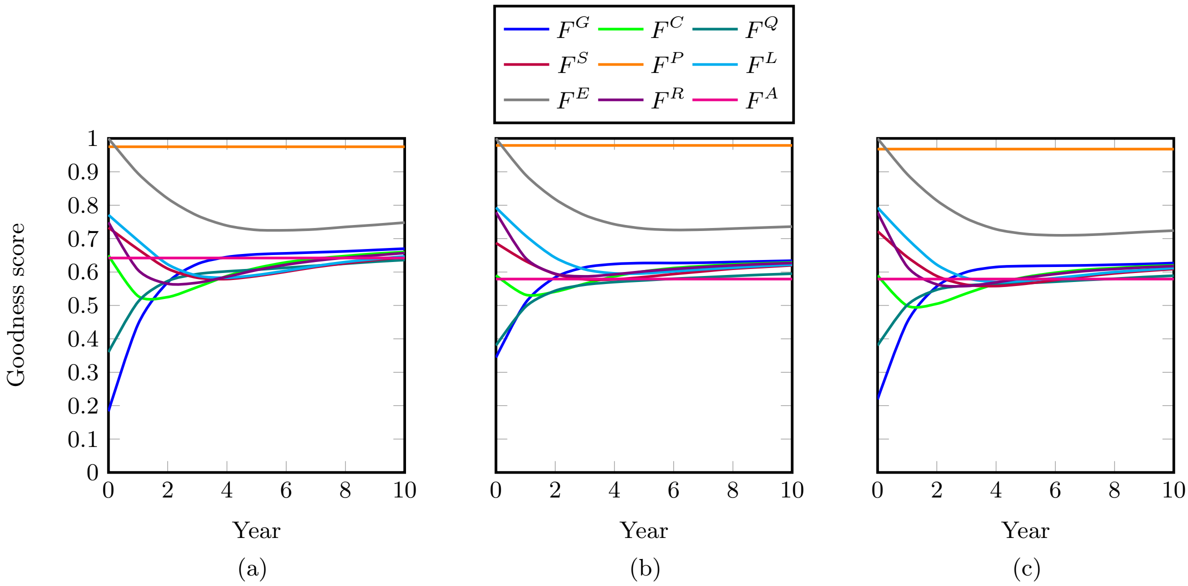



FIGURE 6.19: *Systemic support strengths for BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively for the ten year simulation period for the increased student engagement intervention for systemic support without weight calibration.*

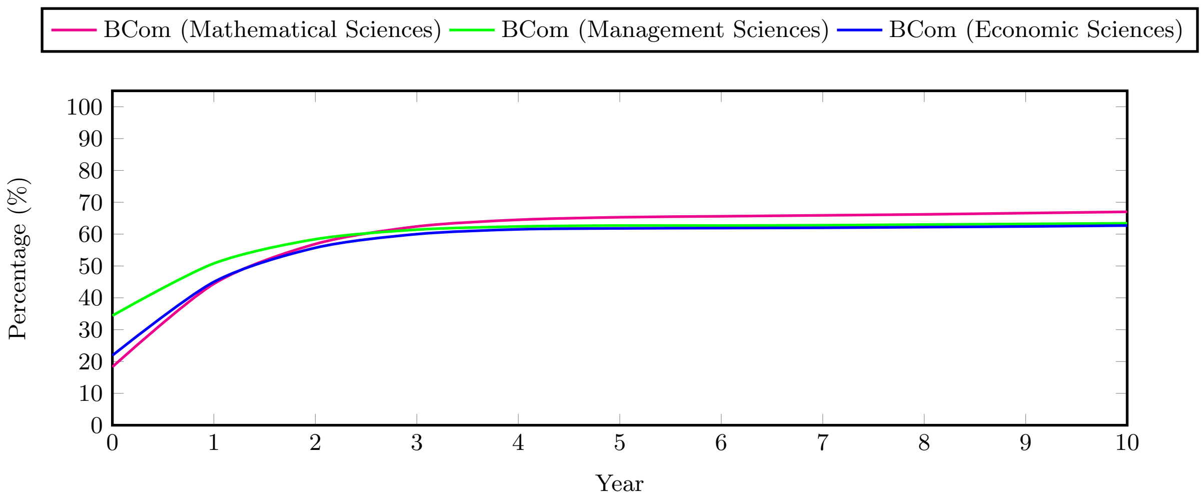



FIGURE 6.20: *The percentage of minimum time graduates per degree programme for the simulation period for the increased student engagement intervention for systemic support without weight calibration.*

### 6.6.4 The influence of a combined intervention

As a fourth intervention the influence of all three interventions combined on minimum pass time is analysed. Figures 6.21 and 6.22 shows the change in graduate success for each degree programme when the initial value for $F^A$, $F^Q$, and $F^E$ is set at the maximum goodness score of 1 at the start of the simulation period. The weights $w_G^E$, $w_G^A$, and $w_G^Q$ is set to a value of 1 so that the impact of these on $F^G$ is comparable to the inertia of $F^G$.

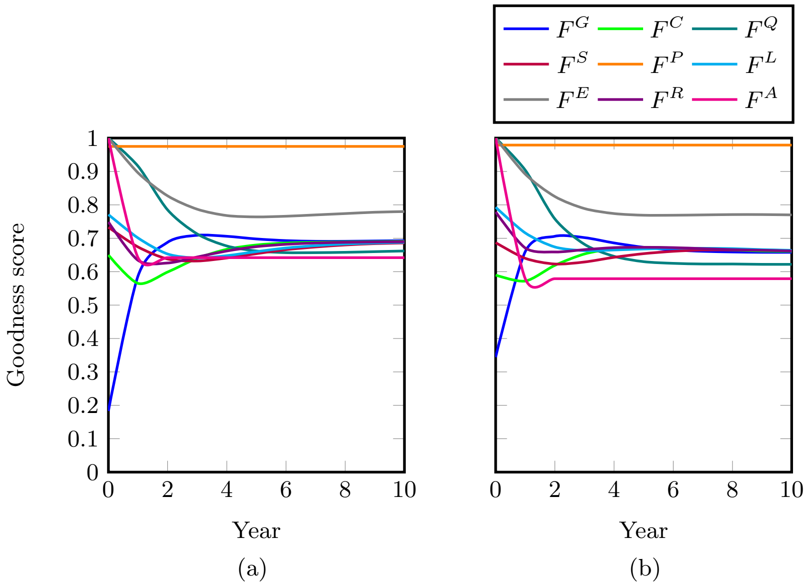



FIGURE 6.21: *Systemic support strengths for BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively for the ten year simulation period for the combination intervention for systemic support without weight calibration.*

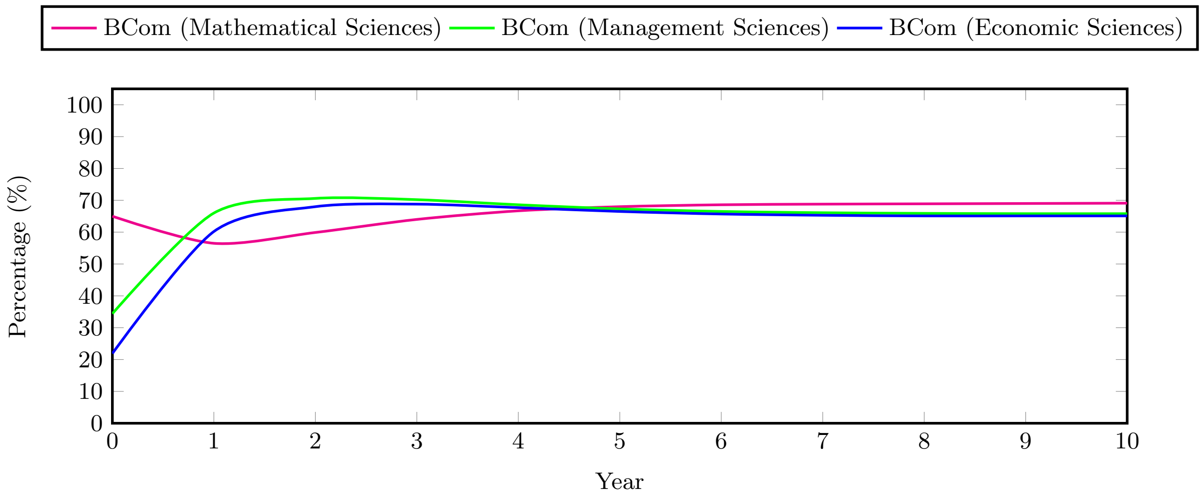



FIGURE 6.22: *The percentage of minimum time graduates per degree programme for the simulation period for the combination intervention for systemic support without weight calibration.*

### 6.6.5 Summary of results

All students appear to be better assisted through the decreased degree difficulty. All students benefit the most from a combination scenario. A summary of each of the intervention strategy is shown in Table 6.5.

| Intervention | BCom(Mathematical Sciences) | BCom(Management Sciences) | BCom(Economic Sciences) |
|---|---|---|---|
| Decreased degree difficulty | 64.4% | 63.4% | 62.0% |
| Increased academic preparedness | 61.0% | 60.6% | 59.2% |
| Increased student engagement | 62.4% | 61.4% | 60.0% |
| Combination | 70.9% | 70.2% | 68.8% |

TABLE 6.5: *The effect of each intervention strategy on the percentage of minimum time graduates after three years where the system strengths converge to equilibrium.*

## 6.7 Chapter summary

This chapter begins by establishing confidence in the SPM. Firstly the output from the student progression submodel, systemic support submodel, and hybrid model are shown to be a valid representation of the observed data. For each degree programme under analysis, the behavioural replication test produced a root mean square error of approximately between 4% to 6% between the averaged observed and simulated values for all three degree programmes. For the systemic support submodel weight calibration produced a unique set of causal inference weights for each degree programme which was then used in the behavioural replication test.

Furthermore an extreme conditions test was conducted on the hybrid model where a perfectly functional system is able to support 100% minimum time graduation in all three programmes under analysis, while 0% graduate in a perfectly dysfunctional system. Additionally a sensitivity analysis was conducted for the determinants that directly influenced graduate success and the resultant behaviour was expected given in the calibrated weights. With these acceptably small approximation errors between the simulation and the observation, the SPM may be assumed to be a valid representation of the system, and may be used to analyse intervention scenarios with confidence.

Four intervention scenarios on three degree programmes where tested to increase the number of minimum time graduates. The first intervention focused on decreasing the influence of the perception of degree difficulty on graduate success, the second was an increase in admission requirements, while the third focussed on the influence of an increase in student engagement, and the final was a combination of these three interventions. These interventions were tested on both cases where the systemic support weights were calibrated to the data, and kept equal for more thorough analysis.

CHAPTER 7

# Conclusion

In this chapter, the model structure and design is summarised and the final results are listed. The achievement of the thesis objectives are described and recommendations for future work arising from this study are presented

## 7.1 Summary of analysis

An introduction of the South African higher education system and the current state of the labour market is provided in Chapter 1. The current methodology of how a student may progress within the SU was examined and a problem statement was formulated. The objectives were set to answer the research questions of which systemic factors influence student progression within higher education and what policies can increase the number of minimum time graduates within the EMS Faculty at SU.

In Chapter 2 a literature review was conducted on the determination of causality in complex systems followed by an analysis of South African basic and higher education system and which techniques and methodology were followed in these analyses. An overview of the simulation techniques considered to model student progression was examined.

In Chapter 3 the various determinants and their influence on graduate success were explored. Later in the chapter the causal relationships between nine determinants were assumed from literature and a causal loop diagramme was constructed. The causal loop diagramme provided a holistic but generic perspective of the determinants that influence graduate success.

A discussion of the different datasets used to populate the initial values and input data for the SPM was presented in Chapter 4. Cleaning and processing of the raw data was discussed along with the quantification of the determinants identified in the causal loop diagramme.

The methodology for a hybrid simulation model is described in Chapter 5. The underlying mechanics and components of each of the submodels is explained as well as hybridisation of the student progression submodel and the systemic support submodel is presented. The model formulation and assumptions were examined.

To increase confidence in the models' results, sensitivity analyses and validity tests were conducted in Chapter 6. A model behaviour test was conducted to compare the simulation results with the real-world observed values. Along with a model behaviour test, a model structure test was examined to test whether the simulation results were as expected. The behaviour and structural validation tests performed as expected to increase confidence in the model. Four intervention scenarios where tested to increase the number of minimum time graduates. These included a decrease in the influence of degree difficulty on graduate success, an increase in admission requirements, an increase in student engagement, an a combination approach of all three interventions scenarios.

## 7.2 Summary of results

The SPM is able to simulate student progression within a higher learning system to answer the research questions. The results discussed in Chapter 6 identified that there is a disparity and bias in the quantification of the initial values for self-reported determinants as these are unreasonably optimistic. This bias necessitated the testing of the four intervention strategies using both the calibrated and non-calibrated determinant weights.

In the calibrated model for Mathematical Sciences students it appears that they are better supported through higher selection criteria while Management and Economic Sciences students appear to be better supported through higher engagement. In the non-calibrated model, all students appear to be better supported through the decreased perception of degree difficulty.

The results identified are specifically identified for the EMS Faculty at SU. The SPM can be populated and calibrated for any other case study (*i.e.,* Engineering at SU or the EMS Faculty at any other higher education institution). While the interventions only focused on the direct determinants that influenced graduate success there are many other interventions that may be applied on any of the other systemic support determinants for insights unique to each use case.

## 7.3 Achievement of objectives

Objective I was achieved in Chapter 2 by listing significant studies and giving an introduction into simulation techniques used for the hybridised approach. In addition, Objective I was further achieved in Chapter 3 by listing the determinants that influence student progression and constructing a generic causal loop diagramme. Objective II was achieved in Chapter 4 by analysing the student feedback forms and the Centre of BI's information to quantify the initial values and rates for each submodel. In addition, Objective III was achieved in Chapter 5 by constructing the SPM. The difference equations and algorithms driving each model and the model schematic is provided in detail. Objective IV is achieved in Chapter 6 through analysis of the validity and sensitivity of the model. Objective V is also achieved in Chapter 6 by testing and analysing the influence of different intervention strategies compared to the base case results. Objective VI is achieved in Chapter 7 where an overview of the findings and future work is discussed.

## 7.4 Future work

The Student Progression Model was modelled as accurately as possible. To construct the model certain assumptions and limitations were observed so that the following suggestions for future expansion is made.

**Proposition I:** This research was focussed on constructing a generic hybrid model of student progression and applying it to the EMS Faculty of SU. The model can be extended to examine various student intersectionalities (*i.e.,* race and gender), programme type (*i.e.,* certificate, associate degree, and bachelor's degree), and attendance type (*i.e.,* part-time or full-time).

**Proposition II:** In addition to expansion, due to limited data from the voluntarily completed student feedback forms, student feedback should be made compulsory to increase the sample size and control for bias. The feedback forms can be handed out during a lecture to ensure that feedback is received from those who attended lectures.

**Proposition III:** This research was focussed on constructing a hybrid simulation model using system dynamics and agent-based modelling. A different simulation approach could be introduced such as discrete event simulation to better model, for example, event driven rather than probability based state changes.

**Proposition IV:** In additional to Proposition I, student progression in other univeristies and faculties could be analysed. For example, comparing a faculty with limited choice modules (*e.g.* Faculty of Engineering) to a faculty that has numerous choice modules (*e.g.* Faculty of EMS).

**Proposition V:** The model could also be adapted to analyse postgraduate student success.

# Bibliography


[1] Adelman C, 1999, *Answers in the Tool Box: Academic Intensity, Attendance Patterns, and Bachelor's Degree Attainment*, U.S. Department of Education.

[2] Angrist J and Pischke JS, 2009, *Mostly Harmless Econometrics: An Empiricist's Companion*, New York: Princeton University Press.

[3] Anylogic, *Simulation Modeling Software Tools and Solutions for Business*, Available from `https://www.anylogic.com/`, Accessed, 18 January 2021.

[4] Astin A, 1984, *Student involvement: A development theory for higher education*, Journal of College Student Development, **40**, pp. 518–529.

[5] Astin A, 1993, *What Matters in College: Four Critical Years Revisited*, San Francisco: Jossey-Bass Publishers.

[6] Bostwick V, Fischer S and Lang M, *Semesters or Quarters? The Effect of the Academic Calendar*, Available from `https://www.k-state.edu/economics/seminars/papers/Bostwick_Valerie_Paper.pdf`, Accessed, 20 February 2021.

[7] Burger A and Naude L, 2019, *Predictors of academic success in the entry and integration stages of students' academic careers*, Social Psychology of Education, **22**, pp. 743–755.

[8] Busato VV, Prins FJ, Elshout JJ and Hamaker C, 2000, *Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education*, Personality and Individual Differences, **29(6)**, pp. 1057–1068.

[9] Cabrera A, Nora A and Castaneda M, 1993, *College persistence: Structural equations modeling test of an integrated model of student retention*, The Journal of Higher Education, **64**, p. 123.

[10] Choy S, 2001, *Students whose parents did not go to college: Postsecondary access, persistence, and attainment. findings from the condition of education*, Available from `https://api.semanticscholar.org/CorpusID:151232779`, Accessed, 28 March 2023.

[11] Coffman JA, 2011, *On causality in nonlinear complex systems: the developmentalist perspective*, pp. 287–309 in Hooker C (Ed), *Philosophy of Complex Systems*, Volume 10 of Handbook of the Philosophy of Science, pp. 287–309. Amsterdam: North-Holland, Amsterdam.

[12] Crede M and Kuncel N, 2008, *Study habits, skills, and attitudes the third pillar supporting collegiate academic performance*, Perspectives on Psychological Science, **3**, pp. 425–453.

[13] Dale SB and Krueger AB, 2002, *Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables*, The Quarterly Journal of Economics, **117(4)**, pp. 1491–1527.

[14] Department of Basic Education, 2022, *ECD School Masterlist Data*, Available from `https://www.education.gov.za/Programmes/EMIS/EMISDownloads.aspx`, Accessed, 10 August 2023.

[15] DesJardins S and Chen R, 2010, *Investigating the impact of financial aid on student dropout risks: Racial and ethnic differences. journal of higher education*, The Journal of Higher Education, **81 (2)**, pp. 179–208.

[16] Donald G, 2013, *Background to the national quintile system*, Available from `https://wcedonline.westerncape.gov.za/comms/press/2013/74_14oct.html`, Accessed, 31 April 2021.

[17] Economics T, 2023, *South Africa Youth Unemployment Rate*, Available from `https://tradingeconomics.com/south-africa/youth-unemployment-rate`, Accessed, 09 April 2023.

[18] Endo J and Harpel R, 1982, *The effect of student-faculty interaction on students educational outcomes*, Research in Higher Education, **16(2)**, pp. 115–138.

[19] Engle J and Tinto V, 2008, *Moving beyond access: College success for low-income, first-generation students*, Washington, D.C.:Pell Institute for the Study of Opportunity in Higher Education.

[20] Forrester J and Senge P, 1980, *Tests for building confidence in system dynamics models*, TIMS Studies in the Management Sciences, **14**, p. 209–228.

[21] Forrester JW, 2009, *Some basic concepts in system dynamics*, Cambridge: Sloan School of Management, Massachusetts Institute of Technology, Accessed 11 June 2021.

[22] Galea S, Riddle M and Kaplan GA, 2009, *Causal thinking and complex system approaches in epidemiology*, International Journal of Epidemiology, **39(1)**, p. 97–106.

[23] Garns R, *Introduction to logic classroom notes: Causation and correlation*, Available from `https://www.nku.edu/~garns/165/ppt9_2.html#:~:text=A%20causal%20generalization%2C%20e.g.%2C%20that,correlation%20between%20the%20two%20properties`, Accessed 11 June 2023.

[24] Gustafsson M and Taylor S, 2018, *Treating Schools to a New Administration: Evidence of the Impact of Better Practices in the System-Level Administration of Schools*, Journal of African Economies, **27**, pp. 515–537.

[25] Hamrick FA, Schuh JH and Shelley MC, 2004, *Predicting higher education graduation rates from institutional characteristics and resource allocation*, Education Policy Analysis Archives, **12**, pp. 19–42.

[26] Hausmann LRM, Schofield JW and Woods RL, 2007, *Sense of belonging as a predictor of intentions to persist among African American and White first-year college students*, Research in Higher Education, **48(7)**, pp. 803–839.

[27] Holland PW, 1986, *Statistics and causal inference*, Journal of the American Statistical Association, **81(396)**, p. 945–960.

[28] Ishitani T, 2006, *Studying attrition and degree completion behavior among first-generation college students in the united states*, The Journal of Higher Education, **77**, pp. 861–885.

[29] Janse van Rensburg R, 2020, *The talent shortages in the South African labour market*, Available from `https://solidariteit.co.za/en/the-talent-shortages-in-the-south-african-labour-market/`, Accessed, 09 April 2023.

[30] Khan K, Ball E, Fox C and Meads C, 2012, *Systematic reviews to evaluate causation: An overview of methods and application*, Evidence-based Medicine, **17**, pp. 137–41.

[31] Kitsantas A, Winsler A and Huie F, 2008, *Self-regulation and ability predictors of academic success during college: A predictive validity study*, Journal of Advanced Academics, **20**, pp. 42–68.

[32] Kuh GD, Cruce TM, Shoup R, Kinzie J and Gonyea RM, 2008, *Unmasking the effects of student engagement on first-year college grades and persistence*, The Journal of Higher Education, **79(5)**, pp. 540–563.

[33] Kuh GD, Schuh JH, Whitt EJ and Associates, 1991, *Involving college*, San Francisco: Jossey-Bass Publishers.

[34] Liefner I, 2003, *Funding, resource allocation, and performance in higher education systems*, Higher Education, **46(4)**, pp. 469–489.

[35] Noyens D, Donche V, Coertjens L, van Daal T and Van Petegem P, 2019, *The directional links between students' academic motivation and social integration during the first year of higher education*, European Journal of Psychology of Education, **34(1)**, pp. 67–86.

[36] OECD, 2020, *Resourcing Higher Education*, Available from `https://www.oecd-ilibrary.org/content/publication/735e1f44-en`, Accessed, 28 August 2023.

[37] Pascarella ET and Terenzini PT, 1991, *How College Affects Students*, San Francisco: Jossey-Bass Publishers.

[38] Pascarella ET and Terenzini PT, 2005, *How College Affects Students (Vol. 2): A Third Decade of Research*, San Francisco: Jossey-Bass Publishers.

[39] Paulsen MB and St John E, 2002, *Social class and college costs: Examining the financial nexus between college choice and persistence*, Journal of Higher Education, **73**, pp. 189–236.

[40] Perrie C and Searle C, 2020, *Analysing the south african junior primary education system with agent-based simulation modelling*, INCOSE International Symposium, **30(1)**, pp. 949–963, Available from `https://onlinelibrary.wiley.com/doi/abs/10.1002/j.2334-5837.2020.00765.x`.

[41] Railsback S and Grimm V, 2019, *Agent-based and individual-based modeling: A practical introduction (2nd ed.)*, New York: Princeton University Press.

[42] Rienties B, Beausaert S, Grohnert T, Niemantsverdriet S and Kommers P, 2012, *Understanding academic performance of international students: The role of ethnicity, academic and social integration*, Higher Education, **63(6)**, pp. 685–700.

[43] Robbins S, Lauver K, Le H, Davis D, Langley R and Carlstrom A, 2004, *Do psychosocial and study skill factors predict college outcomes? A meta-analysis*, Psychological Bulletin, **130**, pp. 261–88.

[44] Roberts N, 2001, *Introduction to computer simulation, A system dynamics modeling approach*, Cambridge: Sloan School of Management, Massachusetts Institute of Technology.

[45] Robledo L, 2013, *A dynamic enrollment simulation model for planning and decision-making in a university*, Doctoral Dissertation, University of Central Florida, Florida.

[46] Schultz TW, 1968, *Resources for higher education: An economist's view*, The Journal of Political Economy, **76(3)**, pp. 327–347.

[47] Shiflet A and Shiflet G, 2015, *Modelling and simulation for the science 2nd Edition*, New York: Princeton University Press.

[48] Shreckengost R, 1985, *Dynamic simulation models: How valid are they?*, National Institute on Drug Abuse Research Monograph, **57**, p. 63–70.

[49] Spaull N and Taylor S, 2014, *Access to what? Creating a composite measure of educational quantity and educational quality for 11 African countries*, Comparative Education Review, **59**, pp. 133 – 165, Accessed, 12 April 2023.

[50] Statista, 2023, *Share of unemployed in South Africa in the third quarter of 2022, by education level*, Available from `https://www.statista.com/statistics/1314504/unemployment-by-education-level-in-south-africa/`, Accessed, 09 April 2023.

[51] Statista, 2023, *Unemployment rate in South Africa from Q1 2016 to Q1 2022, by gender*, Available from `https://www.statista.com/statistics/1129142/unemployment-rate-by-gender-in-south-africa/`, Accessed, 09 April 2023.

[52] Statista, 2023, *Unemployment rate in South Africa from Q1 2019 to Q3 2022, by population group*, Available from `https://www.statista.com/statistics/1129481/unemployment-rate-by-population-group-in-south-africa/`, Accessed, 09 April 2023.

[53] Stellenbosch University, *Annual Integrated Report 2019*, Available from `http://www.sun.ac.za/english/Documents/Annual%20Integrated%20Report/Current/Annual%20Integrated%20Report_eng.pdf`, Accessed, 20 March 2021.

[54] Stellenbosch University, *Faculty of Economic and Management Sciences Academic Programmes and Faculty Information*, Available from `http://www.sun.ac.za/english/Documents/Yearbooks/Current/EMS-Calendar-2021.pdf`, Accessed, 18 February 2021.

[55] Stellenbosch University, *General Policies and Rules*, Available from `https://www.sun.ac.za/english/Documents/Yearbooks/Current/GeneralPoliciesAndRules.pdf`, Accessed, 16 March 2021.

[56] Szafran RF, 2001, *The effect of academic load on success for new college students: Is lighter better?*, Research in Higher Education, **42(1)**, pp. 27–50.

[57] Tamborini CR, Kim C and Sakamoto A, 2015, *Education and lifetime earnings in the united states*, Demography, **52(4)**, pp. 1383–1407.

[58] Tech B, 2020, *Opinion: South Africa battling with shortage of teachers in public schools*, Available from https://businesstech.co.za/news/government/365048/the-best-performing-schools-in-every-province-in-south-africa/, Accessed, 6 May 2021.

[59] The Times Higher Education, *Stellenbosch University World Ranking*, Available from https://www.timeshighereducation.com/world-university-rankings/stellenbosch-university, Accessed, 20 March 2021.

[60] Tinto V, 1975, *Drop-outs from higher education: A theoretical synthesis of recent research*, Review of Educational Research, **45**, pp. 89–125.

[61] Tinto V, 1993, *Leaving college: Rethinking the causes and cures of student attrition (2nd ed.)*, Chicago: University of Chicago Press.

[62] Tinto V, 1997, *Classrooms as communities: Exploring the educational character of student persistence*, **68(6)**, pp. 599–623.

[63] Tinto V and Goodsell A, 1992, *Freshman interest groups and the first year experience: Constructing student communities in a large university*, Washington, D.C.: ERIC Clearinghouse.

[64] Titus M, 2006, *Understanding college degree completion of students with low socioeconomic status: The influence of the institutional financial context*, Research in Higher Education, **47**, pp. 371–398.

[65] UNESCO, *Education sector analysis*, Available from https://learningportal.iiep.unesco.org/en/issue-briefs/plan-for-learning/education-sector-analysis, Accessed 29 July 2023.

[66] University S, *Information Governance About us*, Available from http://www.sun.ac.za/english/InformationGovernance/Pages/About-us.aspx, Accessed, 19 July 2023.

[67] Van der Berg S and Van Broekhuizen H, 2012, *Graduate unemployment in south africa: A much exaggerated problem*, Doctoral Dissertation, Department of Economics, Stellenbosh University, Stellenbosch Economic Working Papers, Accessed, 09 April 2023.

[68] van der Zanden PJ, Denessen E, Cillessen AH and Meijer PC, 2018, *Domains and predictors of first-year student success: A systematic review*, Educational Research Review, **23**, pp. 57–77.

[69] Venter L, 2020, *The Primary school Model*, Doctoral Dissertation, Stellenbosch University, Stellenbosch.

[70] Venter L and Grobler J, 2020, *A systems perspective on teacher training in South Africa*, Working Paper, Stellenbosch: Stellenbosch University.

[71] Venter L and Jansen van Rensburg C, 2018, *Prospective student selection decision support*, Working Paper, Stellenbosch: Stellenbosch University.

[72] Venter L and Passerini A, 2019, *An agent-based simulation model of Early Childhood Development education in South Africa*, Stellenbosch University, Stellenbosch.

[73] Venter L and Rantsane V, 2016, *Assessing student diversity trends with the Student Diversity Target Assessment Model (SDTAM)*, Working Paper, Stellenbosch: Stellenbosch University.

[74] Venter L and Slamang S, 2020, *A systems perspective on public high school management within the Western Cape*, Working Paper, Stellenbosch: Stellenbosch University.

[75] Venter L and Viljoen T, 2020, *A systems perspective on school improvement with a focus on teachers*, Systemic Practice and Action Research, **33**, pp. 265–293.

[76] Venter L and Vosloo M, 2018, *A systems perspective on school improvement approaches*, International Journal of Applied Systemic Studies, **8**, p. 151.

[77] Wakeland WW and Gallaher EJ, 2004, *A comparison of system dynamics and agent-based simulation applied to the study of cellular receptor dynamics*, 37th Hawaii International Conference on System Sciences, p. 10.

[78] Warburton EC, Bugarin R and Nuñez AM, 2001, *Bridging the gap: Academic preparation and postsecondary success of first-generation students. statistical analysis report. postsecondary education descriptive analysis reports.*, Available from `https://ies.ed.gov/pubsearch/pubsinfo.asp?pubid=2001153`.

[79] Wilensky U and Rand W, 2015, *An introduction to agent-based modeling: modeling natural, social, and engineered complex systems with NetLogo.*

[80] Zou T, Zhang Y and Zhou B, 2022, *Does GPA matter for university graduates' wages? New evidence revisited*, PLOS ONE, **17(4)**, pp. 1–16.

APPENDIX A

# Sensitivity analysis

In this section, the sensitivity of determinants that directly influence graduate success was analysed. The various weights of academic preparedness, degree difficulty, and student engagement was tested in 0.1 increments on the interval [0.1,1]. The graphic representation of the influence that each weight has on graduate success is shown in Figures A.1 to A.3.

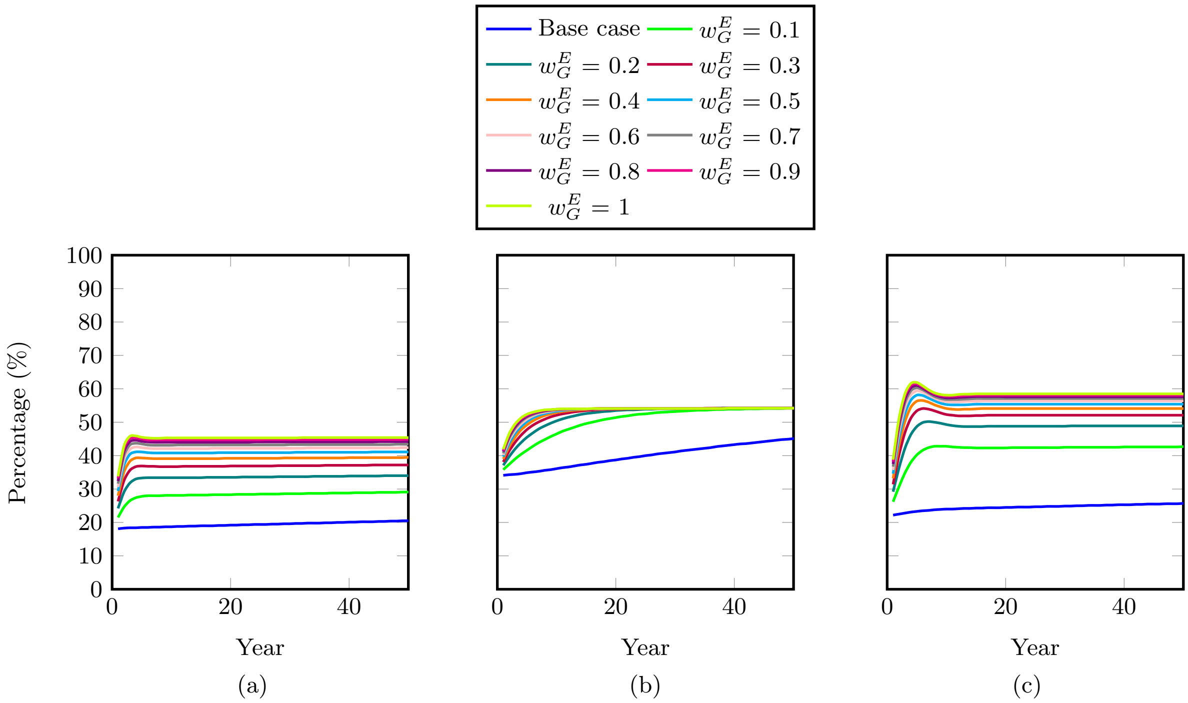



FIGURE A.1: *The sensitivity of the output metric $F^G$ for BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively to the changes in $w_G^E$, all weights stable at the values listed in Table 6.2.*

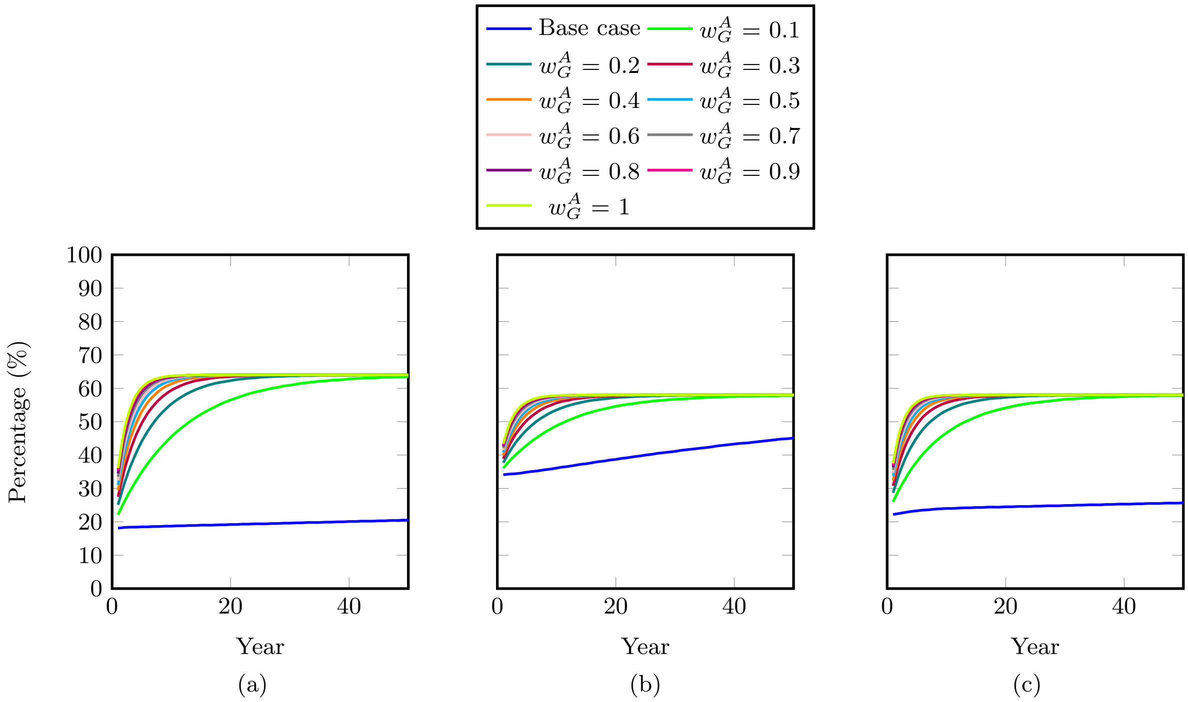



FIGURE A.2: *The sensitivity of the output metric $F^G$ for BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively to the changes in $w_G^A$, all weights stable at the values listed in Table 6.2.*

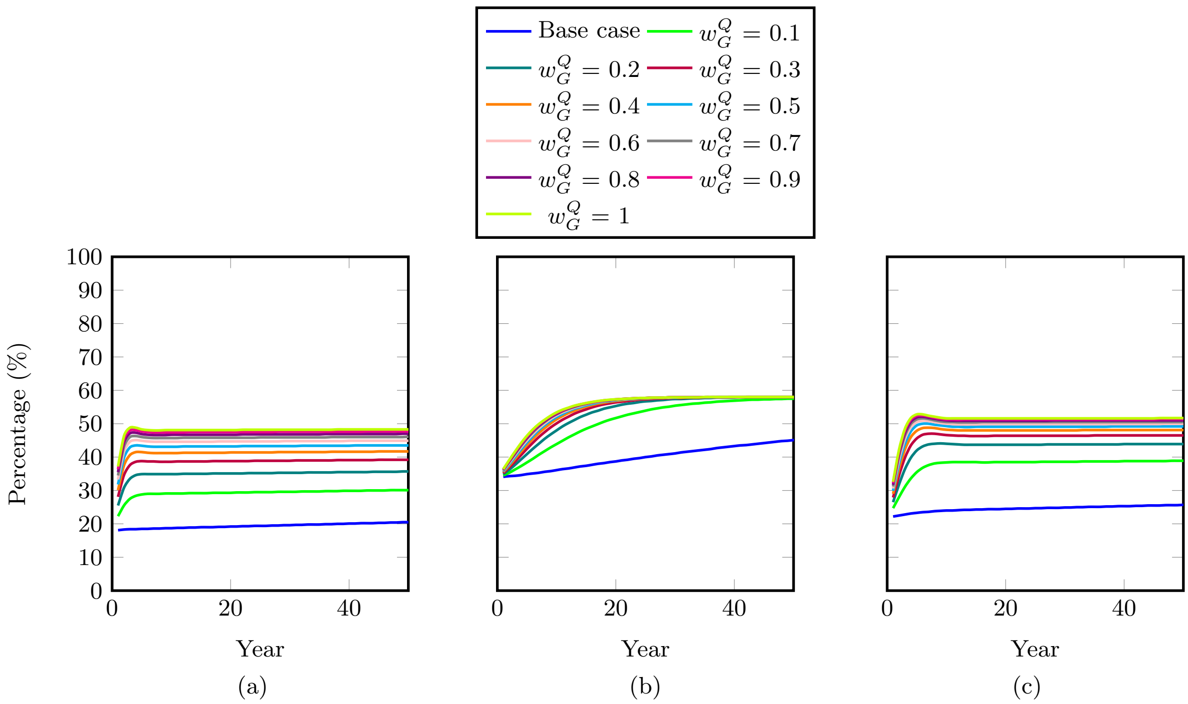



FIGURE A.3: *The sensitivity of the output metric $F^G$ for BCom (Mathematical Sciences) (a), BCom (Management Sciences) (b), and BCom (Economic Sciences) (c), respectively to the changes in $w_G^Q$, all weights stable at the values listed in Table 6.2.*