Using Predictive Risk Analysis to Identify Vulnerable First-Year Students at University: The Importance of NCEA Results

Kamakshi Singh

Department of Economics, Faculty of Business and Law

A dissertation submitted for the degree of Master of Business at Auckland University of Technology, Auckland, New Zealand

2015

Abstract

Much research has highlighted the factors leading to increasing dropouts among first-year undergraduates around the globe. This phenomenon is also an issue in New Zealand. Therefore, this research estimates the importance of various factors, derived from the administrative data provided by the Department of Strategy and Planning at AUT, on the probability of successful course completion at university. Non-completion of first-year courses may form the basis of future non-retention amongst undergraduate students. Efforts to avoid future substantial costs to the university and the government could prove beneficial by identifying factors that result in successful completion of courses as early as possible. This research focuses on first-year students who entered university using valid NCEA Level 3 scores. Majority of the universities in New Zealand, including AUT, have traditionally summarised NCEA results with a composite 'rank' score that arbitrarily assigns weights to Achieved, Merit and Excellence credits without any empirical study supporting the appropriateness of this weighting scheme. This study provides some empirical evidence on the validity of this weighting scheme by estimating the contributions of these different credits in predicting the successful completion of first-year courses. Results from our research also indicate that other factors (e.g., part-time study, gender and the degree programme) may play crucial roles in predicting successful course completion rates. Most importantly, we found that Merit and Excellence credits do not significantly differ in terms of predicting the probability of successful completion of courses. Therefore, we propose an alternative weighting scheme based on this empirical evidence that outperforms the existing NCEA rank score in predicting the successful completion of first-year courses at university.

Table of Contents

1. Introduction 1

2. Literature Review 6

3. Data 14

4. Methodology 21

5. Empirical Results 24

6. Conclusion 34

References 38

Appendix 42

Tables

Table 1: Summary of Some Empirical Studies on Non-Completion and Dropout Amongst University Students 6

Table 2: Definition of all the variables 42

Table 3: Descriptive Statistics 44

Table 4: Model 1 - Probit model parameter estimates and average marginal effects on the dependent variables for all independent variables and best 80 credits broken by Achieved, Merit and Excellence Categories 46

Table 5: Model 2 - Probit model parameter estimates and average marginal effects on the dependent variables for all independent variables and overall score calculated using 2*Achieved, 3*Merit and 4*Excellence points values 48

Table 6: Model 3 - Probit model parameter estimates and average marginal effects on the dependent variables for all independent variables and new overall score calculated using 1*Achieved, 2*Merit and 2*Excellence points values 50

Table 7: Model 4 - Probit model parameter estimates and average marginal effects on the dependent variables for all independent variables and all Achieved, Merit and Excellence credits 52

Table 8: Model 5 - Probit model parameter estimates and average marginal effects on the dependent variables for all independent variables and overall score calculated for all credits using 1*Achieved, 2*Merit and 2*Excellence points values 54

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Kamakshi
26/11/2015

Acknowledgements

First and foremost, I would like to thank my research supervisor Professor Tim Maloney. Without his assistance and dedicated involvement in every step throughout the process, this dissertation would have never been accomplished. I would like to thank you very much for your support and understanding during my MBus studies. The experience of working with you was in all instances interesting, challenging, and invaluable to my life.

My heartfelt thanks also goes to academic staff of the Economics department at AUT University. Particularly, Associate Professor Gail Pacheco, Associate Professor Saten Kumar, Dr Nan Jiang, and Dr Lydia Pikyi Cheung have been very helpful during my MBus programme. Also, thanks to the Department of Strategy and Planning at AUT for providing me with the dataset for my research.

I would also like to thank my parents, Sanjaya and Nirmala, who always believe in me and support me in every possible way in the pursuit of my dream. It is their endless love, constant encouragement, and unconditional support that helps me overcome the obstacles and achieve the seemingly impossible. I am forever indebted to them for the sacrifices they have made to support me throughout the completion of my education.

Finally, to my caring, loving, and supportive brother, Vikrant I offer my deepest gratitude. Your encouragement when times got tough is greatly appreciated and duly noted.

1. Introduction

Over the past decade, worldwide enrolments in tertiary institutions have increased substantially (The World Bank Group, 2015). Increasing enrolments of more marginal students raises the risks of course non-completion at university. Non-completion of courses not only imposes explicit costs such as financial costs to the university and the government (Yorke, 1998), but also implicit costs: opportunity costs are borne by the university and the wider community. From the policy makers' points of view, therefore, being aware of any factors that may result in higher rates of course non-completion amongst university students seems very appealing. Also, course non-completion may eventually lead to non-retention at university, which poses an even greater cost to the institution, government and society. Thus, studying the factors that may lead to such occurrences is of major interest.

Numerous empirical studies carried out around the globe have focused on factors that affect student dropout rates (Araque, Roldán, & Salguero, 2009; Jia & Maloney, 2015; Montmarquette, Mahseredjian, & Houle, 2001; Singell & Waddell, 2010). Results from these studies indicate that factors distributed under three broad categories are some of the key factors that may significantly influence dropout behaviour amongst university students. They include demographic factors such as gender, ethnicity, age etc. (Juhong & Maloney, 2006; Rodgers, 2013), high school background information such as the GPA attained in school, SAT scores, and socioeconomic status (Cohn, Cohn, Balch, & Bradely Jr., 2004; Johnes, 1997; Montmarquette et al., 2001; Murtaugh, Burns, & Schuster, 1999; Vignoles & Powdthavee, 2009), and institutional enrolment information, like study areas and earlier academic performance (Araque et al., 2009; Jia & Maloney, 2015; O'Keefe, Laven, &

Burgess, 2011; Rask, 2010; Singell & Waddell, 2010). Some of the research examining the factors that lead to higher dropout rates have cited issues such as insufficient data and self-reported bias on key variables as compromising their analysis (Byrne & Flood, 2008; Cohn et al., 2004; O'Keefe et al., 2011; Singell & Waddell, 2010). These limitations may have restricted the researchers from using an extensive list of variables to study dropout behaviour. As a result, there may be substantial differences in the magnitude and significance of some variables across studies. Therefore, in order to develop a model that explores the significance of a broad array of explanatory variables on rates of course non-completion, we will use a more comprehensive dataset similar to the one used by Jia and Maloney in their recent study (2015).

Our study uses a comprehensive dataset provided by the Department of Strategy and Planning at Auckland University of Technology (AUT) for the explicit purposes of this research project. This dataset contains a comprehensive record of demographic factors, high school background information and institutional information on all students enrolled at AUT. Therefore, it would be safe to say that our dataset is free from sample selection bias. Relative to survey data, administrative datasets are considered to be more precise and thorough in nature. These types of data are readily available to universities, and could be utilised to make informed decisions concerning the likelihood of student course non-completion and future non-retention behaviour. However, our administrative dataset lacks information on any sort of financial aid provided to the students, parental income etc. These variables have been considered significant in predicting dropout rates in earlier studies (DesJardinsa, Ahlburg, & McCall, 2006; Singell & Waddell, 2010; Strarron, O'Toole, & Wetzel, 2008).

As mentioned previously, we do not have access to information on a few explanatory variables that are considered important in estimating dropouts. Since our study is focusing on just those first-year students who enrolled at AUT for the first time in a Bachelor's degree programme via their NCEA Level 3 rank score (referred as overall composite score in our analysis), having detailed academic information on students is crucial. As reported by Pearl (2013) in the Waikato Times, one-in-five students drop out of university based on information provided by the Tertiary Education Commission (TEC). Therefore, the issue of dropout amongst university students exists in NZ, just like the rest of the world. It is also known that many governments allocate more resources and funds to those tertiary institutions that show greater productivity and higher levels of research (Alexander, 2000; Liefner, 2003). In New Zealand, the Student Achievement Component (SAC) funding is a performance-linked funding allocated to universities based on their performance on the following four major educational performance indicators: successful completion of courses, completion of qualifications, retention at university, and progression to a higher degree programme (Tertiary Education Commission, 2015a). As noted by the Tertiary Education Commission (2015b) successful completion of courses is deemed more important at Bachelor's degree and other pre-degree programmes (e.g., Certificates and Diplomas) than at Masters and Doctorate degree in determining the amount of SAC funding that is made available to universities. Therefore, it would be in the interest of AUT (and other universities) to increase their overall rate of successful course completion in order to achieve more funding from the government. Which, in turn, would increase the university's overall performance. So, there is a need to identify and mitigate the factors that significantly affect course non-completion rates.

Unlike Jia and Maloney (2015), our dataset does not contain some self-reported NCEA scores, but rather those provided by the New Zealand Qualifications Authority (NZQA). Therefore, we have complete and accurate data on the total number of Achieved, Merit and Excellence credits for individual students, and their overall composite score. Due to having detailed academic information, examining the significance of overall composite scores used by majority of NZ universities as entrance criteria can be analysed successfully. New Zealand has a total of eight universities. Historically, six out of these eight universities have subjectively attached specific point values (referred to as weights in our analysis) to these three credit types, where Achieved, Merit and Excellence credits correspond to 2, 3 and 4 points respectively1. Even though the other two universities do not use these weights for the purpose of enrolment into a Bachelor's degree programme, they still use them for the purpose of giving out scholarships to students2. Therefore, we will particularly look at the empirical justification for the weights that are allocated to these credits, and whether these weights are significantly different from one another in predicting successful completion of first-year courses at university. In doing so, our core aim is to provide valuable results to NZ educational institutions that use these weighting schemes to calculate a summary measure of NCEA results for admission into various Bachelor degree programmes.

The remainder of this paper will be presented as follows. Section 2 summarises the existing literature in this field of study. Section 3 describes the data used for the purposes of model

[1] Auckland University of Technology, Massey University, University of Auckland, University of Canterbury, University of Otago and Victoria University of Wellington are the majority of New Zealand universities that use the mentioned weighting scheme for enrolment purposes.

[2] Lincoln University and University of Waikato use the assigned weights to the three types of credit for the purpose of scholarship qualification.

estimation. This is followed by section 4, which develops the methodology used for this analysis, along with some robustness tests. Model estimation results will be presented in Section 5, followed by the conclusion in Section 6.

2. Literature Review

Course non-completion and student non-retention rates are few key indicators of university performance (Tertiary Education Commission, 2015a). This section will empirically examine the vast body of literature that has looked at factors deemed important for such outcomes. In our literature review, these factors will be broadly categorised under demographic factors, high school background information, and enrolment-related information. Some of the main studies in the field are summarised in the following table.

Table 1: Summary of Some Empirical Studies on Non-Completion and Dropout Amongst University Students

Title	Author	Data	Model used	Key Results
The determinants of university dropouts	Montmarquette et al. (2001)	Longitudinal dataset from the Universite´ de Montre´al for students enrolled in three semesters (i.e., fall 1987-fall 1988)	Bivariate Probit model	1. Prior academic achievements play an important role in university retention. 2. Performance in university papers (i.e., those taken in previous semester) significantly predict retention rate.
Ethnicity and academic success at university	Juhong and Maloney (2006)	Administrative data, for the year 2000, from a cohort of first-year students who enrolled at a large New Zealand university	OLS and Probit model	1. Student's ethnicity significantly predicts his or her dropout behaviour. 2. Students' GPA is considered the most important factor that determines their decision to dropout.
The socioeconomic gap in university dropouts	Vignoles and Powdthavee (2009)	Two unique longitudinal administrative datasets were linked for students who entered UK universities during 2004-2006	Probit model	1. Students from high socioeconomic background are less likely to dropout of university. 2. Parents' background in

				terms of their job and financial status is a determinant of students' dropout behaviour.
Modelling retention at a large public university	Singell and Waddell (2010)	First-year students enrolled at the University of Oregon from the year 2001 to 2006	Probit model	1. Financial support, such as scholarships or loans are important in determining the likelihood of student retention. 2. Retention is better predicted by current academic results than those obtained in high school.
High school grades and university performance	Cyrenne and Chan (2012)	Cross-sectional data on students who entered University of Winnipeg from the year 1997 through to 2002	Least Square Dummy Variable estimator (LSDV) and Hierarchical Linear Model (HLM)	1. High school grades strongly predict GPA obtained in university. 2. Students coming from different areas (i.e., high income area or low income area) differed in their university performance over the course of their full degree.
Using predictive modelling to identify students at risk of poor university outcomes	Jia and Maloney (2015)	Administrative data on all first-year students enrolled at a large New Zealand university for the academic year of 2009 to 2012	Probit model and PRM	1. Various demographic, academic and institutional factors significantly predict both course non-completion and student non-retention. 2. PRM can be used as a cost-effective tool to identify students at risk of course non-completion and non-retention.

Demographic factors (e.g., ethnicity, gender and age), and their potential impacts on non-completion and student non-retention have been widely studied (Bradely & Renzulli, 2011;

Cohn et al., 2004; Cyrenne & Chan, 2012; DesJardinsa et al., 2006; Jia & Maloney, 2015; Johnes, 1997; Juhong & Maloney, 2006; Ozga & Sukhnandan, 1998; Rodgers, 2013; Singell & Waddell, 2010; Strarron et al., 2008; Vignoles & Powdthavee, 2009). Most studies have found significant differences between ethnic groups in rates of course non-completion and non-retention at university (Bradely & Renzulli, 2011; DesJardinsa et al., 2006; Jia & Maloney, 2015; Juhong & Maloney, 2006; Rodgers, 2013; Singell & Waddell, 2010; Vignoles & Powdthavee, 2009). For example, Jia and Maloney (2015) in their study carried out in NZ, which used predictive risk modelling, found that Asian and European students have a higher probability of successfully completing first-year courses compared to Maori and Pacifica students once other measurable individual differences were held constant. However, there are several studies that have not found enough evidence to support the impact of membership of different ethnic groups on rates of retention at university (DesJardinsa et al., 2006; Rodgers, 2013; Singell & Waddell, 2010). Take, for example, the results from a large public university in the U.S., which concluded that there is no significant difference in the rate of retention amongst Hispanics, Native American, and non-white students from white students (Singell & Waddell, 2010). Likewise, a study based on a single cohort of undergraduate students in the United Kingdom concluded that the rate of non-completion was relatively similar between white and minority students, given controls for their socioeconomic backgrounds (Rodgers, 2013).

Nowadays, it is known that ratios of undergraduate female students who enrol in university appear to outweigh those of male students. Numerous studies have also supported this and, in addition, concluded that female students tend to have lower rates of course non-completion and university non-retention compared to their male counterparts (Bradely &

Renzulli, 2011; Jia & Maloney, 2015; Juhong & Maloney, 2006; Montmarquette et al., 2001; O'Keefe et al., 2011; Singell & Waddell, 2010). Take, for example, Rodger (2013), who found that female students have a non-completion rate of 34.5% compared to 41.4% for male students. However, the same study also established that female Asian-Muslim students have significantly higher rates of non-completion than male Asian-Muslim students. Similarly, there are some studies that contradict the findings that female students have lower rates non-completion than male students (Belloc, Maruotti, & Petrella, 2010; Rodgers, 2013; Singell & Waddell, 2010). Another such study carried out to analyse the dropout rates amongst undergraduate students enrolled in the faculties of Economics and Business in Rome found lower dropout rates for male students compared to female students (Belloc et al., 2010). Another interesting demographic factor thought to increase the probability of non-completion amongst students is their growing age (Jia & Maloney, 2015; Ozga & Sukhnandan, 1998). Ozga and Sukhnandan (1998) concluded that mature students have a significantly higher probability of non-completion, not because they cannot cope with the study load, but because of external factors such as looking after their family or working in a paid job. However, the authors did not specify which age group was considered to be mature students. Moreover, one would expect full-time students to perform better in university and therefore have lower rate of non-completion and non-retention (Jia & Maloney, 2015; Montmarquette et al., 2001; Triventi, 2014). Jia and Maloney (2015) found significant results to support the former claim. They estimated that the probability of course non-completion for a part-time student increases by 17.23 percentage points compared to their full-time counterparts. Also, Montmarquette et al. (2001), from their study that focused on students enrolled in three semesters, found the dropout rates to be higher amongst part-time students. Also, from a study carried out in

Italy on first-year university students, found significant results that predicted poor academic outcomes for both, student who had high-intensity employment status and those who had low-intensity employment status (Triventi, 2014). These findings do conform to the belief that part-time students have relatively higher rate of non-completion and non-retention at university than full-time students. In addition to these mentioned factors that looked at student demographic features, we also have detailed high school information.

The importance of having data on prior educational attainment (i.e., grades in high-school, school decile, competitive exam score, e.g., SAT, ACT etc.) have been highlighted in numerous studies in the past (Byrne & Flood, 2008; Cohn et al., 2004; Cyrenne & Chan, 2012; Jia & Maloney, 2015; Johnes, 1997; Juhong & Maloney, 2006; Lasselle, McDougall-Bangall, & Smith, 2014; Montmarquette et al., 2001; Singell & Waddell, 2010; Vignoles & Powdthavee, 2009). Intuitively, students who achieve academically in high school (i.e., those who receive higher grades) should have a lower probability of not completing their course and hence dropping out. Cyrenne and Chan (2012) and Montmarquette et al. (2001) concluded that higher grades obtained in high school significantly decrease the risk of course non-completion and university non-retention amongst first-year students. Whereas, a study carried out in Dublin City University that focused on students who studied accounting found that while prior knowledge in the field of accounting did not decrease the risk of course non-completion, it did significantly decrease overall dropouts (Byrne & Flood, 2008). Another study supported the importance of performance in current courses as being more important in predicting dropout and non-completion rates than performance from earlier exams like SAT, ACT etc. (Singell & Waddell, 2010). Moreover, results from an elite institution in Scotland, which looked at students under the age of 21 years, concluded that those students who came from a below average high school, but got three A grades performed better in university than their counterparts who came from an above average high school (Lasselle et al., 2014). So, their study concluded that even though a student came from a below average high school, given he or she achieved high grades in high school, had a higher probability of graduating with a First or Second class degree. As noted from these studies, the importance of prior school grades has been found to be very significant.

Another variable of interest is the socioeconomic status of students and its effect on the dropout rates. Subconsciously, students coming from a higher socioeconomic background should have lower rates of dropout, perhaps because they have enough funds to support the completion of their study or have better self-building opportunities compared to students coming from lower socioeconomic backgrounds. Both Cyrenne and Chan (2012) and Vignoles and Powdthavee (2009) support the former claim and conclude that students who come from a lower socioeconomic status have limited access to external resources, and therefore have higher chances of dropping out of university. Jia and Maloney (2015) also supported this. In their study, they found students coming from low school deciles (and hence from a low socioeconomic background) have a higher risk of course non-completion and university non-retention.

Finally, the importance of different areas of study and the level of courses taken during university in predicting non-completion and dropout behaviour has also been extensively studied (Jia & Maloney, 2015; Pike & Killian, 2001; Rask, 2010; Robst, Keil, & Russo, 1998; Rodgers, 2013). Degrees which are thought to be technically difficult and which result in higher dropout rates could be those in the fields of Engineering and Sciences relative to other fields. A study that included 13 broad subject categories from different universities found medicine to be the most difficult degree (with dropout rates of around 30%), followed by other science-related degrees (Johnes, 1997). Another study carried out in the south of Spain looked at various degrees under three major faculties in a university, and found the Humanities department had the highest rates of dropout (63.5%), followed by the department of Software Engineering (49.6%) and Economics Sciences (43.6%) (Araque et al., 2009). Rask (2010) interestingly found that students who had attained good grades in high school had significant impacts on retention rates in science, technology, engineering and mathematics (STEM) courses in university. Likewise, Students enrolled in the Bachelor degrees of Design, Health Science and Education all had higher probabilities of successful course completion, while the Bachelor degrees of Computer Information Science and Engineering Technology had lower probabilities of successful course completion compared to the Bachelor of Arts (Jia & Maloney, 2015).

Our study is similar to Jia and Maloney (2015), as we will also use a comprehensive dataset and Probit estimation models to analyse the impact of various independent variables on course completion outcomes for first-year students in Bachelor degree programmes. However, it differs in terms of the focus group of students. Our focus group comprises all first-year students who entered university via a valid NCEA level 3 score. In addition, we have more detailed information on these NCEA Level 3 results. Specifically, we have information on the total number of Achieved, Merit and Excellence credits attained in NCEA Level 3 exams on top of the overall score.

Having comprehensive data on demographic factors, high school backgrounds and enrolment-related information will enable us to better predict course completion outcomes. As mentioned by Jia and Maloney (2015), focusing on factors that influence course non-completion can act as an initial indicator of students at risk of future course non-completions and, eventually, university non-retention. Therefore, detailed academic information will not only help us in estimating the individual effect of these three types of credits on the rate of course completion, but also assist in determining the significance of the weights assigned to these credits by NZQA which is used for the purposes of university entrance. After estimating the model, if we find the marginal effects of the three credit types are inconsistent with the weights assumed in the overall score currently used by the majority of universities in NZ, then our research will suggest a possible re-weighting of this overall NCEA score. Empirically analysing the respective weights on Achieved, Merit and Excellence credits for the purposes of university entrance will prove beneficial for NZQA and NZ universities in making more informed decisions. Also, this could eventually be helpful to students when making decisions during their high school education.

3. Data

Auckland University of Technology's Department of Strategy and Planning provided the administrative data that are used in this research. Our dataset has comprehensive information on all first-year Bachelor degree students enrolled at AUT during the academic years of 2013 and 2014. The full sample consists of 64,446 course observations that are used to summarise all of the data on first-year students regardless of their entrance type (i.e., whether a student was enrolled via their Cambridge or International Baccalaureate score, NCEA level 3 score, etc.). Eventually, a sub-sample with a total of 32,423 course observations and a sum of 4,898 student observations is used to address our research question. This sub-sample of 32,423 course observations consists of only those students enrolled at AUT via NCEA level 3 for whom valid NCEA level 3 scores were available. In this study, course observations are used to analyse successful course completion outcomes in first-year courses. Definitions of all the variables used in this research are provided in Table 2 (refer to Appendix). Table 3 (in the Appendix) provides a comparison of descriptive statistics for the full sample of all first-year students and the sub-sample of those just enrolled via NCEA level 3 with a valid score.

We separate the information on students in our dataset into various broad categories (e.g., ethnicity, country of origin, other demographic characteristics, high school backgrounds, entrance types, and other academic information). For students in the dataset, we specifically know their age, gender, school decile, Bachelor degree programme in initial enrolment, and the number of achieved, merit and excellence credits achieved in NCEA level 3.

Successful completion outcomes for all first-year courses are used as the binary dependent variable in this study. This dummy variable takes on a value of one for those students who received a passing grade in a course and therefore successfully completed it; zero otherwise. Our sample of students entering with NCEA level 3 has an 81.12% successful completion rate in first-year courses, whereas the full sample of all first-year students has a 79.10% successful course completion rate during the first-year of study.

In our sample, we can compare successful course completion rates between the two years of 2013 and 2014. The year 2014 is indicated by a dummy variable, while the year 2013 is the omitted category. Sample statistics reported in Table 3 indicate that 50.58% of course observations came from 2014 in the full sample, while 54.47% of similar observations came from the same year in the NCEA level 3 sample. Thus, relatively more students were entering this University with NCEA level 3 results in the latter year.

As mentioned earlier, we have demographic information on every student. These demographic factors include ethnicity, country of origin, gender, part-time study or full-time study, whether a student's first language is English or not, and age. We had six self-reported ethnic groups provided by the students. Five out of these six groups (Asian, Maori, Pacifica, other minority ethnic groups, and not declared ethnicities) were included as dummy variables, with the omitted category being students who reported European as their ethnicity. The overall distribution of mean observations for the ethnic groups was relatively similar between the sub-sample and the full sample (refer to Table 3). However, there were slightly higher relative proportions of students with European and Maori ethnicities in the NCEA level 3 sample. As indicated by Table 3, European (47.52%) and Asian (20.64%)

constituted the top two ethnic groups of first-year students, whereas Maori students accounted for 11.71% and Pacifica students accounted for 12.89% of the observations for the sub-sample of first-year students.

China, New Zealand, India, South Korea and other countries are used as dummy variables for the country of origin category. The omitted category is that set of students for whom no information on their country of origin was provided. In our study, a student's country of origin does not necessarily align with associated ethnic groups. The reason behind this mismatch could be due to the high immigrant population in New Zealand, and where a New Zealander (someone who is a NZ citizen or permanent resident) reports their country of origin as "New Zealand" but categorises themselves under "Asian" or "Maori" when it comes to ethnic group. As observed, for the full sample, 73.74% of the students stated New Zealand as their country of origin followed by China (7.16%), South Korea (1.96%) and India (1.18%). Similarly, for the sub-sample, the highest proportion of students have stated New Zealand (83.38%) as their country of origin, followed by China (2.37%), South Korea (1.11%) and India (0.83%). This indicates that students in the sub-sample are more likely to come from New Zealand than other countries, which conforms to the fact that NCEA system is incorporated in New Zealand high schools only. Having said that, results from Table 3 for the sub-sample also indicated that 1 in every 6 student who entered a New Zealand University with NCEA level 3 did not originally come from New Zealand. Therefore, there seems to be considerable immigrant population attending high schools and further progressing to university education in New Zealand.

The dummy variable for gender took a value of one for female; zero for male. As noted from Table 3, 65.56% of the mean course observations came from female students in the sub-sample, whereas in the full sample, 61.31% of these course observations came from female students. This indicates that female students are more likely to enter university under NCEA level 3 than male students. 60.35% of the students in the sub-sample had reported English as their first language. Though in the full sample, 56.14% students had reported English as their first language. Five dummy variables for age were created in order to allow for possible non-linear effects of age on the probability of course completion in our regression analysis. The five dummy variables were set equal to one for those under the age of 18, 19 years old, 20 years old, 21 years old, and those over 21 years, with the omitted category being those who were 18 years old. The average age of students in our sub-sample is 18.49 years, unlike the mean age of the full sample, which is 20.85 years. The higher mean age for the full sample compared to the sub-sample does not seem surprising, as AUT has a relatively higher intake of students under the special admission category (13.55%) than other universities. More mention will be made of this in subsequent paragraphs that will elaborate on students' high school information. Also, 94.87% and 90.33% of the observations consisted of full-time students in the sub-sample and the full sample, respectively.

Factors included in the high school background information collected on students in our dataset are: entrance type at AUT, school decile, total number of achieved, merit and excellence credits, and overall composite NCEA score. 58.06% of all first-year students at AUT enter via NCEA level 3. The next largest group of admittance at AUT was those under "Special Admission". Students who have not attained University Entrance requirements and are over the age of 20 years can apply for entrance under the "Special Admission" category. Also, other students are present in the full dataset (i.e., those students who have enrolled in Bachelor's degree programmes at AUT due to previously obtained pre-degree certificates or diplomas either from AUT (i.e., "Internal") or from other universities (i.e., "External")). Another group of students enrolled at AUT are under the "Other entrance type" criteria. These are mostly international students who gain admission at AUT based on grades obtained in an equivalent high school overseas. For the purposes of our research, we only include those students who entered AUT via NCEA level 3. In addition, NCEA level 3 students who did not have a valid NCEA score (which constituted 13.35% of the NCEA level 3 students) were excluded from this analysis, and this resulted in the sub-sample that is used to address our research question.

We included 10 dummy variables for school decile (i.e., one each for school decile from 1 through to 5, one each for school decile from 7 through to 10, and one for those who did not report their school decile). The omitted category was school decile 6 as this "middle" school decile category can be useful when comparing students coming from higher decile schools (i.e., above decile 6) and lower decile schools (i.e., below decile 6). In New Zealand, Decile 1 schools include "the 10% of schools with the highest proportion of students coming from low socio-economic communities", whereas the 10% of schools with the lowest proportion of students coming from low socio-economics status corresponds to Decile 10 schools (Education Counts, 2015). The distribution of the mean observations for all school deciles appears to be slightly higher for the sub-sample than the full sample (refer to Table 3). Also, 12.92% of the students in the full sample had not reported their school decile compared to 1.17% of the students in the sub-sample. This could be because the full sample consists of other students than just those who enrolled with NCEA level 3 (for example international students), and hence these students might be unaware of their school decile,

For our research the overall composite NCEA score and the total number of achieved, merit and excellence credits are of utmost importance. For each student, the overall composite score reported takes into consideration only the top 80 credits obtained, multiplied by their respective point values (i.e., every credit for achieved is multiplied by 2, for merit multiplied by 3, and for excellence multiplied by 4). For example, a student who obtained 20 achieved credits, 40 merit credits and 35 excellence credits would have got an overall score of 270 (35*4 + 40*3 + 5*2). Fifteen achieved credits would be 'discarded' in computing this composite score. As indicated by Table 3, the mean overall NCEA score for the sub-sample was 192.04, and the respective means for achieved, merit and excellence credits in the same sample were 41.59, 23.49 and 13.75.

Finally, for enrolment-related information, we had data on the Bachelor degree programme that each student enrolled in and the levels of the courses they undertook during their first year of study. We use 10 dummy variables (BA, BBus, BCIS, BCS, BDes, BEdu, BEngTech, BHs, BIHM, and BSR) for the Bachelor degree programmes, with the omitted category being the other smaller Bachelor degree programmes offered at AUT. Bachelor of Health Science (BHS) seems to be a popular choice amongst the other Bachelor degree programmes, accounting for 20.23% (in the sub-sample) and 21.65% (in the full sample) of first-year students. Also, the observed means for all the degree programmes are relatively similar between the sub-sample and the full sample (refer to Table 3). For course level, we have included 3 dummy variables (one each for levels 4, 6 and 7 courses), with level 5 courses being the omitted category. Ideally, when a student enrols in a Bachelor degree programme, during their first-year of study they undertake level 5 courses. Therefore, 89.35% of those in the sub-sample account for level 5 courses. Having said that, some students have relatively weaker academic backgrounds, and therefore they must take level 4 courses in their first year of study. Moreover, some students can take courses at higher levels (i.e., level 6 and 7) in their first year of study, as they may have relatively stronger academic backgrounds, or because some programmes might require them to take higher-level courses.

4. Methodology

As mentioned earlier, our aim is to predict the effects of various independent variables (including demographic factors, high school backgrounds and enrolment-related information) on the probability of successfully completing first-year university courses. Therefore, in our research we used Maximum Likelihood Probit models to estimate the factors that affect successful completion of first-year courses.

$$Y_i^* = \beta X_i + \varepsilon_i$$

The above equation represents a basic Probit model, where Y_i^* denotes a latent dependent variable. We are only able to observe Y_i, a dummy variable that takes on a value of 1 if the student had successfully completed the first-year course in which they are enrolled; 0 otherwise.

$$Y_i = \begin{cases} 1, & if\ Y_i^* > 0 \\ 0, & if\ Y_i^* \leq 0 \end{cases}$$

All the independent variables (mentioned in Table 3) are represented by the vector of variables X_i. The coefficient vector β is estimated using the maximum likelihood methods, and ε_i denotes the random error term in the model that, by assumption, is independent and normally distributed.

The following equation depicts the probability of successfully completing first-year courses by a student, where $\Phi(.)$ represents the Cumulative Distribution Function (CDF) of the standard normal distribution.

$$P_i = \Pr(Y_i = 1|X_i) = \Pr(Y_i^* > 0) = \Pr(\beta X_i + \varepsilon_i > 0) = \Pr(\varepsilon_i > -\beta X_i) = \Phi(\beta X_i)$$

As is well known, the Probit model is non-linear in nature. Therefore, to provide meaningful interpretations of the effects of the independent variables on the probability of this outcome, we need to compute the average marginal effects of a one-unit change in a given independent variable on this conditional probability across the entire sample.

$$\frac{\partial \Pr(Y_i = 1|X_i)}{\partial X_k} = \beta_k \phi(\beta X_i)$$

$\phi(.)$ represents the Probability Distribution Function (PDF) of the standard normal distribution.

Our sub-sample consists of students who had either an overall of less than 80 credits, exactly 80 credits, or more than 80 credits. For those students who had fewer than or exactly 80 credits, all the credits were used to get their overall composite score (referred to as overall NCEA score in our models). Therefore, all of the Achieved, Merit and Excellence credits attained by these students were left unchanged in models 1, 2 and 3. For a student who obtained more than 80 credits, their best 80 credits were obtained by adding up all of their Excellence credits first, followed by Merit credits, and finally Achieved credits. For example, if a student had an overall total of 120 credits, where they had attained 60 Excellence credits, 40 Merit credits and 20 Achieved credits, then their best 80 credits would be a total of 60 Excellence credits plus 20 Merit credits. In order to get students' overall composite score, these individual credits are multiplied by their respective weights, where 2, 3 and 4 are weights given to Achieved, Merit and Excellence credits respectively.

Universities in NZ subjectively use these weights to calculate students' overall composite scores. These overall scores are then used to either determine the programme a student can get into or for the purpose of scholarship qualification, or both. This is in addition to certain numbers of mandatory literacy and numeracy credits that are required for university entrance. The universities calculate this overall score based on the best 80 credits obtained by a student. Therefore, for Models 1 and 2, we used students' 80 best credits and the arbitrary weights proposed by universities for the three credit types to carry out our analysis. For Model 3, we again used the best 80 credits, but we proposed new weights for Achieved, Merit and Excellence credits based on our estimated marginal effects coefficients for these credits from Model 1. Finally, in Model 4 we used all credits obtained by students to carry out estimation of our Probit regression.

5. Empirical Results

Four Probit models were estimated to analyse the effects of changes in the probability of successful course completion due to changes in an explanatory variable, holding other variables constant. Since Probit models are non-linear in nature, a one-unit change in an explanatory variable does not directly predict changes in the probability of successful completion of first-year courses. Hence, we evaluated the marginal effects of the corresponding explanatory variables by computing these partial derivatives or marginal effects for every observation in our sample and reporting the resulting sample means of these marginal effects. The four different Probit models estimated to address our research question included all the independent variables mentioned in Table 3 for the sub-sample (those students who entered AUT with valid NCEA level 3 scores). More specifically, Model 1 estimated the Probit model for the best 80 credits (refer to Table 4). Model 2 estimated the Probit model for the overall score based on current weights, where Achieved credit is awarded 2 points, Merit credit is awarded 3 points and Excellence credit is awarded 4 points (refer to Table 5). Model 3 estimated the Probit model for overall score based on weights proposed by our research (refer to Table 6). Finally, Model 4 estimated the Probit for all the credits attained by an individual student (refer to Table 7). Tables 4 through 7 are included in the Appendix. Every set of regression results includes coefficient estimates, standard errors on these coefficients estimates, and the estimated mean marginal effects in the sample. The results obtained from these models will be compared in the subsequent paragraphs.

The reason for the inclusion of all four models was to show the slight, though still significant, impact of: breaking down overall scores into their component scores of

Achieved, Merit and Excellence credits; overall score using 2*Achieved, 3*Merit and 4*Excellence criteria; new overall score using the empirically calculated 1*Achieved, 2*Merit and 2*Excellence criteria; and inclusion of all credits obtained for each of the three components. As mentioned earlier, for those students who enter university via NCEA level 3, their best 80 credits were chosen. Therefore, Models 1, 2 and 3 used the maximum Achieved, Merit and Excellence credits (that did not exceed the 80 credits boundary) obtained by students, whereas Model 4 included all credits obtained by an individual student.

From Model 1, holding variables on students' demography, high school background and institutional enrolment information constant, we found the probability of successfully completing first-year courses to be higher in the year 2013 (which was the omitted category) than in the year 2014. Although this decline in the probability of successfully completing first-year courses in the year 2014 compared to the year 2013 does not necessarily indicate any trend in our research.

Various researchers have highlighted the importance of ethnicity on students' course completion and retention in university (Bradely & Renzulli, 2011; DesJardinsa et al., 2006; Jia & Maloney, 2015; Juhong & Maloney, 2006; Rodgers, 2013; Singell & Waddell, 2010; Vignoles & Powdthavee, 2009). Where most of them predicted significant difference in the rates of dropout between different ethnic groups while some contradicted this finding. Likewise, in our research, we found that Maori students, Pacifica students and students who did not declare their ethnicity had a lower probability of successfully completing first-year courses compared to European students. In other words, for every course taken by a

Pacifica student, his or her probability of successfully completing that course goes down by 8.46 percentage points relative to a European student, holding all other variables constant (Refer to Table 4). However, we did not find any significant impact of a student's country of origin on their successful course completion. This could be because our sub-sample focused only on the NCEA entrance type, which resulted in a higher proportion of domestic students in our sample (83.38% of the students reported NZ as their country of origin) and relatively smaller proportion of overseas students.

Female students, compared to males, tend to have lower course non-completion and dropout rates as highlighted in numerous studies discussed earlier (Bradely & Renzulli, 2011; Jia & Maloney, 2015; Juhong & Maloney, 2006; Montmarquette et al., 2001; O'Keefe et al., 2011; Singell & Waddell, 2010). Similarly, results from Model 1 showed that female students had a higher probability of successfully completing first-year courses compared to their male counterparts. Given a student is female, holding other variables constant, the probability of her successfully completing a course goes up by approximately 3.00 percentage points compared to a male student (refer to Table 4). Intuitively, students studying part-time could be expected to have lower successful completion rates possibly due to other commitments that they might have. On the other hand, being a full-time student entails one putting more effort and time into studies, and therefore they may not have substantial amount of time to allocate to other activities. The results from our study indicated that part-time study reduces the probability of successfully completing courses by 3.56 percentage points, holding all other variables constant (refer to Table 4).

Surprisingly, results obtained from Model 1 indicated that students who reported English as their first language had a lower probability of successfully completing courses when compared to their counterparts. For example, a student who reports English as their first language decreases his or her probability of successfully completing first-year courses by 2.62 percentage points, holding other variables constant. Also concluded from a study was that mature students are more likely to drop out of university (Ozga & Sukhnandan, 1998). Compared to our omitted group of students aged 18 years, we found that those students younger than 18 years and those who are 21 years had lower and higher probabilities of successfully completing their courses, respectively (refer to Table 4). Holding all other variables constant, Model 1 concluded that, a student who is under 18 years of age has a lower probability of successfully completing his or her course by 7.27 percentage points compared to an 18-year-old student, whereas the probability of successfully completing a course goes up by 6.96% percentage points given a student is 21 years old. These results may not seem surprising, as students who are below 18 years of age (i.e., 0.47% of the sub-sample) may not have acquired enough subject knowledge to successfully complete the courses they enrolled in during their first year of study, while students aged 21 (i.e., 1.18% of the sub-sample) may have already attained some extra subject knowledge prior to the commencement of their first-years courses. Interestingly, the estimated marginal effect for a student above the age of 21 years on successful course completion was quite high (12.20 percentage points higher) relative to an 18-year-old student, but it also reported a big standard error (refer to Table 4). The statistical non-significance of this estimated marginal effect could be due to the very small proportion of those above the age of 21 years (i.e., 0.14% of the sub-sample).

Our study highlights the importance of students' enrolment-related information (for example, the different Bachelor Degree programmes) on the probability of successfully completing first-year courses. Results from Model 1 found that Bachelor of Communication Studies, Bachelor of Design, Bachelor of Education, Bachelor of Health Science, and Bachelor of International Hospitality Management all have significant positive impacts on the probability of successful completion of first-year courses at 1% level of significance. For instance if a student enrols in a Bachelor of Education programme he or she has a higher probability of successfully completing first-year courses (i.e., 16.01 percentage points higher) compared to a student who enrols in other smaller Bachelor programmes (refer to Table 4). We also found that those students doing a double degree had a higher probability of successful course completion. A student who enrols in a double degree increases his or her probability of successfully completing a first-year course by 18.95 percentage points, holding all other variables constant (refer to Table 4). Students enrolled in a double degree programme may have stronger academic background resulting in higher probability of successfully completing first-year courses. In addition to the importance of Bachelor degree programmes on successful course completion, level of course undertaken during first-year of study is another factor that should be considered. As expected, a student who takes a level 4 course (a lower-level course) increases his or her probability of successful course completion by 6.90 percentage points compared to a student who takes a level 5 course (refer to Table 4).

With regards to students' high school background information, we had data on school decile and their detailed NCEA level 3 results. School decile, which could also be predicted by a student's socioeconomic status, is yet another factor thought to have a significant impact on students' course completion and dropout rates (Cyrenne & Chan, 2012; Vignoles & Powdthavee, 2009). One would expect students from high decile schools to do well in university and vice versa for students coming from low decile schools. Compared to our base group (i.e., students from school decile 6), we found that the probability of successfully completing first-year courses is significantly lower in students coming from school deciles 1 through to 5, holding all other variables constant (refer to Table 4). For example, a student from a decile 1 school has a considerably lower successful course completion rate (10.56 percentage points lower) than a student from a "middle" decile 6 school (refer to Table 4). However, results from Model 1 also indicate that a student from a higher decile school (i.e., school decile 8 and 10) has a lower probability (i.e., 6.80 and 4.35 percentage points lower, respectively) of successfully completing first-year courses compared to a student from a decile 6 school, at 1% level of significance. This might be consistent with the finding from Lasselle et al. (2014) study, which could imply that a student from a decile 6 school might have performed better in their high school exam (i.e., NCEA level 3 in our case) and therefore had higher probability of successful course completion at university than a student from either a decile 8 or decile 10 school. Thus, the former possible explanation brings us back to the importance of high school grades on successful course completion at university.

Similar to Model 1, the rest of the estimated models (i.e., Models 2 through to 4) all reported the same explanatory variables to have a significant marginal effect on the probability of successful course completion. Though, the magnitude of these estimated marginal effects differed slightly from Model 1 (refer to: Table 5 for Model 2, Table 6 for Model 3, and Table 7 for Model 4).

More importantly, crucial to our research question is the impact of having detailed information on NCEA level 3 results and the appropriateness of the points value assigned to the three types of credits on the probability of successful course completion. Therefore, the following four paragraphs will elaborate on the results obtained from Probit Models 1 through to 4 on the estimated marginal effect of detailed NCEA level 3 results on the dependent variable.

As mentioned earlier, Model 1 estimated the Probit model using the best 80 Achieved, Merit and Excellence credits. Results from this model indicated that, holding all other variables constant, the probability of successfully completing a course goes up by 0.29 percentage points for every additional Achieved credit. This probability goes up, but remains very similar, for a student who obtains one extra Merit credit (i.e., 0.58 percentage points) and / or an extra Excellence credit (i.e., 0.60 percentage points) (refer to Table 4). Since, the estimated marginal effects of Merit and Excellence credits on successful course completion was found to be very similar, we conducted a F-test to check whether the marginal effects of these three credits differ from one another. The results from the F-test showed that an Achieved credit is significantly different to a Merit and an Excellence credit ($p < 0.01$), whereas Merit and Excellence credits did not statistically differ from one another ($p = 0.67$). As a result, new weights for Achieved, Merit and Excellence credits were proposed. More on these new weights will be discussed in the paragraph that elaborates on Model 3.

The second Probit model we estimated was using the overall composite score obtained using the best 80 credits. The overall composite score for the best 80 credits was obtained by using the current weights attached to Achieved (2 points), Merit (3 points) and Excellence (4 points) credits. Model 2 indicated that given a student obtains an additional point for their composite score (for example, an overall composite score of 161 rather than 160), then his or her probability of successfully completing a course increases by 0.17% percentage points, holding all other variables constant. Additionally, for the purpose of comparison, we estimated another Probit model that had no information on students NCEA results on the probability of successful course completion. This model yielded a Pseudo R^2 of 0.0888. Therefore, the importance of having information on high school academic records (like that in Model 2) increases the explanatory power of the regression model by 37.02%.

The new proposed weights were used to get the "New Overall NCEA Score", which was used as an explanatory variable in our third Probit model. From the F-test carried out after the estimating Model 1, we found that Merit and Excellence credits did not statistically differ from one another and thus using different weights for these two types of credit did not seem statistically correct. Therefore, new weights for Achieved, Merit and Excellence credits, as proposed by our research, were 1, 2 and 2 respectively. Model 3 concluded that for an additional overall composite score that a student gets, his or her probability of successful course completion increases by 0.30 percentage points, holding other variables constant. The Pseudo R^2 for Model 3 was 0.1385 (refer to Table 6) and that for Model 2 was 0.1371 (refer to Table 4). As expected, using the new weights on Achieved, Merit and Excellence credits, on top of all other independent variables, resulted in a 1.01% increase in the explanatory power of successful course completion than using the old weights. Thus, if a university uses overall composite scores as their entrance criteria, they can better predict the probability of successful course completion given they use the new proposed weights for the three types of credits.

Finally, Model 4 was estimated using all of the Achieved, Merit and Excellence credits obtained by students. So far, no research has shown the significance of using the best 80 credits rather than all credits obtained by any student. One could expect a high deserving student to obtain more credits than his or her counterpart student. Therefore, predicting students' successful completion of first-year courses merely based on their best 80 credits might not be optimal. In including all of the credits for every student, in Model 4, we aim to better predict the probability of successful course completion. As observed from Table 7, every additional Achieved, Merit and Excellence credit a student obtains, his or her probability of successful course completion increases by 0.26, 0.50 and 0.48 percentage points respectively. It is also interesting to note that an additional Merit credit increases the probability of successful course completion by a slightly greater percentage point (i.e., 0.02 percentage points more) than an additional Excellence credit. Furthermore, the overall explanatory power of our regression analysis increases by 1.77% (an increase in Pseudo R^2 from 0.1385 in Model 1 to 0.1410 in Model 4) given we use all of the credits obtained by students to predict their probability of successfully completing first-year courses. Again, for the purpose of comparison, we ran another Probit model that included an explanatory variable for the overall composite score for all the credits obtained multiplied by their new proposed weights (refer to Table 8). And, we found the Pseudo R^2 for Model 5 to be exactly the same as Model 4. Hence, using the new weights to obtain the overall composite score, in addition to all of the other independent variables, increases the overall predicative power of successful completion of first-year courses at university by 2.77% (an increase in Pseudo R^2 of 0.1410 in Model 5 from 0.1371 in Model 2).

6. Conclusion

The aim of our study was to examine the factors that lead to course non-completion, and to what extent, amongst first-year Bachelor's degree students who enrolled at university with a valid NCEA Level 3 score. In addition, another vital aim of this study was to empirically examine the weights associated to the three types of credits (Achieved, Merit and Excellence). This weighting scheme helps in determining a student's overall composite 'rank' score, which is used by all the universities in NZ for the purpose of enrolment in different Bachelor's degree programmes or scholarship qualification, or both. Majority of NZ universities have used this weighting scheme subjectively. Conducting an empirical study to determine the significance of the effect of these different credits (with their respective point values) on completion of courses would prove to be beneficial overall. These universities would, therefore, get a clear indication of whether these weights should be the same as the ones used in the past or if they should be different based on our study. As a result of this, it will benefit both policy makers and universities in strategising better entrance criteria that could result in higher rates of course completion.

Our study, which utilised administrative data provided by AUT, concluded that there are various factors that significantly govern a student's probability of completing courses in their first year of study. These factors, in our study, are broadly categorised as: demographic factors, high school backgrounds and enrolment-related information. A student's ethnicity, gender, time of study (part-time or full-time), whether they speak English or not, and age were significant in influencing the probability of course completion. Our study showed that European students have a higher probability of course completion compared to Maori and Pacifica students. Moreover, female students tend to have a higher rate of completing first-year courses compared to their male counterparts. The intuition that full-time students should have a higher rate of course completion was also proved to be significant in our study. Interestingly, however, we found that students whose first language was English had a relatively lower probability of completing first-year courses. We also found that students under the age of 18 are less likely, and those aged 21 are more likely, to complete first-year courses they enrolled in relative to those students aged 18.

High school background information proved to be a crucial category that significantly estimated the probability of successfully completing first-year courses. The importance of having detailed academic records to estimate successful course completion of first-year university students was highlighted in our research. We found that a student who gets an extra Achieved, Merit and/or Excellence credit is likely to have a higher probability of successful course completion. The same was found to be true for the overall score variable, which predicted higher first-year course completion rates for students who get an extra point for their overall score. Also, compared to students from school decile 6, the lower school decile students tend to have lower rates of course completion, but even those students from higher school deciles also tend to not successfully complete their courses.

The results for the significance of the weights of 2, 3 and 4 for Achieved, Merit and Excellence credits respectively gave us some interesting findings. We found Achieved credits significantly differed from Merit or Excellence credits. Those students who got an extra Achieved credit had a higher probability of successful course completion, but those who had an extra Merit or Excellence credit had an even higher probability of course completion. However, no significant difference was found in the estimation level of first-year course completion rates amongst students who got either an extra Merit or Excellence credit. In order words, a student's probability of successful completion of first-year courses is similar regardless of whether they obtain an extra Merit or Excellence credit. Therefore, our findings diverged from the weighting scheme that is currently being used by the majority of universities in NZ for enrolment purposes. Until now, these universities have been subjectively using the mentioned weights for the three types of credits, where Achieved credit is awarded the lowest weight, followed by Merit credit and then Excellence credit. This implies that Excellence credits might have thought to result in the highest rates of successful course completion relative to other credits. However, our study concludes that Merit and Excellence credits predict the same level of successful course completion, and therefore should be given the same weights. Hence, we suggested using 1, 2 and 2 as new weights for Achieved, Merit and Excellence credits respectively. Also, we conducted analysis with these new weights on not only the best 80 credits (which has been historically used by the majority of universities), but also for all the credits a student attained in their high school. We found that using the new weights and all the credits to obtain the overall composite score, in addition to all of the other independent variables, we are better able to predict the probability of successful course completion.

Finally, institutional enrolment information factors have also proved to be significant at estimating the probability of successful completion of first-year courses. In particular, students who enrol in Bachelor of Communication Studies, Bachelor of Design, Bachelor of Education, Bachelor of Health Science, and Bachelor of International Hospitality

Management have a relatively higher chance of successful course completion. This is also true for students who enrol in a conjoint (or double degree programme) and for those who take level 4 courses.

Therefore, our study could prove to be beneficial to both, universities and the government, although there are a few recommendations that may better predict the probability of successful completion of first-year courses. Firstly, inclusion of family factors, for example parental financial status, their educational background, their work experience, etc., could prove to be beneficial (Araque et al., 2009; Rodgers, 2013; Vignoles & Powdthavee, 2009; Wintre & Yaffe, 2000). Secondly, having information on financial aid, (did the enrolled student get any kind of scholarship, are they studying by taking a loan, are they working extra hours to pay for their fees etc.) have proved to be significant in past studies (Belloc et al., 2010; Cohn et al., 2004; Cyrenne & Chan, 2012; DesJardinsa et al., 2006; Garwe & Manganga, 2015; Singell & Waddell, 2010; Strarron et al., 2008)**.** Finally, having even more detailed academic information, i.e., knowing a student's grade in every course they took in high school might be very useful when predicting outcomes. Even though various recommendations have been made to make this study better, significant results from our study still shed light on various aspects that could be paid more attention to in order to minimise the rate of course non-completion.

References

Alexander, K. F. (2000). The changing face of accountability: Monitoring and assessing institutional performance in higer education. *The Journal of Higher Education, 71*(4), 411-431.

Araque, F., Roldán, C. n., & Salguero, A. (2009). Factors influencing university drop out rates. *Computers & Education, 53*, 563-574.

Belloc, F., Maruotti, A., & Petrella, L. (2010). University drop-out: An Italian experience. *Higher Education, 60*, 127-138.

Bradely, C. L., & Renzulli, L. A. (2011). The complexity of non-completion: Being pushed or pulled to drop out of high school. *Social Forces, 90*(2), 521-545.

Byrne, M., & Flood, B. (2008). Examining the relationships among background variables and academic performance of first year accounting students at an Irish University. *Journal of Accounting Education, 26*, 202-212.

Cohn, E., Cohn, S., Balch, D. C., & Bradely Jr., J. (2004). Determinants of undergraduate GPAs: SAT scores, high-school GPA and high-school rank. *Economics of Education Review, 23*(577-586).

Cyrenne, P., & Chan, A. (2012). High school grades and university performance: A case study. *Economics of Education Review, 31*, 524-542.

DesJardinsa, S. L., Ahlburg, D. A., & McCall, B. P. (2006). The effects of interrupted enrollment on graduation from college: Racial, income, and ability differences. *Economics of Education Review, 25*, 575-590.

Education Counts. (2015). *New Zealand Schools*. Retrieved from https://http://www.educationcounts.govt.nz/data-services/directories/list-of-nz-schools

Garwe, E. C., & Manganga, E. (2015). The effect of student financial constraints on university non-completion rates. *International Journal of Education, 7*(2), 322-336.

Jia, P., & Maloney, T. (2015). Using predictive modelling to identify students at risk of poor university outcomes. *Higher Education, 70*, 127-149.

Johnes, J. (1997). Inter-university variations in undergraduate non- completion rates: A statistical analysis by subject of study. *Journal of Applied Statistics, 24*(3), 343-361.

Juhong, B., & Maloney, T. (2006). Ethnicity and academic success at university *New Zealand Economic Papers, 40*(2), 181-213.

Lasselle, L., McDougall-Bangall, J., & Smith, I. (2014). School grades, school context and university degree performance: evidence from an elite Scottish institution. *Oxford Review of Education, 40*(3), 293-314.

Liefner, I. (2003). Funding, resource allocation, and performance in higher education systems. *Higher Education, 46*, 469-489.

Montmarquette, C., Mahseredjian, S., & Houle, R. (2001). The determinants of university dropouts: A bivariate probability model with sample selection. *Economics of Education Review, 20*.

Murtaugh, P. A., Burns, L. D., & Schuster, J. (1999). Predicting the retention of university students *Research in Higher Education, 40*(355-371).

O'Keefe, M., Laven, G., & Burgess, T. (2011). Student non-completion of an undergraduate degree: wrong program selection or part of carrer plan? *Higher Education Research & Development, 30*(2), 165-177.

Ozga, J., & Sukhnandan, L. (1998). Undergraduate non-completion: Developing an explanatory model. *Higher Education Quarterly, 52*(3), 316-333.

Pearl, H. (2013). University first-year drop-out rates clear in new data. *Waikato Times*. Retrieved from http://www.stuff.co.nz/waikato-times/news/8988104/University-first-year-drop-out-rates-clear-in-new-data

Pike, G. R., & Killian, T. S. (2001). Reported gains in student learning: Do academic disciplines make a difference? *Research in Higher Education, 42*(4), 429-454.

Rask, K. (2010). Attrition in STEM fields at a liberal arts college: The importance of grades and pre-collegiate preferences. *Economics of Education Review, 29*, 892-900.

Robst, J., Keil, J., & Russo, D. (1998). The effect of gender composition of faculty on student retention. *Economics of Education Review, 17*(4), 429-439.

Rodgers, T. (2013). Should high non-completion rates amongst ethnic minority students be seen as an ethnicity issue? Evidence from a case study of a student cohort from a British University. *Higher Education, 66*, 535-550.

Singell, L. D., & Waddell, G. R. (2010). Modeling retention at a large public university: Can at-risk students be identified early enough to treat? *Research in Higher Education, 51*, 546-572.

Strarron, S. L., O'Toole, M. D., & Wetzel, N. J. (2008). A multinomial logit model of college stopout and dropout behaviour. *Economics of Education Review, 27*, 319-331.

Tertiary Education Commission. (2015a). *Key features of the PLF framework for the SAC funded sector*. Retrieved from http://www.tec.govt.nz/Funding/Policies-and-processes/Performance-linked-funding/Key-features-of-the-performance-linked-funding-framework-for-the-SAC-funded-sector/

Tertiary Education Commission. (2015b). *Successful completion of courses*. Retrieved from http://www.tec.govt.nz/Learners-Organisations/Learners/performance-in-tertiary-education/what-the-indicators-mean/Successful-completion-of-courses/

The World Bank Group. (2015). *School enrollment, tertiary (% gross)*. Retrieved from http://data.worldbank.org/indicator/SE.TER.ENRR/countries?display=map

Triventi, M. (2014). Does working during higher education affect students' academic progression? *Economics of Education Review, 41*, 1-13.

Vignoles, A. F., & Powdthavee, N. (2009). The socioeconomic gap in university dropouts *The B.E. Journal of Economic Analysis & Policy, 9*(1).

Wintre, M. G., & Yaffe, M. (2000). First-year students' adjustment to university life as a function of relationships with parents. *Journal of Adolescent Research, 15*(1), 9-37.

Yorke, M. (1998). Non-completion of full-time and sandwich students in English higher education: Costs to the public purse, and some implications. *Higher Education, 36*, 181-194.

Appendix

Table 2: Definition of all the variables

Variable	Definition
Dependent variable	
Successful completion	1 if paper was successfully completed
Year of Cohort	
Year 2014	1 if sudent enrolled in the year 2014
Year 2013	*Omitted category for students enrolled in the year 2013*
Ethinicity	
Asian	1 if student reported Asian under ethnicity
Maori	1 if student reported Maori under ethnicity
Pacifica	1 if student reported Pacifica under ethnicity
Other	1 if student reported none of the above
Not declared	1 if student did not declare their ethinic group
European	*Omitted category for those students who reported European under ethnicity*
Country of Origin	
China	1 if student's country of origin was China
India	1 if student's country of origin was India
New Zealand	1 if student's country of origin was NZ
Korea	1 if student's country of origin was Korea
Others	1 if student's country of origin was none of the above
Unknown	*Omitted category for those students whose country of origin was not known*
Demographic features	
Female	1 if student was female
Part-time	1 if student was studying part-time
LanEnglish	1 if student's first language was English
Age	1 if student's age was in one of the five dummy variables categories: Under 18, Age 19, Age 20, Age 21 and above 21. The omitted category was Age 18.
High school background	
Overall_NCEA_Score	Student's overall NCEA score for best 80 credits using points values of 2, 3 and 4 for Achieved, Merit and Excellence credits, respectively.
achievedcredits	Total Achieved credits obtained by a student
meritcredits	Total Merit credits obtained by a student
excellencecredits	Total Excellence credits obtained by a student
sc1	1 if school decile was 1
sc2	1 if school decile was 2
sc3	1 if school decile was 3
sc4	1 if school decile was 4
sc5	1 if school decile was 5

sc7	1 if school decile was 7
sc8	1 if school decile was 8
sc9	1 if school decile was 9
sc10	1 if school decile was 10
scUnk	1 if school decile for a student was not known
sc6	*Omitted category of school decile wac decile 6*
Entrance type	
External	1 for those students who had a pre-degree qualification from another university in NZ
Internal	1 for those students who had a pre-degree qualification from AUT
Bursary	1 if student's entrance type was Bursary
NCEA Level 3	1 if student's entrance type was NCEA Level 3
Other entrance type	1 if student's entrance type was not reported
Special admission	1 if student entered with special admission
Academic information	
BA	1 if student enrolled in Bachelor of Arts
BBus	1 if student enrolled in Bachelor of Business
BCIS	1 if student enrolled in Bachelor of Computer Information Science
BCS	1 if student enrolled in Bachelor of Communication Studies
BDes	1 if student enrolled in Bachelor of Design
BEdu	1 if student enrolled in Bachelor of Education
BEngTech	1 if student enrolled in Bachelor of Engineering Technology
BHS	1 if student enrolled in Bachelor of Health Science
BIHM	1 if student enrolled in Bachelor of International Hospitality Management
BSR	1 if student enrolled in Bachelor of Sports and Recreation
Other smaller programmes	*Omitted category for all other small Bachelor's degree programmes*
Double Degree	
Level 4	1 if student took a level 4 course
Level 6	1 if student took a level 6 course
Level 7	1 if student took a level 7 course
Level 5	*Omitted category for those students who took level 5 course*

Table 3: Descriptive Statistics

Variable	All first-year undergraduate students Mean	All first-year undergraduate students Std. Deviation	NCEA students with valid score Mean	NCEA students with valid score Std. Deviation
Dependent variable				
Successful Competition	0.7910	0.4066	0.8112	0.3914
Year				
Year 2014	0.5058	0.5000	0.5447	0.4980
Year 2013	0.4942	0.5000	0.4553	0.4980
Ethnicity				
Asian	0.2341	0.4234	0.2064	0.4047
Maori	0.1067	0.3088	0.1171	0.3215
Pacifica	0.1323	0.3388	0.1289	0.3351
Other	0.0619	0.2409	0.0553	0.2285
Not declared	0.0635	0.2439	0.0172	0.1302
European	0.4016	0.4902	0.4752	0.4994
Country of Origin				
China	0.0716	0.2579	0.0237	0.1521
India	0.0118	0.1082	0.0083	0.0909
New Zealand	0.7374	0.4400	0.8338	0.3723
Korea	0.0196	0.1388	0.0111	0.1049
Others	0.1555	0.3624	0.1201	0.3251
Unknown	0.0039	0.0625	0.0030	0.0546
Demographic features				
Female	0.6131	0.4870	0.6556	0.4752
Part-time	0.0967	0.2956	0.0513	0.2207
LanEnglish	0.5614	0.4962	0.6035	0.4892
Age	20.8493	4.9260	18.4898	0.7114
High School Backgrounds				
Overall NCEA Score*	178.2833	64.2512	192.0403	54.0991
Achieved Credits*	39.7563	16.9936	41.5899	16.4710
Merit Credits*	21.3993	14.6954	23.4922	14.2655
Excellence Credits*	12.2827	16.1328	13.7475	16.6566
sc1	0.0390	0.1937	0.0349	0.1835
sc2	0.0399	0.1958	0.0434	0.2037
sc3	0.0624	0.2418	0.0685	0.2526
sc4	0.0942	0.2921	0.1005	0.3006
sc5	0.0577	0.2332	0.0682	0.2521
sc7	0.0854	0.2795	0.0914	0.2882
sc8	0.0807	0.2724	0.0966	0.2954
sc9	0.1323	0.3388	0.1613	0.3678
sc10	0.2144	0.4104	0.2434	0.4291
scUnk	0.1292	0.3354	0.0117	0.1075
sc6	0.0647	0.2461	0.0802	0.2716
Entrance Type				
External	0.1443	0.3514	-	-
Internal	0.0961	0.2948	-	-
Bursary	0.0352	0.1842	-	-
NCEA_LEVEL_3	0.5806	0.4935	1.0000	0.0000
Other Entrance Type	0.0083	0.0909	-	-
Special Admissions	0.1355	0.3422	-	-
Academic Information				
BA	0.1049	0.3064	0.1041	0.3053
BBus	0.1973	0.3979	0.1759	0.3807
BCIS	0.0553	0.2286	0.0455	0.2085

BCS	0.0720	0.2584	0.1137	0.3175
BDes	0.0632	0.2433	0.0892	0.2850
BEdu	0.0349	0.1835	0.0263	0.1601
BEngTech	0.0385	0.1923	0.0344	0.1822
BHS	0.2165	0.4119	0.2023	0.4017
BIHM	0.0434	0.2037	0.0364	0.1872
BSR	0.0561	0.2302	0.0649	0.2464
Others	0.1180	0.3226	0.1074	0.3096
Double Degree	0.0103	0.1010	0.0132	0.1141
Level 4	0.0048	0.0691	0.0045	0.0672
Level 6	0.1177	0.3223	0.1000	0.3000
Level 7	0.0050	0.0705	0.0019	0.0440
Level 5	0.8725	0.3335	0.8935	0.3085
Number of Observations	64,446		32,423	

*Number of observations for the sample of all first-year undergraduate students who had reported Overall NCEA Score, number of Achieved, Merit and Excellence credits was 37,122.

Table 4: Model 1 - Probit model parameter estimates and average marginal effects on the dependent variables for all independent variables and best 80 credits broken by Achieved, Merit and Excellence Categories

Variable	Coefficient	Std. Error	dy/dx
Constant	-0.4760*	0.2751	-
Year of Cohort			
Year 2014	-0.1305***	0.0303	-3.04%***
Ethnicity			
Asian	0.0363	0.0448	0.84%
Maori	-0.2132***	0.0502	-4.96%***
Pacifica	-0.3633***	0.0498	-8.46%***
Other	-0.0945	0.0651	-2.20%
Not Declared	-0.2371*	0.1247	-5.52%*
Country of origin			
China	0.0776	0.2729	1.81%
India	0.2876	0.3011	6.69%
New Zealand	0.2562	0.2517	5.96%
Korea	0.1715	0.2883	3.99%
Others	0.2234	0.2554	5.20%
Demographic features			
Female	0.1288***	0.0351	3.00%***
Part-time	-0.1531***	0.0572	-3.56%***
LanEnglish	-0.1125***	0.0318	-2.62%***
Under 18	-0.3126*	0.1647	-7.27%*
Age 19	0.0065	0.0328	0.15%
Age 20	0.0187	0.0607	0.44%
Age 21	0.2992**	0.1330	6.96%**
Above 21	0.5240	0.4232	12.20%
High School backgrounds			
Achieved Credits	0.0124***	0.0016	0.29%***
Merit Credits	0.0251***	0.0016	0.58%***
Excellence Credits	0.0259***	0.0017	0.60%***
sc1	-0.4539***	0.0965	-10.56%***
sc2	-0.1727**	0.0878	-4.02%**
sc3	-0.1899**	0.0804	-4.42%**
sc4	-0.1724**	0.0740	-4.01%**
sc5	-0.1779**	0.0791	-4.14%**
sc7	-0.0697	0.0790	-1.62%
sc8	-0.2923***	0.0742	-6.80%***
sc9	-0.1026	0.0700	-2.39%
sc10	-0.1870***	0.0658	-4.35%***
scUnk	-0.0091	0.1489	-0.21%
Academic information			

BA	0.0758	0.0656	1.76%
BBus	-0.0041	0.0579	-0.10%
BCIS	0.0700	0.0786	1.63%
BCS	0.4707***	0.0798	10.96%***
BDes	0.3354***	0.0749	7.81%***
BEdu	0.6877***	0.1347	16.01%***
BEngTech	-0.0247	0.0896	-0.57%
BHS	0.2684***	0.0593	6.25%***
BIHM	0.6070***	0.0983	14.13%***
BSR	-0.0379	0.0723	-0.88%
Double Degree	0.8141***	0.1624	18.95%***
Level 4	0.2965***	0.1086	6.90%***
Level 6	0.0083	0.0384	0.19%
Level 7	0.1503	0.2061	3.50%
Number of Observations			32,423
Number of Independent Variables			46
Pseudo R^2 Statistic			0.1385
Log Pseudo likelihood			-13,533.592

Note: *** denotes significance at 1% level, **denotes significance at 5% level, * denotes significance at 10% level.

Table 5: Model 2 - Probit model parameter estimates and average marginal effects on the dependent variables for all independent variables and overall score calculated using 2*Achieved, 3*Merit and 4*Excellence points values

Variable	Coefficient	Std. Error	dy/dx
Constant	-0.5542**	0.2674	-
Year of Cohort			
Year 2014	-0.1268***	0.0302	-2.96%***
Ethnicity			
Asian	0.0351	0.0446	0.82%
Maori	-0.2169***	0.0505	-5.06%***
Pacifica	-0.3696***	0.0498	-8.62%***
other	-0.1000	0.0650	-2.33%
Not Declared	-0.2440*	0.1257	-5.69%*
Country of origin			
China	0.0738	0.2725	1.72%
India	0.2819	0.3002	6.57%
New Zealand	0.2519	0.2510	5.87%
Korea	0.1644	0.2880	3.83%
Others	0.2136	0.2548	4.98%
Demographic features			
Female	0.1359***	0.0351	3.17%***
Part-time	-0.1551***	0.0571	-3.62%***
LanEnglish	-0.1144***	0.0318	-2.67%***
Under 18	-0.3240*	0.1661	-7.55%*
Age 19	0.0022	0.0328	0.05%
Age 20	0.0127	0.0611	0.30%
Age 21	0.2918**	0.1334	6.80%**
Above 21	0.5239	0.4313	12.22%
High School backgrounds			
Overall NCEA Score	0.0075***	0.0003	0.17%***
sc1	-0.4710***	0.0964	-10.98%***
sc2	-0.1827**	0.0878	-4.26%**
sc3	-0.1888**	0.0804	-4.40%**
sc4	-0.1789**	0.0739	-4.17%**
sc5	-0.1828**	0.0791	-4.26%**
sc7	-0.0624	0.0788	-1.46%
sc8	-0.2919***	0.0742	-6.81%***
sc9	-0.0997	0.0699	-2.32%
sc10	-0.1830***	0.0657	-4.27%***
scUnk	0.0126	0.1498	0.29%
Academic information			
BA	0.0755	0.0656	1.76%
BBus	-0.0089	0.0579	-0.21%
BCIS	0.0576	0.0786	1.34%
BCS	0.4776***	0.0800	11.14%***
BDes	0.3295***	0.0756	7.68%***

BEdu	0.6776***	0.1346	15.80%***
BEngTech	-0.0224	0.0894	-0.52%
BHS	0.2676***	0.0594	6.24%***
BIHM	0.6056***	0.0977	14.12%***
BSR	-0.0387	0.0724	-0.90%
Double Degree	0.8033***	0.1622	18.73%***
Level 4	0.2932***	0.1078	6.84%***
Level 6	0.0105	0.0385	0.25%
Level 7	0.1613	0.2061	3.76%
Number of Observations			32,423
Number of Independent Variables			44
Pseudo R^2 Statistic			0.1371
Log Pseudo likelihood			-13,555.346

Note: *** denotes significance at 1% level, **denotes significance at 5% level, * denotes significance at 10% level.

Table 6: Model 3 - Probit model parameter estimates and average marginal effects on the dependent variables for all independent variables and new overall score calculated using 1*Achieved, 2*Merit and 2*Excellence points values

Variable	**Coefficient**	**Std. Error**	**dy/dx**
Constant	-0.4954*	0.2670	-
Year of Cohort			
Year 2014	-0.1300***	0.0302	-3.02%***
Ethnicity			
Asian	0.0354	0.0447	0.82%
Maori	-0.2124***	0.0502	-4.94%***
Pacifica	-0.3642***	0.0498	-8.47%***
other	-0.0948	0.0652	-2.21%
Not Declared	-0.2363*	0.1247	-5.50%*
Country of origin			
China	0.0738	0.2729	1.72%
India	0.2823	0.3008	6.57%
New Zealand	0.2520	0.2516	5.86%
Korea	0.1671	0.2885	3.89%
Others	0.2191	0.2553	5.10%
Demographic features			
Female	0.1300***	0.0352	3.03%***
Part-time	-0.1544***	0.0572	-3.59%***
LanEnglish	-0.1124***	0.0318	-2.62%***
Under 18	-0.3117*	0.1653	-7.25%*
Age 19	0.0065	0.0328	0.15%
Age 20	0.0193	0.0608	0.45%
Age 21	0.3013**	0.1331	7.01%**
Above 21	0.5318	0.4246	12.37%
High School backgrounds			
New Overall NCEA Score	0.0128***	0.0006	0.30%***
sc1	-0.4530***	0.0964	-10.54%***
sc2	-0.1717**	0.0878	-3.99%**
sc3	-0.1899**	0.0804	-4.42%**
sc4	-0.1711**	0.0739	-3.98%**
sc5	-0.1788**	0.0790	-4.16%**
sc7	-0.0690	0.0790	-1.61%
sc8	-0.2915***	0.0742	-6.78%***
sc9	-0.1025	0.0700	-2.38%
sc10	-0.1866***	0.0658	-4.34%***
scUnk	-0.0082	0.1491	-0.19%
Academic information			
BA	0.0754	0.0657	1.75%
BBus	-0.0045	0.0579	-0.10%
BCIS	0.0701	0.0785	1.63%
BCS	0.4713***	0.0797	10.97%***
BDes	0.3392***	0.0748	7.89%***

BEdu	0.6880***	0.1348	16.01%***
BEngTech	-0.0263	0.0897	-0.61%
BHS	0.2682***	0.0593	6.24%***
BIHM	0.6083***	0.0983	14.15%***
BSR	-0.0383	0.0723	-0.89%
Double Degree	0.8186***	0.1613	19.05%***
Level 4	0.2971***	0.1086	6.91%***
Level 6	0.0087	0.0384	0.20%
Level 7	0.1505	0.2061	3.50%
Number of Observations			32,423
Number of Independent Variables			44
Pseudo R^2 Statistic			0.1385
Log Pseudo likelihood			-13,534.026

Note: *** denotes significance at 1% level, **denotes significance at 5% level, * denotes significance at 10% level.

Table 7: Model 4 - Probit model parameter estimates and average marginal effects on the dependent variables for all independent variables and all Achieved, Merit and Excellence credits

Variable	**Coefficient**	**Std. Error**	**dy/dx**
Constant	-0.3600	0.2639	-
Year of Cohort			
Year 2014	-0.1265***	0.0303	-2.94%***
Ethnicity			
Asian	0.0226	0.0450	0.53%
Maori	-0.2062***	0.0502	-4.79%***
Pacifica	-0.3654***	0.0499	-8.49%***
Other	-0.0942	0.0652	-2.19%
Not Declared	-0.2488*	0.1250	-5.78%*
Country of origin			
China	0.0801	0.2680	1.86%
India	0.2603	0.2979	6.05%
New Zealand	0.2568	0.2466	5.97%
Korea	0.1767	0.2838	4.11%
others	0.2231	0.2504	5.19%
Demographic features			
Female	0.1372***	0.0351	3.19%***
Part-time	-0.1507***	0.0573	-3.50%***
LanEnglish	-0.1127***	0.0319	-2.62%***
Under 18	-0.3183*	0.1660	-7.40%*
Age 19	0.0051	0.0328	0.12%
Age 20	0.0112	0.0604	0.26%
Age 21	0.2938**	0.1303	6.83%**
Above 21	0.5045	0.4154	11.73%
High School backgrounds			
Achieved Credits	0.0112***	0.0012	0.26%***
Merit Credits	0.0215***	0.0013	0.50%***
Excellence Credits	0.0205***	0.0015	0.48%***
sc1	-0.4596***	0.0968	-10.68%***
sc2	-0.1765**	0.0877	-4.10%**
sc3	-0.1912**	0.0806	-4.44%**
sc4	-0.1754**	0.0743	-4.08%**
sc5	-0.2026**	0.0799	-4.71%**
sc7	-0.0757	0.0794	-1.76%
sc8	-0.3097***	0.0747	-7.20%***
sc9	-0.1008	0.0701	-2.34%
sc10	-0.1817***	0.0660	-4.22%***
scUnk	-0.0291	0.1466	-0.68%
Academic information			
BA	0.0821	0.0656	1.91%
BBus	-0.0003	0.0579	-0.01%
BCIS	0.0680	0.0784	1.58%

BCS	0.4772***	0.0802	11.09%***
BDes	0.3430***	0.0751	7.97%***
BEdu	0.6836***	0.1340	15.89%***
BEngTech	-0.0284	0.0904	-0.66%
BHS	0.2641***	0.0594	6.14%***
BIHM	0.6081***	0.0976	14.13%***
BSR	-0.0300	0.0723	-0.70%
Double Degree	0.7571***	0.1674	17.60%***
Level 4	0.2979***	0.1085	6.92%***
Level 6	0.0000	0.0386	0.00%
Level 7	0.1683	0.2050	3.91%
Number of Observations			32,423
Number of Independent Variables			46
Pseudo R^2 Statistic			0.1410
Log Pseudo likelihood			-13,494.41

Note: *** denotes significance at 1% level, **denotes significance at 5% level, * denotes significance at 10% level.

Table 8: Model 5 - Probit model parameter estimates and average marginal effects on the dependent variables for all independent variables and overall score calculated for all credits using 1*Achieved, 2*Merit and 2*Excellence points values

Variable	**Coefficient**	**Std. Error**	**dy/dx**
Constant	-0.3303	0.2607	-
Year of Cohort			
Year 2014	-0.1277***	0.0302	-2.97%***
Ethnicity			
Asian	0.0248	0.0448	0.58%
Maori	-0.2077***	0.0503	-4.83%***
Pacifica	-0.3635***	0.0498	-8.45%***
other	-0.0933	0.0651	-2.17%
Not Declared	-0.2493*	0.1251	-5.80%*
Country of origin			
China	0.0863	0.2678	2.01%
India	0.2708	0.2977	6.30%
New Zealand	0.2639	0.2464	6.14%
Korea	0.1840	0.2838	4.28%
Others	0.2307	0.2502	5.36%
Demographic features			
Female	0.1342***	0.0350	3.12%***
Part-time	-0.1488***	0.0573	-3.46%***
LanEnglish	-0.1126***	0.0319	-2.62%***
Under 18	-0.3185*	0.1648	-7.41%*
Age 19	0.0056	0.0328	0.13%
Age 20	0.0110	0.0604	0.26%
Age 21	0.2911**	0.1306	6.77%**
Above 21	0.4930	0.4109	11.46%
High School background			
Overall NCEA score	0.0105***	0.0005	0.24%***
sc1	-0.4596***	0.0965	-10.69%***
sc2	-0.1773**	0.0877	-4.12%**
sc3	-0.1913**	0.0806	-4.45%**
sc4	-0.1768**	0.0742	-4.11%**
sc5	-0.1993**	0.0796	-4.64%**
sc7	-0.0769	0.0795	-1.79%
sc8	-0.3100***	0.0747	-7.21%***
sc9	-0.1013	0.0702	-2.36%
sc10	-0.1830***	0.0661	-4.25%***
scUnk	-0.0308	0.1461	-0.72%
Academic information			
BA	0.0823	0.0656	1.91%

BBus	0.0003	0.0579	-0.01%
BCIS	0.0689	0.0783	1.60%
BCS	0.4758***	0.0802	11.06%***
BDes	0.3370***	0.0755	7.84%***
BEdu	0.6838***	0.1339	15.90%***
BEngTech	-0.0258	0.0904	-0.60%
BHS	0.2648***	0.0594	6.16%***
BIHM	0.6061***	0.0974	14.09%***
BSR	-0.0297	0.0723	-0.69%
Double Degree	0.7546***	0.1677	17.55%***
Level 4	0.2968***	0.1085	6.90%***
Level 6	-0.0002	0.0386	0.00%
Level 7	0.1661	0.2051	3.86%
Number of Observations			32,423
Number of Independent Variables			44
Pseudo R^2 Statistic			0.1410
Log Pseudo likelihood			13,495.486

Note: *** denotes significance at 1% level, **denotes significance at 5% level, * denotes significance at 10% level.