Understanding Ethnic Differences in Student Success at Universities in New Zealand Using Linked Administrative Data

Zhaoyi Cao

A thesis submitted to Auckland University of Technology in fulfilment of the requirements of Doctor of Philosophy

School of Economics, Faculty of Business, Economics and Law

2021

Abstract

Using administrative data provided by a New Zealand university, Cao and Maloney (2018) examined the academic differences in first-year course completion and GPA achievement between the main ethnic minority groups (Māori, Pasifika and Asian) and European. The authors found that roughly a quarter of the Māori/Pasifika-European gaps in academic outcomes could be explained by observable factors that included personal characteristics, high school backgrounds and university enrolment patterns. The current thesis extends Cao and Maloney (2018) in two ways. First, it uses more data on personal, family and school backgrounds that stored in a large database, the IDI, which is maintained and operated by Statistics New Zealand. Second, this is national-level analysis on a broader range of university outcomes including participation, first- and second-year course completions, and, finally, qualification completion.

The original study sample in this thesis is comprised of approximately 180,000 individuals who turned 15 years old in the years 2010 to 2012 and were enrolled in a school in New Zealand. These students are split into three age-15 cohorts. Second-year course analysis is restricted to the 2010 and 2011 cohorts, while degree completion analysis is further limited to the 2010 cohort. Based on ethnicity prioritisation, the population is ethnically broken down into 57.1% European, 21% Māori, 10.3% Asian, 9.1% Pasifika, 1.8% MELAA and 0.8% from other ethnicities. Probit with Maximum Likelihood, Fractional Probit and decomposition techniques are used to answer the research questions. Other personal characteristics, school characteristics, parent-related factors and university factors are included in our analysis.

The regression results show clear patterns that are relatively in line with the literature. First, there are some sizeable overall differences between the ethnic minority and the European students. The most academically disadvantaged groups are Māori and Pasifika, who, when compared to European, are much less likely to undertake bachelor studies at university, complete their courses in the first and second years, or complete a degree qualification. Asian are initially more likely to participate in university relative to European, but to have lower levels of performance in the later stages of university studies. Our decomposition results indicate that giving Māori and Pasifika the same characterises as European could close most of these ethnic gaps in university participation, no more than half of the gaps in course completions and about one quarter of the gaps in degree completion.

Disclaimer

Access to the data used in this thesis was provided by Statistics New Zealand under conditions designed to give effect to the security and confidentiality provisions of the Statistics Act 1975. The results presented in this study are the work of the author, not Statistics New Zealand or individual data suppliers.

The results are based in part on tax data supplied by Inland Revenue to Statistics New Zealand under the Tax Administration Act 1994 for statistical purposes. Any discussion of data limitations or weaknesses is in the context of using the Integrated Data Infrastructure IDI for statistical purposes, and is not related to the data's ability to support Inland Revenue's core operational requirements.

These results are not official statistics. They have been created for research purposes from the Integrated Data Infrastructure (IDI). which is carefully managed by Statistics New Zealand. For more information about the IDI please visit https://www.stats.govt.nz/integrated-data/

Table of Contents

List of Figures

List of Tables

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.

Signed: _____ Date: 31/05/2021

Acknowledgements

I would first like to thank my supervisors, Professor Tim Maloney and Associate Professor Peer Skov, who provided me with a great deal of support, assistance and advice throughout my PhD journey. The completion of this thesis would not have been possible without my supervisors' help. A debt of gratitude is also owed to Statistics New Zealand for approving access to the data used in this thesis and for checking and releasing the analytical outputs. I would like to also thank my colleagues for their support, guidance and encouragement. All remaining errors and omissions are solely my own.

Chapter 1. Background

1.1. Introduction

In today's rapidly changing world, many people understand the importance of university education. Human capital theory states that a better education leads to higher potential earnings through increases in productivity (Becker, 1993). For instance, the New Zealand Ministry of Education reported in 2014 that 54% of undergraduates were employed one year after study, while the employment rate for those who completed a level 1 to 3 certificate was 35%. The median earnings for people who received a Bachelor's qualification was 46% higher than the national level median earnings and 45% above the median incomes of those who completed a certificate at levels 1 to 3 (Ministry of Education, 2014). In addition to providing more job opportunities and higher earnings, university education also improves a person's quality of life. Studies showed that when compared to those with any school qualification, or none at all, university graduates have healthier behaviours including not smoking, a greater engagement in civil activities, greater access to health care and, in general, happier lives (e.g., Hout, 2012; Pampel, Krueger & Denney, 2010).

Despite university education being associated with such benefits, significant differences in academic achievement with respect to ethnicity continue to exist in university education across many OECD countries including New Zealand. The primary subject of this thesis is to explore ethnic differences in a series of academic outcomes at all universities in New Zealand.

New Zealand is a highly multicultural country and now includes individuals who identify with over 200 ethnicities originating from nearly 200 countries, according to the 2018 Census. Given this diverse population, New Zealand is an ideal place in which to conduct a study of this kind, which looks at ethnic differences in academic outcomes at university education level. The five largest ethnic groups in the country are New Zealand European, Māori, Pasifika, Asian, and Middle Eastern, Latin American and African (MELAA). Māori are the native Polynesian people of Aotearoa (New Zealand), who arrived in New Zealand from Eastern Polynesia sometime between 1250 and 1300 AD (Howe, 2003). Pasifika peoples are from the South Pasifika regions and include people born in New Zealand who have close family and cultural connections to Pasifika Island countries, such as the Cook Islands, Tonga, Niue, Samoa, Tuvalu, and Tokelau (Ministry of Education, 2006). Asian immigration to New Zealand began in the 1970s when the immigration policy was relaxed (Brawley, 1993). The six largest Asian immigrant populations currently are from China, India, the Philippines, Korea, Japan and Sri Lanka.

The data from the most recent 2018 Census reveals that 70.2% of the population in New Zealand identified as European, which makes up the largest ethnic group in the country. The second largest ethnic group is Māori (16.5%), followed by Asian (15.1%), Pasifika people (8.1%) and MELAA (1.5%). Other ethnicities account for the remaining 1.2% of the population. It should be noted that the sum of the proportions for the ethnic groups mentioned above is greater than 100%. This is because ethnicity in New Zealand is self-identified and people are often allowed to report multiple ethnicities (as they were in the 2018 Census). Thus, the total responses are more than the total number of respondents in the survey.

1.2. Government involvement in tertiary education

In New Zealand, tertiary education covers all post-secondary school education, thus it contains both higher formal education and vocational training. The tertiary education providers in the country include universities, institutes of technology and polytechnics (ITPs), Wānanga[1], private training establishments (PTEs) and workplace training. There are eight universities that are all funded by the government. Degree level or higher education is mainly offered by the universities. The programmes are usually research-led and more academic than the ones taught in vocational education. The Tertiary Education Commission is in charge of funding tertiary education and implementing tertiary education strategies and priorities.

The recent history of tertiary education reforms in New Zealand can be traced back to the late 1980s (Crawford, 2016). The fourth Labour Government carried out the first major tertiary education reform in 1989, which was driven by the belief that tertiary education providers were reacting to the needs of the rapidly changing economy and the growing demand for skilled workers. This required an expanded participation in tertiary education, especially for the disadvantaged and low-income groups such as Māori, Pasifika and women. As a result of the reform, tertiary education participation rapidly increased during the 1990s. The fast growth in participation created serious problems for government spending that resulted in a second reform. The second wave of reforms in tertiary education was conducted during the late 1990s and early 2000s, where the main purpose was to balance public and private funding for tertiary education and to ensured participation results in successful qualification completion. A few years after the second reforms, the latest tertiary education reforms were initiated in the mid-

[1] Wānanga is a special state-owned tertiary institution in the New Zealand education system, which offers education in Māori cultural traditions. For more information about Wānanga, see https://www.nzqa.govt.nz/audience-pages/wananga/.

2000s with a focus on increasing the contribution of tertiary education to economic development and to improving graduates' labour market outcomes.

Over time a series of government and institutional policies have been made with the aim of improving tertiary outcomes for two key groups of learners, Māori and Pasifika. A typical example of government strategy is one of the six strategic priorities of the Tertiary Education Strategy 2014-2019 to boost the achievement of Māori and Pasifika students (Ministry of Education, 2014). In terms of the strategies related to Māori education, Ka Hikitia — Managing for Success 2008–2012 set the goal to improve how the education system performs for Māori leaners, and its successor Ka Hikitia — Accelerating Success 2013-2017 focused on improving how the education system performs to ensure Māori students enjoy and achieve education success (Ministry of Education, 2013). Also, Tū Māia e te Ākonga strategy sought to deal with the participation and completion rates for Māori leaners in tertiary education at higher levels (Tertiary Education Commission, 2016).

With regard to Pasifika education strategies, the Pasifika Education Plan of 2013-2017 was targeted to bringing Pasifika students' participation and achievement, at all educational levels, up to at least on par with other students (Ministry of Education, 2013), and the goal of the Pasifika Operational Strategy 2017-2020 was to ensure that more Pasifika students were successful in tertiary education (Tertiary Education Commission, 2017).

There are also institutional actions supporting Māori and Pasifika students. Examples include the Tuākana programme, which enhances the academic success of Māori and Pasifika students at the University of Auckland, the Wānanga series that aims to build a culture of support for AUT (Auckland University of Technology) Māori and Pasifika Postgraduate students across the entire University, and the Āwhina and Pasifika Student Success academic mentoring programmes offered at the Victoria University of Wellington, which aims to help Māori and Pasifika students achieve better grades.

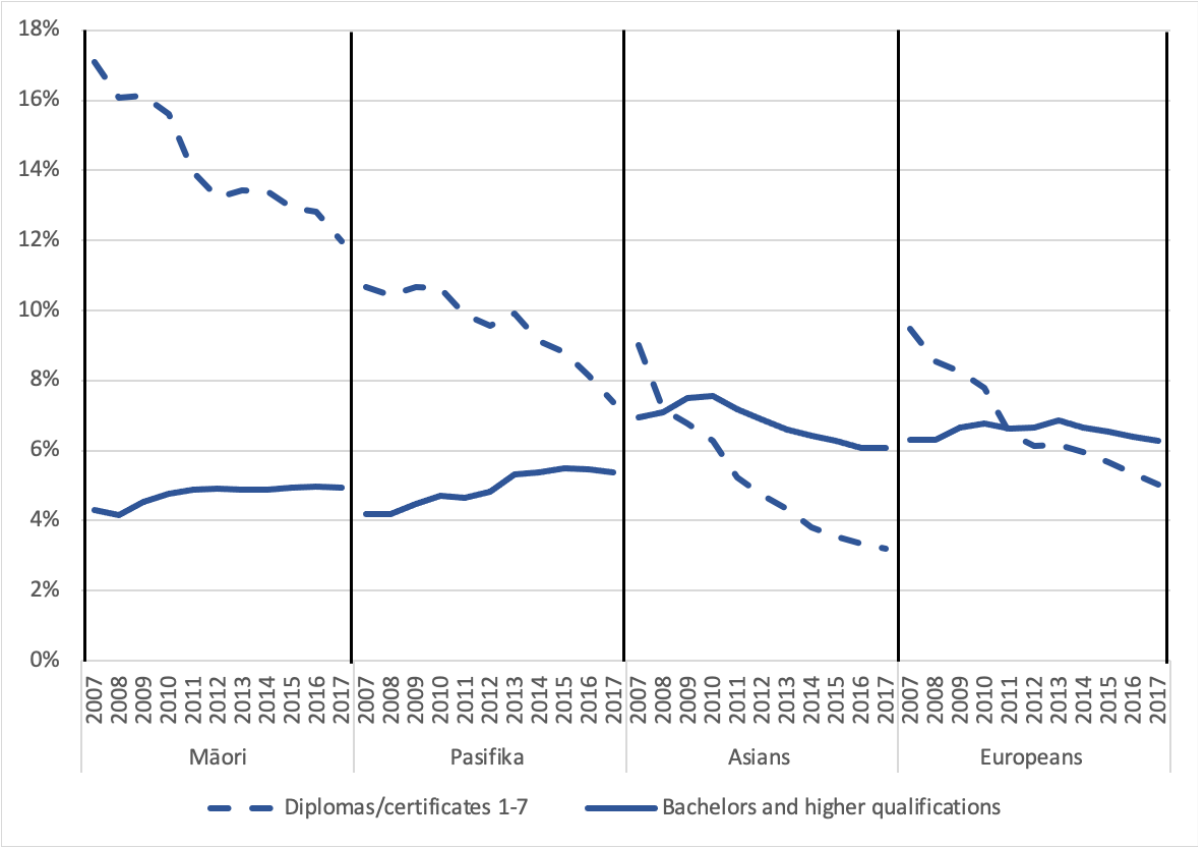
1.3. Ethnic differences in the New Zealand tertiary education system

There is some evidence that targeted educational reforms and strategies have been successful in improving the academic performance of tertiary students, especially among Māori and Pasifika leaners. For example, Māori have participated in tertiary level education at the highest rate compared to any other ethnicities since 2006 (Ministry of Education, 2017). Also, the Māori population with a tertiary qualification increased by about five percent over the period from 1991 to 2005 (Ministry of Education, 2006).

Although these results are very encouraging, recent educational data indicate that Māori students are still not as likely to succeed as their European peers, especially in bachelor and higher-level studies. For instance, as shown in Figure 1, Māori had the highest participation rates (age-standardised) among any ethnic group in diplomas or certificates 1-7 qualifications for the years between 2007 to 2017. However, they also had the lowest participation rates among any ethnic group in bachelor's and higher qualifications during the same ten year period.

Additionally, it can be seen from Figure 2 that, for the students who began a tertiary qualification between 2005 and 2018, Māori students had a higher average first-year retention rate in diplomas or certificates for 1-7 level qualifications, but were the least likely to continue on to second-year study in bachelors and higher-level qualifications. Furthermore, Figure 3 shows that Māori tertiary students starting a qualification during the years 2005 to 2012 had a lower average of eight year completion rates in bachelors and higher-degree qualifications than that of European, even though the probability of them completing diplomas/certificates levels 1-7 qualifications were higher.

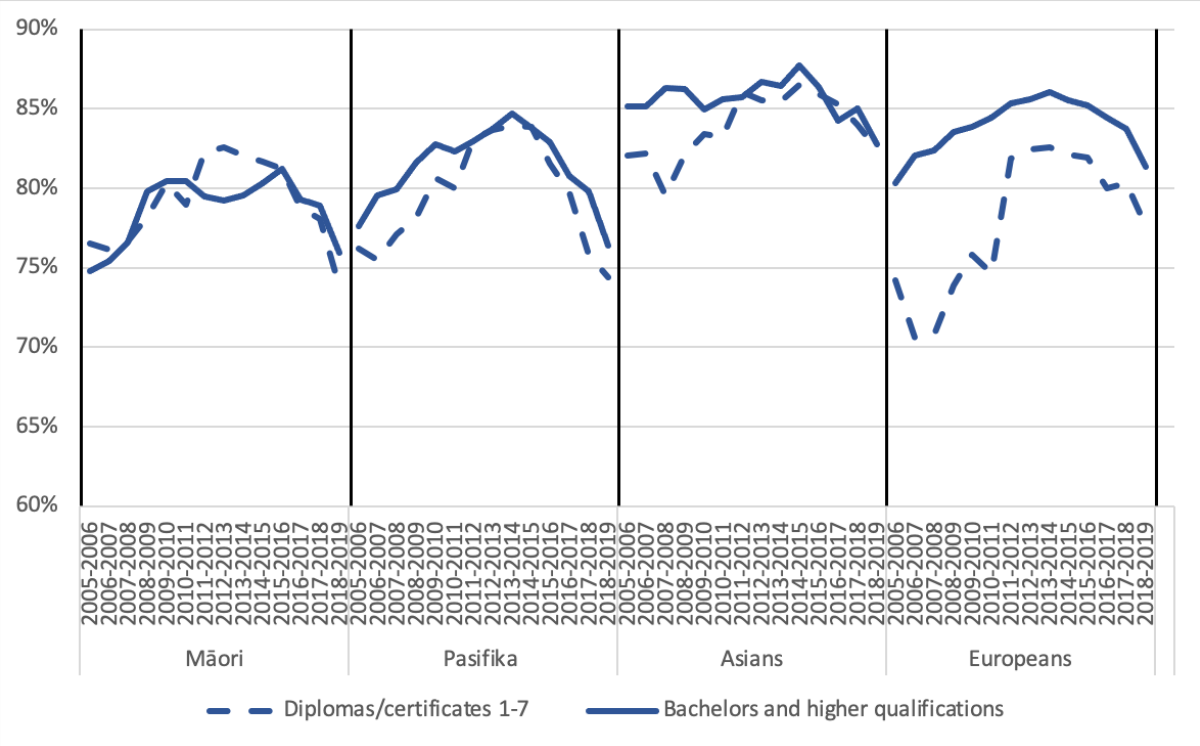
Figure 1.1. Participation rates (age-standardised) at higher education by qualification levels and ethnicity



Data source: Education counts https://www.educationcounts.govt.nz/home.

Pasifika participation rate in tertiary education also increased from 1997 to 2000 (Ministry of Education, 2001). Moreover, tertiary enrolments by Pasifika students grew at a faster rate than that of European between 1996 and 2001 (Ministry of Education, 2002). Despite these improvements, participation rates in bachelor's and higher qualifications for Pasifika students continued to be lower than the rate for European during 2007 to 2017 (see Figure 1). Pasifika also had average first-year retention rates in bachelor or higher-level qualifications that were lower than the rates for European for the years 2005 to 2018 (see Figure 2). Similarly to Māori, Pasifika students on average were less likely than European to completed qualifications at bachelors and higher levels within eight years since the start of the qualification during 2005 and 2012 (see Figure 3).

Figure 1.2. Average first-year retention rates at higher education by qualification level and ethnicity

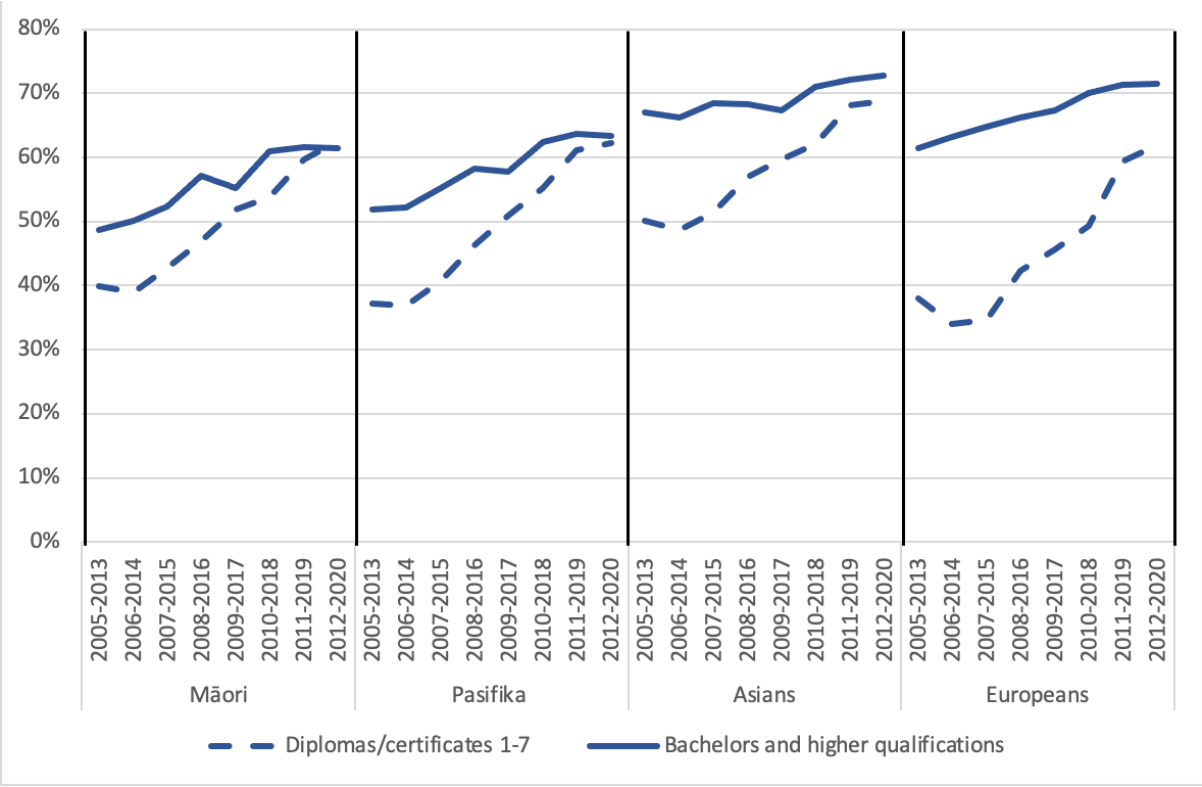


Data source: Education counts https://www.educationcounts.govt.nz/home.

There has been much concern over the historical underachievement of Māori and Pasifika students in tertiary education system. It has been estimated by the Tertiary Education Commission that Māori and Pasifika communities will make up 30% of the labour force in 2020 (Tertiary Education Commission, 2018), so their educational success is important to the

country's future prosperity. If Māori and Pasifika learners performed equally well as other equivalent students in tertiary study, there would be another 2,700 Māori and 1,100 Pasifika completing a degree qualification each year. It is reasonable to assume that highly educated Māori and Pasifika could earn as much as non-Māori and non-Pasifika. As a consequence, an additional $123 million and $55 million could be generated for Māori and Pasifika households respectively (Universities New Zealand, 2018). These numbers indicate that lifting the academic performance of the two ethnic groups to the same levels of other students in tertiary education can not only effectively promote economic growth but also help to build a more socially cohesive and equitable society. Therefore, improving Māori and Pasifika's tertiary outcomes has been a top priority for the Ministry of Education and for tertiary organizations.

Figure 1.3. Average eight-year qualification completion rates at higher education by qualification level and ethnicity



Data source: Education counts https://www.educationcounts.govt.nz/home.

Researchers have attempted to look at tertiary educational disparities between ethnic minority groups and European students in New Zealand. For example, Scott (2005a) used unit record level data on student enrolments and degree completions from 1998 to 2002 to examine their retention, attrition, completion and progression status. The author found that participation and completion rates for Māori and Pasifika at degree level and above were significantly lower than

the rates for European and Asian students during the sample period. In the same vein, Earle (2007) noted that Māori students were significantly underrepresented in degree level studies.

Instead of examining ethnic differences in participation, Juhong and Maloney (2006) had an interest in students' drop-out behaviour that can be the opposite indicator of qualification completion. The authors found that for a cohort of students enrolled in Arts and Science degrees at a large urban university in the year 2000, Māori and Pasifika students were more likely to discontinue their study compared with their European counterparts. The differences in dropout rates between Māori or Pasifika and European could be largely explained by the differences in GPA achievement.

A recent study by Meehan, Pacheco, & Pushon (2019) used unit-level administrative data from all New Zealand tertiary education providers to explore ethnic differences in degree participation for a sample of individuals born between 1990 and 1994. The authors found that Māori and Pasifika were 19.84 and 11.45 percentage points respectively, less likely than European school leavers to undertake a bachelor-level qualification before the age of 20. If Māori were accorded the same observed factors as that for European, it could reduce the ethnic gap in participation by no more than 87%. Removing the differences in the measurable factors between Pasifika and European could fully explain this ethnic gap for those enrolling in a bachelor's programme.

1.4. Motivation for the current thesis

Another New Zealand study, Cao and Maloney (2018) was sought to explain ethnic gaps in first-year course completions and in GPAs in bachelor's degrees. Following a cohort of students who enrolled in bachelor's degree programmes at an urban New Zealand university from 2012 to 2015, the authors examined first-year academic differences in course completion rates and GPAs between European students and three other ethnic minority groups - Māori, Pasifika and Asian. The authors found that Māori and Pasifika, on average, lagged behind European students in terms of both course completion and GPA outcomes. Their decomposition results suggested that high school academic achievement proxied by NCEA scores was the most significant factor in explaining the relatively poorer first-year academic outcomes of both Māori and Pasifika.

It was found that eliminating differences in all measurable factors (personal characteristics, high school backgrounds and university enrolment patterns) between Māori/Pasifika and European students could explain at most one-quarter of the observed ethnic gaps in the first-

year university performance. These results were statistically significant at a 1% level. It was speculated that the substantial unexplained ethnic gaps could be accounted for by information such as family backgrounds (e.g., parental education, household income and family structure) and school histories (e.g., number of schools attended and truancy records) that was unavailable in the dataset provided by the university used for this study. Therefore, we want to know how much of the unexplained ethnic differences in university outcomes remains once we are able to control for a wider range of personal and household characteristics, and for past educational experiences.

The current project is an extension of this previous research by Cao and Maloney. Specifically, this thesis is designed to investigate ethnic differences in a series of academic outcomes across all universities in New Zealand, taking into account more detailed student information drawn from the mainly administrative data stored in the Integrated Data Infrastructure (IDI).

The analysis will cover four academic outcomes from the beginning to the end of university studies. These are participation, first- and second-year course completion, and qualification completion. The IDI is a large research database that contains data collected by government agencies, Statistics New Zealand surveys and non-government organisations. This thesis uses data from the Ministry of Education, the Ministry of Social Development, the Ministry of Justice, the 2013 Census survey, and the Inland Revenue Department. The use of these data allows us to control for demographic characteristics, school characteristics, school histories, university factors and parental backgrounds when examining the ethnic differences in the university outcomes. We can also look at the impact of employment while studying on course and qualification completions by using the IDI data. Maximum Likelihood Probit and Fractional Probit estimations, together with linear and non-linear decomposition techniques, are utilized to answer the four research questions, which are:

(1). How large are the overall ethnic disparities in academic outcomes at the different stages of university study?

(2). How much of these disparities can be explained by a wide range of relevant factors?

(3). Which factors are the most important determinants in explaining these ethnic disparities?

1.5. Findings of current thesis

In our analysis, students are ethnically identified by using both self-reported prioritised ethnicity and all reported ethnicities. It is found that there is very little difference in the regression results when using these alternative ethnic identifications. By comparing the means of the academic outcomes of interest, we find that Māori students have rates in university participation, first and second-course completion, and degree completion that are 22.30, 13.14, 12.43 and 19.15 percentage points, respectively, lower than that of their European peers. Our decomposition results show that eliminating differences in measurable factors between Māori and European could reduce these ethnic gaps by over 85% in university participation, no more than 50% in individual course completions and 25% in degree completion. The most important factors explaining the European-Māori gap in university participation are highest school qualification, school deciles and parents' highest level of education. The most important factors explaining the gaps in course completions are school deciles, parental education attainment, part-time enrolment and school qualifications. The most important factors explaining the gap in the qualification completion are school deciles, prior activity, other school factors and school qualifications.

The rates of university enrolment, first- and second-year course completion, and degree completion for Pasifika students are 13.48, 26.07, 24.85 and 27.25 percentage points, respectively below their European counterparts. If Pasifika had the same observed characteristics as that for European, this could eliminate approximately the entirety of this ethnic gap in university enrolment, about one third in course completions and around one quarter in degree completion. The primary factors contributing to this ethnic gap in university enrolment are school qualifications, school deciles and parental education attainment. The primary factors contributing to the gaps in course completions are school deciles, part-time enrolment, parental education attainment and school qualifications. The primary factors contributing to the gap in degree completion are school deciles, the highest school qualification, other school factors and prior activity.

The European-Asian gap in university attendance, in first- and second-year course completion rates, and in degree completion respectively are –19.33, 2.01, 2.21 and 1.77 percentage points. If Asian and European had the same observable factors, we could explain about 65, more than 60, less than 35 and more than 80% of this ethnic difference in university attendance, first-year course completion, second-year course completion and degree completion, respectively. The main contributors to the gap between European and Asian in university attendance are the highest secondary school qualification, other school factors and

parental education. The main contributors to the gap in first-year course completion are New Zealand citizenship, other school factors, gender/school deciles and enrolment status. The main contributors to the gap in second-year course completion include New Zealand citizenship status/part-time enrolments, school deciles, gender and other parental/school factors. The main contributors to the gap in degree completion are New Zealand citizenship status, other parental factors, female representation/school deciles and university distribution.

1.6. Research implications

This research has some limitations, including a lack of information on course letter grades and incomplete data for some variables, while employment hours are not provided for a large proportion of students. Despite the limitations, we hope that our findings will be helpful for various communities to better understand the determinants associated with university outcomes in particular, understanding what influences university performance is important to three communities. First, students who may rely on the information in making decisions about whether to study at university. Second, university administrators who attempt to design appropriate programmes to support Māori and Pasifika students. Third, government policy makers who develop targeted interventions and policies aimed at reducing overall academic gaps across the ethnic groups. Furthermore, our findings will provide a unique perspective on the determinants of university outcomes that are only available with detailed, linked administrative data that might be of interest to wider international academic and policy communities.

The rest of thesis is organized as follows. Chapter 2 provides a review of selected relevant New Zealand and international literature. Chapter 3 discusses data and methodologies used in this thesis along with providing descriptive statistics for the variables used in the participation analysis. In Chapter 4, we report our regression results. Finally, discussion and conclusion are provided in Chapter 5.

Chapter 2. Literature review

2.1. Introduction

Previous literature has commonly shown that students from different sociodemographic subgroups who are normally defined by ethnicity, gender and income, academically perform differently at university across the world. This chapter will focus on discussing the differences in terms of ethnicity. Studies on ethnic differences in university success have been conducted in countries such as the US between African American, Hispanic or Asian students relative to white students, and in the UK between Black, Caribbean or Asian students relative to white students, in the Netherlands between Surinamese, Moroccan, Westerner or Non-westerner students relative to Dutch students, and in New Zealand between Māori, Pasifika or Asian students relative to European students. These studies are all similar in terms of having concerns over ethnic inequities in education outcomes in their contexts. They have come up with diverse findings due to the fact that ethnic minorities are not the same in a variety of different countries. Therefore, this chapter reviews both international and New Zealand empirical literature on this particular topic.

The remainder of this chapter is comprised of eight sections. Section 2.2 describes Becker's human capital theory of the demand for education, which is the theoretical framework of our analysis. Causality will be briefly discussed in Section 2.3. Section 2.4 discusses factors influencing university success. Sections 2.5 to 2.7 review relevant overseas and domestic observational studies on ethnic differences in university outcomes including participation, first-year and during-study performance, and qualification completion in terms of the data used, the research method(s), the outcome variable(s), and key explanatory variables and main findings. Section 2.8 then surveys how working while studying affects university achievement. Meanwhile, a comparison and contrast of the studies are provided in the in-text summary tables (Tables A, B and C). Finally, Section 2.9 provides the motivation for this current thesis.

2.2. Human capital theory in relation to the demand for education

In economics, the decision to invest in education can be viewed from the perceptive of the human capital theory of Gary Becker. The origin of Becker's human capital theory can be traced back to the 1950s. As opposed to physical capital such as real estate, machinery or equipment, human capital is an intangible asset such as education and health that people possess which yields benefits over much of their lifetime (Becker, 1993, pp.15). From Becker's view, taking part in education is essentially an investment in human capital that is similar to a

business investment in equipment. It is costly now, but provides a stream of benefits in the future.

If education is treated as an economic product, demand for education is determined by whether the expected returns of education are greater than the costs of education. Returns of education are, for example, higher earnings, more job options and a happier life. The cost of education includes both direct and indirect (opportunity) costs. Examples of direct costs are tuition fees, accommodation and travel expenses; while time and foregone income from delayed entry into labour market are examples of opportunity costs.

For the purpose of simplicity, an assumption needs to be made, which is that the direct cost of education is paid at the zero period, denoted by C. If an individual chooses not to invest in education, let their current earnings to be denoted by W_t^0, $t = 0,1,2, \ldots$; If the person chooses to invest in education, let their expected (future) earnings to be denoted by W_t^1, $t = 0,1,2, \ldots$; Let i to be the market interest rate. According to human capital theory, a rational individual will undertake education when

$$\sum_{t=0}^{\infty} \frac{W_t^1}{(1+i)^t} - \sum_{t=0}^{\infty} \frac{W_t^0}{(1+i)^t} \geq C \quad (1)$$

If the return from investment in education is assumed to be constant per period, then the return is denoted by $r = W_t^1 - W_t^0$ for $t = 1,2, \ldots$ and assuming that in period zero the person's income is zero if choosing education and is W_t^0 if not choosing education, so that the person will invest in education if.

$$C + W_t^0 < \frac{r}{i} \quad (2)$$

Equations adapted from McCall, Smith & Wunsch, 2016, pp.489.

Equation (2) above provides important predictions of the demand for education. First, an individual is more likely to invest in education if the direct costs (C) are lower. For example, free or more affordable education could encourage many more students from low income and ethnic minority backgrounds to go to university. Second, a person will choose to invest in education if they face the lower opportunity cost (W_t^0) of doing that. In the New Zealand context, the opportunity cost of attending university instead of working (e.g., forgone income)

is relatively lower for Māori and Pasifika populations, as they on average earn a lower income. This, however, does not explain the lower Māori and Pasifika enrolments in university.

Third and, more importantly, an individual's education decision is subject to their expectations of returns (r) to the investment. Investing in education is associated with risk and the risk is linked to uncertainty. People choose a level of education to maximise the present value of their lifetime returns (or earnings). However, there is usually uncertainty about the returns on the investment so people have to form expectations. Thus, expectations on the returns are crucial when deciding whether or not to choose an additional level of education. Ethnic minority groups, for example, Māori and Pasifika populations in New Zealand, tend to believe education is less profitable (expect lower r) because they are less future oriented (have a higher discount rate), which results in the fact that these groups are less likely to participate (invest) in university education.

Fourth, an individual will have a greater incentive to invest in education if the (borrowing) interest rate (i) is lower. Accessing capital markets is undeniably difficult for low socioeconomic and ethnic minority groups as lenders are more unsure about their capacity to pay back their loans on time and exhibit a lack of trust in them. Thus, there is a lower chance for these groups to borrow an amount to invest in education at a lower interest rate. Also, for ethnic minority groups, internal financing seems to be less unlikely given that they often come from lower income and less wealthy families. Because of these capital and financial impediments, the New Zealand government set up the Student Loan programmes so domestic students could apply for interest free loans to pay the related costs of studying. Such financial support, however, does not eliminate all the possible ethnic differences in university participation. This could be because immediate and sudden drops in household incomes may lead younger people to seek out employment to offset such income losses and thus discourage university enrolments. Apart from capital market imperfections, young people usually having an inadequate knowledge about the importance of such long-term investment opportunities, thus pre-empting an investment in education.

2.3. Causality

In explaining the demand for education, human capital theory assumes that education raises an individual's income and productivity by providing knowledge and skills. Differently, job-market signalling theory argues that educational qualifications can be a signal to the employer, revealing a level of innate ability that someone may possess and is correlated with, but that does not necessarily cause higher wages, because people participating in higher levels of education are not a random group of individuals; they are high-ability people who are inherently more likely to attend university and be productive (Spence, 1973). Causality and correlation are the key themes the two theories focus on. Spence showed that, under certain conditions, innately high-ability individuals would have an incentive to invest in education while low-ability persons would not have a similar incentive. This would be true even if education itself had no impact on eventual productivity. Both human capital and signalling theory predict that earnings will increase with education, but for very different reasons.

Causality should not be mixed up with correlation. Causality means that a statistical relationship exists between two occurrences (Russell, 1945). Correlation is a measure of covariance that describes the size and direction of a mutual relationship/association between two or more variables (Parker, 1901). Correlation can be positive, negative or non-existent. Positive correlation between variables A and B simply means B generally increases in value as A increases. For example, more years in education is associated with a higher return in wages. However, it does not necessarily imply that more years in education causes higher wage rates. In other words, the higher wages may not directly be caused by spending more time in education. That is because there could be a mediator variable, for example intelligence or excellent problem-solving skills that directly or indirectly affect both qualification attainment and earnings. For instance, an innately intelligent student is more likely to attend additional years of education (perhaps due to lower costs) and earn more than others. Under signalling theory, this is the only way that individuals can demonstrate that they are inherently more productive.

Distinguishing causality from correlation is important. Most of the empirical results based on the observational studies will be discussed below but only suggest a correlation rather than causation between the determinants and university academic outcomes of interest, given not all the relevant factors can be controlled for, and students are not randomly selected in these studies. In other words, the analysis of the discussed studies and this thesis just demonstrate the partial correlations between the study variables and the university outcomes, which may

not identify the true impacts. However, uncovering the causal effects of factors on outcomes is not the focus of the current thesis.

2.4. Factors influence university academic success

There are a variety of factors that influence academic performance in universities. First of all, according to human capital theory, the demand for education is determined by opportunity cost, expectations on returns, and capital market imperfections (discussed above). Moreover, other common factors affecting academic outcomes in universities have been discussed in past studies that cover other demographic characteristics, pre-university academic achievement, and family-related backgrounds. Other demographic characteristics include, for example, gender (e.g., Rodgers, 2013; Scott, 2005a), age (e.g., Keith et al., 2006; Mutuku & Kiilu, 2016) and socioeconomic status (Boyd, Chalmers & Kumekawa, 2001). Pre-university academic performance are usually proxied by the GPA received in high school (Ting & Bryant, 2001). Family-related backgrounds include such things as parental educational attainment (e.g., Ahlburg, McCall & Na, 2002; Cingano & Cipollone, 2007) and parents' occupations (Johnes & McNabb, 2004).

University factors include first-year university performance (Olani, 2009), fields of study (e.g., Scott, 2005b; Urban et al., 1999), enrolment nature (e.g., Richardson & Woodley, 2003; Smith & Naylor, 2001), and class attendance (Credé, Roch & Kieszczynka, 2010), which are all thought to be important factors in determining later-stage university success. University achievement is sometimes affected by the institution (both at school and university level) and other factors. Institutional factors involve the type of school attended (Hoare & Johnston, 2011; Smith & Naylor, 2005), the classroom environment (Geisinger & Raman, 2013), teaching quality (e.g., Darling-Hammond & Synder, 2001; Heinesen, 2010), student support services (Benseman et al., 2006) and the teacher/student relationship (Osaikhiuwu, 2014). Academic self-efficacy (Richardson, Abraham & Bond, 2012), and the reasons for attending university (e.g., Guiffrida et al., 2013; Kennett, Reed & Stuart, 2013) are also relevant. Labour market conditions (e.g., Cappellari & Lucifora, 2009; Di Pietro, 2004) is an example of other factors.

In addition to the aforementioned factors, academic performance in a university also varies by ethnicity. Students from ethnic minority groups are on average lagging way behind the students from the dominant ethnic group in university education across countries such as the USA, the UK, Australia, the Netherlands and New Zealand. In terms of how students' ethnicity makes a difference to their university achievement, Powdthavee & Vignoles (2009) and Connor et al. (2004) have commented that ethnicity itself has a direct influence on university

academic outcomes, whereas Modood (2006) and Richardson (2012) have argued that ethnicity affects university success through other related factors such as socioeconomic status. Other similar studies highlight that a lower level of performance for ethnic minority students may result if they tend to have insufficient information on the steps that are needed to prepare for university study (Tym et al., 2004), or they may face relatively more financial pressures during enrolment, which consequently distracts them from study (Clotfelter, Ladd & Vigdor, 2015).

In New Zealand, the educational underachievement of students from ethnic minority backgrounds (e.g., Māori and Pasifika) was found to be closely correlated to their family socioeconomic disadvantages (e.g., Marie et al., 2008; Fergusson and Woodward, 2000). Also, as discussed above, Māori and Pasifika are less likely to undertake university study because of their perception of lower returns to investment in education (human capital theory). Ethnic differences in university participation, first-year and during-study performance, and qualification completion will be discussed next.

2.5. Ethnic differences in university participation

The idea that university enrolment varying by ethnicity is common in many countries around the world. For example, in the USA, although the total college enrolment rates increased from 63% in 1993 to 66% in 2019, the enrolment rate of white school leavers (67%) were higher than for Hispanic (63%) and black (51%) school graduates in 2019 (US Department of Labour, 2020). As mentioned in the introduction, the official data shows that in New Zealand participation rates in tertiary education by Māori and Pasifika grew at more than twice the rate of other ethnicities from 1999 and 2007, but Māori and Pasifika leaners still remain significantly underrepresented at the degree level and above relative to non-Māori and non-Pasifika students.

2.5.1. Findings from selected international studies

In the US, university attendance is closely linked with ethnicity. Many past studies focused on whether the lower enrolments of ethnic minority groups can be explained by differences in other determinants of university participation. Clotfelter, Ladd & Vigdor (2015) examined the disparities in four university outcomes including participation between African American and non-Hispanic white students at the University of North Carolina (UNC) System. The authors used administrative school and university data on two cohorts of students who completed 8th grade tests in the spring of 1999 and 2004 at the state-owned schools in North Carolina.

University participation was defined as enrolling in a UNC university immediately after high school graduation. The explanatory variables that were used included gender, ethnicity, parental education (indicated socioeconomic status) and 8th grade test scores.

Academic disadvantage was consistently observed for African American students. For example, the results of basic OLS regression showed that African American students in the 1999 cohort was 1.7 percentage points less likely to enrol in a UNC university than white students if only controlled for gender. The ethnic gap was tripled to 4.6 percentage points for the 2004 cohort. However, the gaps reversed, meaning the black-white differences in university enrolment can be fully explained, once parent education and 8th grade test scores were also statistically controlled for. It should be pointed out that the data used did not include students who studied in private schools and those who moved into North Carolina as high school students. The data also ignored the students enrolling in universities outside the UNC system.

In the 1960s, one of the key findings of the report on Equality of Educational Opportunity was that although a smaller number of African Americans than white students reported "wanting to go no further than high school in each region of the country", they enrolled in college at a lower rate than white students (Coleman et al., 1966). The higher college aspirations of African American students have continued until today, but there are still lower proportions of black students in colleges nowadays. To address this question, Schneider and Saw (2016) used survey data from the High School Longitudinal Study of 2009 (HSLS:09) to exam the enrolment in a four-year bachelor's programme, a two-year programme, or a certificate or other training programme following high school graduation. The sample contained more than 15,000 ninth and eleventh-grade students. Multinomial logit regressions were conducted, taking into account individual, family and school factors.

The results showed that the proportion of black students enrolling in a two-year college programme right after high school completion has dramatically increased so it achieved the rate as white students. The situation, however, was different for four-year university enrolments. In particular, black students were about 15 percentage points less likely than white students to attend a four-year college. The difference reduced to 3.6 percentage points when personal demographic and family characteristics were added to the regression, suggesting that 76% of the black-white gap in four-year college participation can be explained by holding the measured factors constant. In addition, the authors found that school interventions to increase students' knowledge about tertiary education, such as meeting with college admission consultants and taking college classes while in school, have had a small impact on

college enrolment behaviour. Instead, advanced academic preparation, including completing a set of college preparation courses or earning Advanced Placement (AP) and International Baccalaureate (IB) credits was a powerful predictor of college participation. The results need to be intercepted with cautions due to the fact that the outcome measure used was restricted to immediate university enrolment following finishing high school. Also, the survey data lacked family income information, which can have a strong impact on university enrolment. Lastly, the analyses could have been extended to cover university persistence and graduation.

In the other countries, such as the UK, students from ethnic minority groups do not necessarily participate in university at lower rates if the sample is restricted to only male students. Stockfelt (2018) intended to explain ethnic variation in higher education participation amongst only male students in the UK. The study was based on about 3,000 male students from 596 schools. These students were aged 13 or 14 in 2004 and were respondents in the Next Step survey. The Next Step was administered by the Centre for Longitudinal Studies (CLS) in the UK. The survey used random sampling. Probit regression models were used to assess the contributions of a range of factors including attitude to school, prior attainment, parental expectation and social class in explaining ethnic differences in tertiary education participation. The enrolment outcome in the study was defined as undertaking a course in a UK tertiary education institution at age 18 or 19 by March 2010 as the person's primary activity.

The author found that white British boys have a significantly lower propensity to move into higher education than African, Indian and Pakistani/Bangladeshi male students, except for male students from the Caribbean. After controlling for prior attainment, social class, parental expectation and attitude to school, Africans were no longer significantly more likely than white British males to participate in higher education, while the figures for Indian-White-British and Pakistani/Bangladeshi-White-British respectively reduced by 75 and 15%. The study sample was, however, made up of only male students. Also, the variable attitude to school was difficult to measure because children's attitude changes with age.

Another study provided more insight into ethnic disparities in the progression onto university study for school students (both male and female) in the UK. Jackson (2012) aimed to examine ethnic inequalities in educational attainment in England and Wales. The author focused on two important educational transitions; the move from compulsory to post-compulsory education at age 16, and the move from school to a university degree at age 18. The data were drawn from the Youth Cohort Studies survey. To ensure a large enough number of ethnic minority students, the author combined three of the survey datasets containing three cohorts of students (more than 45,000) who were aged 16 in 1998, 2000 and 2002. Ethnicity, parents'

occupations (proxy of social class) and academic performance at ages 16 and 18 were used as the explanatory variables. The sample was classified into 11 ethnic groups. Logit regression was utilised to assess the relative importance of the variables in explaining the ethnic inequalities.

The findings suggested that all ethnic minority groups, except the Black Caribbean and Other Black groups had a higher transition rate to degree-level study compared to white students. For example, Black Caribbean (OR=0.91) and Other Black (OR=0.94) were 9% and 6%, respectively less likely than white students to transit to degree-level study, whereas Chinese (OR=3.11) were over three times more likely than whites to make the transition to a degree qualification. Controlling for social class background recued the white-Black Caribbean gap by 67%, had no effect on the gap between whites and Other Black, but increased the white-Chinese gap by 16%.

Although ethnic minority students may not have lower university participation rates than whites, they are less likely to receive an offer from prestigious universities in the UK. Boliver (2013) explored the extent of access to the UK's more prestigious universities, Russell Group universities in particular. The access included applying to and receiving an offer of admission from a Russell Group university. The study data drew on individuals who applied to university through the Universities and Colleges Admissions Service (UCSA) in each even year between 1996 to 2006. The UCSA is responsible for processing nearly all applications for full-time study at the higher education level in the UK. The study sample included about 50,000 students but, did not cover part-time applicants. This could potentially generate a bias in the results given about one third of UK undergraduates are enrolled as part-time students.

The independent variables of interest were ethnicity, social class origin, school background and prior academic attainment (grades at A-levels or equivalent qualifications). It is important to note that the grades of the school qualifications were predicted other than actual due to the application data could not be linked to school data, which may not be completely reliable. Black Caribbean and Black African, and Pakistani and Bangladeshi were combined into two groups due to small subsample size.

Logit regression models were used. The results of the basic model (with controlling for social class and school background) showed that Black Caribbean/Black African (OR=0.24), Pakistani/Bangladeshi (OR=0.33) and Indian (OR=0.59) were statistically significantly 76%, 67% and 41% respectively, less likely than comparable whites to receive an offer from a Russell Group university. The results of the extended model indicated that further controlling

for prior attainment could reduce the gaps of Caribbean/Black African (OR=0.53), Pakistani/Bangladeshi (OR=0.57), relative to whites by 38% and 36%, respectively, and resulted in the White-Indian gap being statistically insignificant.

2.5.2. Findings from selected New Zealand studies

The gaps between ethnicities in university enrolment have drawn many researchers' attention in New Zealand. Shulruf, Hattie & Tumen (2008) intended to identify the factors that influence students' enrolment and success in university education. The data used in the study were received from the New Zealand Ministry of Education and a university in Auckland. The data concerning school factors were retrieved from the Ministry of Education website, while the data on students were provided by the planning office of the university. The study sample is comprised of about 10,000 students who studied in secondary schools within the Auckland region in 2004. Forty percent of them were European, 32% Asian, 14% Pasifika, 6% Māori and 7% others. The analysis was divided into stages. The first stage was about students' pathways from secondary schools to university. The second stage was about students' first-year GPA performance at the university, which will not be discussed in this section.

For the first stage analysis, the outcomes were comprised of three binary variables, which were (i) the acquisition of UE (university entrance qualification); (ii) enrolment at the university; and (iii) becoming an active student (i.e., having achievement records on the database of the university). Logit regressions were used to identity the likelihood of students belonging to each of the groups. School administrative determinants (i.e., school authority), socio-geographic determinants (i.e., school decile), demographic determinants (i.e., gender and ethnicity) and peer determinants (i.e., school achievements) were included in the regression analyses.

The authors found that Māori (OR=0.25) and Pasifika (OR=0.16) students were statistically significantly less likely to gain UE than their European peers. Adjusting for other predictor variables reduced the lower achievement of Māori (OR=0.31) and Pasifika (OR=0.23) by about 8%. Māori (OR=0.29) and Pasifika (OR=0.51) were also less likely than European to enrol in university unconditional on acquisition of the UE. The European-Māori gap closed by 38% and the European-Pasifika gap became statistically insignificant after controlling for all other independent variables. It is important to note that the study did not identify to what extent ethnicity affects students' educational outcomes in terms of course and programme completion and it measured the effects of ethnicity on the outcomes at only one university. This thesis attempts to look at a series of academic outcomes, with a focus on the students at all the universities in the country.

A 2012 study by Strathdee & Engler attempted to assess propensity to progress to bachelor-level study for a sample of more than 45,000 students. Data for the study were drawn from two sources. Data on the students' school academic performance were provided by the New Zealand Qualifications Authority (NZQA). The tertiary enrolment data were supplied by the New Zealand Ministry of Education. These data were linked via the national student number.[2] The authors focused on the students who passed the NCEA Level 3 exams and gained University Entrance (UE). These students were split into three cohorts, each cohort comprising students who were 17 years old in the years 2004 to 2006. The authors had interest in which of these students attended bachelor-level study. Attending bachelor-level study included those progress directly and indirectly (took one or two years' break after high school) on to university. Logit regression was used to model the likelihood of achieving the academic outcome with including ethnicity, socioeconomic status, gender and school performance as the explanatory variables.

Their regression results indicated that when controlling for school achievement, there were few differences in the likelihood of enrolling in bachelor-level study between ethnic groups for the students from higher deciles (9 and 10) schools. For the students from lower-decile (1 and 2) schools, Māori students were found to have a statistically significantly lower likelihood of progressing on to bachelor-level study than European students. The authors claimed that the best explanation of the likelihood of progressing from school to bachelor study is provided by the pair-wise interactions of ethnicity, school deciles and school achievement. It should be borne in mind that these findings are based on students who achieved an NCEA Level 3 qualification and obtained UE. In other words, those with lower NCEA qualifications were arbitrarily excluded from the analysis. Such selection bias does not occur in this thesis as we consider all types of high school qualifications.

A more recent study by Meehan, Pacheco & Pushon (2019) examined the underachievement in bachelor's degree participation by Māori and Pasifika relative to European by using individual-level linkable administrative data from a large database (the IDI), operated by Statistics New Zealand. Their data are similar to the data that is used in the current thesis, but our data contain more recent information. Detailed information about the IDI data will be discussed in the beginning of next chapter of this thesis. The participation outcome was defined by the authors as enrolling in a bachelor's qualification at a tertiary education provider

[2] For more information about national student numbers, see https://www.education.govt.nz/school/managing-and-supporting-students/national-student-number-nsn-for-schools/#sh-national%20student%20number%20.

before age 20. This variable is measured slightly differently in our study. In particular, we do not consider students from non-university organizations because bachelor's qualifications are mainly offered by universities and the ethnic composition of students can be different between the two types of organizations. We also allow the age of participation to be extended to 21. The sample of interest for the study is comprised of about 138,000 individuals born between 1 July, 1990 and 30 June, 1994. They were divided into four July-to-June year cohorts. The sample did not include foreign students and those who primarily lived overseas when making decisions to undertake a bachelor's degree in New Zealand before the age of 20.

Using the data, the authors could control for factors in their analysis, including individual characteristics, socioeconomic status, school characteristics, school performance and engagement, and parental education. Ethnicity is one of the most important individual factors, but there are challenges when dealing with ethnicity. First, ethnicity is self-reported, and people can report multiple ethnicities. Second, ethnicity can vary overtime and can be different in different data sources. The authors handled these issues by applying ethnic prioritisation rules relative to the school data. The prioritisation order is Māori, Pasifika, Asian, Other and European. The same ethnicity classification method is used in the current thesis, but we also consider multiple ethnicities, while Meehan, Pacheco and Pushon did not do so in their analysis. After using prioritised classification, the population of the study could be broken down into 21% Māori, 8% Pasifika, 7% Asian, 2% Other and 62% European. Academic performance at school was indicated by NCEA Level 1 qualification, which is the lowest level of NCEA qualification, usually gain in Year 11 (age 15 to 16). According to the authors, there were two reasons for using NCEA Level 1 as the indicator of school performance. First, it was available for the majority of their sample. Second, using the earliest NCEA qualification could minimize the endogeneity issue, rather than using the later qualifications. This is because students who planned to go to university were more likely to complete the higher levels of the qualification. Differently, we take into account all types of school qualifications awarded in New Zealand schools in our analysis.

In the study, socioeconomic status was proxied by the deprivation index derived from the 2006 census. The deprivation index measures the level of deprivation of people in the area they are living in. It is shown as deciles ranging from 1 to 10. Decile 1 indicates areas with the lowest deprivation scores, while decile 10 indicates areas with the highest deprivation scores. These scores can change from census to census. The deprivation index is a substitute variable for school deciles because both are socioeconomic indicators. In this thesis, we choose to use school deciles because information on deprivation scores is only available for a small proportion of our sample. Parents' education was the only family factor that was considered

by the authors. However, other more parental-related factors, including family composition, household income and parents' criminal and benefit histories are taken into account in our analysis.

The research questions the authors attempted to address were:

(1). How do degree-level participation rates vary by ethnicity?

(2). what is the relative importance of socioeconomic status, school performance, and parents' educational attainment in explaining the ethnic differences?

Separate Probit Maximum Likelihood estimations were run for Māori, Pasifika, Asian and European. Fairlie decompositions were then used to explain the ethnic gaps. It was found that the actual participation rate was 34.75% for European, only 14.91% for Māori, 23.31% for Pasifika and 71.73% for Asian, indicating that European-Māori, European-Pasifika and European-Asian gaps were 19.84, 11.45 and negative 36.97 percentage points, respectively.

The decomposition analysis results suggest that about 87% of the Māori-European gap in participation could be closed if the ethnic groups had the same individual characteristics, socioeconomic status, school characteristics, school performance and engagement and parental education, while the Pasifika-European gap could be fully explained by the differences in the observable factors. The primary drivers of these ethnic gaps are school performance (NCEA Level 1), socioeconomic status and parental education. Among the three factors, school played the greatest role in explaining the ethnic gaps in the participation outcome.

Table 2.1. Summary of selected international studies on ethnic differences in university academic performance

Author(s), date	Country	Data	Method	Outcome variable(s)	Explanatory variables	Relevant findings	Comments
Clotfelter, Ladd & Vigdor (2015)	US	Administrative school and university data.	OLS	(1) Participation: enrolling in a UNC institution immediately after high school graduation, (2) GPA: GPA of 3.0 after 45 credit hours, (3) Choosing a STEM major: majoring in a STEM subject after 60 credit hours, (4) Graduation: graduating within four years.	Ethnicity, gender, parental education (measure of socioeconomic status) and high school test scores.	African American students were less likely than white students to enrol in and graduate form a UNC university. The disparities can be fully explained by controlling for gender, high school test scores and parental education.	The study ignores private school students and students enrolling outside UNC system. More independent variables can be used.
Fletcher & Tienda (2009)	US	Longitudinal administrative data from the University of Texas at Austin.	OLS. Year and high school fixed effects used control for time-invariant high school factors.	(1) GPA: first semester GPA, (2) Two-year: Enrolled in the university after two years.	Ethnicity, school tests scores, class rank, gender, maternal educational attainment and the number of high school classmates who enrolled at the university at the same time.	After adjusting for other factors, GPAs for blacks and Hispanics were 0.038 and .0023 points, respectively, lower than that of whites. Adding school fixed effects could fully eliminate the ethnic gaps in the GPA performance.	The study was based on the students at one public university, and not able to exam graduation rates due to data unavailability.

Fletcher & Tienda (2010)	US	Ten years of administrative data at four Texas public universities.	OLS with year and school fixed effects.	(1) Average GPA for 1st semester, (2) Average GPA for 6th semester, (3) Four-year graduation.	Ethnicity, class rank, school academic performance.	Black and Hispanic students had lower rates of graduation than white students at UT, TAMU and TECH. Controlling for school quality, class rank and school test scores reduced the disparities in graduation rates by about 40 to 50%.	Although the study looked at both most selective and less selective universities in Texas it did not consider students at private and community universities/colleges.
Jackson (2012)	UK	Data were drawn from the Youth Cohort Studies survey.	The method descried in Erikson *et al*. (2005) and applied in Jackson *et al*. (2007).	(1) Transition from compulsory to post-compulsory education at age 16. (2) Transition from school to university degree at age 18.	Ethnicity, parental occupations, academic performance at ages 16 and 18.	All ethnic groups other than the Black Caribbean and Other Black groups had a greater transition rate to university than the whites. For example, Black Caribbean and Other Black were respectively 9 and 6%, and Chinese were more than three times to do so than whites. The white-Black Caribbean gap	Small number of control variables.

						decreased by 67%, the white-Other Black gap did not change, and the white-Chinese gaps increased by 16%, once controlling for social class background.	
Light & Strayer (2002)	US	Data were drawn from the 1979 National Longitudinal Survey of Youth (NLSY79).	Binomial Probit.	(1) College attendance: first attended a four-year college, (2) College completion: received a degree from the college originally enrolled in.	Ethnicity, age, whether received financial aid, whether lived at home, tuition fee and the area unemployment rate. Also controlling for college quality quartile.	Minority students overall were less likely than their white peers to complete college. Controlling for other observed characteristics closed the ethnic gap by about 18% for student enrolled in QQ1 colleges with lowest quartile school test scores, and reversed the ethnic gap for those enrolled in QQ4 colleges with highest quartile school test scores.	The study was based on survey data. The outcome graduation excluded those who received a degree from a college other than the one previously enrolled in.
Lorah & Ndum (2013)	US	First-year college grade	Logit regressions.	First-year success: obtaining a B or higher	Ethnicity, gender, family income and self-	For the 1998 cohort African American	The analysis relied on self-

		data for ACT-tested students. They were self-reported data.		grade (or a C or higher grade) in the first-year credit-bearing college courses including English Composition, College Algebra, Social Science and Biology.	reported course grades.	students were 23, 13 and 37% less likely to achieving a B or better grade in English Composition, Social Science and Biology, respectively, than Whites. Hispanics performed as better as Whites except they were 25% less likely than Whites to gain a c or better grade in Biology.	reported student data. The study sample contained relatively fewer African American and Hispanic students.
Massey (2006)	US	The National Longitudinal Survey of Freshmen (NLSF) data.	OLS	First-year GPA: GPA earned during first three terms.	Ethnicity, school/neighbourhood segregation, academic preparation, social/psychological preparation, exposure to stress, demographic background and parental socioeconomic status.	GPAs for blacks and Latinos were more than a third (0.37) of a point and almost a quarter (0.24), respectively, lower those of whites. Controlling for family background, socioeconomic status and minority concentration reduced the black-white differential by 30% and the Latino-	The study sample was hardly representative of all students in the US, as the respondents were coming from the country's most prestigious colleges/universities.

						white differential by 33%.	
Rodgers (2013)	UK	Data from a new UK university.	Probit regression.	Degree completion: completing a three-year degree within three year or a four-year degree within four years.	Academic ability, subject of study, personal characteristic (ethnicity, gender, disability, whether was a mature student and whether lived in parents' home) and socioeconomic backgrounds, and a series of interactions.	Caribbean, Pakistani and Bangladeshi students were 40.4, 53.5 and 17.9%, respectively, less likely to complete a degree than Whites within six years. If other factors were controlled for, the White-Caribbean and White-Pakistani gaps in degree completion closed by about 10 and 31%, respectively, and the gap between White and Bangladeshi could be fully explained.	The student sample was drawn from a new UK university where tend to recruit relatively more ethnic minority students.
Schneider & Saw (2016)	US	Survey data from the High School Longitudinal Study of 2009 (HSLS: 2009).	Multinomial logit regression	Participation: enrolling in a four-year bachelor's programme, in a two-year program, or in a certificate or other training programme	Individual, family and school factors.	Black students were as likely as whites to attend tow-year colleges. However, enrolment rate in four-year colleges was about 15 percentage points	The analyses did not include outcomes measures on university persistence and graduation.

				following high school graduation.		lower for blacks than whites. The difference reduced by 76% to 3.6 percentage points when added personal demographic and family characteristics.	The data lack measures of family income.
Stegers-Jager et al. (2012)	Netherlands	Data on individual backgrounds from Statistics Netherlands. Student data were provided by the university. Data on socio-demographic factors collected through survey.	Logit regressions.	(1) Year 1 course completion: passed all Year 1 exams within one year, (2) Pre-clinical course completion: passed all pre-clinical exams within four years, (3) Good clinical performance: achieved at least three of grades of ≥8.0 from internal medicine, surgery, paediatrics, psychiatry and neurology.	Socio-demographic characteristics (first-generation immigrant, first language, first-generation university student, parents as medical doctor and urban background), gender, ethnicity and pre-university GPA.	Surinamese/Antillean and Asian students respectively were 67 and 45% less likely to complete first-year courses than Dutch students. After controlling for age, gender school GPA and socio-demographic variables, the Surinamese/Antillean-Dutch and Asian-Dutch disparities reduced by 10 and 38%, respectively.	The analysis was based on the students at one medical university. Data on socio-demographic factors were available for a restricted number of students.
Stockfelt (2018)	UK	Data from the Next Step survey. [a]	Probit regressions	Participation: enrolling on any course in a UK tertiary institution at age 18/19 in March	Ethnicity, attitude to school, school attainment at age 16,	White-British male students were less likely move into higher education.	The lower higher education participation

				2010 and where this is the main activity.	parental expectation and social class.	Accounting for prior attainment, social class, parentela expectation and attitude to school explained the entire African-White-British gap, and 75 and 15% of Indian-White-British and Pakistani/Bangladeshi-White-British gaps, respectively.	could be caused by looking only male students. The variable attitude to school is difficult to be measured as it is time varying.
Boliver, V. (2013)	UK	The data supplied by the Universities and Colleges Admissions services (UCAS).	Logit regression.	(1) Application to a Russell Group university, (2) Receiving an offer of admission from a Russell Group university.	Ethnicity, social class origin, school background and prior academic attainment.	Black Caribbean/African, Pakistani/ Bangladeshi and Indian students were 76, 67 and 41%, respectively, less likely than their Whites peers to receive an offer from a Russell Group university. Controlling for prior attainment reduced the gaps of Black Caribbean/African, Pakistani/ Bangladeshi relative to Whites by 38 and 36%, respectively, and resulted in the gap between Whites	The study did not include part-time students. Prior attainment was predicted rather than actual.

						and Indian statistically insignificant.	
Zorlu (2013)	Netherlands	Administrative individual-level data.	Non-parametric and parametric survival analyses.	(1) Graduation: completing a degree within ten year since start, (2) Duration of study: duration of completing a degree in months.	Gender, age, ethnicity, preliminary training, subject of study, parents' economic position in 1999.	For both students enrolled in higher vocational education and university, ethnic minority groups were statistically significantly less likely to graduate within ten years compared to Dutch students. About 9-19% of the ethic gaps in graduation could be closed if parents' socioeconomic status were additionally controlled for.	The data did not contain attitudinal and ability variables that can affect university performance.

[a] The Next Step is a panel study organised by the Centre for Longitudinal Studies in the UK that was based on a survey of more than 15,000 children who were 13/14 years old in 2004.

2.6. Ethnic differences in the first-year university study

The first year is the crucial stage for university students, because first-year experiences play a key role in determining future success in university studies (e.g., Allen & Robbins, 2008; Brahm, Jenert & Wagner, 2017). At some universities, first-year students must gain a required number of credits in order to be eligible for enrolling in second-year studies (Bijsmans & Schakel, 2018). Substantial disparities in first-year university achievement have been found between students from different ethnicities.

2.6.1. Findings from selected international studies

An American study conducted by Massey in 2006 documented how academic preparation and achievement for higher education was affected by the degree of segregation experienced by minority students in schools and neighbourhoods while growing up. The data were drawn upon the National Longitudinal Survey of Freshmen (NLSF) that randomly sampled more than 4,500 students who entered 28 selective colleges and universities in the fall of 1999. The final sample used in the study was made up of slightly less than 4,000 students, including 959 whites, 998 Asian, 916 Latinos and 1,051 African Americans. The author wanted to address the issue of whether the experience of segregation had any implications for college achievement. College achievement was measured by the GPA received during the first three terms in college/universities. A series of explanatory variables included were ethnicity; school/neighbourhood segregation (the average proportion of ethnic minority people present in the schools and neighbourhoods in which the students grew up); academic preparation; social/psychological preparation; exposure to stress; and demographic background and parental socioeconomic status.

The OLS estimation was used to regress the GPA on the factors. The coefficients indicated that GPAs earned by blacks and Latinos were more than a third of a point (0.37) and nearly a quarter point (0.24) respectively, lower than those of whites. Once controlled for family background, socioeconomic status and minority concentration, the black-white difference reduced by 30% to 0.26 points and the Latino-white difference closed by 33% to 0.16 points. For all ethnicities, the GPA statistically significantly decrease by 0.13 points if shifting from absolute integration to absolute segregation. It, does however, need to be mentioned that the students involved in the NLSF were hardly representative of all the students in the US, as they were all students attending the country's most prestigious tertiary educational institutions.

Like Massey, Fletcher & Tienda (2009) also only focused on public universities. Fletcher and Tienda considered whether high school characteristics influenced college achievement. The study used longitudinal administrative data from the University of Texas in Austin that included more than 66,000 students who applied to the university from the early 1990s through to 2003. The sample included almost two-third whites, 4% blacks, 15% Hispanics and 17% Asian. The college achievement of interest was measured by first semester college GPA and two-year persistence (still enrolled at the university after two years). Ethnicity and the number of high school classmates who enrolled at the university at the same time along with school test scores, class rank and maternal educational attainment were controlled for in the analysis. Classical OLS estimation was used. Due to the fact that the data were pooled for over 10 years, year fixed effects was also used to deal with any time specific effects.

The results showed that the adjusted GPAs for blacks and Hispanics were 0.038 and 0.023 points respectively, below that of whites, while Asian's GPA was 0.035 higher (relative to whites). To control for high school characteristics that could directly affect the university GPA, school fixed effects was also added to the estimation. After controlling for time and school characteristics (number of high school classmates), both black-white, Hispanic-white and Asian-white gaps in GPAs were closed. These results together suggested that differences in high school factors was the main contributor to the university GPA differences between the ethnic groups. The reader should be aware that these results were based on the students attendance at a single public university. Also, due to data unavailability the authors were unable to examine graduation performance.

Ethnic gaps in the academic success of first-year college students were also examined by Lorah & Ndum (2013). The authors studied gaps in the first-year academic success of student subgroups that were defined by ethnicity, income and gender over time. The study was based upon the data of more than 330,000 students who were enrolled in 101 colleges from 1998 through to 2009. Success was defined as obtaining grades of B or higher or C or higher in four first-year credit-bearing courses. These were English Composition I, College Algebra, Social Science and Biology. The study aimed to address the following research questions:

(1). Do achievement gaps exist in first-year college courses by ethnicity, income and gender subgroups?

(2). Did the sizes of the gaps change over the period 1998 to 2009?

Logit regressions were used to model the success outcomes on ethnicity, high school grades, income and gender. It is important to note that the data on the predictor variables are self-reported, which means these variables may have been reported with error that could bias the analytic results discussed next. The authors found that African American students had lower success rates than white students in all courses, apart from College Algebra, after controlling for high school achievement and family income. For example, in the 1998 cohort, African American students were 23%, 13% and 37% less likely to achieve a B or better grade in English Composition, Social Sciences and Biology, respectively, than whites. The results regarding the gaps in obtaining a C or better grade in the courses between African American and white students were similar. The African American-white gaps in English Composition remained unchanged between 1998 and 2009 but the gaps in Social Science became wider over the period. There were no statistically significant achievement gaps between Hispanics and whites, except that Hispanics were 25% less likely to achieve a C or better grade in biology than whites (in 1998). The Hispanic-whites achievement gaps did not significantly change for that decade. The study, however, did not provide evidence of how much of the ethnic gaps can be explained after certain factors being controlled for. Another two limitations in the study were that the analytical sample contained relatively few African American and Hispanic students, and some relevant factors, such as parents' education were, not included.

Underperformance of students from ethnic minority groups is also found in Europe. Stegers-Jager et al. (2012) investigated the underachievement of ethnic minority students, compared to those from the ethnic majority, at a medical school in the Netherlands. The study was carried out at the Erasmus MC Medical school in Rotterdam, the Netherlands, which was considered to have a relatively higher number of ethnic minority students. The analytical sample included more than 2,000 students who enrolled in the school between 2002 and 2007. The data used in the study came from three sources. Data on ethnicities were made available from Statistics Netherlands. Student data on pre-university GPA, age and gender were provided by the university and other data on social background were collected through survey.

Three university outcomes of interest were (1) Year 1 course completion: passed all Year 1 exams within one year, (2) pre-clinical course completion: passed all pre-clinical exams within four years and (3) good clinical performance: achieved at least three of grades of $\geq$8.0 for internal medicine, surgery, paediatrics, psychiatry and neurology. The grades ranged from 5 to 10. The students were classified into six groups based on their ethnicity; Dutch (the ethnic majority), Surinamese/Antillean, Turkish/Moroccan/African, Asian, Western and Other. Logit

regression was used to estimate the probability of achieving the outcomes holding constant age, gender, and pre-university GPA, along with other socio-demographic factors, including first-generation immigrant, first-generation university student, first language, medical doctor as parent and urban background.

The regression results show that Surinamese/Antillean (OR=0.33) and Asian (OR=0.55) students were 67 and 45% respectively, less likely to complete a Year 1 course within one year, compared to Dutch students. Adjusting for other individual factors, pre-university GPA and socio-demographic characteristics could explain the 10% disparity between Surinamese/Antillean and Dutch, and 38% of the Dutch-Asian difference. It needs to be noted that the information on socio-demographic factors was available for a limited number of the students, which may bias the effect of these factors on the academic outcome. Another limitation is that the analysis was based on the students enrolled at a single medical university.

2.6.2. Findings from selected New Zealand studies

Juhong & Maloney (2006) provided evidence of the ethnic differences in academic performance within the New Zealand tertiary education. The study used a dataset from a cohort of students who first enrolled in either the Bachelor of Science (BSc) or Bachelor of Arts (BA) at a large New Zealand university in the academic year 2000. Relying on these data, the authors could track the academic outcomes of this group of students through the academic year 2003. These administrative data were provided by Starpath. Starpath was an educational research project in New Zealand that aimed to improve Māori and Pasifika success in education.

The dataset had information on a sample of more than 3,000 students and Māori (7%) and Pasifika (7.7%) students were almost equally represented. Asian and European or Other accounted for 30.9% and 54.3%, respectively, of the sample. Apart from ethnicity, information on gender, previous academic performance, age, school decile and school qualification were also available in the dataset. The authors, however, lacked information on the school factors and family backgrounds of the students. However, such variables are used in our analysis.

The authors wanted to consider two university outcomes. The first one was the GPA, that is the average GPA across all courses in which these students were enrolled at the university over the four-year period (2000 to 2003). The GPA ranged from zero for a failed grade of D+ or lower to a highest of nine for a grade of A+. The second outcome variable was dropout, which was defined as individuals who discontinued their studies in the initial programme by the academic year 2003 without completing a qualification at the university. However, the authors could not further identify those who left the university and may have enrolled in other institutions or may have changed to different degree programmes within the same university. Therefore, the analytical results on dropout should be interpreted with great caution. These two outcomes were also examined at different stages throughout the university studies, but these results will not be discussed in this section.

The analysis began by measuring the mean differences in GPA and dropout outcomes across the ethnic groups. The authors found that the first-year GPA for Māori and Pasifika students was more than 1 point and more than 2 points respectively, lower than the GPA of European students. Māori and Pasifika were also 7.4 and 9.1 percentage points, respectively, more likely than their European peers to discontinue their studies. Asian students had a slightly higher first-year GPA to Māori and the lowest dropout rate after first year among the ethnic groups.

The authors next wanted to address:

(1). What is the relationship between the ethnic minority groups' GPA and dropout behaviour?

(2). How much of the ethnic gaps in university outcomes can be explained by the differences in other observed factors?

Classical OLS and maximum likely Probit were used to answers these questions. The results of OLS regression showed that about 32% of the gaps in the first-year gap between Māori or Pasifika and European could be explained if controlling for other explanatory variables. According to the Probit regression results, controlling for other factors could reduce the higher dropout rates after the first year of Māori and Pasifika studies, relative to European, by 62 and 33% respectively. The ethnic gaps were reversed if further controlling for first-year GPA. These results suggest that GPA achievement was the most important factor in explaining the ethnic differences in the first-year academic outcomes. All results were based on students undertaking two bachelor programmes at one university. In contrast, the current thesis relies

on students who undertake different bachelor level qualifications across all the universities in this country.

Scott (2008) tested the relationship between secondary and tertiary achievement. The data used in the study came from two sources: the New Zealand Qualifications Authority (NZQA) and the New Zealand Ministry of Education. These sources of data were linked via the national student number. The study sample included more than 24,000 students who left school in 2004 and enrolled in a tertiary institution in 2005.

The tertiary outcomes included degree students' first-year course completion and the attrition rate. First-year course completion was defined as passing all the courses, while first-year attrition was defined as leaving tertiary education after the first-year of study without completing any qualification. A potential issue for the definition of first-year attrition was that it excluded students who took a year's break (which is very common in New Zealand) and who entered the market place the following year. Academic performance at school was measured by the highest school qualification (whether NCEA Level 3 and/or University Entrance). In addition to school achievement, more independent variables were used, including gender, age, ethnicity, school factors (school roll, school decile, etc.), tertiary field of study, full-time/part-time enrolment and the number of courses enrolled in. However, the study had a limited ability to capture the important effects of parental factors. In our analysis, we are able to take into account several family background factors, such as household income and parental education.

Logit regression was used to model the two outcomes. The analysis results showed that being Māori (OR=0.66) or Pasifika (OR=0.58) statistically significantly lowered the probability of passing all first-year courses by 34 and 42%, respectively, if other factors were kept constant. Māori students were also about 50% more likely to drop out of degree study after the first year. The author found that factors other than school performance, such as school deciles, enrolment status and the major subject enrolled in had relatively stronger predictive power on the academic outcomes for Māori learners.

Another study by Jia & Maloney (2015) aimed to estimate the determinants of first-year course non-completion outcomes and second-year non-retention outcomes. The data used in the

study were provided by a large public university in New Zealand. The data included all first-year students who enrolled in a bachelor's degree programme for the first time at the university during the academic years of 2009 to 2012. The full study sample consisted of about 16,000 students and about 88,500 course-level observations. The course observations were used to exam first-year course non-completion, which was defined as did not successfully complete a course in the first year. The student observations were used to study second-year non-retention rates, defined as not returning to enrol at the university at the beginning of the second year.

According to the definition of second-year non-retention rates, it is obvious that the authors ignored those who switched to other tertiary educational providers and did not return to the original university in the second year. The dataset contained detailed information on the students such as year of entry, demographic characteristics (gender, ethnicity and age), pre-university academic performance, and course and programme enrolment information. However, information on other relevant factors, such as parental education and family finances was not accessible. Another data limitation was that records on academic performance in high school were only available for some of the students in the sample. This thesis uses data from a large research database so we can include family background variables, and school information was available for the majority of our sample.

Maximum Likelihood Probit was utilised to estimate the influences of the explanatory variables on the two university outcomes. The authors found the outcomes vary by ethnicity. For example, holding other factors constant, the probabilities of first-year course non-completion for Pasifika and Māori were about 11 and nearly 7 percentage points, respectively, statistically significantly higher than European. There was no statistical evidence of a difference in the course non-completions between Asian and European students. Māori students were also more than 5 percentage points statistically significantly more likely to not return to the university in the beginning of the second year than European. Similar to many other studies, the analytical sample was drawn from one university. While the advantage of our data is that they cover students enrolled in all the universities in New Zealand.

Table 2.2. Summary of selected New Zealand studies on ethnic differences in university academic performance

Author(s), date	Data	Method	Outcome variable(s)	Explanatory variables	Relevant findings	Comments
Cao & Maloney (2018)	Data were provided by a university in New Zealand.	OLS. Maximum Likelihood Probit. Decomposition techniques.	(1) First-year course completion: successfully completed each first-year course, (2) First-year grade: grade received in each first-year course (converted to nine-point scale).	Gender, enrolment status, years of enrolled, ethnicity, age, school decile, NCEA results, university entrance type, type of bachelor's programme and course level.	About a quarter of the relatively underperformance of Māori and Pasifika could be eliminated if these ethnic groups had the same observed factors as European.	The results drawn from the study confined to one university. The university dataset lacked potentially important variables such as family income and parents' education attainment.
Jia & Maloney (2015)	Administrative data provided by a large public university in New Zealand.	Maximum Likelihood Probit.	(1) First-year course non-completion: did not successfully complete a course in the first year, (2) Non-retention in the second year: did not return to re-enrol at the university at the beginning of the second year.	Year of enrolment, demographic characterises (ethnicity, gender and age), high school academic performance, and course and program enrolment information.	The probabilities of first-year course non-completion for Pasifika and Māori were about 11 and nearly 7 percentage points, respectively, statistically significantly higher than European, after controlling for other factors. Māori were also more than 5	The study sample come from one university. The data had no information on parental education, family finances and community characteristics. Only partial information on student academic performance in high school was accessible.

					percentage points more likely than European to not return to second-year study.	
Juhong & Maloney (2006)	Administrative data provided by a university.	OLS. Maximum Likelihood Probit.	(1) GPA: GPA over every course over four-year period, (2) Dropout: discontinuation of studies in the initial programme by the year 2003 academic year without receiving a qualification at the university.	Ethnicity, age, gender, school decile, school qualification, Bursary Score.	Māori and Pasifika had lower first-year GPA and higher dropout rates after first year, compared to European. Controlling for other expandatory variables reduced the ethnic gaps in GPA by about 32%. The ethnic gaps in dropouts after first year become opposite when controlling for other factors and first-year GPA.	The study had limited data on school factors and family backgrounds. It also lacked data on individuals who switched programmes and universities, thus the results on dropout should be interpreted with caution.
Meehan, Pacheco & Pushon (2017)	Same as Meehan, Pacheco & Pushon (2019).	Same as Meehan, Pacheco & Pushon (2019).	Degree completion: completing a degree qualification within five since the start.	Same as Meehan, Pacheco & Pushon (2019), expect for school qualification was excluded and first-year course	The gaps between Māori, Pasifika or Asian and European were 20, 29 and -3 percentage points, respectively.	Excluded NCEA Level 1 from the regression abridgedly assign more predicative power to first-year course pass rate that biased

				pass rate was added.	Eliminating the differences in the factors between the ethnic minorities and European could close 73, 78 and more than 100% of the European-Māori, European-Pasifika and European-Asian gaps, respectively.	upward the explained gaps for Māori or Pasifika relative to European.
Meehan, Pacheco & Pushon (2019)	Individual-level administrative data.	Maximum Likelihood Probit. Fairlie decomposition.	Participation: enrolling in a bachelor's qualification before the age of 20.	Ethnicity, individual characteristics, socioeconomic status, school characteristics, school performance and engagement and parents' education.	Māori and Pasifika participation rates were 19.84 and 11.45 percentage points, respectively, lower compared to European. About 87% of European-Māori gap and the whole European-Pasifika gap in university participation can be explained if the minorities were given the same covariates as European.	The study looked at only one educational outcome. The lower NCEA qualification (NCEA Level 1) was used as the indicator for school performance, because the information was available for the majority of the student sample and the concern of endogeneity bias.

Scott (2005a)	Unit-level data collected by the New Zealand Ministry of Education.	Simple mathematical computation.	(1) Retention: enrolled in tertiary education more than one year, (2) Completion: completed a tertiary qualification with 5 years (up to 2002) since 1998, (3) Progression: progression in the year following completion.	The outcomes were looked at by level of qualification, ethnicity, age and gender.	Māori and Pasifika degree completion rates were 8 and 14 percentage points, respectively, lower than the rate of European.	Completion rates in the study underrepresented the true completion rates because they were not adjusted for part-time study and for students who took a break year from study. Lack of discussion about the size of the ethnic differences in completion rates can be explained by the differences in relevant explanatory variables.
Scott (2008)	The data came from the New Zealand Qualification Authority and the New Zealand Ministry of Education.	Logit regression.	(1) First-year course completion: passing all first-year courses, (2) First-year attrition: leaving tertiary education after first year without any qualification.	School academic performances, school characteristics, gender, age, ethnicity, tertiary field of study, number of courses taken.	Māori and Pasifika were 34 and 42%, respectively, less likely than non-Māori and non-Pasifika to pass all first-year courses once other factors were kept constant.	The study was not able to capture parental influences on tertiary achievement. The measure of attrition ignored the students who will return tertiary study in

					Māori students were also about 50% more likely to drop out of degree study after first year.	subsequent years.
Scott & Smart (2005)	Administrative data provided by the New Zealand Ministry of Education.	Logit regression.	Completion: successfully completed a bachelor's degree within six years after start (by 2003).	Gender, ethnicity, disability, domestic or international, highest school qualification, school decile, prior activity, intramural or extramural and field of study.	The actual degree completion rate for European was 1.2 and 1.4 times higher than that of Māori and Pasifika, respectively. Adjusting for other factors widened the European-Māori gap by about 8% and narrowed the European-Pasifika gap slightly more than 7%.	Although the data contained a range of study and demographic variables, they lacked potentially important information on parental backgrounds.
Shulruf, Hattie &Tumen (2008)	Data coming from the Ministry of Education and a university in Auckland	Logit regressions.	(1) Student pathway: moving from secondary schools to the university, (2) Student achievement: first-year GPA.	School administrative determinants, socio-geographic determinants, demographic determinants and peer determinants.	Māori and Pasifika students were statistically significantly less likely to gain UE and enrol in university than their European peers. Controlling for other	The study measured the outcomes at a single university only. The extent to ethnicity affected the educational outcomes such as course and

					covariates could close about 8% of Māori/Pasifika-European gap in gaining UE, and 38% and the entire of European-Māori and European-Pasifika gaps, respectively, in university enrolment.	qualification completion was not discussed.
Strathdee & Engler (2012)	Data provided by NZQA and the Ministry of education	Logit regressions.	Participation: progressing directly or indirectly (took a one- or two-year break) to bachelor level study.	Ethnicity, school achievement, school decile and gender.	For students from higher decile (9 and 10) schools, no statistically significant ethnic differences were found by controlling for school achievement. For students from lower decile (1 and 2) schools Māori students had lower likelihoods of progression on to bachelor-level study. Neither ethnicity, nor school decile, nor school	The results were only for students who achieved NCEA Level 3 and gained UE. This is a potential selection issue.

					achievement alone provided the best explanation of the outcome.	
Tumen, Shulruf & Hattie (2008)	Data came from a New Zealand university	Logit regression models	(1) Completion: gained the qualification at the end of third to fifth, (2) Departure: left the program at the end of first to fourth year, (3) Continuation: remained in the program at the second to fifth year.	Students background factors: age, ethnicity, gender, school decile, SES, residency status and bursary exam scores; program-related factors: intensity of study, breaks from the program, academic progress and GPA; program type.	The probability of completion did not statistically significantly differ among ethnicities, when tertiary and school achievements, intensity of study and other background factors were controlled for. The results should be interpreted with great caution.	The study was based on students enrolling in a bachelor program at a university. Completion in third and fourth years was compared to continuation in the following year. Small number of Māori and Pasifika students remained in study after third year.

2.7. Ethnic differences in degree completion

In addition to university participation and first-year course completions, successful degree completion is another common measure that is used to assess the long-term performance of both universities and students themselves. University completion benefits individuals and society through job satisfaction, civil engagement and job market access, which eventually boosts the country's economic development and global competitiveness (DeAngelo et al., 2011).

2.7.1. Findings from selected international studies

An early study by Light & Strayer (2002) investigated how white and minority students differ in their academic performance in college. The study used data from the 1979 National Longitudinal Survey of Youth (NLSY97), which started in 1979 with a sample of more than 12,000 respondents born between 1957 and 1964. The final sample included more than 7,000 respondents who satisfied the authors' sample selection criteria. The sample consisted of about 25% black, 16% Hispanic and 59% white students. Male and female students were almost equally represented in the sample.

The two academic outcomes of interest were college attendance and college completion. College attendance was defined as first attending a four-year college. College completion was defined as receiving a degree qualification from the college the students were enrolled in. According to the definition of college completion, it is obvious that those who switched to another college and eventually graduated from that college were excluded from the college completion outcome. This would bias downward the probability of overall college completion. Binomial Probit model was used to estimate the probability of college completion. The covariates used in the analysis included ethnicity, age, high school test scores (ranging from quartile 1 to 4), whether the student received financial aid, whether he or she lived at home, the tuition fees charged and the area's unemployment rate, while controlling for college quality quartile (from QQ1 represented bottom-quartile to QQ4 represented top-quartile).

The results suggested that white students overall had a higher probability of college completion than black and Hispanic students. For example, for students who chose to attend bottom-quartile (QQ1) colleges and had school test scores in the lowest quartile, the graduation probability was 23% for whites and 14.6% for the minorities, creating a gap of 8.4 percentage points. Controlling for other observed factors closed the ethnic gap in graduation by about 18% (the graduation probabilities were 22.4 and 15.5%, respectively, for whites and

minorities). For those attending QQ4 colleges with school test scores in the top quartile, the graduation probability for whites (50.9%) was 5.3 percentage points higher than for the minorities (45.6%). The ethnic gap changed direction (whites with 48.8% and minorities with 50.9%) if other observed factors were held constant. It should be noted that the study was based entirely on survey data that contained much less information about students and may suffer from measurement error in key variables.

Fletcher & Tienda (2010) examined the determinants of ethnic differences in university achievement using 10 years of administrative data for the students at four public universities in the State of Texas. The four institutions included the two most selective universities – the University of Texas at Austin (UT) and Texas A&M University (TAMU) and two less selective universities – the University of Texas-San Antonio (UTSA) and Texas Tech University (TECH).

Apart from the selectivity of admissions, the four universities are also different in terms of the demographic composition of their student bodies and in geographic location. For example, Hispanic students comprised 9% of the student body at TECH, 11% at TAMU, and 15% at UT, but 43% at UTSA, while Black students represented between 3 and 5% at TECH, TAMU and UTSA, but 17% at UT. UT is located in the capital city where there is a high level of diversity ethnically, politically and socially.

The study sample contained approximately 200,000 students enrolled in the four universities from the early 1990's to 2003. The authors looked at three university academic outcomes. These were 1st semester and 6th semester average GPA, and 4th year graduation. The key academic and demographic variables available in the dataset were ethnicity, gender, high school class rank and high school test scores. OLS with year and school effects were used. The authors found large ethnic disparities in the four-year graduation rates for all of the universities except UTSA. In particular, Hispanic students had lower graduation rates than their white peers at TAMU (13 percentage points), TECH (15 points) and UT (18 points). Similarly, Black students were less likely to graduate than whites within four years from TAMU/TECH (both 11 points) and UT (12 points). Applying school-level controls (school quality, class rank and school test scores) could reduce Hispanic-white and Black-white gaps in graduation rates by about 40 to 50%, which left around half of the differences unexplained. Although the study considered different type of students at the universities in Texas, it did not include students at private and community universities/colleges.

Similar ethnic gaps in qualification completion were also found in the UK. Rodgers (2013) examined the how lower degree completion rates of ethnic minority students in the UK's higher

education system should be interpreted. The study used data from a new (post-1992) UK university. The analytical sample included about 2,400 students who enrolled in a three- or four-year degree at the university in 2001/2002. Post-1992 universities were formally polytechnics and colleges of higher education before being granted university status. Usually, these universities were more likely to recruit students from poorer socioeconomic backgrounds and ethnic minority groups. This means ethnic minority students were overrepresented in the sample, which in turn could bias the results. Thus, the findings discussed below maybe not be applicable to the students at other universities.

Degree completion was defined as graduating within six years for those enrolled on a three- or four-year qualification. Explanatory variables included academic ability, subject of study, personal characteristics (ethnicity, gender, disability, whether the student was mature or lived with their parents) and socioeconomic backgrounds, and a series of interactions. Probit regression was used to identify the effects of the key variables on the probability of degree completion. The author found that Caribbean, Pakistani and Bangladeshi students were 40.4, 53.5 and 17.9%, respectively, less likely to complete a degree than Whites within six years. After controlling for other measurable factors, the gap in degree completion between White and Caribbean students reduced by about 10% (to 36.1%). The White-Pakistani gap decreased by around 31% (to 36.8%). More interestingly, the White-Bangladeshi gap reversed, indicating that the ethnic differences in degree completion could be fully explained by keeping other characteristics constant.

In addition to the ethnic disparities found in British and American universities, the underachievement in degree completion for minority students has also received public attention in European countries such as the Netherlands. Using individual-level administrative data, Zorlu (2011) addressed ethnic differences in the academic achievement of students in the Dutch tertiary education system, including both university (WO) and higher vocational education (HBO). The author focused on a sample of more than 77,000 students who started tertiary study in Dutch institutions for the first time in 1996 and looked at their academic outcome in 2005 to determine whether they successfully completed their study. The data were drawn from two sources. Student data on study information (e.g., subject of study, type of education, institution of study, and month and year of graduation) and personal characteristics (e.g., ethnicity, age and gender) came from the Central Register of Higher Education (CRIHO), while data about parental backgrounds (e.g., parental economic position in 1999) came from the Social Statistical Database (SSB). However, the author had no access to data about the attitudes of students that can potentially influence their university performance. Similarly, such variables are also unavailable in the data used in this thesis.

Dutch is the dominant ethnic group in the Netherlands. Ethnic minority students in the study were categorised into four main groups, based on their country of origin. These groups are Mediterranean, Caribbean, non-Westerners and Westerners. It is important to note that the majority of the ethnic minority students had been exposed to the mainstream norms and values in Dutch society as they were born in the Netherlands or moved to the country at a young age. Ethnic disparities were estimated in two tertiary outcomes – completion of a degree within ten years since it was started, and duration of completion in months. Parametric duration models were used to assess the sources of ethnic disparities in the university outcomes.

The results of the basic model (including personal demographic factors and the subject of study) showed that for the students enrolled in higher vocational education, Mediterranean, Caribbean, Westerners and non-Westerners were statistically significantly less likely to graduate within ten years compared to the Dutch students. Adding variables measuring parents' socioeconomic position recused the lower graduation rates of Mediterranean, Caribbean, Westerners and non-Westerners relative to the Dutch by 8.5, 16.3, 18.8 and 11.4%, respectively. For those enrolled in university, the ethnic minorities also had a lower probability of graduating than the Dutch students. Further controlling for parents' socioeconomic status increased the Dutch-Mediterranean gap but decreased the Dutch-Caribbean, Dutch-Westerner, Dutch-non-Westerner gaps by 9.2, 19.4 and 7.0%, respectively.

2.7.2. Findings from selected New Zealand studies

Concern is commonly expressed about the ethnic gaps in degree completion across New Zealand universities. Scott (2005a) completed the first comprehensive longitudinal study to describe the results of qualification retention, completion and progression in New Zealand tertiary education. The study sample included all domestic students (New Zealand citizens, permanent residents or Australian citizens) who enrolled in a formal qualification in tertiary providers in 1998. The study focused on two groups of students. For those who started a qualification at a tertiary provider in 1998, the author examined their retention and completion. Retention was defined as enrolled in tertiary education for more than one year, while completion was defined as completed a tertiary qualification within five years (by the end of 2002). It is important to note that the completion rates computed in the study would underrepresent actual completion rates because the rates were not adjusted for part-time study or for taking a break year during study. For students who completed a qualification in 2001, the author looked at their progression to higher level study in the following year.

Table 6. Five-year qualification completion rates, by level studied and ethnic group

Domestic students starting a qualification in 1998 at public providers	Percentage of Students Successfully Completed by end of 2002 by Qualification Level						
	Certificates	Diplomas	Degrees	Postgrad cert/dips	Honours/ master's	Doctorates	All levels
Māori	36	35	39	42	50	16	45
Pasifika	29	26	33	44	53	–	31
Asian	38	35	55	49	66	33	52
European/Pākehā*	29	32	47	50	58	26	39
All students	30	32	46	49	59	26	40

Source: Scott (2005a), Table 6, p12

It was found that completion rates varied by qualification level, gender, age and ethnicity (as shown in the Table 6 of Scott (2005a)). For example, Asian students had the highest five-year qualification completion rates at all levels. Although the rates for Māori were highest at lower qualification levels, they remained low at degree level and higher. Pasifika students had the lowest completion rates over all levels. The numbers reported in the table indicated that the Māori and Pasifika completion rates in degree qualification were 8 and 14 percentage points, respectively, lower than that of European. However, we find that these ethnic gaps have become much wider nowadays based on our analytical results. Scott did not explain the ethnic gaps his study. In the current thesis, we investigate how much of the ethnic gaps could be closed once a range of relevant factors are controlled for.

In addition to Scott (2005a), Scott & Smart (2005) aimed to provide a better understanding of the factors that influence degree completion in New Zealand. The analytical sample included about 38,000 domestic students who enrolled in a bachelor's degree at a New Zealand tertiary institution in 1998. Of the sample, European were the largest ethnic group (64%), followed by Māori (12), Asian (11), Other and Unknown (9%), and Pasifika (4%). It is important to note that our sample contains higher proportions of Māori (21%) and Pasifika (9%) students, which allows our regression models to yield more reliable results with greater precision and knowledge about ethnic differences in university outcomes. The study was based on student enrolment and completion data that were provided by tertiary institutions to the New Zealand Ministry of Education. Completing, in the study, was defined as successfully completing a bachelor's degree at a tertiary institution within six years of starting it (by 2003).

The authors took into account a range of demographic and study-related factors, including gender, ethnicity, disability, whether the student was domestic or international, highest school qualification, school decile, prior activity, intramural or extramural and field of study. The data, however, did not include information on parental backgrounds, which can potentially affect

degree completion. Unlike this previous work, we are able to include several parental related factors into our regression analysis. The authors utilised logit regression to analyse the impact of the considerable factors on the probability of completion of a degree qualification. It was found that the actual degree completion rate for European was 1.2 and 1.4 times higher than the rates of Māori and Pasifika students, respectively. The European-Māori gap in degree completion increased by about 8% and the gap between European and Pasifika closed by slightly more than 7%, once other factors were controlled for.

Tumen, Shulruf & Hattie (2008) intended to understand who was more likely to complete a bachelor's degree programme and to identify the predictors of university outcomes. The study relied on bachelor degree student enrolment and completion data of more 7,000 individuals, provided by a large New Zealand university, that covered the period from 2001 to 2005. Of the student sample, 43% were European, 38% were Asian, 6% were Pasifika, 5% were Māori, and the remaining 8% were other ethnicities. The cohort of students was defined by student programme that was eventually used in the analysis. The sample therefore comprised the more than 7,800 programmes taken by more than 7,000 students.

The authors looked at three university outcomes. (1) Completion: gained the qualification at the end of the third to fifth year; (2) Departure: left the programme at the end of first to fourth year; (3) Continuation: remained in the programme until the second to fifth year. The completion outcome at the end of the third and fourth years was compared with continuation in the following year. The completion outcome by the end of fifth year was compared to noncompletion. However, the fifth-year completion analysis did not include ethnicity because of insufficient numbers of Māori and Pasifika programmes. Therefore, the related results are not of interest to us. Students who finished their degree qualification within two years were excluded because they were assumed to be undertaking previous studies before 2001. Students who took a break year were considered as continuation and a variable was used to control for such situations. The independent variables used in the study were personal background factors (age, ethnicity, gender, school decile, SES, residency status and bursary scores), programme-related factors (study load, taking breaks from the programme, academic progress and annual GPA) and programme type (51 programmes were classified into 13 categories). Logit regression was used to identify the predictors of the university outcomes.

The regression results indicated that the probability of degree completion by the end of third and fourth years did not differ statistically significantly among the ethnic groups, once controlling for other observed factors. These findings, however, should be interpreted with great caution. First, as mentioned above, completion was compared to continuation, so the

results did not show ethnic differences in the 'actual' degree completion. Second, it is not ideal that GPA achievements were used to predict degree completion, given they are variables of the same kind (both outcome variables). If someone achieved a higher GPA, they in theory were very likely to complete their qualification. This may be why the authors did not find ethnic differences in degree completion when they controlled for university GPA results and other relevant factors. Third, the authors mentioned that there a small number of Māori and Pasifika students kept studying after the third year, which will have affected the comparison of the mean probabilities of degree completion between ethnic groups and European.

Meehan, Pacheco & Pushon (2017) was the earlier version of Meehan, Pacheco & Pushon (2019). Their studies used the same data and methodologies (see Section 2.5.2 for the discussion of the data and methodologies). They also looked at ethnic disparities in degree completion. The analysis used a sample that was restricted to the students enrolled in a three-year bachelor's programme. Degree completion was defined as completing a bachelor's qualification within five years of first enrolling.

The authors found that the degree completion rate for European was 20 and 29 percentage points, respectively, higher than that of Māori and Pasifika, but 3 percentage points below Asian. Their decomposition results indicated that if Māori, Pasifika or Asian had the same measurable factors as that for European, this could eliminate 73, 78 and more than 100%, respectively, of the European-Māori and European-Pasifika and European-Asian gaps in degree completion. First-year course pass rate was the most important factor when explaining the lower completion rates for Māori and Pasifika, relative to European. Eliminating differences in first-year course pass rates between Māori/Pasifika and European could close the European-Māori gap by 62% and the European-Pasifika gap by 63%. It should be noted that school qualification (NCEA Level 1) was not used in the degree completion analysis, although it was included in the participation analysis. This arbitrarily allocates more explanatory power to first-year course pass rates that will eventually bias upwards the estimated coefficients for the total explained gaps in degree completion between Māori/Pasifika and European. In this thesis, course completions in first and second year are used as dependent variables, in addition to participation and qualification completion. We consider different types of school qualifications and use the same variables, indicating school qualifications across all of our analyses.

2.8. Impact of paid employment on university outcomes while studying

There is a growing interest in the impact of university students' paid employment on their educational outcomes at university. On the one hand, working while studying can adversely affect university performance, because students engaged in employment tend to spend less time on academic activities (e.g., Bozick, 2007; Kalenkoski & Pabilonia, 2012). Reduced time spent on studying can subsequently harm academic achievement (Arulampalam, Naylor & Smith, 2012). On the other hand, educational attainment can benefit from employment for several reasons. First, working gives students an opportunity to apply what they have learned in class into practice (Geel & Backes-Gellner, 2012). Second, employment helps to develop transferable skills such as time management, teamwork and problem solving (e.g., Buscha, Maurel & Speckesser, 2012; Staff & Mortimer, 2007). Third, having a part-time job does not necessarily mean student workers reduce the time spent on studying (Triventi, 2014). Neyt et al. (2019) reviewed related literature and summarised that the effect of employment on university performance, depending on type of academic outcomes (GPA, continuation or graduation, etc.), the analysis method used, the country of analysis, job characteristics and student characteristics. This is why the findings on this topic have proven to be so inconsistent in the literature.

Callender (2008) explored the impact of students' term time work on their academic attainment at university. The study was based on survey data in a sample of more than 1,000 final year bachelor students from six new and established UK universities. The data were collected through postal self-completion questionnaires sent to the students by the universities. The students' academic attainment was measured by their average final year marks and by achieving a good degree. Final year marks were standardised in the study because the universities the students came from had different marking schemes and scales for their courses. A good degree was defined as a first or upper second class.

Employment was measured as average number of hours worked per week during term time. Logit analysis was used to estimate the overall relationship between term-time employment and the two university outcomes, taking into account a number of other factors, including institution, school qualifications, gender, age and subject area of study. The regression results showed that there was a statistically significant and negative relationship between term time

working and final year marks or degree classification. More hours spent working during the final year was associated with a poorer achievement in academic outcomes. For example, working one more hour reduced the final-year marks by 0.014 points, while the odds of achieving a good degree decreased by more than 3% by working an additional hour, after controlling for other factors. However, the author did not find evidence that working a very high number of hours was associated with a more negative effect on the academic outcomes.

In Beffy, Fougère, & Maurel (2010), the authors investigated the impact of part-time work on tertiary academic outcomes. The data used in the study were drawn from the annual Labour Force Surveys (LFS), administrated by the French National Institute of Statistics and Economic Studies (INSEE) from 1992 to 2002. The final study sample contained slightly more than 1,600 students who were enrolled in university studies for initial education and were prepared for an associate, bachelor's or a master's degree qualification. The sample was also restricted to those who were younger than 29 when participating in the survey and who were born in mainland France.

Two outcomes of interest were graduation (passing the year-end exam) and staying in education (continuing to study in the following year). Part-time employment was measured as no working, working less than 16 hours and working 16 hours or more. Probit models were used to estimate the probability of graduation and continuing to study. The findings suggested that, on average, holding a regular part-time job statistically significantly reduced the probability of graduation by 35 percentage points, compared to not working. When considering the number of hours worked, the authors found that working 16 hours or more per week had a statistically significant and larger effect (54 percentage points) on the graduation probity relative to not working. However, the additional analytical results suggested that being employed part-time had no statistically significant effect on the probability of continued studying.

Darolia (2014) analysed the effect of work on grades and credit completion across subgroups (gender, ethnicity, four- and two-year college, and full- and part-time study) for bachelor students in the US. The data were from the National Longitudinal Survey of Youth 1997 (NLSY97). More than 4,000 undergraduate students who were 12 to 16 years old in 1996 were included in the study sample. Grade outcome was measured by yearly average GPAs, ranging from 0 to 4. Credit completion included credits toward bachelor's and associate's degrees in

each year. Work hours were measured as average hours worked in each year during the second week of February to the second week of October.

OLS regression was used to estimate the influence of work on the college outcomes. Student-level fixed effects were also added to control for unobserved factors that may affect both academic achievement and work behaviour. The author found no evidence that additional work reduced the students' GPAs for either full-time or part-time students after using student fixed effects to control for personal characteristics. The possible reason was that increasing work hours did not necessarily decrease the amount of time allocated to study. Their results, however, showed that working had a detrimental effect on credit completion. For example, each hour of an increase in work statistically significantly lowered credits completed for four-year-college, male and female students by 0.11, 0.10 and 0.15 points, respectively.

The study by Wenz & Yu (2010) sought to understand the link between term-time employment and university academic performance. The study used survey data collected from Winona State University. The data contained a sample of more than 4,000 students who attended the university (which has two campuses in Southwestern Minnesota), during the years 2004 to 2008. Most students were 21 years old or younger. It is important to note that students who worked more than 30 hours weekly were not included in the sample. This could bias the results because those who worked intensively are the group mostly likely to be harmed by their employment. The outcome variable of interest was the average GPA being restricted to between 0 and 4. Employment was defined as number of hours worked per week. Fourty-nine percent of the students in the sample reported that they engaged in term-time employment.

OLS estimation was used to find the effect of employment on the academic outcome. College fixed effects was also added to control for college-related differences. The regression results indicated that working an additional hour had a very weak negative (0.007 points) effect on the GPA performance. The effect was statistically significant at 5% level. The authors also found that the students' GPA varied by their motivation to work. In particular, students working for career-specific skills tended to obtain higher GPAs than those who worked for financial reasons and those who worked for general experience.

A New Zealand-based study by Richardson et al. (2013) investigated the effect of hours worked on course letter grades. The study was based on students who enrolled in a bachelor's

course in 2010 at a New Zealand university (University of Canterbury), located in the South Island. The students were invited to participate on an online survey, and the data were collected. The survey was developed based on the one used by Haultain (2009). Data on a final sample of less than 1,900 participants was used in the analysis. The key dependent variable in the study was average GPAs on first semester courses. The GPA can run from 9 (A+) to 0 (D) and -1 (E). D and E were fail grades. Employment was measured as hours worked per week. Of the sample, 57.3% were employed and 42.7% were unemployed or out of the labour force. Those employed, on average, worked 13.80 hours per week and were paid a mean wage of $14.64 per hour. The explanatory variables included hours worked, study attitude, debt score, major, time management skills, personal characteristics and study-related factors.

There are some differences between Richardson et al and the current thesis. First, they used survey data and the respondents reported their working hours, whereas the current thesis relies on administrative data. However, we also need to estimate the hours worked due to the fact that they are not directly provided in the data. Second, the study controlled for variables, such as study attitude, debt score and the time management skills of the students, and we have no access to such variables. Third, the study looked at the effect of work hours on first semester GPAs. In the present thesis, we examine the effects of first- and second-year course completions through to qualification completion.

Their analytical results first showed that the mean GPAs between those employed and those unemployed are not statistically significantly different, which could be interpreted as that employment had no effect on GAP achievement. However, when the authors used Ordinary Least Squares to regress GPA on the mentioned variables for the working subsample, they found that working an additional hour statistically significantly reduced GPA by 0.03 points.

Table 2.3. Summary of selected studies of the impact of employment on university academic performance

Author(s), date	Country	Data	Method	Outcome variable(s)	Employment-related explanatory variable	Relevant findings
Callender (2008)	UK	Survey and university data.	Logit model	(1) Final year marks, (2) Achieving a good degree.	Average number of hours a week during term-time.	Term-time working negatively affected the outcomes. Controlling for other factors, working one more hour reduced the final-year marks by 0.014 points and the odds of achieving a good degree by more than 3%.
Beffy, Fougère & Maurel (2010)	Franch	data from the annual Labour Force Surveys (LFS).	Probit models.	(1) Graduation: passing the year-end exam, (2) Stay on in education: continuing in the following year.	Less or more than 16 hours per week.	Holding a regular part-time job statistically significantly reduced the probability of graduation by 35 percentage points, while combining work and study had no statistically significant effect on the probability of continuation.
Darolia (2014)	USA	Data came from the 1997 National Longitudinal Survey of Youth (NLSY97).	OLS estimation with student-level fixed effects.	(1) Yearly average GPA ranged from 0 to 4, (2) Yearly credits competed.	Average hours worked during second weeks of February and October.	There had no evidence that average GPA were harmed by additional work hours. However, four-year-

						college, male and female students gained 0.11, 0.10 and 0.15 fewer credits, respectively, for one hour increase in work.
Richardson et al. (2013).	New Zealand	Survey data.	OLS.	Average GPA of first semester courses.	Hours per week.	There had no statistically significant difference in GPAs between being employed and being non-employed. For the working subsample, GPA statistically significantly decreased 0.03 points for one hour increase in work.
Wenz & Yu (2010)	USA	Survey data collected by the university.	OLS with college fixed effects. Tobit.	GPA, bounded between 0 to 4.	Hours per week.	GPA fell only by 0.007 points for working one more hour. This effect was statistically significant at 5% level.

2.9. Motivation for the current thesis

Cao & Maloney (2018) used individual level administrative data records from 2012 to 2015, that were provided by a large urban New Zealand university, to tracked the success of students during their first year of study in a bachelor's degree programmes at that university. The sample contained more than 181,000 courses taken by more than 45,000 first-year students. The authors looked at two first-year academic outcomes. Course completion was defined as successfully passing a first-year course. The letter grade received in a first-year course was converted into a nine-point scale ranging from 9 (A+) to 1 (C-) and 0 (D). D (or below) is a failing grade.

Using the university data allowed the authors to control for factors that included gender, enrolment status, years enrolled, ethnicity, age, school decile, NCEA results, university entrance tyep, type of bachelor's programme and course level. The sample was comprised of 10.5% Māori, 15.1% Pasifika, 28.3% Asian, 38.2% European and 7.9% other ethnicity. Ordinary Least Squares and Probit regressions along with formal decomposition techniques were used to identify differences in the university outcomes between Māori/Pasifika/Asian and European students.

It was found that the average paper completion rate and GPA for Māori students was 9.95 and 0.804 percentage points, respectively, lower than that of European. The same first-year gaps were wider between European and Pasifika students, which were 21.36 and 1.886 percentage points respectively in course completion and GPAs. All of the differences were statistically significant at 1% level. Regression results revealed that, when controlling for other individual factors, high school backgrounds and university-related factors could reduce the Māori-European gaps by 25.7% in paper completion rates, and 25.9% in GPAs. The same observed characteristics constant could eliminate 39.2% and 28.2% of the differences in course completion rates and GPAs, respectively, between European and Pasifika students.

In addition, the decomposition results illustrated that if Māori and Pasifika had the same personal characteristics, high school backgrounds, and university factors as European, this could eliminate about 25% of the overall ethnic gaps in course completions and average grade points. This left the majority of differences in the university outcomes unexplained. The substantial unexplained differences could be accounted for by information such as family background (e.g., family income, parental education attainment and family structure) and school histories (e.g., school type, number of times switching school, and unjustified absence from school) that were unavailable in the administrative dataset used in this study. Two other

limitations were that the information on NCEA results was available for only about half (48.3%) of the sample and the results drawn from this study were confined to one university.

The current thesis extends the past literature, mainly Cao & Maloney (2018), by using administrative data in the Integrated Data Infrastructure (IDI) to explore the differences in academic achievement between ethnic groups across all universities in New Zealand. The data in the IDI contains much more comprehensive histories on students relative to the university datasets. This research investigates students' academic disparities, from participation through to graduation (which was not possible in the previous study), when studying for a bachelor's degree. In particular, university outcomes of interest, including participation, first- and second-year course completion and degree completion, are examined by taking into account a wider range of factors including demographic characteristics, school characteristics, school histories, university-related factors, term-time employment and parental backgrounds. A major drawback of the IDI data is that they do not contain information on course letter grade (GPA), which prevents us from looking at GPA performance.

Regression analyses are used to discover the size of the remaining ethnic differences in university outcomes from our previous study, once we control for a broader set of relevant factors. In addition, modern statistical decomposition techniques are employed to estimate the importance of the distinct categories of variables in explaining pairwise differences in the academic outcomes between the main ethnic minority group (Māori, Pasifika and Asian), and European who are the largest ethnic group in New Zealand.

Chapter 3. Data and research methods

3.1. Data and descriptive statistics

This thesis makes use of individual-level linked administrative and survey data that are securely kept in the IDI.[3] The IDI is a large and frequently used research database maintained and operated under strict privacy and confidentiality protocols by a public service government department, Statistics New Zealand. The IDI contains microdata on millions of people and households in New Zealand. The data come from a range of government agencies, Statistics New Zealand surveys and non-government organisations.

When Statistics New Zealand collects data from the source agencies, they first make sure the data contain all the information they should have and then prepare the data for linking. Information about the same individual in different datasets is linked together by using both 'deterministic' and 'probabilistic' methods, subject to the relevant variables contained in each dataset.[4] If common personal identifier variables such as National Health Index numbers or passport codes are available, records are allowed to be exactly linked together. This is the 'deterministic' linking method. In cases where personal identifiers are unavailable, 'probabilistic' linking is done using demographic variables, such as name, gender and date of birth. Although the linking methods are highly robust it does not completely rule out the possibility of some records being incorrectly linked and missed. However, these linking errors can be minimised through frequently preforming quality checks.

Before making the data available for research analysis, all the identifiable personal information like names, dates of birth and IRD numbers has been removed or replaced to minimise the risk of individuals being identified. The deidentified data in the IDI are generally classified into eight broad data categories: health, education and training, benefits and social services, people and communities, population, income and work, housing, and criminal justice. Researchers can link the different categories of data through a personal unique identifier variable (snz_uid) by putting some simple codes into statistical software packages such as

[3] For detailed information about the IDI data see https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/.

[4] For more information about how data are linked in the IDI see http://archive.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure/idi-how-it-works.aspx#gsc.tab=0.

Stata, SAS, R Studio or SQL. Statistics New Zealand regularly (usually once a quarter) carries out quality checks and data refreshes to ensure the IDI data are high quality and up to date. Only approved researchers who have a public interest project can access the IDI data in a secure data lab, which is either directly or indirectly managed by Statistics New Zealand. Users can only access the data needed for their specific project. Research results must be checked by Statistics New Zealand before they can be released to the data lab to ensure no individual can be identified. Any results with risk that means individuals could be identified will not be released.

The data that this thesis primally relies on are administrative data. They include education, tax, justice, benefit, personal details and the 2013 Census data. Similarly to the data from the surveys such as the Household Labour Force Survey (HLFS), the Survey of Family Income and Employment (SoFIE) and the Household Economic Survey (HES), the 2013 Census data were also collected and processed by the government. Moreover, the Census data are taken from the entire resident population, given it solicits information from everyone in the country at a particular time.

From our perspective, there are two advantages in terms of using administrative data contained in the IDI. First, compared to survey data, administrative data may provide more accurate and detailed information on, for example, educational qualifications, income and benefit histories. It means that this may result in much lower levels of measurement error with the administrative data. Second, the education datasets in the IDI contain richer records on students, such as their school histories and characteristics, that are not available in the dataset provided by the university that were used in Cao and Maloney (2018), as mentioned before in Chapter 2. The next section provides detailed discussions on the different types of data that are used in this thesis.

3.1.1. Data sources

Education and training data are collected by the New Zealand Ministry of Education which covers data on all the levels of education and industry training. Based on the type of education, these data are grouped into early childhood education, primary and secondary school, tertiary education, and industry/targeted training data. This thesis mainly uses the data from primary through to tertiary education. Data in the primary and secondary education table covers all students who have enrolled since 2007 and are currently enrolled at schools in the country,

and those who have studied within the National Standards programmes. The data are about the students' personal details, enrolment information, support interventions, academic performance, including standards studied and qualifications, and attendance. Meanwhile, the data are also provided with detailed information about the associated school(s) that includes school decile, gender composition, authority, area and region, and more. However, the data on specific grades from a school qualification, for example, NCEA Level 3 achieved, or merit or excellence credits, are only available for a small subset of students.

The data about tertiary education cover all students who have enrolled in formal and non-formal tertiary qualifications at government funded Tertiary Educations Organizations (TEOs) within New Zealand from the year 1994. These data contain details of tertiary-level course enrolment and completion, qualification enrolment and completion, characteristics of TEOs where courses and qualifications being undertaken, as well as student characteristics. One potential limitation with the data is they do not include course grades. Tertiary data used by the Ministry of Education and the Tertiary Education of Commission (TEC) for the purposes of funding and monitoring the TEOs' performance. It should be noted that school qualification records are also contained in the tertiary data tables and we choose to link school qualifications to our sample by using tertiary data as it allows for a higher proportion of linking to be made. But specific grades from a school qualification are not provided in the tertiary data, which restricts us from looking at the impacts of school qualification grades on the university outcomes.

A variety of tax data related to both individual and business activities that have occurred within this country are collected by the Inland Revenue Department (IRD). The collection of the data is to ensure New Zealand's taxation system functions efficiently over time. Meanwhile, the IRD tax data have been used in a wide range of research projects conducted by individual researchers, research institutes and governmental originations. The data are comprised of more than 50 million taxpayer records from 1 April 1999 to the current period. Individuals' employment income tax data are drawn from the Employer Monthly Schedule (EMS). In New Zealand, employers are required by the IRD to submit an EMS form for each of their employees on a monthly basis. A month is identified as the final day of the calendar month. A monthly EMS form for an employee reports their name, IRD number, gross earnings, total PAYE deductions, tax code and employment start and end date, if the employee started or left the job in that month. The employee's name cannot be seen by the 'external' data users. Although the EMS form is very informative, it does not provide the associated employment

work hours. The lack of information forces us to estimate the work hours ourselves in some situations, because we want to know the weekly number of hours that the first- and second-year students may have spent on their employment work during the academic months.

Justice data are provided by the Ministry of Justice (MOJ) to Statistics New Zealand for use in the IDI. Data of court charges contain all criminal charges for people that have been disposed since 1992. A charge being disposed means an outcome has been determined by the criminal courts in New Zealand. The Ministry of Justice also records charges for organisations. However, the data on charges for organisations cannot be linked to the IDI, but this does not create any issue for us because the data are not relevant to this thesis. The information of each charge for people being recorded in the table includes the date that the offence happened, when the charge was filed and by whom, and when the charge was finalised, with what outcome. A charge usually refers to one criminal offence that can be filed by police officers, local authorities, Corrections or other government agencies. However, there could be more than one charge for one crime. For example, someone may face three charges of financial fraud and one charge for money laundering (four charges for two crimes). The outcome of a charge can be a conviction, not proved (i.e., dismissed, discharged, withdrawn or acquitted), other proven (i.e., discharge without conviction or diversion), or other. The availability of this justice data allows us to obtain the number of convicted criminal charges for the interested people during a certain period of time.

Benefit dynamics data are sourced from the Ministry of Social Development. The data are about all people who received working-age social welfare benefit(s) for the period beginning from 1 January 1993 to the date of the most recent data update (July 2020 for the purposes of this study). These data provide related details on the primary benefit recipient. These details include their demographic characteristics, benefit start and end date, benefit type, and benefit status, as well as the history of partner(s) and dependent child(ren) included in the benefit. We use the data is to estimate the proportion of days that the students' parent(s) have been living on benefits during a selected range of ages of the students and to derive the main type of benefit the parent(s) relied on for that particular period of time.

The 2013 Census was a national survey of the entire population and dwellings, which was collected by Statistics New Zealand during the period starting on 5 March 2013 (The National Census Day) and ending on 10 April 2013. The original data were checked, evaluated and

edited by Statistics New Zealand to make sure they were suitable for different uses, including this current research project. That is one of the main advantages of using Census data from a researcher's perspective. The modes in which the survey was administrated were both online and on hardcopy paper. All New Zealanders, including those who usually reside overseas, were expected to take part in the survey. They were asked to answer a series of questions. The questions generally were divided into topics that are about population structure, location, culture and identity, education and training, work, income, families and households, housing, transport, and finally health and disability.

The collection of Census data can be used to assess how government funding was spent on society in the past, and to help to make a short to medium term budget for the near future. Apart from that, the Census data are important because they allow us to identify the students' parent(s) within the same household, and then to link parental factors such as educational attainment, benefit histories, household type (i.e., one-parent household) and household income to the students. We can identify more than 72% of the parents of the students in the sample through using the 2013 Census. The unlinked parents are most likely to be the parents of the students from the earlier cohorts (cohorts 2010 and 2011) who became 17-18 years old in the year 2013 and some of whom moved out from their parents' house, which makes it impossible to identity these students' parents if we rely only on the 2013 Census. More parents could have been identified if we had made use of the earlier Census data from 2006, or had used later age-15 cohorts (e.g., cohorts 2012 or 2013) of students in our sample. However, both potential strategies cannot be adopted in this study. The first reason is that we cannot use the 2006 Census data because these data have not been integrated into the IDI and the unique personal identifier variable, snz_uid, is not available in any of the data tables. The second reason is that if we used later cohorts of students, this would prevent us from examining the university outcomes of qualifications completed due to the unavailability of tertiary education data after 2019 in the IDI by the time our analysis was carried out.

The table of personal details covers all persons who have a record in any of the datasets in the IDI. The data are derived from available data sources such as the Census, the Department of Internal Affairs (DIA) and the Ministry of Health (MoH). If a person was found to have mixed information in the different data sources, Statistics New Zealand has their way to derive the most trusted information. Demographic information available in the personal details table includes gender, year of birth and month of birth. The day of birth, however, is unavailable due to privacy protection. Because of that, we therefore assume the day of birth is the 15th in the

month reported when estimating the age of the students in our sample, which is the way we can on average minimise the variance between the true age and the estimated age for a random, sufficient size group of students. Ethnicity information is also available in the personal details table, but it is not used to determine ethnicity. The source of ethnicity data will be discussed later in this chapter.

3.1.2. Dependent variables

This thesis focuses on young adults who turned 15 years old in the years 2010 to 2012 while they were studying in a New Zealand school. The total sample comprises 179,034 students. Among the sample there is a small number of foreign students, who are treated in the same way as domestic students. This study sample is split into three cohorts, based on the year they turned 15 years old (cohorts 2010, 2011 and 2012). For the cohorts of students, we set out to track the ethnic differences in four university academic outcomes, from participation in a bachelor's qualification by age 21 to qualification completion, while controlling for a broad range of factors suggested in previous studies.

Table 3.1 provides the definitions of both the dependent and independent variables used in this thesis. Participation is the first outcome of interest to be looked at, which is defined as enrolling in a formal bachelor's qualification of more than 0.03 of an equivalent full-time student (EFTS)[5] with a New Zealand university by age 21. It is a dummy variable that takes the value of 1 indicating positive participation outcome; 0 otherwise.

Conditional on those who have enrolled in a bachelor's qualification, we next examine their first- and second-year academic performance in terms of course completion. First- and second-year course completion is defined as the proportion of courses taken during the first and second year that were successfully completed. They are fractional/proportional variables with the minimum value of zero and maximum value of one. Zero means successfully completing none of courses enrolled in during the academic year, while 1 indicates successfully completing all of the courses enrolled in during that same year. For example, if a

[5] 0.03 EFTS is equivalent to one week's duration as a full-time student.

first-year student successfully passed six out of eight courses they were enrolled in during their first year, then their course completion rate was 75%. For second-year course completion analysis, the sample of interest is restricted to the 2011 and 2012 cohorts, because the tertiary data needed for cohort 2013 onwards are not yet available.

We then want to understand which individuals are more likely to complete their qualification. Thus, our last dependent variable is degree completion. For degree completion analysis, the sample is further restricted to the cohort 2010 students who were enrolled in a three-year bachelor's qualification[6], again due to data unavailability. Degree completion is defined as successfully completing a three-year bachelor's degree by age 24. It is also a dummy variable that takes the value of 1 indicating successfully completing a three-year bachelor's qualification by the selected age, 0 otherwise. In this thesis, analysis of participation, first- and second-year course completion and qualification completion are sometimes referred to as first, second, third and fourth stage analysis, respectively. The first-stage analysis is considered as the fundamental assessment as it is the analysis that sets up the second to fourth stages.

[6] A typical three-year bachelor's qualification consumes a total of 3 EFTS, in other word, three full-time years' study.

Table 3.1. Variables used in this research

Variable names	Definitions	Used in first stage analysis	Used in second stage analysis	Used in third stage analysis	Used in forth stage analysis
Dependent variables					
Participation	Equals 1 if enrolled in a bachelor's qualification with a New Zealand university by age 21; 0 otherwise	✓			
First-year course completion	Proportion of first-year successful course completion ranging in value between 0 and 1 inclusive		✓		
Second-year course completion	Proportion of second-year successful course completion ranging in value between 0 and 1 inclusive			✓	
Qualification completion	Equals 1 if enrolled in and successfully completed a three-year bachelor's qualification at a New Zealand university by age 24; 0 otherwise				✓
Independent variables					
Female	Equals 1 if being female; 0 otherwise	✓	✓	✓	✓
Cohort years					
Cohort 2010	Equals 1 if turned 15 years of age in 2010; 0 otherwise	Reference group	Reference group	Reference group	✓
Cohort 2011	Equals 1 if turned 15 years of age in 2011; 0 otherwise	✓	✓	✓	
Cohort 2012	Equals 1 if turned 15 years of age in 2012; 0 otherwise	✓	✓		
Ethnicities					
Māori	Equals 1 if prioritized ethnicity is Māori; 0 otherwise	✓	✓	✓	✓
Pasifika	Equals 1 if prioritized ethnicity is Pasifika; 0 otherwise	✓	✓	✓	✓

Asian	Equals 1 if prioritized ethnicity is Asian; 0 otherwise	✓	✓	✓	✓
MELAA[a]	Equals 1 if prioritized ethnicity is MELAA; 0 otherwise	✓	✓	✓	✓
European	Equals 1 if prioritized ethnicity is European; 0 otherwise	Reference group	Reference group	Reference group	Reference group
Other	Equals 1 if prioritized ethnicity is none of the above; 0 otherwise	✓	✓	✓	✓
School deciles[b]					
Decile 1	Equals 1 if last studied in decile 1 school; 0 otherwise	✓	✓	✓	✓
Decile 2	Equals 1 if last studied in decile 2 school; 0 otherwise	✓	✓	✓	✓
Decile 3	Equals 1 if last studied in decile 3 school; 0 otherwise	✓	✓	✓	✓
Decile 4	Equals 1 if last studied in decile 4 school; 0 otherwise	✓	✓	✓	✓
Decile 5	Equals 1 if last studied in decile 5 school; 0 otherwise	✓	✓	✓	✓
Decile 6	Equals 1 if last studied in decile 6 school; 0 otherwise	Reference group	Reference group	Reference group	Reference group
Decile 7	Equals 1 if last studied in decile 7 school; 0 otherwise	✓	✓	✓	✓
Decile 8	Equals 1 if last studied in decile 8 school; 0 otherwise	✓	✓	✓	✓
Decile 9	Equals 1 if last studied in decile 9 school; 0 otherwise	✓	✓	✓	✓
Decile 10	Equals 1 if last studied in decile 10 school; 0 otherwise	✓	✓	✓	✓
Decile unknown	Equals 1 if last studied school decile is unknown; 0 otherwise	✓	✓	✓	✓
School authority					
State	Equals 1 if last studied in state owned school; 0 otherwise	Reference group	Reference group	Reference group	Reference group
State integrated	Equals 1 if last studied in state integrated school; 0 otherwise	✓	✓	✓	✓
Private	Equals 1 if last studied in private school; 0 otherwise	✓	✓	✓	✓

Unknown	Equals 1 if last school's ownership is unknown; 0 otherwise	✓	✓	✓	✓
School gender composition					
Co-education	Equals 1 if last studied in co-education school; 0 otherwise	Reference group	Reference group	Reference group	Reference group
Single sex	Equals 1 if last studied in single sex school; 0 otherwise	✓	✓	✓	✓
Unknown	Equals 1 if last school's gender is unknown; 0 otherwise	✓	✓	✓	✓
School area					
Main Urban Area	Equals 1 if last school studied is in main urban area; 0 otherwise	Reference group	Reference group	Reference group	Reference group
Minor Urban Area	Equals 1 if last school studied is in minor urban area; 0 otherwise	✓	✓	✓	✓
Secondary Urban Area	Equals 1 if last school studied is in secondary rural area; 0 otherwise	✓	✓	✓	✓
Rural Area	Equals 1 if last school studied is in rural area; 0 otherwise	✓	✓	✓	✓
Unknown	Equals 1 if last school studied is in unknown area; 0 otherwise	✓	✓	✓	✓
School region					
Northland	Equals 1 if last school studied is in Northland; 0 otherwise	✓			
Auckland	Equals 1 if last school studied is in Auckland; 0 otherwise	Reference group			
Waikato	Equals 1 if last school studied is in Waikato; 0 otherwise	✓			
Bay of Plenty	Equals 1 if last school studied is in Bay of Plenty; 0 otherwise	✓			
Gisborne	Equals 1 if last school studied is in Gisborne; 0 otherwise	✓			
Hawkes Bay	Equals 1 if last school studied is in Hawkes Bay; 0 otherwise	✓			
Taranaki	Equals 1 if last school studied is in Taranaki; 0 otherwise	✓			

Manawatu-Whanganui	Equals 1 if last school studied is in Manawatu-Whanganui; 0 otherwise	✓			
Wellington	Equals 1 if last school studied is in Wellington; 0 otherwise	✓			
West Coast	Equals 1 if last school studied is in West Coast; 0 otherwise	✓			
Canterbury	Equals 1 if last school studied is in Canterbury; 0 otherwise	✓			
Otago	Equals 1 if last school studied is in Otago; 0 otherwise	✓			
Southland	Equals 1 if last school studied is in Southland; 0 otherwise	✓			
Tasman	Equals 1 if last school studied is in Tasman; 0 otherwise	✓			
Nelson	Equals 1 if last school studied is in Nelson; 0 otherwise	✓			
Marlborough	Equals 1 if last school studied is in Marlborough; 0 otherwise	✓			
Unknown	Equals 1 if last studied in correspondence school or region of last studied school is unknown; 0 otherwise	✓			
Highest school qualification					
No formal qualification	Equals 1 if did not complete a formal school qualification; 0 otherwise	Reference group	Reference group	Reference group	Reference group
NCEA Level 1	Equals 1 if highest school qualification is NCEA Level 1; 0 otherwise	✓	✓	✓	✓
NCEA Level 2	Equals 1 if highest school qualification is NCEA Level 2; 0 otherwise	✓	✓	✓	✓
NCEA Level 3	Equals 1 if highest school qualification is NCEA Level 3; 0 otherwise	✓	✓	✓	✓
Overseas qualification[c]	Equals 1 if highest school qualification is overseas qualification; 0 otherwise	✓	✓	✓	✓
Other qualification[d]	Equals 1 if highest school qualification is neither NCEA nor overseas qualification; 0 otherwise	✓	✓	✓	✓

Unknown qualification	Equals 1 if highest school qualification is unknown; 0 otherwise	✓	✓	✓	✓
Switching school					
0	Equals 1 if did not switch school during the period of school education; 0 otherwise	Reference group	Reference group	Reference group	Reference group
1	Equals 1 if switched school once during the period of school education; 0 otherwise	✓	✓	✓	✓
2	Equals 1 if switched school twice during the period of school education; 0 otherwise	✓	✓	✓	✓
3	Equals 1 if switched school three times during the period of school education; 0 otherwise	✓	✓	✓	✓
4 or more	Equals 1 if switched school four times or more during the period of school education; 0 otherwise	✓	✓	✓	✓
Truancy, suspension and/or unjustified absence	Equals 1 if being truant, suspended and/or unjustified absent from school; 0 otherwise	✓	✓	✓	✓
Age					
Under 18	Equals 1 if age under 18 in the university academic year; 0 otherwise		Reference group	Reference group	
18	Equals 1 if age 18 in the university academic year; 0 otherwise		✓	✓	
19	Equals 1 if age 19 in the university academic year; 0 otherwise		✓	✓	
20	Equals 1 if age 20 in the university academic year; 0 otherwise		✓	✓	
21	Equals 1 if age 21 in the university academic year; 0 otherwise		✓	✓	
22 or older	Equals 1 if age 22 or older in the university academic year; 0 otherwise		✓	✓	
Hours of employment work (1)[e]					

Zero	Equals 1 if on average worked zero hours per week during the university academic year; 0 otherwise;		Reference group	Reference group	
Low	Equals 1 if on average worked between zero and less than 10 hours per week during the university academic year; 0 otherwise		✓	✓	
Median	Equals 1 if on average worked between 10 and less than 20 hours per week during the university academic year; 0 otherwise		✓	✓	
High	Equals 1 if on average worked 20 hours or more per week during the university academic year; 0 otherwise		✓	✓	
Unknown	Equals 1 if records on employment work during the university academic year are unknown; 0 otherwise		✓	✓	
Hours of employment work (2)					
Zero	Equals 1 if worked zero hours per week during the university years completing the first degree qualification by age 24; 0 otherwise;				Reference group
Low	Equals 1 if worked between zero and less than 10 hours per week on average during the university years completing the first degree qualification by age 24; 0 otherwise				✓
Median	Equals 1 if worked between 10 and less than 20 hours per week on average during the university years completing the first degree qualification by age 24; 0 otherwise				✓
High	Equals 1 if worked 20 hours or more per week on average during the university years completing the first degree qualification by age 24; 0 otherwise				✓
Unknown	Equals 1 if records on employment work during the university years completing the first degree qualification by age 24 are unknown; 0 otherwise				✓
Prior activity					

Secondary school student	Equals 1 if being secondary school student prior to enrolling in a bachelor's degree at university; 0 otherwise		Reference group	Reference group	Reference group
Unemployed or beneficiary	Equals 1 if being unemployed or beneficiary prior to enrolling in a bachelor's degree at university; 0 otherwise		✓	✓	✓
Employed	Equals 1 if being wage/salary worker or self-employed prior to enrolling in a bachelor's degree at university; 0 otherwise		✓	✓	✓
Other types of student	Equals 1 if being below-bachelor-level student at university, polytechnic, college, private training establishment and Wānanga prior to enrolling in a bachelor's degree at university; 0 otherwise		✓	✓	✓
House person	Equals 1 if being house person prior to enrolling in a bachelor's degree at university; 0 otherwise		✓	✓	✓
Overseas	Equals 1 if being overseas prior to enrolling in a bachelor's degree at university; 0 otherwise		✓	✓	✓
University[f]					
University A	Equals 1 if studied at University A in the academic year; 0 otherwise		Reference group	Reference group	Reference group
University B	Equals 1 if studied at University B in the academic year; 0 otherwise		✓	✓	✓
University C	Equals 1 if studied at University C in the academic year; 0 otherwise		✓	✓	✓
University D	Equals 1 if studied at University D in the academic year; 0 otherwise		✓	✓	✓
University E	Equals 1 if studied at University E in the academic year; 0 otherwise		✓	✓	✓
University F	Equals 1 if studied at University F in the academic year; 0 otherwise		✓	✓	✓
University G	Equals 1 if studied at University G in the academic year; 0 otherwise		✓	✓	✓
University H	Equals 1 if studied at University H in the academic year; 0 otherwise		✓	✓	✓
New Zealand citizenship	Equals 1 if being New Zealand citizen; 0 otherwise		✓	✓	✓

Part-time	Equals 1 if being part-time student in the academic year; 0 otherwise		✓	✓	
Single-parent household	Equals 1 if lived in single-parent household; 0 otherwise	✓	✓	✓	✓
Parents highest qualification					
No qualification	Equals 1 if parents did not complete any qualification; 0 otherwise	Reference group	Reference group	Reference group	Reference group
Level 1-3 Certificate	Equals 1 if parents' highest qualification is Level 1 to 3 Certificate; 0 otherwise	✓	✓	✓	✓
Level 4 Certificate	Equals 1 if parents' highest qualification is Level 4 Certificate; 0 otherwise	✓	✓	✓	✓
Level 5-6 Diploma	Equals 1 if parents' highest qualification is Level 5 to 6 Diploma; 0 otherwise	✓	✓	✓	✓
Bachelor's degree or level 7 qualification	Equals 1 if parents' highest qualification is bachelor's degree or other level 7 qualification; 0 otherwise	✓	✓	✓	✓
Postgraduate	Equals 1 if parents' highest qualification is postgraduate degree, honours, master's or doctorate degree; 0 otherwise	✓	✓	✓	✓
Overseas school qualification	Equals 1 if parents' highest qualification is overseas school qualification; 0 otherwise	✓	✓	✓	✓
Unknown qualification	Equals 1 if parents' highest qualification is unknown; 0 otherwise	✓	✓	✓	✓
Known parents' benefit and crime histories	Equals 1 if parents' benefit and crime histories during ages 6 and 16 of associated student(s) are known; 0 otherwise	✓	✓	✓	✓
Parents' proportion of days living on benefit	Average parents' proportion of days living on benefit support during ages 6 and 16 of the associated student(s) ranging from zero to one.	✓	✓	✓	✓
Parents' number of convicted charges	Average parents' number of convicted charges during ages 6 and 16 of the associated student(s).	✓	✓	✓	✓

Household income[g]					
Zero	Equals 1 if household income is zero; 0 otherwise	✓	✓	✓	✓
1 to 30,000	Equals 1 if household income is $1 to $30,000; 0 otherwise	✓	✓	✓	✓
30,001 to 50,000	Equals 1 if household income is $30,001 to $50,000; 0 otherwise	Reference group	Reference group	Reference group	Reference group
50,001 to 70,000	Equals 1 if household income is $50,001 to $70,000; 0 otherwise	✓	✓	✓	✓
70,001 to 100,000	Equals 1 if household income is $70,0001 to $100,000; 0 otherwise	✓	✓	✓	✓
100,001 or more	Equals 1 if household income is $100,001 or more; 0 otherwise	✓	✓	✓	✓
Unknown	Equals 1 if household income is unknown; 0 otherwise	✓	✓	✓	✓

[a] MELAA is the acronym for Middle Eastern/Latin American/African.

[b] Decile 1 schools are the bottom 10% of schools with the highest proportion of students from low socioeconomic regions, whereas decile 10 schools are the top 10% of schools with the lowest proportion of these students.

[c] Overseas qualifications include mainly Cambridge International Exams (CIE) and International Baccalaureate (IB), and other overseas qualifications.

[d] Other qualifications include School Certificate, Other University Entrance, Sixth Form Certificate and University Bursary.

[e] In the IRD datasets, information on taxpayers' earnings from employment is given on a monthly basis, but the associated work hours and hourly rate both are not available. Given that, the weekly average employment hours during the academic year (March to November) of the students are approximated by assuming they were paid minimum wage.

[f] Particular university is not allowed to be identified due to the confidentiality requirements of Statistics New Zealand.

[g] Household income is before-tax earnings received by a household between 1 April 2012 to 31 March 2013. It included wages and salaries, self-employment income, property and rental income, dividends and investments, social insurance, superannuation, government assistance payments and private transfers such as child support.

3.1.3. Independent variables and descriptive statistics

Independent variables used in our regression analyses are grouped into six categories. They are demographic factors, school characteristics, school histories, university factors, employment workload during university study and parental backgrounds. Again, definitions of these explanatory variables and the uses of the analysis can be found in Table 3.1. For example, university factors and employment workload are not taken into account in the participation analysis.

Given that this thesis aims to explain ethnic differences in educational outcomes, ethnicity is the primary explanatory variable of interest. According to the Ministry of Education and Statistics New Zealand, ethnicity is defined as a cultural affiliation that relates to the ethnic group or groups that people identify with or feel they belong to (Ministry of Education, 2014). It is not difficult to understand this conceptual definition of self-reported ethnicity, but there are challenges when dealing with ethnicity data in the IDI. That is because ethnicity identity information for the same individual can be different in different data sources and can vary over time. To overcome the issues of mixed and inconsistent ethnic information we make the choice to use the steady ethnic information reported in the school enrolment dataset.[7]

Students are allowed to self-report up to three distinct ethnic identities when enrolling in school. The students who self-reported only one ethnic identity will be assigned to that ethnic group. For those who identify with more than one ethnic group, we first apply the prioritisation rules to designate their ethnicity. The ethnicity prioritisation system has been used by government and researchers from 2004/2005 in New Zealand (see Boven et al., 2020; Cormack & Robson, 2011; Reid, Bycroft & Gleisner, 2016, for more discussion). The order of priority is Māori, Pasifika, Asian, MELAA (Middle Eastern, Latin American and African), Other and European (Pākehā). For example, a student who reports Māori as one of their ethnicities is considered to be Māori. Non-Māori students are deemed to be Pasifika if one of their self-reported ethnicities is Pasifika. Non-Māori and non-Pasifika students are classified as Asian if Asian is one of their self-reported ethnicities. The remaining students are allocated across the MELAA, Other, and European categories in a similar way. Students with unknown ethnicity are combined into the Other category. Throughout the first to fourth stage analysis European has been used as the control ethnic variable. All the mentioned ethnicity variables are used in

[7] We compare the ethnic information provided by the Ministry of Education to those reported by other agencies and find that this information in fact is not much different across the data sources.

Probit analyses. However, because of their relatively small sizes, students with MELAA and other ethnic identities will be ignored in our decomposition analyses.

Table 3.2 reports means of the variables used in the university enrolment analysis. By using the prioritized ethnic designations, the whole sample consists of 57.1% European, 21% Māori, 10.3% Asian, 9.1% Pasifika, 1.8% MELAA and 0.8% other ethnicity.[8] It should be note that there are lower proportion of European, a similar amount of Māori, but higher proportions of Pasifika and Asian in our sample, compared to the sample used in a similar study, by Meehan, Pacheco & Pushon (2019). The Ethnicities section in Table 3.2 also shows the percentages of the students who self-reported multiple ethnicities. For example, students whose prioritised ethnic identify is Māori also self-reported themselves as Pasifika (5.4%), Asian (0.6%), MELAA (0.1%), European (18.6%) and Other (0.3%). Similarly, those who designated as Pasifika under prioritization who also self-reported themselves as Asian (2.2%), MELAA (0.1%), European (9%) and Other (0.2%). No students whose prioritised ethnicity is Pasifika can also report themselves as Māori under this ethnicity prioritisation system. The percentages for Asian, MELAA, European and Other using prioritised ethnic identification are interpreted in the same way. Both the 'prioritised' and 'all ethnicity' identifications are used in our subsequent analyses. All ethnicity of a student includes all the ethnicities they self-reported in the data.

Gender, in addition to ethnicity, is another demographic factor. It is represented by the dummy variable Female, which is used to capture gender differences among students. Female and male are the only gender identities recorded in the IDI. We therefore follow this way to measure students' gender. It can be seen from Table 3.2 that female (48%) students are slightly less likely to be represented in our sample. This situation remains true for the subsamples of students from each of the ethnic groups. Three dummies are generated to separate students in the sample into three cohorts, depending on the year in which they reached the age of 15. The dummies play the role of examining differences in university performance between the cohorts of students. Cohort 2010 students comprise 34% of the sample, while students in cohorts 2011 and 2012 each make up 33%. Cohort 2010 is set as the reference category.

Apart from the demographic factors, we are able to control for school characteristics, such as last school's decile. The decile of school last attended is used as a proxy for socioeconomic status. School deciles show the proportion of students from low-income communities in the

[8] Using all ethnicities, the sample is made up of 58.42% European, 19.57% Māori, 9.91% Asian, 9.59% Pasifika, 1.68% MELAA and 0.83% other ethnicity.

school.[9] It potentially reflects the socioeconomic condition of the students who attended the school. Decile 1 schools are in the bottom 10% of schools from the poorest socioeconomic regions, whereas decile 10 schools are in the top 10% of schools from the best socioeconomic areas. School deciles are used by the Ministry of Education to target the funding schools should receive, which is used for addressing the problems faced by students from lower socioeconomic background.

We create 11 dummy variables including, deciles 1 to 10 as well as an unknown decile. Decile 6 is used as the benchmark. According to Table 3.2, just 1.7% of the students' last school decile is unidentified (i.e., unknown), due to attending correspondence school or being home schooled.[10] Out of the sample, more than 38% of students previously attended upper decile schools (deciles 8 to 10) and less than 18% of them came from schools in the lower deciles (deciles 1 to 3). The data also reveals that Māori and Pasifika are severely underrepresented in schools with upper decile rankings. For instance, 56.1% of Asian and 46.3% of European attended schools in the top three deciles, compared to only 16.0% of Māori and 15.6% of Pasifika students from these upper decile rankings. Māori (37.9%) and Pasifika (48.5%) students are found to be more likely to attend lower decile schools compared to Asian (12.3%) and European (7.0%).

Other school characteristics that we can use as predictors include authority type (state or privately owned), gender composition (coeducation or single sex), and the area in which the school is located (urban or rural). State-owned schools, also known as public schools, are funded by the government; these schools provide free education to domestic students. Unlike state schools, private schools charge fees for the school term or year. Fees can cost as much as $20,000 NZD per year. In total, 81.4% of students across our cohorts attended state schools, while 5.1% were from private schools. Asian (7.4%) and European (6.7%) are six to seven times more likely to attend private schools relative to their Māori (1.2%) and Pasifika (1.5%) peers.

In New Zealand, most schools cater for both male and female students, which is why the majority of the students in our sample come from coeducational schools, regardless of their ethnicity. Māori and Pasifika students are found to have a higher likelihood of attending

[9] School deciles may change overtime as the population of various areas change. However, these changes are unlikely to affect the findings of this thesis that relate to the variables. For more information about how school deciles are calculated and used, see *Inquiry into decile funding in New Zealand State and integrated schools* by Donnelly (2003).

[10] Correspondence school, the former name of Te Kura, is a state-owned school providing distance education for all ages. For more information about correspondence school see https://www.tekura.school.nz.

coeducational schools and a lower likelihood of attending single-sex schools in comparison to their European and Asian counterparts. Among all the ethnic groups, Māori (4.6%) were most likely to have attended schools located in rural areas. Pasifika, however, attended rural schools at a lower proportion (0.9%), which is because the majority of Pasifika people reside within the Auckland region, which contains only a few rural schools.

Geographic region is another school factor that is taken into account. The schools last attended by 98.0% of the students in our cohorts are located in 16 different geographical regions and the rest of the 2.0% come from 'schools' of an unknown region. The schools locate in the unknown geographic region include the Correspondence School and home schooling. We create 17 dummy regional variables and set the Auckland region as the reference category. The top three most popular regions for the sample were Auckland (33.3%), Canterbury (11.9%) and Wellington (10.1%). This is not a surprising finding because Auckland and Wellington are the two largest cities in the North Island and Canterbury is the biggest city in the South Island. It is worth noting that the proportions of Pasifika (71.4%) and Asian (65.8%) students coming from the Auckland schools appears to be higher than the proportions of European (24.7%) and Māori (22.2%), due to the fact that the Pasifika and Asian populations are concentrated in the Auckland region. The mean proportions coming from Auckland schools for each of the ethnic minority groups (i.e., Pasifika, Asian, Māori) are statistically significantly different from the mean proportion for European at a 1% level, using t-test.

Secondary school academic achievement is believed to have a significant impact on university academic performance. School academic achievement is measured by the highest school qualification attained before leaving school. The National Certificate of Educational Achievement (NCEA) has been the primary secondary school qualification in New Zealand since it was introduced between 2002 and 2004.[11] The NCEA system has three certificates that can be awarded, from Levels 1 to 3. Students need to gain a required amount of credits at or above a specified threshold to pass each level. An NCEA Level 3 qualification can be used as University Entrance, depending on the requirement set by the universities. Also, NCEA certificates can be used to apply to universities and colleges in English-speaking countries such as the United States, Australia and the United Kingdom. In addition to NCEAs, other types of school qualifications offered by New Zealand schools include School Certificate, Other University Entrance, Sixth Form Certificate and University Bursary, which have been largely replaced by the NCEA system since its introduction. Besides these national school

[11] For more information on NCEA qualifications, see https://www.nzqa.govt.nz/ncea/.

qualifications, some schools, mainly private schools, also offer overseas qualifications such as Cambridge International Exams (CIE) and the International Baccalaureate (IB).

We generate seven dummy variables to capture the highest secondary school qualification gained. These dummies include students with NCEA Levels 1 to 3, overseas school qualifications (mainly CIE and IB), and other school qualifications as well as unknown school qualifications. The students who left school with no formal school qualifications are set as the benchmark group and accounted for 7.6% of the sample, according to Table 3.2. The statistics in Table 3.2 also show that NCEA Level 3 is the most common secondary school qualification and is held by 35.7% of the sample.

Attention should be drawn to the ethnic differences in these school qualifications. It is found that 14.9% of Māori and 8.0% of Pasifika students do not gain a formal qualification before leaving school. The same is true of 1.6% of Asian and 6.1% of European students. Moreover, Māori and Pasifika school students lag behind their Asian and European peers in attaining the highest NCEA qualifications. For example, between 19 to 27% of Māori and Pasifika completed NCEA Level 3, whereas the comparable percentages are approximately double for Asian (50.1%) and more than 1.5-times higher for European (40.1%).

Asian (7.6%) are the group found to be most likely to possess an overseas school qualification, while Māori and Pasifika (both no more than 1.5%) are least likely to have such qualifications. The proportion holding other school qualifications (including School Certificate, Other University Entrance, Sixth Form Certificate and University Bursary) for all ethnic groups of students varies between 10 and 14%. Ethnic differences in the highest school qualification attainment using all ethnicities definition are similar to the differences reported using prioritised ethnicity.

In terms of school history factors, changing school, not including transferring to home schooling, is one to be considered. Changing school here is defined as the number of times an individual switched schools during the duration of their schooling from age six. We use six dummy variables to capture the number of times. The control group is comprised of students who had never changed school. Switching school once is commonly found for nearly a half of the sample, which is primarily due to the transition from primary to secondary school. However, changing school more often might be related to negative family environments, such as parental divorce or financial pressures. Among all ethnic groups, Māori and Pasifika students switched schools more frequently than Asian and European. For instance, 6.2% of Pasifika

and 13.7% of Māori changed school at least four times, compared to only 2.9% and 4.6% of Asian and European, respectively. In addition to moving schools, temporary absence from school is the other school history factor that we can control for. Only 1.6% of the sample has been found truant, suspended and/or had an unjustified absent from school. There are found to be few differences in the means of temporary absence from school between ethnic groups.

We now consider factors that are related to parents. Single-parent household status was only measured at the time of the 2013 Census. Table 3.2 shows that approximately 20% of the sample came from a single-parent household. The proportion of students living in a one-parent household are higher for Māori (25.5%) and Pasifika (19.4%) when compared to European (18.6%) and Asian (13.8%).

We create a series of dummy variables that indicates the parents' highest qualification. The qualification is classified into eight categories: Level 1-3 Certificate, Level 4 Certificate, Level 5-6 Diploma, bachelor's degree or other Level 7 qualification, Postgraduate qualification (including postgraduate, honour's, master's and doctorate degree), an overseas qualification and unknown qualification. The control group is the students whose parents who had no qualifications. Students who have parents with no qualifications account for slightly over 8% of the sample. The equivalent number varies between the ethnicities. Pasifika students have the highest proportion of parents with no formal educational qualifications, at 14.2%. These figures are lower for others, at 13.1% for Māori, 7.8% for European and 5.7% for Asian students, respectively. In terms of obtaining a bachelor's degree or other Level 7 qualification, Asian are found to have the highest proportion of parents with a bachelor's degree or other level 7 qualification at 15.7%, compared to 13.5% for European, 6.4% for Māori, and 5.2% for Pasifika students. The proportion of parents whose highest qualification is a postgraduate qualification are relatively lower for Māori and Pasifika (both no more than 3%), compared to 9.6% for Asian and 8.4% among European students.

The highest proportion of days parents received a social welfare benefit during the period of ages six to sixteen of their child(ren) is also included in the regression analysis. This is a continuous variable. It can be seen in Table 3.2 that the benefit histories are available for 72.4% of the sample. The proportion with known benefit records are higher for European (77.8%) and Asian (74.0%), relative to Pasifika (67.1%) and Māori (58.8%). Conditional on data being available, the aggregated mean proportion of days receiving benefit payments is

4.8%. There is no considerable ethnic differences in the means of the variable (Māori with 4.7%, Pasifika with 4.9%, Asian with 4.8% and European with 4.9%).

Another parental factor is the highest number of charges parents were convicted for during ages six to sixteen of their child(ren). Again, we have this criminal history information available for 72.4% of the sample. Based on those students with available parents' conviction records, the overall average number of charges of parents is 0.39 times. The comparable numbers for Māori, Pasifika, Asian and European respectively are 0.44, 0.41, 0.36 and 0.38 times. It is important to note that those means are not in line with similar findings reported by the Ministry of Justice, which stated that Māori and Pasifika are about twice more likely to receive at least one conviction by age 38 than European and Asian for individuals born in 1978 (Ministry of Justice, 2019). This may be because there is incomplete data on the variable, especially for Māori and Pasifika students.

Household income refers to before-tax income received by a household from 1 March 2012 to 31 March 2013. It includes wages and salaries, self-employment income, property and rental income, dividends and investments, social insurance, superannuation, government assistance payments, and private transfers such as child support. Household income is been grouped into six income categories, there are zero, $1-$30,000, $30,001-$50,000, $50,001-$70,000, $70,001-$100,000 and $100,001 or more. We also create a dummy variable to reflect whether household income information is unknown. The proportion having zero household income is less than 1% for the sample over all the ethnic groups. In the sample, 6.1% have a household income of $1-$30,000 (considered as low income) and 15.5% have a household income between $30,001-$70,000 (considered as middle income), while a household income of $70,001-$100,000 and $100,001 or more (all considered as high income) account for 12.3 and 26.1% respectively. European students (5.0%) are the least likely to be from low-income households, while Asian (8.8%) have the highest proportion of low household income. Fourteen percent of the Māori and nearly 13.4% of Pasifika live in households earning a middle income. The percentage jumps to about 20% for Asian and 15.6% for European. The proportion of households receiving a high income for European (48.2%) is more than twice as high as the proportions of Māori and Pasifika (both 22.3%), and is about 1.5 times higher relative to the proportion of Asian (31.6%).

Table 3.2. Variable means and sample sizes for participation analysis[a]

Variables	All ethnicities	Māori	Pasifika	Asian	MELAA	European	Other
Dependent variable							
Participation	0.3157	0.1308	0.2190	0.5471	0.4252	0.3538	0.3333
Independent variables							
Female	0.5155	0.5147	0.5100	0.5216	0.5091	0.5154	0.5380
Cohort years							
Cohort 2010	0.3389	0.3386	0.3316	0.3318	0.3213	0.3415	0.3671
Cohort 2011	0.3324	0.3286	0.3383	0.3302	0.3289	0.3331	0.3481
Cohort 2012	0.3288	0.3328	0.3299	0.3380	0.3499	0.3254	0.2848
Ethnicities[b]							
Māori	0.2097	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Pasifika	0.0914	0.0543	1.0000	0.0000	0.0000	0.0000	0.0000
Asian	0.1030	0.0056	0.0215	1.0000	0.0000	0.0000	0.0000
MELAA	0.0176	0.0010	0.0011	0.0013	1.0000	0.0000	0.0000
European	0.5705	0.1861	0.0902	0.0563	0.0772	1.0000	0.1076
Other	0.0080	0.0026	0.0015	0.0013	0.0057	0.0000	1.0000
school deciles							
Decile 1	0.0470	0.0968	0.2500	0.0124	0.0076	0.0039	0.0316
Decile 2	0.0665	0.1535	0.1262	0.0395	0.0305	0.0313	0.0422
Decile 3	0.0657	0.1283	0.1087	0.0714	0.0667	0.0351	0.0422

Decile 4	0.0922	0.1276	0.1381	0.0950	0.1192	0.0704	0.0970
Decile 5	0.1019	0.1267	0.0847	0.0517	0.0705	0.1058	0.0865
Decile 6	0.1293	0.1188	0.0636	0.0695	0.1125	0.1549	0.1350
Decile 7	0.0990	0.0677	0.0684	0.0968	0.1087	0.1154	0.1055
Decile 8	0.1341	0.0747	0.0699	0.1336	0.1640	0.1655	0.1287
Decile 9	0.1134	0.0528	0.0499	0.1629	0.1354	0.1361	0.1224
Decile 10	0.1339	0.0329	0.0358	0.2644	0.1802	0.1617	0.1414
Decile unknown	0.0170	0.0201	0.0050	0.0024	0.0067	0.0200	0.0696
School authority							
State	0.8135	0.9098	0.8074	0.7971	0.8208	0.7825	0.7679
State integrated	0.1272	0.0735	0.1757	0.1282	0.1554	0.1383	0.1118
Private	0.0506	0.0121	0.0152	0.0737	0.0210	0.0668	0.0717
Unknown	0.0087	0.0046	0.0017	0.0008	0.0038	0.0124	0.0464
School gender							
Co-education	0.6976	0.7649	0.7288	0.6749	0.7073	0.6714	0.7131
Single sex	0.2916	0.2219	0.2683	0.3238	0.2888	0.3159	0.2405
Unknown	0.0108	0.0133	0.0029	0.0015	0.0038	0.0127	0.0464
School area							
Main Urban Area	0.7608	0.6607	0.9351	0.9390	0.9056	0.7333	0.7405
Minor Urban Area	0.1156	0.1746	0.0202	0.0293	0.0439	0.1274	0.0823
Secondary Urban Area	0.0777	0.0842	0.0273	0.0262	0.0410	0.0936	0.0886
Rural Area	0.0218	0.0458	0.0090	0.0026	0.0029	0.0189	0.0274

Unknown	0.0241	0.0345	0.0086	0.0028	0.0057	0.0267	0.0633
School region							
Northland	0.0362	0.0792	0.0077	0.0073	0.0133	0.0308	0.0359
Auckland	0.3330	0.2215	0.7136	0.6581	0.5624	0.2469	0.3586
Waikato	0.0971	0.1386	0.0343	0.0558	0.0953	0.0993	0.1160
Bay of Plenty	0.0647	0.1162	0.0147	0.0260	0.0276	0.0622	0.0422
Gisborne	0.0116	0.0340	0.0026	0.0013	0.0029	0.0069	0.0063
Hawkes Bay	0.0401	0.0657	0.0176	0.0104	0.0114	0.0408	0.0211
Taranaki	0.0253	0.0277	0.0044	0.0070	0.0153	0.0314	0.0190
Manawatu-Whanganui	0.0533	0.0786	0.0189	0.0213	0.0210	0.0565	0.0380
Wellington	0.1008	0.0866	0.1190	0.0911	0.1058	0.1044	0.1203
West Coast	0.0062	0.0048	0.0011	0.0007	0.0067	0.0084	0.0042
Canterbury	0.1190	0.0605	0.0396	0.0830	0.0925	0.1605	0.1203
Otago	0.0404	0.0205	0.0106	0.0203	0.0172	0.0569	0.0359
Southland	0.0214	0.0177	0.0046	0.0047	0.0076	0.0292	0.0084
Tasman	0.0091	0.0049	0.0009	0.0015	0.0038	0.0137	0.0042
Nelson	0.0130	0.0076	0.0028	0.0070	0.0048	0.0180	0.0063
Marlborough	0.0087	0.0079	0.0028	0.0023	0.0048	0.0112	0.0084
Unknown	0.0202	0.0281	0.0049	0.0020	0.0038	0.0229	0.0569
Highest school qualification							
No formal qualification	0.0764	0.1491	0.0800	0.0159	0.0381	0.0612	0.0781
NCEA Level 1	0.0713	0.1022	0.0759	0.0259	0.0524	0.0682	0.0612

NCEA Level 2	0.1562	0.1648	0.1751	0.0957	0.1420	0.1616	0.1414
NCEA Level 3	0.3572	0.1925	0.2668	0.5011	0.4414	0.4036	0.3650
Overseas qualification	0.0231	0.0072	0.0105	0.0755	0.0286	0.0211	0.0338
Other qualification	0.1120	0.1391	0.1159	0.1107	0.1106	0.1017	0.1097
Unknown qualification	0.2038	0.2451	0.2760	0.1754	0.1868	0.1826	0.2152
Switching school							
0	0.0956	0.0540	0.0875	0.1772	0.1192	0.0954	0.1899
1	0.4361	0.3585	0.4240	0.4148	0.4185	0.4714	0.4008
2	0.3001	0.3042	0.3140	0.2959	0.3127	0.2973	0.2616
3	0.1030	0.1461	0.1122	0.0828	0.1020	0.0895	0.0907
4 or more	0.0652	0.1372	0.0620	0.0291	0.0477	0.0464	0.0591
Truancy, suspension and/or unjustified absence	0.0159	0.0153	0.0147	0.0153	0.0143	0.0165	0.0190
Single-parent household	0.1956	0.2551	0.1942	0.1375	0.1783	0.1855	0.1688
Parents highest qualification							
No qualification	0.0830	0.1308	0.1421	0.0778	0.0887	0.0569	0.0696
Level 1-3 Certificate	0.1848	0.1968	0.1878	0.0506	0.0467	0.2096	0.1013
Level 4 Certificate	0.0904	0.0719	0.0442	0.0272	0.0429	0.1179	0.0633
Level 5-6 Diploma	0.0839	0.0559	0.0482	0.0369	0.0601	0.1092	0.0738
Bachelor’s degree or level 7 qualification	0.1144	0.0641	0.0523	0.1572	0.1239	0.1347	0.1224
Postgraduate	0.0673	0.0255	0.0213	0.0955	0.1068	0.0835	0.0865
Overseas school qualification	0.0752	0.0068	0.1174	0.2700	0.2650	0.0511	0.1857
Unknown qualification	0.3010	0.4484	0.3866	0.2849	0.2650	0.2372	0.3017

Household income							
Zero	0.0022	0.0014	0.0024	0.0093	0.0076	0.0010	0.0000
1 to 30,000	0.0606	0.0742	0.0633	0.0877	0.0753	0.0498	0.0696
30,001 to 50,000	0.0723	0.0698	0.0627	0.0991	0.0734	0.0699	0.0759
50,001 to 70,000	0.0822	0.0703	0.0712	0.0983	0.0763	0.0857	0.0823
70,001 to 100,000	0.1232	0.0909	0.0904	0.1247	0.1077	0.1406	0.1245
100,001 or more	0.2607	0.1317	0.1330	0.1914	0.2641	0.3411	0.2511
Unknown	0.3987	0.5618	0.5769	0.3899	0.3966	0.3119	0.3945
Known parents' benefit and crime histories	0.7238	0.5884	0.6705	0.7401	0.7712	0.7775	0.7300
Highest proportion of days living on benefit	0.0484	0.0473	0.0487	0.0477	0.0569	0.0485	0.0491
Highest number of convicted charges	0.3891	0.4353	0.3632	0.4055	0.4438	0.3770	0.2572
Number of observations (n)	179,034	37,536	16,359	18,435	3,147	102,135	1,422
Sample proportions	1.0000	0.2097	0.0914	0.1030	0.0176	0.5705	0.0079

[a] The values reported in this table were based on the numbers of observations that have been randomly rounded to base 3 or supressed if the original observation count is fewer than 6 due to the confidentiality requirements from Statistics New Zealand.

[b] The values in italics reported in this section show the proportions of students who self-identify with more than one ethnicity. For example, the proportion of students being prioritised as Māori (accounting for 20.79% of the sample) who also report as Pasifika, Asian, MELAA, European and Other are 5.4%, 0.6%, 0.1%, 18.6% and 0.3%, respectively. Those designated as Pasifika under prioritization who also self-report as Asian (2.2%), MELAA (0.1%), European (9%) and Other (0.2%). No Pasifika students can also report themselves as Māori under the ethnicity prioritization system. The values for the students being officially defined as Asian, MELAA, European and Other are interpreted in the same way.

3.1.4. Ethnic differences in university outcomes

The full set of differences in achievements in the four university outcomes between the ethnic groups using both prioritised and all ethnicity definitions are displayed in Table A.3.3 located in the Appendix of this thesis. For each university outcome, the first column is based on prioritised ethnicity. The first column shows that 31.57% of the students in the sample were enrolled in a bachelor's degree at one of the universities in New Zealand by age 21. European school leavers participate in university at a rate of 35.38%. The participation rates are comparatively lower for Māori (13.08%) and Pasifika (21.90%) but higher for Asian students (54.71%). In other words, the ethnic gaps in university participation between Māori, Pasifika or Asian and European are 22.30, 13.48 and -19.33 percentage points, respectively, as shown in Table 3.3.

Table 3.3. Overall ethnic differences in the university outcomes: two different ways of measuring ethnicity

Ethnic differences	Participation		First-year course completion		Second-year course completion		Qualification completion	
	Prioritized ethnicity	All Ethnicities	Prioritized ethnicity	All ethnicities	Prioritized ethnicity	All ethnicities	Prioritized ethnicity	All ethnicities
European – Māori	0.2230	0.2149	0.1314	0.1273	0.1243	0.1202	0.1915	0.1845
European – Pasifika	0.1348	0.1361	0.2607	0.2534	0.2485	0.2431	0.2725	0.2723
European – Asian	-0.1933	-0.1947	0.0201	0.0178	0.0221	0.0189	0.0177	0.0121

This table only reports the overall ethnic differences in the university outcomes. The detailed ethnic differences in the university outcomes are provided in Table A.3.3 in the Appendix.

The second column of each university outcome itemizes all of the self-reported ethnicities of these students. Therefore, those who self-identified with multiple ethnicities are counted in different ethnic groups. For example, a student who self-reported as Māori, Pasifika and European will be counted in all of these three ethnic groups. The sample size thus increases by 7.11% from 179,034 in the first column, to 191,766 in the second column of Table A.3.3 in the Appendix.[12] However, using all ethnicities does not change the sample size and mean participation rate of Māori because Māori are the first ranked ethnicity under the prioritisation

[12] In the sample, 166,854 (or 93.20%) of students reported one ethnicity, 11,631 (or 6.50%) reported two ethnicities and 552 (or 0.31%) reported three ethnicities.

system. Making use of multiple ethnicities, the European mean (34.57%) is still higher than the mean for Māori (13.08%) and Pasifika (20.96%) while below the Asian mean (54.04%). Thus, by using all ethnicities, the ethnic gap in university participation between Māori, Pasifika or Asian and European respectively are 21.49, 13.61 and –19.47 percentage points (see the second column of Table 3.3). It can be seen that using all ethnic designations slightly narrows the European-Māori gap but widens the European-Pasifika and European-Asian differences.

The third and fourth columns in Table 3.3 report differences in first-year course completion rates between the three ethnic minority groups and European, using the same measures of ethnicity as mentioned above. Making use of prioritised ethnicity, we find that the average first-year course completion rates for Māori, Pasifika and Asian are 13.14, 26.07 and 2.01 percentage points, respectively, lower, relative to European. When focusing on the students reported multiple ethnicities, the sample size increases from 56,223, by 5.49%, to 59,307. The percentage change (5.49%) in the sample size is smaller than the percentage increase (7.11%) in sample size with all ethnicities in the sample of school leavers who could have enrolled at university, which suggests that those who did enrol at university were generally less likely to report multiple ethnicities. There are minor changes on the ethnic differences. Māori, Pasifika and Asian pass their first-year courses at a rate that is 12.73, 25.34 and 1.78 percentage points, respectively, below European.

Similar ethnic differences in second-year course completion rates are provided in the fifth and sixth columns of Table 3.3. The average completion rates for all the ethnic groups improved compared to the rates in first year. There are two possible reasons. First, the majority of the lower-performing first-year students are ‘filtered out’ because they are more likely to drop out of university. Secondly, second-year students may take their academic achievement more seriously than first-year students because their goal is clearer (Zaitseva, Milsom & Stewart, 2013). In spite of these improvements, Māori, Pasifika and Asian continue to have a second-year course pass rate that respectively is 12.43, 24.85 and 2.21 percentage points below European, using prioritised ethnicity.

The gaps between Māori or Pasifika and European in second-year course completions decline slightly compared to the gaps in first-year course completions, while the European-Asian gap become slightly wider in second year than that in first year. If we look at all ethnicities, the sample size increase by 5.11% from 31,722 to 33,342. Using all ethnicities, European-Māori, European-Pasifika and European-Asian gaps change a little to 12.02, 24.31 and 1.89 percentage points, respectively. It is also important to note that, based on all ethnicities, the

gaps between Māori or Pasifika and European are smaller, relative to the gaps in first-year course completions. However, the second-year European-Asian gap become a little wider.

The overall ethnic differences in degree completion using prioritised and all ethnicities are shown in the last two columns of Table 3.3. Prioritised Māori, Pasifika and Asian are 19.15, 27.25 and 1.77 percentage points, respectively, less likely than European to complete a degree qualification. It is important to note that these Māori/Pasifika-European degree completion gaps are relatively larger than similar ethnic gaps in both university participation and course completions, implying that dropout behaviour is a particular issue for Māori and Pasifika students. The same is not true for the Asian-European gaps. There is a 5.55% increase in the sample size from 15,843 to 16,722, if we focus on the students who reported being of more than one ethnicity. Using this alternative ethnic measure, we see that the ethnic gaps shrinks slightly. In particular, European-Māori, European-Pasifika and European-Asian gaps respectively are 18.45, 27.23 and 1.21 percentage points.

Overall, the ethnic gaps in university outcomes are similar when using the different ethnic identification schemes. Māori school leavers are least likely to attend degree-level study at university among the ethnic groups regardless of the measures of ethnicity. Pasifika students have the poorest performance once at university in both course and qualification completions using either prioritised or all ethnicities.

3.2. Methodology

Maximum Likelihood Probit is used to estimate the marginal effects of the relevant explanatory variables on participation and degree completion, which are binary variables. The basic Probit model developed from linear model can be expressed as:

$$Z_i = X_i\beta + \varepsilon_i \quad (3)$$

where ε_i is a random error term assuming to have a standard normal distribution. X_i is a vector of the explanatory variables included in the regression analysis, while β is a vector of coefficients. The equation (3) above is also known as a latent variable equation, because a given university outcome, for example participation, denoted as y_i is linked to the latent variable Z_i by the following relationship:

$$y_i = \begin{cases} 1 \text{ if } Z_i \geq 0 \\ 0 \text{ if } Z_i < 0 \end{cases} \quad (4)$$

the condition probability of participation, y_i, is equal to 1, given the explanatory variables taken into account, and is expressed as:

$$P(y_i = 1|X_i) = F(X_i\delta) \quad (5)$$

where F(·) is the cumulative distribution function of the standard normal.

Unlike participation and degree completion, first- and second-year course completion rates are continuous proportional dependent variables that take values from 0 to 1 inclusive. Given that, Fractional Probit, which was developed by Papke and Wooldridge (1996), is more appropriate than standard Probit when calculating the marginal effects of the associated explanatory variables on course completion rates. The mean proportion of course completion, c_i , is conditional on the explanatory variables, which can be expressed as:

$$E(c_i|v_i) = \Phi(v_i\theta) \quad (6)$$

where v_i is a vector of the explanatory variables considered and θ is a vector of the coefficients to be estimated. $\Phi(\cdot)$ is the standard normal cumulative density function, which is restricted in [0,1]. This ensures the predicted values of the proportion of course completion fall into [0,1].

Probit and Fractional Probit regression analyses indicate how much of the sizeable ethnic differences in the university outcomes can be closed when controlling for the relevant explanatory variables. We next want to measure to what extent these ethnic differences can be explained by measurable differences in the considered factors. For example, will the lower university participation rates for Māori and Pasifika relative to European be accounted for by the differences in school deciles, school qualification and parental education background? This can be done through using formal decomposition methods.

Decomposition technique was first introduced by the well-known economist Ronald Oaxaca in 1973. The technique was used to explain the gaps in the mean wage rates between male and female workers in the labour market (see Oaxaca, 1973). In the same year, but after the publication of Oaxaca's decomposition paper, an almost identical decomposition technique was employed by Blinder (1973) to explain the average wage differences between race (e.g., blacks and whites) and gender (e.g., male and female). Since then, the so-called Blinder-

Oaxaca linear decomposition has been widely used in the study of labour market discrimination, as well as in studies of other forms of discrimination.

The Oaxaca (linear) decomposition method is chosen to analyse the ethnic gaps in first- and second-year course completion, due to the linearity quality of the two outcome variables. Using this method allows us to estimate how much of these ethnic differences in first- and second-year course completions could be explained if a minority group was given the same covariates of European. An example of the Oaxaca decomposition for the gap in the proportions of first-year course completions between Māori and European can be written as:

$$\overline{C}^{E} - \overline{C}^{M} = \left[b^{M}\left(\overline{X}^{E} - \overline{X}^{M}\right)\right] + \left[\overline{X}^{E}(b^{E} - b^{M})\right] \quad (7)$$

An equivalent expression can be written as:

$$\overline{C}^{E} - \overline{C}^{M} = \left[b^{E}\left(\overline{X}^{E} - \overline{X}^{M}\right)\right] + \left[\overline{X}^{M}(b^{E} - b^{M})\right] \quad (8)$$

Where $\overline{C}^{E}$ in Equation (7) is the mean proportions of first-year course completion for European students and $\overline{C}^{M}$ in Equation (8) is the mean proportion for Māori students. b^{E} and b^{M} are a vector of estimated coefficients for European and Māori, respectively.

The first term on the right-hand side of Equations (7) and (8) reflects the ethnic difference in first-year course completion between European and Māori students because of differences in the observable factors (the independent variables included). The term is called an 'explained' component that can be weighted in three different ways. First, it can be weighted by the Māori coefficient estimates, b^{M}, as shown in Equation (7). Second, it can be weighted by the European coefficient estimates, b^{E}, as reflected in Equation (8). The third possibility is to weight the term by using coefficient estimates from a pooled sample of European and Māori. The third weighting method is chosen for this thesis, primarily because it simultaneously takes into account both European and Māori characteristics (see Neumark 1988; Oaxaca and Ransom 1994 for more discussion on this approach).

The second term on the right-hand side of Equations (7) and (8) provides the ethnic difference between European and Māori students in the first-year university outcome, which is attributable to differences in the estimated coefficients of the observed characteristics, also known as the 'unexplained' component. The 'unexplained' component in Equation (7) is

weighted by the European's explanatory variables (X^E), while the explanatory variables of Māori (X^M) have been used to weight the component in Equation (8). The 'unexplained' term also captures the ethnic difference due to group differences in factors that are unobserved in our regression analysis.

The 'explained' component measures the ethnic gap in the academic outcome between European and Māori, which can be explained by all the independent variables together. Apart from that, the decomposition method also identifies how much of the ethnic gap can be contributed to a specific individual or group of independent variables. To do that, we group the independent variables into categories. In this way, we will be able to understand the relative importance of each of the category variables in explaining the gap between the two ethnic groups. The same linear decomposition approach is used to explain the gaps between Pasifika/Asian and European, and is repeated for second-year course completion.

Fairlie (1999, 2005) extended the Oaxaca (linear) decomposition method to non-linear models; in other words, models with a binary independent variable. Given that university participation and qualification completion are both binary outcome variables, Fairlie nonlinear decomposition technique was adopted to explain the ethnic gaps in these two university outcomes. For example, if we want to decompose the mean values of university participation between European and Māori, the decomposition equation can be expressed as:

$$\overline{O}^E - \overline{O}^M = \left[\sum_{i=1}^{N^E} \frac{F\left(X_i^E b^E\right)}{N^E} - \sum_{i=1}^{N^M} \frac{F\left(X_i^M b^E\right)}{N^M}\right] + \left[\sum_{i=1}^{N^M} \frac{F\left(X_i^M b^E\right)}{N^M} - \sum_{i=1}^{N^M} \frac{F\left(X_i^M b^M\right)}{N^M}\right] \quad (9)$$

The above decomposition equation can be alternatively written as:

$$\overline{O}^E - \overline{O}^M = \left[\sum_{i=1}^{N^E} \frac{F\left(X_i^E b^M\right)}{N^E} - \sum_{i=1}^{N^M} \frac{F\left(X_i^M b^M\right)}{N^M}\right] + \left[\sum_{i=1}^{N^E} \frac{F\left(X_i^E b^E\right)}{N^E} - \sum_{i=1}^{N^E} \frac{F\left(X_i^E b^M\right)}{N^E}\right] \quad (10)$$

Where $\overline{O}^E$ and $\overline{O}^M$ in Equations (9) and (10) are the mean probabilities of university participation for European and Māori students, respectively. N^E is the sample size for European students and N^M is the sample size for Māori students. b^E and b^M are a vector of the estimated coefficients for European and Māori, respectively. F(·) is the cumulative distribution function of the standard normal distribution.

The first term on the right-hand side of Equations (9) and (10) is known as an 'explained' component. It can also be weighted in the three different coefficients as discussed above. Again, we use the coefficient estimates from the pooled regression of European and Māori to weight this component. The second term on the right-hand side of the two equations is the 'unexplained' component. The term in Equation (9) is weighted by the Māori's explanatory variables X^M, while the term in Equation (10) is weight by the explanatory variables of European X^E. The nonlinear decomposition analysis is repeated for Pasifika-European, Asian-European and qualification completion.

Chapter 4. Empirical results

4.1. Participation in a bachelor's degree by age 21

4.1.1. Probit regression analysis

The complete set of regression results on maximum Likelihood Probit analysis of participation using both prioritised and all ethnicities are provided in Table A.4.1 in the Appendix of this thesis. The table reports the estimated coefficients, along with their standard errors. Because these estimated coefficients in this non-linear regression are difficult to directly interpret, we also report the average marginal effects from this Maximum Likelihood estimation in the same table.[13] The regression results from adopting prioritised ethnicity are fairly similar to those when using all ethnicities. For the rest of this thesis, discussions on analytical results are based on prioritised ethnicity unless otherwise stated.

Most of the estimated coefficients are statistically different from zero at the conventional test levels and have the expected signs. For example, female students are on average 5.50 percentage points more likely to participate in a bachelor's degree by age 21 than male students, regardless of the ethnicity designations and other factors that have been held constant in the regression. These effects are statistically significant at a 1% level. Compared to the students in cohort 2010, cohorts 2011 and 2012 are respectively 1.31 and 1.24 percentage points more likely to enrol in a bachelor's qualification, no matter which ethnicity designation is used. These effects are both highly statistically significant.

It has been commonly found in past literature that school deciles play a critical role in determining university success (e.g., Juhong and Maloney, 2006; Shulruf, Hattie & Tumen, 2008). Our results provide similar findings. As can be seen in Table A.4.1, the marginal effects of all the deciles of the last school attended are highly statistically significant, except that decile 4 is statistically significant at a 10% level and decile 7 is statistically significant at a 5% level, using either way of measuring ethnic identity. Coming from schools with deciles 1 to 5 negatively affects the probability of participation relative to attending decile 6 schools. The probability of participation significantly increases as decile ranking (from decile 7) increases. For instance, compared to coming from decile 6 schools, attending schools in decile 7 slightly increases the likelihood of participation by 0.82 percentage points, while attending schools in decile 10 lifts the likelihood by more than sixfold, to 5.24 percentage points. It should be noted

[13] Average marginal effect for a variable such as being female is the average of the estimated marginal effects of a study sample when the variable changes from zero to one, while holding all other independent variables constant.

that the results on school deciles also indicates that the probability of enrolling in bachelor's study increases by nearly 13.00 percentage points, on average, in moving from a decile 1 to decile 10 school. These decile impacts may not be due to the school characteristics in the deciles per se, but are related more to underlying factors, such as household and community attributes, that allow students to attend schools in different deciles.

Attending a state-integrated school results in an almost 1 percentage point statistically significantly more likely to enrol in a university bachelor's qualification, compared to the omitted variable of being from a state school. Choosing private schools increases the probability of attending university by more than three times to 3.54 percentage points, and this effect is statistically significant at a 1% level. Students who come from single-sex schools tend to be 3.50 percentage points statistically significantly more likely to move into university than those attend coeducated schools, in terms of participation performance.

In comparison to attending urban schools, studying in schools located in secondary urban areas slightly lower the probability (by 0.72 percentage points) of enrolling in a bachelor's qualification. This marginal effect is statistically significant at a 5% level. Being from minor urban and rural schools appears to have no measurable effects on the probability of enrolment, as the estimated effects are statistically insignificant.

Our regression results show that a school's geographical region matters for the students' probability of university participation. We find that students coming from schools that are not in the Auckland region have statistically significantly lower probabilities of enrolling in a bachelor's degree at university. For instance, Wellington and Canterbury school students are 3.60 percentage points and 4.22 percentage points, respectively, less likely than Auckland school students to participate in university. Schooling within the Tasman and Marlborough regions reduces the probability by 6.43 to 7.53 percentage points, while attending schools that were not assigned to any specific region (i.e., correspondence school and home education) lowers the probability by about 26.83 percentage points.

Students' academic performance in high school is expected to be one of the main factors influencing their university participation. We used a series of dummies to capture the impact of the highest school qualification on the bachelor's degree qualification enrolment. The results in Table A.4.1 indicate that the propensity of participation increases as the highest level of NCEA qualification increases, and that having an NCEA Level 3 as the highest school qualification leads to the highest propensity of participation among the types of school qualifications. In particular, compared to those with no formal school qualifications, having an

NCEA Level 1 as the qualification increases the probability of participation by 11.00 percentage points, while obtaining a NCEA Level 2 raises the probability by about twice to slightly more than 20 percentage points, and receiving an NCEA Level 3 further increases the probability by around fivefold to about 53 percentage points. Holding other school and overseas qualifications (e.g., Cambridge International Exams (CIE) and the International Baccalaureate (IB)) lift the likelihood of studying a bachelor's programme by approximately 33 and 46 percentage points, respectively. All of these marginal effects are statistically significant at a 1% level.

Switching schools, and temporary absence from school, are other school factors we included in the participation analysis. Relative to the omitted category of never switching schools, switching schools once mainly because of the shift from primary to secondary school positively and statistically significantly affects the probability of enrolling in bachelor-level study. However, switching schools more than once has an adverse impact on the participation probability and the more frequent the switching of schools the more severe the adverse impact. All of the effects are statistically significant except for switching schools twice. Truancy, suspension and/or unjustified absence do not have statistically significant impacts on the likelihood of attending university.

Our results suggest that parental factors do influence children's university participation. For example, those students who lived in a single-parent household at the time of the 2013 Census have a statistically significantly lower probability of participation, although this effect is not large in magnitude (-0.48 percentage points). However, the effect of parents' highest educational qualification on participation appears to be much stronger. An individual having either parent whose highest qualification is Level 1 to 3 certificate has a probability of participating in university that is 3.23 percentage-points higher than someone whose parents have no educational qualifications. The impact, however, drops slightly to about 2.75 percentage points when having either parent with a Level 4 certificate. This may be because parents with a Level 4 certificate as the highest qualification tend to be technical or skills-based workers (i.e., builders or plumbers), who do not prepare their children for higher education. The impact then increases to more than 5.06 percentage points when either parent has a Level 5 to 6 diploma, 7.23 percentage points when either parent has a bachelor's degree or equivalent, and 8.70 percentage points when at least one parent has a postgraduate qualification (including postgraduate, honours, master's and doctorate degrees). For the students who have parents with an overseas qualification, their probability of enrolling at university is 5.39 percentage-points higher than the similar probability for students whose parents have no educational qualifications.

Students living in households with a $30,001 to $50,000 income were used as the reference group when examining how household earnings contributes to students' university participation. Relative to the reference group, coming from households with a zero to $100,000 income has no measurable effects on the university participation. However, having a household income of $100,001 or more increases the probability of university attendance by 1.60 percentage points, relative to having a household income of $30,001 to $50,000. This effect is statistically significant at a 1% level. The proportion of days spent by parents reliant on benefit support, or the number of convictions, are found to have no statistically significant impact on their children's participation in university.

Now consider the estimated marginal effects on the ethnic minority dummy variables for Māori, Pasifika and Asian, which are highlighted in Table 4.1. Because the students who ethnically identified as MELAA or Other will be excluded in our subsequent decomposing analysis, the marginal effects of being from these two ethnic minority groups are therefore not discussed in this chapter.

Table 4.1. Key results on ethnicity from the Probit regression analysis for participation in a bachelor's degree with a New Zealand university by age 21

	Prioritised ethnicity			All ethnicities		
Variables	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Ethnicities						
Māori	-0.2845***	0.0136	-0.0503***	-0.2863***	0.0134	-0.0507***
Pasifika	-0.0311*	0.0185	-0.0055*	-0.0414**	0.0174	-0.0073**
Asian	0.3309***	0.0147	0.0586***	0.3277***	0.0144	0.0580***
MELAA	0.1530***	0.0304	0.0271***	0.1411***	0.0300	0.0250***
Other	0.0369	0.0464	0.0065	0.0506	0.0446	0.0090
Number of observations (n)	179,034			191,766		
Likelihood function	-56617.074			-56612.861		
Pseudo R^2	0.4929			0.4929		

This table only reports the results of identifying with ethnic minority groups for the university outcome. The full set of regression results are provided in Table A.4.1 in the Appendix, along with the full set of covariates that are held constant in this regression analysis.

* denotes statistical significance at a 10% level
** denotes statistical significance at a 5% level
*** denotes statistical significance at a 1% level

It can be seen in Table 3.3 that Māori university participation rates on average were 22.30 and 21.49 percentage points, respectively, lower than that of European using prioritised and all ethnicities. The estimated coefficients of marginal effects for Māori suggest that Māori students are 5.03 and 5.07 percentage points, respectively, less likely to attend university than European using prioritised and all ethnicities once other measurable factors are held constant. These coefficients indicate that there are 5.03 percentage points of the prioritised European-Māori gap and 5.07 percentage points of the all ethnicities European-Māori gap that remained unexplained. This means that keeping constant all the personal characteristics, school factors, prior performance and parental factors, can eliminate 77.44% of the gap in participation between prioritised European and Māori, and 76.41% of the gap between European and Māori if using all ethnicities. These remaining unexplained ethnic gaps are statistically significant at a 1% level. A short summary for these effects will be that more than three-quarters of the gap in university participation between Māori and European are eliminated by controlling for other measurable factors.

It has been shown that Pasifika university participation rates were 13.48 and 13.61 percentage points, respectively, below the rate of European using prioritised and all ethnicities. The estimated marginal effect for being Pasifika under ethnicity prioritisation indicates that the unexplained difference declines to 0.55 percentage points, closing approximately 96% of the differences between prioritised European and Pasifika students in the participation outcome. The remaining ethnic difference is statistically significant at a 10% level. If using all ethnicities, the unexplained difference decreases to 0.73 percentage points, which eliminates nearly 95% of the gap in participation between all ethnicities European and Pasifika students. The remaining ethnic difference is, however, statistically significant at a 5% level.

Asian students were 19.33 and 19.47 percentage points, respectively, more likely to attend university than their equivalent European peers using both prioritised and all ethnicities. The unexplained difference reduces to 5.86 percentage points when using prioritised ethnicity, and to 5.80 percentage points when using all ethnicities, once the categories of observed factors are kept constant, which eliminates about 70% of the initial differences in participation between prioritised and all ethnicities European and Asian students. The ethnic differences remaining are both statistically significant at a 1% level.

4.1.2. Decomposition analysis

The Probit regression analysis above provides an understanding of the impacts of different factors on university participation, but does not identify how important a factor is in accounting for the overall gaps in the mean university participation rates between two ethnic groups. Formal decomposition analysis allows us to estimate the separate contributions of individual or groups of factors in the gaps between the ethnic minority groups and European. For example, this lets us to say how much of the differences in mean participation rates between European and Māori could be eliminated if the ethnic minority group had the same average covariates of European.

Table A.4.2 in the Appendix provides the full set of results from the nonlinear (Fairlie) decomposition on the gaps in participation between Māori, Pasifika and Asian ethnic minority groups and European. The decomposition analysis for one ethnic minority group (e.g., Māori) and the European group, started with running a pooled regression for these two subsamples of students. The estimated coefficients from the pooled regression were then used in the decomposition analysis. The process was repeated for Pasifika-European and Asian-European comparisons. Fairlie (2005) suggested that a random sample must be drawn from the majority ethnic group to match the sample of the ethnic minority group when conducting decomposition, given that the sample size the majority ethnic group is usually larger than the sample size of the ethnic minority group. The decomposition results reported in this thesis are based on 100 replications of this process. More discussion about this decomposition technique can be found in Section 3.2.

Table 4.2. Key results on ethnicity from the nonlinear decomposition of participation in a bachelor's degree with a New Zealand university by age 21

	Prioritised ethnicity			All ethnicities		
	European vs Māori	European vs Pasifika	European vs Asian	European vs Māori	European vs Pasifika	European vs Asian
Total ethnic difference	0.2230	0.1348	-0.1933	0.2149	0.1361	-0.1947
Total explained differences	0.1911*** (0.0009) [85.70%]	0.1310*** (0.0020) [97.18%]	-0.1242*** (0.0020) [64.22%]	0.1921*** (0.0009) [89.39%]	0.1305*** (0.0017) [95.89%]	-0.1264*** (0.0019) [64.92%]
Unexplained differences	0.0319 [14.30%]	0.0038 [2.82%]	-0.0692 [35.78%]	0.0228 [10.61%]	0.0056 [4.11%]	-0.0683 [35.08%]
Number of observations (n)	139,671	118,494	120,567	149,565	130,422	131,025

This table only reports the overall decomposition results in the university outcome. The full set of decomposition results are provided in Table A.4.2 in the Appendix.

* denotes statistical significance at a 10% level
** denotes statistical significance at a 5% level
*** denotes statistical significance at a 1% level

4.1.2.1. European–Māori gap

As shown earlier, prioritised Māori were, on average, 22.30 percentage points less likely than prioritised European to participate in university. The separate and total contributions of the eight individual and categories of variables in explaining the prioritised European-Māori gap are shown in the first column of Table A.4.2. The other school factors include school authority, gender composition, area and region, and the number of times switching schools, as well as truancy, suspension and/or unjustified absence. Single-parent households, parents' benefit history and crime histories are grouped into other parental factors. The results in Table 4.2 emphasize hat if Māori had the same personal characteristics, school factors, highest school qualification and parental factors as that of European, we could explain 19.11 percentage points or 85.7% of the overall European-Māori gap in university participation. This result is statistically significant at a 1% level. The result also indicates that Māori are still 3.19 percentage points less likely than identical European to undertake bachelor-level study at university.

As shown in Table A.4.2 in the Appendix, the most important factors in explaining the gap in probability of participation between European and Māori from the largest to the smallest are the highest school qualification obtained (12.80 percentage points 57.40%), school decile (3.09 percentage points or 13.86%), parents' highest educational qualification (1.66 percentage points or 7.44%), other school factors (1.57 percentage points or 7.04%) and household income (0.47 percentage points or 2.11%). However, if Māori had the same female representation as European students, this would widen the gap in participation between European and Māori by 0.01 percentage points, or 0.04%. Gave Māori the same female representation, cohort year and other parental factors such as European would widen the participation gap between these two ethnic groups. In other words, eliminating the three factors together widens the European-Māori gap by 0.48 percentage points or 2.14%.

We repeated the decomposition analysis for students who self-identified with multiple ethnicities. These decomposition results are reported in the fourth column in Table 4.2. It can be seen that decomposition results are generally similar to the previous decomposition results found based on prioritised ethnic identify. The gap between European and Māori students in

participation was 21.49 percentage points when using all ethnicities. If Māori had the same measurable factors as European, this could reduce the ethnic gap in university attendance by 19.21 percentage points, or approximately 90%. The highest school qualification performance, school decile ranking, parental educational attainment, other school factors and household income are still the most important explanatory factors, and in that order. Gave Māori and European the same female representation, cohort year and other parental factors would increase this ethnic gap university participation.

4.1.2.2. European–Pasifika gap

The second column of Table 4.2 reports the similar decomposition results between prioritised European and Pasifika students. It has previously been shown that prioritised Pasifika had a participation rate that was 13.48 percentage points lower than the rate for prioritised European. These indicate that eliminating the differences in all observed factors can account for 13.10 percentage points of the gap between European and Pasifika. This means that if we gave Pasifika the same measurable covariates as European, we could strikingly eliminate 97.18% of this ethnic gap in the probability of university participation.

The factor that explains the largest proportion of the European-Pasifika gap is the highest school qualification achievement (8.31 percentage points or 61.65%). The next most important factor is school decile (3.96 percentage points or 29.38%), followed by parental highest education attainment (1.57 percentage points or 11.65%) and household income (0.45 percentage points or 3.34%). If Pasifika and European had the same female representation, this could lead the gap in participation between the ethnic groups to statistically significantly rise by 0.16 percentage points, or 1.19%. Eliminating the differences in female representation, cohort year, other school factors and other parental factors between European and Pasifika widens the ethnic gap by 1.19 percentage points, or 8.83%.

Using all ethnic identities, the European-Pasifika gap in enrolling in a bachelor's qualification slightly increases to 13.61 percentage points. As shown in the fifth column of Table 4.2, eliminating the ethnic differences in all observable categories of factors between Pasifika and European could explain 13.05 percentage points or 95.89% of the European-Pasifika gap. The most import factors explaining the ethnic gap, in order, are the highest school qualification achievement, school decile, parental education and household income. Removing the ethnic differences in female representation, cohort year, other school factors and other parental related factors would make the gap between the ethnic groups wider.

4.1.2.3. European–Asian gap

Similar decomposition results in participation between prioritised European and Asian are displayed in the third column of Table 4.2. It has been noted previously that Asian are the ethnic minority group whose participation rate was 19.34 percentage points higher than the participation rate of European. The decomposition results for European and Asian are a little different to the results discussed above. Eliminating the ethnic differences in all the observable characteristics can account for 12.42 percentage points of the total gap in participation between European and Asian, suggesting that if we gave Asian and European the same individual characteristics, school factors, highest school qualifications and parental factors we could explain 64.22% of this ethnic gap. This will leave Asian with a university participation rate that is still 6.92 percentage points higher than otherwise observationally equivalent for European.

The highest school qualification (8.66 percentage points or 44.78%), other school factors (2.74 percentage points or 14.71%), parents' highest education attainment (0.99 percentage points or 5.12%), school decile (0.46 percentage points or 2.38%) and cohort year (0.02 percentage points or 0.10%) are the most important factors, in that order, in explaining the European-Asian gap. Given Asian students the same female representation as that of European could result in an 0.17 percentage points, or an 0.88% increase in the gap in participation between European and Asian. Removing the differences in female representation and household income could jointly widen the European-Asian gap by 0.42 percentage points, or 2.17%.

The gap in undertaking bachelor-level study between European and Asian students was 19.47 percentage points when using all ethnicities. Given Asian the same measured characteristics of European would reduce the gap in participation outcome between European and Asian by 12.64 percentage points, or 64.92% (see the sixth column in Table 4.2). The highest school qualification received, other school factors, parental highest qualifications, school decile and cohort year remain the most important factors in explaining the European-Asian gap. If Asian and European were given the same female representation and household income, as well as other parental factors, this would increase the gap between these ethnic groups.

4.2. First-year course completion

4.2.1. Probit regression analysis

Based on the sample of 56,223 students who had enrolled in a bachelor's qualification at a New Zealand university by the age of 21, we wanted to exam the ethnic differences in first-year course completion performance. This regression analysis used almost the same independent variables as from the previous analysis. The only changes were that the school region information was replaced by specific university in which the individual studied, along with new variables including age, employment hours, activity prior to university enrolment, university studied at, the country of citizenship and the nature of enrolment (i.e., part-time) have been added (see Table 3.1 for the variables used).

Table A.4.3 in the Appendix reports the complete set of regression results including estimated coefficients, standard errors and mean marginal effects from this Maximum Likelihood estimation for first-year course completion, using both prioritised and all ethnic identities. These results are quite comparable. What is important to note is that the current Pseudo R^2 statistic (0.0820) reported is substantially smaller than the R^2 reported in the previous regression analysis (0.4929), which suggests that the control variables used in the first-year course completion regression overall have relatively a much lower predictive power on the university outcome of successfully completing courses once an individual is enrolled at university.

It was found that female students were not just more likely to participate in university, they also perform better in terms of first-year course completion than male learners. For example, being female increases the course pass rate by 5.50 percentage points. This result is statistically significant at a 1% level. Students in the later cohorts tend to have a higher level of course completion in their first year. Particularly, students in the cohort 2011 have a completion rate that is about 1 percentage point higher compared to the omitted group of cohort 2010 students. The effect increases to 1.41 percentage points for the cohort 2012 students. These findings are all highly statistically significant.

Previous studies have found that higher school deciles have a strong positive influence on first-year GPA performance (e.g., Juhong and Maloney, 2006; Sopoaga et al., 2013) and persistence to second-year study (Jia and Maloney, 2015) in tertiary education. Our results also confirm the positive impact of school deciles on first-year course completion achievement. In particular, compared to the benchmark group of last attending a school in the 6^{th} decile,

studying in lower decile schools leads to poorer first-year university performance. For example, participating in a decile 5 school is associated with an average decrease of 1.36 percentage points in first-year course completion, while coming from a decile 1 school lowers the course completion rate by more than 12 percentage points. These effects are all statistically significant at a 1% level. Being deciles 7-8 school students has no statistically significant impact on the course completion performance. However, studying in schools in the 9^{th} or 10^{th} deciles lifts the course completion rate by 1.12 and 1.16 percentage points, respectively. These effects are statistically significant at a 5% level. Also, first-year course pass rates increase by more than 13 percentage points if attending a decile 10 compared to a decile 1 school, irrespective of ethnicity and other the measurable factors included in this regression.

We previously found that attending a private school resulted in a statistically significantly higher probability of university participation. The results in Table A.4.3, however, suggest that studying in state-owned schools tends to result in a higher rate of first-year course completion. For instance, coming from a state integrated school statistically significantly lowers the course completion rate by just over 1 percentage point, compared to attending a state-owned school. Choosing to attend private schools has no measurable impact on the course completion outcome. Although being a student in a single-gender school was formerly found to have a positive and statistically significant impact on the propensity to undertake bachelor-level study, school gender composition now has no statistically significant effect on the first-year course completion rate. Coming from a school in a minor urban area increases the rate of course completion by 2.11 percentage points, compared to attending schools located in a main urban area. Attending secondary urban area or rural area schools has no statistically significant impact on university achievement.

Students' school academic performance has always been of significant relevance to first-year academic success at university (e.g., Ferrão & Almeida, 2019; van Herpen et al., 2017; Tickell and Smyrnios, 2005). According to our results, it should not be surprising then that obtaining NCEA Level 1 and 2 as the highest qualification has a statistically insignificant impact on first-year course completion outcome. This is because NCEA Level 1 and 2 are the two lower school qualifications and are usually done at Years 11 (ages 15-16) and 12 (ages 16-17), so are deemed to be less relevant to university outcomes beyond participation. Although NCEA Level 3 was previously found to have a relatively stronger impact than overseas qualifications such as Cambridge International Exams (CIE) and the International Baccalaureate (IB) in terms of university participation, the relationship is reversed when it comes to first-year course completions. For example, compared to leaving school without a qualification, completing

NCEA Level 3 raises the rates of a pass for first-year courses by 10.28 percentage points, while owning an overseas qualification leads to a 16.50 percentage-point increase in the rate. These results are statistically significant at a 5 and 1% level, respectively. Having other school qualifications does not statistically significantly influence the course completion performance.

The two patterns that can be seen from our results are that students who never switched schools are likely to achieve better first-year course completion outcome, which more frequent switchers tend to have poorer performance in this completion outcome. Switching schools once has no effect on the course completion. However, changing school twice statistically significantly lowers the course completion rate by 1.59 percentage points. This negative effect increases to 3.64 and 7.89 percentage points, respectively, for switching school three times and four times or more. These results are all statistically significant at a 1% level. The marginal effect of truancy, suspension and/or unjustified absence on the course completion is statistically insignificant and small in magnitude.

Age in the first year enrolled at university is an additional independent variable that has been considered in this completion analysis. A series of dummy variables were created to indicate single ages from 18 to 21 and for two age groups (under 18 and 22 or older). Age 18 students were set as the benchmark group. Students who were younger than 18 tend to have a first-year course completion rate that is 11.52 percentage points above the benchmark. This positive effect is statistically significant at a 1% level. Those aged 19 to 22 or older are found to have no statistically significant associations with the course pass rate.

The main activity that was engaged prior to enrolling in bachelor's study at university was another new factor we could control for in our regressions. Relative to the omitted group of studying in high school, being non-employed or a beneficiary, employed and other types of students (including polythionic, college, PTE and Wānanga students, and university students who were below-degree study) decrease the rate of course completion by 3.52 to 4.74 percentage points. However, students who were overseas have a completion rate that is 4.31 percentage points higher compared to being secondary school student. All the mentioned effects are statistically significant at a 1% level.

New Zealand has eight public universities, which are all funded and monitored by the Tertiary Education Commission, which is accountable to the Ministry of Education. These universities differ from each other in terms of location, size, student population and structure. Therefore, it is highly possible that studying in different universities could affect students' academic outcomes. Using the IDI data allows us to take into account the identities of individual

universities attended by the students in our analyses. However, due to confidentiality reasons, the universities names are not provided in the datasets. We thus designated these universities by using capital letters from A to H (i.e., University A to University H). University A was used as the reference group. We find that the course completion outcomes change across the universities. This finding is statistically significant at a 1% level when using the Chi-square test.[14]

Being a New Zealand citizen has a positive and statistically significant impact (3.60 percentage points) on first-year course completion. Enrolling as a part-time student dramatically reduces the course completion rate during the first year by approximately 22 percentage points, which is the single strongest negative impact found in this analysis. Students who lived in a one-parent household at the time of the 2013 Census have a nearly 2-percentage-points lower course pass rate. These findings are all statistically significant at a 1% level.

It was previously found that university participation was positively related to the parents' highest level of educational attainment. This positive correlation is also found in this first-year course completion analysis. For example, compared to the reference group of parents with no qualifications, having either parent with a Level 1 to 3 certificate increases the rate of course completion by more than a percentage point. Having one parent with a Level 4 certificate has a statistically insignificant impact on first-year course pass rate. Again, that is because those parents are more likely to work at technical or skills-based jobs and may provide insufficient learning support to their children. The greatest positive impact on the first-year performance is found to be having either parent with a postgraduate qualification (6.74 percentage points), followed by a one having a bachelor's or other Level 7 qualification (4.62 percentage points). While the effect of having one parent with level a 5 to 6 diploma (3.11 percentage points) is greater than that of having either parent with an overseas qualification (1.92 percentage points). All these findings are statistically significant at the conventional levels. The other observed parental factors, including household income, parents' proportion of days living on benefit support or number of criminal convictions are found have no statistically significant effect on the course completion outcome.

[14] This is a test of the null hypothesis that all of these seven coefficients are simultaneously equal to zero. For example, holding other factors constant, university identities effect the course completion outcome.

Table 4.3. Key results on ethnicity from the Probit regression analysis for first-year course completion

Variables	Prioritised ethnicity			All ethnicities		
	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Ethnicities						
Māori	-0.3143***	0.0164	-0.0796***	-0.2743***	0.0162	-0.0695***
Pasifika	-0.5654***	0.0194	-0.1432***	-0.5303***	0.0185	-0.1344***
Asian	-0.0255*	0.0144	-0.0065*	-0.0100	0.0141	-0.0025
MELAA	-0.2209***	0.0294	-0.0560***	-0.2043***	0.0291	-0.0518***
Other	-0.0896*	0.0489	-0.0227*	-0.0587	0.0471	-0.0149
Number of observations (n)	56,223			59,307		
Likelihood function	-25595.507			-25602.173		
Pseudo R^2	0.0820			0.0818		

This table only reports the results of identifying with ethnic minority groups for the university outcome. The full set of regression results are provided in Table A.4.3 in the Appendix, along with the full set of covariates that are held constant in this regression analysis.

* denotes statistical significance at a 10% level
** denotes statistical significance at a 5% level
*** denotes statistical significance at a 1% level

Table 4.3 primarily reports the estimated marginal effects of identifying with an ethnic minority group using prioritised ethnicity. It was shown previously in Table 3.3 that Māori's average first-year course completion rate was 13.14 percentage points lower than that of European using prioritised ethnicity and was 12.73 percentage points lower if using all ethnicities. The estimated marginal effects on Māori indicate that the remaining gap between prioritised European and Māori is approximate 8 percentage points and nearly 7 percentage points when using all ethnicities. These numbers suggest that 39.42% of the prioritised European-Māori difference and 45.40% of the all ethnicities European-Māori difference in first-year course completion rates can be eliminated if all the personal characteristics, school factors, prior performance, university factors and parental factors were identical between the two groups. The remaining ethnic differences are statistically significant at a 1% level. More than half of the ethnic gaps in first-year course completion continue to exist even after adjusting for the observable characteristics.

Pasifika students' first-year course pass rate was 26.07 percentage points lower when compared to the rate of European, using prioritised ethnicity. The difference slightly decreased to 25.34 percentage points when using all ethnicities. Holding constant all the factors can close the European-Pasifika gap in course completion rate by 45.07%, from 26.07 to 14.32 percentage points using prioritised ethnicity, and can reduce the gap by 46.96% from 25.34

to 13.44 percentage points, based on all ethnicities. The remaining ethnic differences are also statistically significant at a 1% level.

It was previously found that prioritised Asian's rate of first-year course completion was 2.01 percentage points below prioritised European and 1.78 percentage points under European when using all ethnicities. The ethnic difference reduces to 0.65 percentage points, eliminating 67.66% of the gap between prioritised European and Asian. The remaining difference is statistically significant at a 5% level. If using all ethnicities, the ethnic difference declines to 0.25 percentage points, removing about 86% of the gap between the ethnic groups. This remaining difference, however, is not statistically significant at the conventional levels.

4.2.2. Impact of employment hours on first-year course completion

As mentioned in the literature review, several past studies have found negative relationships between university students' employee workload and their academic performance (e.g., Bozick, 2007; Kalenkoski & Pabilonia, 2012), while other studies claimed that there was no relationship or a non-negative relationship between the two variables (e.g., Darolia, 2014; Parent, 2006).

The availability of income tax data (sourced from the IRD) in the IDI allowed us to test how employment, measured by average weekly hours of work during the academic year from March to November, affects academic outcomes at different stages of the university journey. An advantage of using these data is that we can link monthly earnings information from the academic year to course completion outcomes in the same period. Initially, we expected that having a regular part-time job during term-time might adversely affect university performance, and that working a higher number of hours (e.g., 20 hours or more a week) might result in poorer academic performance, simply because working limits the time available to study.

A limitation of the tax data is that they only provide information on wage and salary earnings on a monthly basis, without providing data on the hours that were worked to generate these earnings. Given that, we chose to estimate employment hours for the students by assuming they were paid the legal minimum wage. The formula for the estimation can be found in the footnote.[15] It should be borne in mind that the work hours for some students may be either over- or under-estimated, and hours on employment outside of the academic months were not taken into account in the analysis.

[15] $\text{Avearge weekly work hours} = \frac{\text{Total reported wages and salaries during the academic year}}{\text{legal minimum wage in the year}} / \text{number of weeks during the academic year}$

Based on the number of estimated weekly employment hours, the full first-year sample is broken down into 10.93% working zero hours, 3.08% working between zero and less than 10 hours, 1.89% working between 10 and less than 20 hours, 13.68% working 20 hours or more, and 70.43% with unknown hours (see Table A.4.4 in the Appendix). The high proportion of the sample with unknown hours of work should catch the reader's eye. The data could be missing in a non-random. For example, those who worked during the academic months and whose earnings information was somehow not recorded (i.e., the data are missing at random). Another possibility is they just did not engage in employment in that period of time so there were no income records in the tax data (i.e., the missing data are not random). Having a large proportion of observations with missing data is a potential issue that could introduce bias to our analytical results. How to handle such a problem and the related regression results is discussed in the paragraph after next.

For the sake of simplicity, between zero and less than 10, between 10 and less than 20, and 20 hours or more are referred to as low, medium and high hours of work, respectively. We created a series of dummy variables for each group of students. All comparisons are made to working zero hours (see Table 3.1 for variables used). The full-sample regression results reported in Table A.4 suggest that work during the term time overall negatively affects first-year course completion performance, but working high hours is not associated with a stronger detrimental effect on the course completion rate. In particular, working low hours reduces the first-year course completion rate by 1.74 percentage points, compared to working zero hours. This result is statistically significant at a 5% level. Spending a medium number of hours on employment has no measurable impact on the course completion rate. Working high hours lowers the course pass rate by 0.79 percentage points and this finding is statistically significant at a 10% level.

To remove the potentially disruptive effect of the missing data on the analysis, we dropped those students with unavailable work hours information. The restricted student sample contains 36.95% working zero hours, 10.43% working low hours, 6.39% working a medium number of hours and 46.25% working a high number of hours. For those students, the regression results reported in Table A.4.5 in the Appendix show that taking on a part-time job with a low number of hours reduces the first-year course completion rate by 1.74 percentage points, compared to working zero hours. This result is statistically significantly different from zero at a 5% level. Working a medium number of hours has no statistically significant impact on the course completion rate. Spending a high number of hours in employment leads to the course completion rate going down by 0.77 percentage points relative to having no

employment commitments. This finding is statistically significant at a 10% level using prioritised ethnicity, but it becomes statistically insignificant when using all ethnic designations. Comparing these effects to the results discussed in the above paragraph, we find that including the students with missing employment data in the full-sample analysis generates very few differences to the impact of the hours worked on first-year course completion.

As mentioned above, part-time enrolment status has a great impact on the outcome of first-year course completion. Also, theoretically speaking, the academic performance of students who were enrolled full-time was more sensitive to the number of employment hours. To test that, we confined the analysis to only full-time students. In this particular subsample, there are 10.96% zero, 3.09% low, 1.88% medium, 13.70% high and 70.37% unknown hours student employees. However, we find that the pattern of the results found in the above analyses remain almost the same among full-time students. For example, students who worked low hours have a course completion rate that is 1.71 percentage points lower than those who did not hold a part-time job. This effect is different from zero at a 5% level. Working a medium number of hours have no statistically significant impact on the first-year course completion outcome. High-hour student employees are 0.80 percentage points less likely to complete first-year courses than their zero-hour-working peers. This result is statistically significant at a 10% level.

Female students were found on average to perform better than their male counterparts in terms of completing first-year courses. Given that, it is possible that the number of hours worked affects first-year course completion rates in different ways for the gender groups. Hence, we repeated the analysis separately for female and male students. For the subsample of female students, the proportions with zero, low, medium, high and unknown employment hours were 10.84, 3.02, 1.86, 13.67 and 70.61%, respectively. The regression results based on the female students show that spending a low number of hours on part-time employment reduce course pass rates in the first year by 1.84 percentage points, compared with not working. This effect is statistically significant at a 5% level. Working a medium or high number of hours have no measurable impact on the course completion.

In the male subsample, 11.06% of the students worked zero hours, 3.17% worked low hours, 1.93% worked a medium number of hours and 13.67% worked a high number of hours, with 70.18% having no employment information. Unlike the female students, male first-year learners working high hours lower the course completion rate by 1.31 percentage points, relative to working zero hours. This effect is statistically significant at a 10% level. Working a low or medium number of hours both have a statistically insignificant effect on the course

completion outcome. These empirical results provide evidence that the point estimates of working while studying on first-year course pass rates are different between female and male students.

Because ethnicity is the focus in this thesis and our earlier results showed that first-year course pass rates vary by ethnicity, it is necessary to individually analyse the employment impacts on the first-year outcome for the ethnic groups. The analysis was first conducted for the Māori and Pasifika groups. These two ethnicities were grouped together due to the fact that these ethnic minorities having the poorest performance in academic outcomes. We could hypothesize that the effects of term-time employment on the academic outcome may be more detrimental for these two groups of students. The combined subsample consist of 11.46% working zero hours, 2.79% working low hours, 1.97% working a medium number of hours and 14.25% working a high number of hours. For 69.53% no work hour information is available. Interestingly, we find that spending a medium number of hours on employment increase first-year course pass rates by about 4.88 percentage points. This finding is statistically significant at a 10% level. Working other numbers of hours have no measurable impact on the course completion performance. One possible reason for these results is that spending a medium amount of time on employment provides Māori and Pasifika students with an income that helps them with their living expenses while studying. In other words, there is a reverse causality between working medium hours and first-year course completion for this merged group of students.

We then wanted to know whether the positive effect of working a medium number of hours on first-year course completion were different between Māori and Pasifika. When only considering Māori students, 10.95% of them had no part-time job, while 2.99% were working low hours, 1.93% were working a medium number of hours, and 14.0% working high hours. And 70.19% had missing information on the variable (see Table 4.4). Surprisingly, it is found that hours worked have no statistically significant effect on the first-year outcome for Māori students (see Table 4.5). An alternative interpretation for this statistically insignificant relationship is that working a medium number of hours is not helpful in increasing the course pass rate of Māori students.

Table 4.4. Proportions of ethnic subsamples based on first-year employment hours: prioritised ethnicity

	Sample proportion
Māori students	
Employment hours	
Zero	0.1095
Low (>0 and 10 hours)	0.0299
Medium (≥10 and <20 hours)	0.0193
High (≥20 hours)	0.1400
Unknown	0.7019
Pasifika students	
Employment hours	
Zero	0.1214
Low (>0 and 10 hours)	0.0253
Medium (≥10 and <20 hours)	0.0202
High (≥20 hours)	0.1459
Unknown	0.6863
Asian students	
Employment hours	
Zero	0.1043
Low (>0 and 10 hours)	0.0316
Medium (≥10 and <20 hours)	0.0176
High (≥20 hours)	0.1302
Unknown	0.7163
European students	
Employment hours	
Zero	0.1104
Low (>0 and 10 hours)	0.0310
Medium (≥10 and <20 hours)	0.0190
High (≥20 hours)	0.1367
Unknown	0.7029

This table only reports the proportions of the ethnic groups. The full set of proportions of all the subsamples are provided in Table A.4.4 in the Appendix.

When we focused on Pasifika students, the proportions working zero, low, medium and a high number of hours, or having no work hour information are, respectively, 12.14, 2.53, 2.02, 14.59 and 68.63%. We find that the positive relationship between working a medium number of hours and first-year course pass rates reoccur when solely looking at Pasifika students. For instance, Pasifika students who worked medium hours have a course pass rate that is 7.55 percentage

points higher than their Pasifika peers who did not have a part-time job. This finding is different from zero at a 10% level.

The analysis was also repeated for ethnic groups other than Māori and Pasifika. The non-Māori and non-Pasifika group covers the European, Asian, MELLA and Other ethnicity students. The students in this combined subsample is separated into 10.84% with zero working hours, 3.13% with a low number of working hours, 1.87% with medium working hours, 13.58% with a high number of working hours and 70.59% with an unknown number of working hours (see Table A.4.4 in the Appendix). The regression results in Table A.4.5 in the Appendix indicate that spending a low or high number of hours at work decrease first-year course completion rates by 1.64 and 0.88 percentage points, respectively, compared to not working. The first and second effects are statistically significant at 5 and 10%, respectively. Working a medium number of hours have no measurable impact on the course completion. These results are similar to those found in the full sample analysis.

Our next focus was on how part-time employment influenced the first-year course completions of Asian and European. In the subsample of Asian, 10.43% of worked zero hours, 3.16% worked low hours, 1.76% worked medium hours, and 13.02% worked a high number of hours as student employees. And 71.63% had missing work hour data. Differently to the above results, we find no evidence that work hours affected Asian's course pass rate in first-year. Unobserved variable(s) and causality (see Section 2.3 for more discussion about causality) might be the factors to explain the findings.

European students in the analysis worked 11.04, 3.10, 1.90, 13.67 and 70.29% at zero, low, medium, high and unavailable work hours, respectively. The story is a bit different for European students. Relative to having no job, working low hours lower the first-year course completion rate by 2.01 percentage points. This result is statistically significant at a 5% level. The negative effect decrease to 1.70 and 1.08 percentage points for working medium and high hours, respectively. Both effects are statistically significant at a 10% level. These results are consistent with the results found based on the full sample in terms of the statistically significantly negative effect does not become stronger with the number of work hours.

Another analysis used the full sample and new thresholds for the number of hours worked. Low, medium and high number of hours were redefined as working between zero and less than 15 hours, between 15 and less than 30 hours, and 30 or more hours, respectively. After redefining the hour ranges, working zero, low, medium or high hours, and having unknown hours account for 10.93, 3.08, 1.89, 13.68 and 70.43% of the sample, respectively. The results

in Table A.4.5 in the Appendix indicate that spending a low (between zero and less than 15) number of hours at work reduce first-year course completion rates by 1.32 percentage points relative to working zero hours, while working high (30 or more) hours decrease the rates by 0.91 percentage points. The former effect is statistically significant at a 5% level and the latter one is different from zero at a 10% level. The marginal effect of spending medium (15 and less than 30) hours in part-time jobs on the first-year outcome is not statistically significant at the conventional levels. These findings also suggest that working a high number of hours does not have a more adverse impact on first-year course completion achievement, which is similar to the results found in the full sample with the initial hour cut-offs.

Table 4.5. Key results on ethnicity from the Probit regression analysis for employment hours on first-year course completion

	Prioritised ethnicity			All ethnicities		
	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Māori students						
Employment hours						
Low (>0 and 10 hours)	-0.0514	0.0965	-0.0165	-	-	-
Medium (≥10 and <20 hours)	0.0898	0.1167	0.0288	-	-	-
High (≥20 hours)	-0.0380	0.0588	-0.0122	-	-	-
Unknown	-0.0525	0.0479	-0.0169	-	-	-
Pasifika students						
Employment hours						
Low (>0 and 10 hours)	-0.0802	0.1179	-0.0290	-0.1530	0.1129	-0.0552
Medium (≥10 and <20 hours)	0.2092*	0.1205	0.0755*	0.1989*	0.1167	0.0717*
High (≥20 hours)	0.0099	0.0614	0.0036	-0.0150	0.0595	-0.0054
Unknown	0.0140	0.0498	0.0051	-0.0139	0.0483	-0.0050
Asian students						
Employment hours						
Low (>0 and 10 hours)	-0.0308	0.0667	-0.0076	-0.0316	0.0663	-0.0078
Medium (≥10 and <20 hours)	0.0442	0.0874	0.0108	0.0636	0.0868	0.0157
High (≥20 hours)	-0.0501	0.0432	-0.0123	-0.0519	0.0427	-0.0128
Unknown	-0.0387	0.0347	-0.0095	-0.0387	0.0343	-0.0095
European students						
Employment hours						
Low (>0 and 10 hours)	-0.0866**	0.0378	-0.0201**	-0.0829**	0.0364	-0.0196**
Medium (≥10 and <20 hours)	-0.0730*	0.0438	-0.0170*	-0.0763*	0.0424	-0.0180*
High (≥20 hours)	-0.0464*	0.0241	-0.0108*	-0.0466**	0.0232	-0.0110**

Unknown	-0.0414**	0.0192	-0.0096**	-0.0423**	0.0186	-0.0100**

This table only reports the results of the ethnic subsamples. The full set of regression results for all the subsamples are provided in Table A.4.5 in the Appendix.

* denotes statistical significance at a 10% level
** denotes statistical significance at a 5% level
*** denotes statistical significance at a 1% level

Overall, our regression results suggest mixed findings. On the one hand, working low hours reduces completion rates in first-year courses by no more than two percentage points, but these effects are not consistently statistically significantly different from zero in the regressions. On the other hand, the hypothesis of the higher number of hours worked (either more than 20 or 30 hours) having a larger detrimental impact on the course completion outcomes is consistently not confirmed by the regressions. There are some reasons for the inconsistent results. First, hours worked may have the expected negative impact on an outcome variable like letter grade on first-year courses. Compared to course completion, course grade may be a better indicator of overall course performance. It is possible that term-time work might lead to a lower course letter grade, without reducing the probability of successfully completing the course. However, information about students' grades for courses was unavailable in the datasets so that we cannot test this hypothesis.

Second, hours are not randomly assigned to students. In other words, the number of hours a student chooses to work is affected by a mediator variable and this variable may influence academic performance. Examples of such variables are study attitudes and organizational skills. Students with poor study attitudes may be more likely to combine study with work, which could lead to an overestimate of the detrimental effects of work hours on the course completion. The detrimental effects could also be underestimated in the case where employers tend to offer work to students who have good organizational skills. But we again were not able to control for such kinds of variables in our analysis due to the fact that they were not provided in the data.

Third, our findings also indicate that engaging in employment has different impacts on the first-year university outcome among gender and ethnic groups, which is supported by Neyt et al. (2019), who suggested that the impact of working while studying on university performance is subject to student characteristics, job characteristics, the analysis method(s) used and the country of analysis.

4.2.3. Decomposition analysis

Linear (Blinder-Oaxaca) decomposition was adopted in this section, because this technique is more appropriate than nonlinear (Fairlie) decomposition when used to decompose the gaps between two ethnic groups in a more 'linear' outcome variable, such as the proportion of first-year course completions. Similar to nonlinear decomposition, linear decomposition analysis used the estimated coefficients from the pooled Ordinary Least-Squares (OLS) regression on each pair of comparison groups (i.e., European and Māori, European and Pasifika, and European and Asian). Table A.4.6 in the Appendix summaries the complete set of results from the decompositions in first-year course completion rate between the three ethnic minority groups (Māori, Pasifika and Asian) and European, using both prioritised ethnicity and all ethnicities.

Table 4.6. Key results on ethnicity from the linear decomposition of first-year course completion

	Prioritised ethnicity			All ethnicities		
	European vs Māori	European vs Pasifika	European vs Asian	European vs Māori	European vs Pasifika	European vs Asian
Total ethnic difference	0.1314	0.2607	0.0201	0.1273	0.2534	0.0178
Total explained differences	0.0403*** (0.0025) [30.67%]	0.0829*** (0.0042) [31.80%]	0.0126*** (0.0024) [62.69%]	0.0404*** (0.0024) [31.74%]	0.0813*** (0.0040) [32.08%]	0.0122*** (0.0023) [68.54%]
Unexplained differences	0.0911*** (0.0051) [69.33%]	0.1778*** (0.0070) [68.20%]	0.0075** (0.0036) [37.31%]	0.0869*** (0.0051) [68.26%]	0.1721*** (0.0067) [67.92%]	0.0056 (0.0036) [31.46%]
Number of observations (n)	40,791	39,525	46,038	43,368	42,372	48,792

This table only reports the overall decomposition results in the university outcome. The full set of decomposition results are provided in Table A.4.6 in the Appendix.

* denotes statistical significance at a 10% level
** denotes statistical significance at a 5% level
*** denotes statistical significance at a 1% level

4.2.3.1. European–Māori gap

The full decomposition results based on prioritised ethnicity are reported in first three columns in Table A.4.6. It should be noted that other parental factors include single-parent household composition, household income and parents' benefit and crime histories. We previously found that the first-year course pass rate for Māori students was 13.14 percentage points lower than that of their European peers. The decomposition results in the first column of Table 4.6 highlights that if Māori had the same observed characteristics as European, this would close the first-year course completion gap between European and Māori by 4.03 percentage points, or 30.67%. This is statistically significant at a 1% level. The results also indicate that the Māori first-year course pass rates continue to be 9.11 percentage points below that of European, even after we account for the impacts of all measurable factors in this regression analysis.

The most important factors in explaining the completion rate between the ethnic groups, in descending order, are school deciles (1.99 percentage points or 15.14%), parents' highest education qualifications (0.83 percentage points or 6.32%), part-time enrolment (0.74 percentage points or 5.63%), the highest school qualification (0.56 percentage points or 4.26%), other school factors (0.33 percentage points or 2.51%), prior activity (0.22 percentage points or 1.67%), age (0.03 percentage points or 0.23%), cohort years and other parental factors (both 0.01 percentage point or 0.08%). If gave Māori and European the same female representation, university distribution and New Zealand citizenship status, this would widen the gap in course completion rates between the ethnic groups by 0.69 percentage points or 5.25% (see Table A.4.6 in the Appendix).

Using all ethnic identities, the European-Māori gap in the completion of first-year courses was 12.73 percentage points. The results in the fourth column of Table 4.6 suggest that if Māori had the same measured covariates as European, this could reduce 4.04 percentage points or 31.74% of the ethnic gap in course completion. The factors that play primary role in explaining the European-Māori gap in course pass rate are school decile scores, parental qualifications, being a part-time student, the highest school qualification, other school factors, prior activity, other parental factors, age and cohort years. Eliminating the ethnic differences between Māori and European in female representation, university distribution and New Zealand citizenship status would widen the gap in course completion between the ethnic groups.

4.2.3.2. European–Pasifika gap

The first-year course pass rate for prioritised Pasifika was 26.07 percentage points lower than that of prioritised European. As shown in the second column of Table 4.6, eliminating the ethnic differences in all measurable factors could explain 8.29 percentage points of first-year course completion gaps between European and Pasifika, which suggests that if Pasifika had the same observable factors as European, it would eliminate 31.80% of the ethnic gap in the course completion. This implies that the Pasifika course completion rate would still be 17.78 percentage points lower than the European rate once we eliminated ethnic differences in the measured covariates used in the regression.

The primary factors that contribute to explaining the European-Pasifika gap in first-year course completion performance are school deciles (4.66 percentage points or 17.87%), part-time enrolment (1.88 percentage points or 7.21%), parents' highest educational qualifications (1.28 percentage points or 4.91%), the highest school qualification attainment (0.79 percentage points or 3.03%), other school factors (0.37 percentage points or 1.42%), New Zealand citizenship (0.18 percentage points or 0.69%), prior activity (0.16 percentage points or 0.61%), other parental factors (0.10 percentage points or 0.38%) and age (0.03 percentage points or 0.12%). Removing the ethnic differences in university distribution, female representation, cohort years and work hours between European and Pasifika together would increase the ethnic gap in course pass rate by 1.16 percentage point, or 4.45%.

Comparing the results, we find that factors including parental education, school qualifications, other school factors, prior activity and age can each explain the relatively higher proportion of the gap between European and Māori, while school deciles, being a part-time student and other parental factors separately has more explanatory power on the European-Pasifika gap.

The gap in first-year course completion was 25.34 percentage points between European and Pasifika, using all ethnicities. If Pasifika had the same observable factors as those of European, this could explain the ethnic gap in course completion in the first year by 8.13 percentage points, or 32.08%. School decile is the most important factor in closing the European-Pasifika gap in the first-year course completion outcome (see the fifth column of Table A.4.6 in the Appendix). The next most important factor is part-time enrolment, followed by parental educational qualifications, the highest school qualification, other school factors, activity engaged in before enrolment, other parental factors, being a New Zealand citizen and age. However, assigning the same university distribution, female representation, employment

hours and cohort years could result in the first-year course completion gap between European and Pasifika being wider.

4.2.3.3. European–Asian gap

It was previously found that the first-year course completion rate for Asian was 2.01 percentage points below European when using prioritised ethnicity. The results provided in the third column of Table 4.6 show that removing the ethnic differences in all the factors considered can account for 1.26 percentage points of the European-Asian gap in first-year course completion, meaning that if we gave Asian and European students the same observed factors we could explain 62.69% of the gap in course completion between the ethnic groups. This result is statistically different from zero at a 1% level, and would leave 0.75 percentage points of the European-Asian gap unexplained.

Differently to Māori and Pasifika, the main factors that shrink the European-Asian course completion gap are, in order, New Zealand citizenship status (0.72 percentage points or 35.82%), other school factors (0.33 percentage points or 16.42%), female representation and school decile scores (both 0.25 percentage points or 12.44%), part-time enrolment (0.13 percentage points or 6.47%), parental qualifications (0.09 percentage points or 4.48%), other parental-related factors (0.02 percentage points or 1.00%) and employment hours (0.01 percentage points or 0.50%). Eliminating the ethnic differences in other university distribution, prior activity, school qualifications, age and cohort years between European and Asian together would increase this ethnic gap in first-year course completion by 0.54 percentage points, or 26.87%.

The all ethnicities European-Asian gap in the first-year pass rate was 1.78 percentage points. If we gave Asian and European the same factors observed, this could eliminate the European-Asian gap in first-year course completion by 1.22 percentage points, or 68.54% (see sixth column in Table 4.6). The most important factors that contribute to explaining the European-Asian gap in the course completion outcome are ordered as New Zealand citizenship status, other school factors, school decile scores, female representation, being part-time, other parental factors/parents' highest educational qualifications and hours of employment. Factors widening the gap between the ethnic groups include distribution of university attendance, prior activity, secondary school qualifications, age and cohort years.

4.3. Second-year course completion

4.3.1. Probit regression analysis

The first year of university is the period when students settle into university life. The courses taught in the first year are usually fundamental. However, second-year courses are more challenging, and students moving into second year might be more focused on their academic achievement. Given the differences in terms of study motivation and course difficulty between the two stages, it is worth looking at second-year course completion outcome in addition to course completion performance in the first year. It should be noted that second-year course completion can be considered as a similar performance indicator to second-year retention as having the course outcome records are subject to persisting at study in the second year. Hence, there may be less reason to also exam second-year retention in this thesis.

The sample used in second-year course completion analysis is restricted to the 2011 and 2012 cohorts because of the current unavailability of tertiary data after 2019 in the IDI. This analysis uses the same independent variables that were used in the first-year course completion regression but the age variables are replaced by the dummy variables, indicating age in the second year.

The full set of regression results for second-year course completion using prioritised and all ethnicities are provided in Table A.4.7 in the Appendix. Making use of the alternative ethnic designations does not change the regression results qualitatively. The Pseudo R^2 statistics of 0.1184 indicates that the independent variables used can better explain this second-year course completion outcome than the variables did in first-year course completion analysis. However, some previously found statistically significant school factors lose their significance in this regression. First, the regression results show that female students still have a statistically significantly greater performance (5.84 percentage points) than their male counterparts in terms of completing second-year courses. The average pass rate in second-year courses for the students in the 2011 cohort is slightly higher (0.66 percentage points) than that of those in the 2010 cohort.

We find that school deciles have a similar impact on first- and second-year course completion outcomes. Coming from a decile 1-4 school is associated with a negative impact on second-year course completion compared to the omitted group of attending decile 6 schools. Also, attending schools with lower decile scores result in a stronger negative impact. Studying in a decile 5 school have no statistically significant impact on the course completion, which is the same as coming from deciles 7 and 8 schools. Being from the schools in the top two deciles

(9 and 10) statistically significantly lift the course pass rate by 1.29 to 1.49 percentage points. Furthermore, the results indicate that the second-year course pass rate goes up by nearly 10 percentage points when moving from a decile 10 to a decile 1 schools. It is found that changing from the top to the bottom decile schools have a weaker positive effect on second-year over first-year course completion rates. In addition, our results suggest that the effects of school deciles on course completions have declined over time.

Not surprisingly, secondary school academic achievements and other school-related factors, including ownership type, gender composition and geographic area are statistically insignificant in this regression analysis. That is because these school characteristics may be considered as 'remote' factors that have little or no connection to the academic performance for those who have participated and remained in second year of university study. This is supported by the findings of Shulruf, Hattie and Tumen (2008), who emphasized that some characteristics of school such as size, gender mix and type of funding do not affect students' success once they entered university. We, however, find that switching schools continues to be an important predictive variable. For instance, relative to never switching school, changing school twice, three times or four times and more reduces the proportion of completing second-year courses by 2.17, 3.23 and around 6.61 percentage points, respectively. All of these marginal effects are different from zero at a 1% level. Again, truancy, suspension and/or unjustified absence is found have no observable effect on second-year course pass rate.

Age in the second year also loses its statistical significance in predicting the university outcome. Compared to studying in high school before enrolling in a bachelor's qualification, being unemployed or a beneficiary decrease the second-year course completion rate by 2.03 percentage points. This result is statistically significant at a 10% level. The negative effect slightly increases to 2.42 percentage points for being employed, and further to nearly 4 percentage points for being other type of student. These effects are statistically significant different from zero at a 1% level. The Chi-square test results provide evidence that second-year course completion rates are statistically significantly different among the universities at a 1% level.

Being a New Zealand citizen boost the second-year course pass rate by 2.07 percentage points. Enrolling as a part-time student is associated with the greatest adverse impact of approximately 22 percentage points in the completion rate. These results are both highly statistically significant.

Coming from a single-parent household at the time of the 2013 Census lower the second-year course completion rate by 1.35 percentage points. Having parents whose highest level of education was a Level 1 to 3 certificate and Level 4 certificate have no statistically significant effect on the course completion relative to having parents who had not completed a qualification. The pass rate for students whose parents had a Level 5 to 6 diploma is 2.52 percentage points higher than for those having parents with no educational qualifications. The positive effects of having parents whose highest level of education was a bachelor's degree or equivalent, or an overseas qualification, is 3.57 and 2.37 percentage points, respectively. The highest beneficial effect of parental education on the second-year course completion of 3.76 percentage points come from those whose either parent had a postgraduate qualification. These effects are different from zero at a 1% level. Similar to the findings from the analyses of university participation and first-year course completion, household income in 2013, parents' benefit and crime histories are found to have no measurable impact on second-year course completion.

Table 4.7. Key results on ethnicity from the Probit regression analysis for second-year course completion

	Prioritised ethnicity			All ethnicities		
Variables	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Ethnicities						
Māori	-0.3734***	0.0237	-0.0784***	-0.3279***	0.0237	-0.0689***
Pasifika	-0.6053***	0.0269	-0.1271***	-0.5599***	0.0258	-0.1176***
Asian	-0.0716***	0.0197	-0.0150***	-0.0470**	0.0194	-0.0099**
MELAA	-0.3110***	0.0404	-0.0653***	-0.2889***	0.0400	-0.0607***
Other	-0.0644	0.0674	-0.0135	-0.0168	0.0658	-0.0035
Number of observations (n)	31,722			33,342		
Likelihood function	-12115.468			-12121.923		
Pseudo R^2	0.1184			0.1179		

This table only reports the results of identifying with ethnic minority groups for the university outcome. The full set of regression results are provided in Table A.4.7 in the Appendix, along with the full set of covariates that are held constant in this regression analysis.

* denotes statistical significance at a 10% level
** denotes statistical significance at a 5% level
*** denotes statistical significance at a 1% level

Table 4.7 reports the key marginal effects of being from an ethnic minority. Māori second-year course completion rate was 12.43 percentage points lower than that of European when using prioritised ethnicity. The ethnic difference slightly decreases to 12.02 percentage points if

using all ethnicities. It can be seen that the remaining unexplained differences decrease to 7.84 and 6.89 percentage points, respectively, based on prioritised ethnicity and all ethnicities. These numbers imply that holding all the other personal, school and university factors as well as employment hours constant can eliminate 36.93% and 42.68% of the differences in second-year course completion rates using prioritised ethnicity and all ethnicities, respectively. The remaining differences are both statistically significant at a 1% level. Comparing the explained differences in first-year course completions between European and Māori, and controlling for almost the same factors can explain narrower of the European-Māori gap in second-year course pass rates.

We saw earlier that the second-year course pass rate for Pasifika was 24.85 percentage points lower than the European rate when using prioritised ethnicity. When using all ethnicities the difference was slightly lower at 24.31 percentage points. The difference between prioritised Pasifika and European reduce 48.85% to 12.71 percentage points, while the all ethnicities European-Pasifika difference decrease 51.62% to 11.76 percentage points, once other measured factors were held constant. Both the remaining differences are statistically significant at a 1% level.

The rate of second-year course completion for prioritised Asian students was 2.21 percentage points lower than it is for prioritised European. The ethnic difference shrinks slightly to 1.89 percentage points if making use of all ethnicities. Keeping all other factors constant closes 32.13% of the difference between prioritised Asian and European, leaving 1.50 percentage points of the difference unexplained. The remaining difference is statistically significant at a 1% level. Holding constant the other independent variables reduces 47.62% of the difference between Asian and European when using all ethnicities, leaving 0.99 percentage points of the difference unexplained. This remaining unexplained difference is statistically significant at a 5% level.

4.3.2. Impact of employment hours on second-year course completion

Our findings from the previous analysis suggested that there was some empirical evidence that combining work with study had a detrimental effect on the first-year university performance, but the effect did not get stronger as the number of employment hours increase. We also wanted to estimate how market work while studying affects second-year course completion outcomes. One hypothesis is that detrimental effects of being a part-time employee on course completion might become stronger as students progress to second year

and study loads increase. An alternative hypothesis is that the detrimental effects of employment on course completion might weaken as students mature and figure out how to combine work with their studies. As before, workload was measured by the average weekly work hours during the academic months (March to November) in the second year, assuming students that were paid the current minimum wage.

As shown in Table A.4.8 in the Appendix, the full sample of students is comprised of 11.20% working zero hours, 3.21%, 1.83% and 14.24% working low, medium and high weekly hours, respectively. Employment information was unavailable for 69.54% of the sample. The ranges of low, medium and high hours are between zero and less than 10, between 10 and less than 20, and 20 or more, respectively. Again, students who did not engage in employment were used as the benchmark group in the regression analysis. Unexpectedly, we find that the signs of the estimated marginal effects are positive but not statistically significant at the conventional levels, meaning there is no measurable effect of hours worked on second-year course completion rate (see Table A.4.7 in the Appendix).

Dropping those with unknown work hour information, the subsample contains 36.76, 10.52, 5.99 and 46.72% students who worked zero, low, medium and high hours, respectively. The analysis results reported in Table A.4.9 in the Appendix show that the positive and statistically insignificant effects of work on the second-year outcome persist even when focusing on only the students with valid work hour records.

When the analysis was limited to full-time students, 11.22% of them worked zero hours, 3.18% worked a low number of hours, 1.82% worked a medium number of hours and 14.27% worked a high number of hours, while 69.50% had unknown hours worked. Still, we find that there are positive effects of work hours on second-year course completion, but no evidence to support these effects are different from zero.

The analysis was conducted separately for female and male students. In the female subsample, proportions having zero, low, medium, high and unknown hours are 10.88, 3.22, 1.75, 14.34 and 69.83%, respectively. The signs on the estimated coefficients are all positive. There are 11.68% of students with zero hours, 3.19% with low hours, 1.93% with medium hours, 14.09% with high hours and 69.14% with unknown work hours in the subsample of males. We find negative sign on the estimated coefficient of working low hours and positive signs on the estimated coefficients of working medium and high hours. However, using these subsamples of students does not change the statistically insignificant correlation between

work hours and course completion in second year. In other words, the statistically insignificant impacts do not vary by gender.

Table 4.8. Proportions of ethnic subsamples based on second-year employment hours: prioritised ethnicity

	Sample proportion
Māori students	
Employment hours	
Zero	0.1113
Low (>0 and 10 hours)	0.0325
Medium (≥10 and <20 hours)	0.0113
High (≥20 hours)	0.1400
Unknown	0.7050
Pasifika students	
Employment hours	
Zero	0.1175
Low (>0 and 10 hours)	0.0294
Medium (≥10 and <20 hours)	0.0196
High (≥20 hours)	0.1648
Unknown	0.6672
Asian students	
Employment hours	
Zero	0.1111
Low (>0 and 10 hours)	0.0322
Medium (≥10 and <20 hours)	0.0176
High (≥20 hours)	0.1388
Unknown	0.7009
European students	
Employment hours	
Zero	0.1131
Low (>0 and 10 hours)	0.0322
Medium (≥10 and <20 hours)	0.0190
High (≥20 hours)	0.1414
Unknown	0.6943

This table only reports the proportions of the ethnic groups. The full set of proportions of all the subsamples are provided in Table A.4.8 in the Appendix.

We then examined the effects of part-time employment on second-year course pass rates with respect to ethnicity. The analysis started by using a conjoined group of Māori and Pasifika

students. The breakdown of the subsample is based on work hours at 11.40, 3.12, 1.56, 15.08 and 68.91%, respectively, with working zero, low, medium and high hours, and with missing data. Relying on this conjoined subsample of students, we find positive but statistically insignificant estimated coefficients. Similar results are found even when we restricted the analysis to only Māori students (11.13% with zero hours, 3.25% with low hours, 1.13% with medium hours, 14.00% with high hours and 70.50% with unknown hours) and only Pasifika learners (11.75% with zero hours, 2.94% with low hours, 1.96% with medium hours, 16.48% with high hours and 66.72% with unknown hours). Those proportions are reported in Table 4.8 and the regression results can be found in Table 4.9. However, little differences are found in the signs on the estimated coefficients between the two subsamples. For the Māori subsample, the signs on the estimated coefficients were positive except the one for unknown. For the Pasifika regression, the sign on the estimated coefficient of working low hours is positive but the coefficients of working medium and high hours both have a negative sign.

The analysis was repeated, using all the ethnic groups apart from Māori and Pasifika. This subsample is split into 11.17, 3.22, 1.87, 14.11 and 69.64%, respectively, working zero, low, medium and high hours, and had no useful work hour information. There is no support for the contention that being a part-time employee reduces the rate of second-year course completion. Such findings persist when we alternatively concentrated on each of the Asian (11.11% with zero hours, 3.22% with low hours, 1.76% with medium hours, 13.88% with high hours and 70.09% with missing hour data) and the European (11.31% with zero hours, 3.22% with low hours, 1.90% with medium hours, 14.14% with high hours and 69.43% with missing hour data) groups. The signs for the estimated coefficients are all positive for the subsamples.

The last attempt was to carry out an analysis using the full sample and alternative measures of hours worked. We redefined low hours as working between zero and less than 15 hours per week, medium hours as working between 15 and less than 30 hours per week and high hours as working 30 hours a week or more. In this way, the proportion having zero, low, medium, high and unavailable information for work hours are 11.20, 3.21, 1.83, 14.24 and 69.54%, respectively. The estimated coefficients for working low and high hours are positive, and the one for working medium hours is negative. But once again, these effects are all statistically insignificant.

Table 4.9. Key results on ethnicity from the Probit regression analysis for employment hours on second-year course completion

	Prioritised ethnicity			All ethnicities		
	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Māori students						
Employment hours						
Low (>0 and 10 hours)	0.1273	0.1228	0.0361	-	-	-
Medium (≥10 and <20 hours)	0.2169	0.1927	0.0614	-	-	-
High (≥20 hours)	0.0147	0.0874	0.0042	-	-	-
Unknown	-0.0185	0.0685	-0.0053	-	-	-
Pasifika students						
Employment hours						
Low (>0 and 10 hours)	0.0207	0.1490	0.0068	-0.0311	0.1434	-0.0103
Medium (≥10 and <20 hours)	-0.0971	0.1909	-0.0320	-0.1420	0.1908	-0.0468
High (≥20 hours)	-0.0119	0.0837	-0.0039	-0.0377	0.0819	-0.0124
Unknown	0.1026	0.0699	0.0338	0.0563	0.0680	0.0186
Asian students						
Employment hours						
Low (>0 and 10 hours)	0.0567	0.0951	0.0119	0.0516	0.0942	0.0108
Medium (≥10 and <20 hours)	0.1031	0.1061	0.0216	0.1103	0.1058	0.0231
High (≥20 hours)	0.0445	0.0558	0.0093	0.0360	0.0553	0.0075
Unknown	0.0481	0.0455	0.0101	0.0446	0.0451	0.0093
European students						
Employment hours						
Low (>0 and 10 hours)	0.0456	0.0514	0.0085	0.0454	0.0496	0.0087
Medium (≥10 and <20 hours)	0.0554	0.0654	0.0104	0.0752	0.0634	0.0144
High (≥20 hours)	0.0004	0.0331	0.0001	-0.0107	0.0318	-0.0020
Unknown	0.0109	0.0262	0.0020	0.0165	0.0253	0.0032

This table only reports the results of the ethnic subsamples. The full set of regression results for all the subsamples are provided in Table A.4.9 in the Appendix.

* denotes statistical significance at a 10% level
** denotes statistical significance at a 5% level
*** denotes statistical significance at a 1% level

The regression results using the full sample, the subsamples and the full sample with alterative hour measures consistently show that second-year performance in terms of course completion is not statistically significantly affected by the number of hours at work. This finding does not support the hypotheses that working while studying has a more or less negative effect on second-year course completion. There are several reasons for this with the first possibility

having a regard for causality. Those students who are retained into the second year tend to have better time management skills and care more about their academic achievement. Possessing such capacities help the students to choose the right number of hours to work while ensuring that work do not reduce the time and energy that should be spent on study. The second possibility is that second-year course completion may not capture overall course performance. Instead, second-year course letter grade may be a better indicator of the aspect of academic performance that can be influenced by work hours (see Section 4.2.2 for a detailed discussion). However, these variables cannot be controlled for due to their unavailability, which is a limitation of this thesis.

4.3.3. Decomposition analysis

The linear (Blinder-Oaxaca) decomposition method is used to summarise the abilities of the covariates to account for the overall ethnic gaps in this second-year course completion outcomes. The linear decomposition analysis also required the estimated coefficients from the pooled OLS regression on two comparison groups (i.e., European and Māori, European and Pasifika, and European and Asian). The entire set of decomposition results from the second-year university outcome between Māori/Pasifika/Asian and European based on prioritised ethnicity and all ethnicities are provided in Table A.4.10 in the Appendix.

Table 4.10. Key results on ethnicity from the linear decomposition of second-year course completion

	Prioritised ethnicity			All ethnicities		
	European vs Māori	European vs Pasifika	European vs Asian	European vs Māori	European vs Pasifika	European vs Asian
Total ethnic difference	0.1243	0.2485	0.0221	0.1202	0.2431	0.0189
Total explained differences	0.0311*** (0.0033) [25.02%]	0.0826*** (0.0059) [33.24%]	0.0077*** (0.0029) [34.84%]	0.0312*** (0.0032) [25.96%]	0.0813*** (0.0055) [33.44%]	0.0076*** (0.0028) [40.21%]
Unexplained differences	0.0932*** (0.0066) [74.98%]	0.1659*** (0.0087) [66.76%]	0.0144*** (0.0042) [65.16%]	0.0890*** (0.0066) [74.04%]	0.1618*** (0.0083) [66.56%]	0.0113*** (0.0041) [59.79%]
Number of observations (n)	22,902	22,338	26,469	24,270	23,841	27,924

This table only reports the overall decomposition results in the university outcome. The full set of decomposition results are provided in Table A.4.10 in the Appendix.

* denotes statistical significance at a 10% level
** denotes statistical significance at a 5% level
*** denotes statistical significance at a 1% level

4.3.3.1. European–Māori gap

In Table A.4.10, other parental factors include single-parent household composition, household income and parents' benefit or crime histories. As mentioned earlier, Māori students' second-year course pass rates on average were 12.34 percentage points behind European when using prioritised ethnicity. The key decomposition results for this ethnic difference can be found in the first column of Table 4.10. In the first instance, the results suggest that eliminating the ethnic differences in all covariates can account for 3.11 percentage points of the overall gap in the course pass rate, meaning that if we gave Māori the same measurable characteristics as European we could explain 25.02% of this gap in second-year course completion. This result is statistically significant at a 1% level. The results also suggest that Māori who have a second-year course pass rate that is still 9.32 percentage points lower than for their European peers. The proportion of the explained European-Māori gap in course completion is lower in the second-year than in the first-year, after removing the ethnic differences in the same factors.

According to the results in Table A.4.10 in the Appendix, the most important factors explaining the ethnic gap are, in order, school deciles (1.31 percentage points or 10.54%), part-time enrolment (1.08 percentage points or 8.69%), parents' highest level of education (0.57 percentage points or 4.59%), school qualifications (0.31 percentage points or 2.49%), other school factors (0.21 percentage points or 1.69%), prior activity (0.10 percentage points or 0.80%), age (0.08 percentage points or 0.64%) and cohort year 2011 and employment hours (both 0.01 percentage points or 0.08%). Some major changes should be noted to compare the similar results in the first-year. First, part-time enrolment individually explains a smaller proportion of the ethnic gap in the second year. Second, being part-time has more explanatory power in the ethnic gap in the second-year.

Assigning Māori and European the same female representation, university distribution, New Zealand citizenship and other parental factors together would widen the gap in the second-year course pass rate between the ethnic groups by 0.57 percentage points, or 4.59%. It is found that other parental factors change from a positive explanatory factor in the first year to a negative explanatory factor in the second year, although the magnitudes are not large.

The European-Māori gap in second-year course completion rates narrowed slightly to 12.02 percentage points, when using all ethnicities. The fourth column in Table 4.10 shows that if

Māori had the same covariates as European, this would reduce the ethnic gap in the second-year course completion rate by 3.12 percentage points, or 25.96%. The main factors closing the ethnic gap are school decile scores, part-time enrolment, parental educational qualifications, secondary school qualifications, other school factors, prior activity, age, cohort year of 2011 and other parental factors. Other parental factors returns to a positive explaining factor when using all ethnicities. Factors that increase the gap are female representation, university distribution and New Zealand citizenship.

4.3.3.2. European–Pasifika gap

We found earlier that the second-year course pass rate for Pasifika was 24.85 percentage points lower, relative to the rate of European when using prioritised ethnic designation. If we gave Pasifika and European the same observed characteristics we could explain 8.26 percentage points or 33.24% of the total ethnic gap (see the second column in Table 4.10). This finding is statistically different from zero at a 1% level.

The factors contributing to the most of the gap are part-time status (4.36 percentage points or 17.55%), school decile rankings (3.33 percentage points or 13.4%), parents' highest level of education (0.75 percentage points or 3.02%), school qualifications (0.47 percentage points or 1.89%), other parental factors (0.27 percentage points or 1.09%), other school factors/New Zealand citizenship (both 0.09 percentage points or 0.36%), prior activity (0.07 percentage points or 0.28%) and age/employment hours (both 0.01 percentage points or 0.04%). Compared to the results from the first year, the factors collectively could explain the slightly wider European-Pasifika gap in second-year course completion, where school deciles and part-time enrolment turn out to be the most important explanatory factors, with the proportion of the explained gap being increased by more than twofold in the second year.

Removing the ethnic differences in the remaining three factors (university distribution, female representation and cohort year 2011) would increase the European-Pasifika gap in the course completion rate by 1.19 percentage points, or 4.79%. Employment hours are no longer a negative explanatory factor in contrast to the results in the first year.

The all ethnicities European-Pasifika gap in second-year course completion was only slightly lower at 24.31 percentage points. Eliminating the differences in all the observable characteristics could close the ethnic gap by 8.13 percentage points, or 33.44% (see the fifth column in Table 4.10). Part-time enrolment is still the most relevant factor in reducing the ethnic gap between European and Pasifika. The second most important factor is school decile

scores, followed by parental education, the highest school qualification, other parental factors, other school factors, prior activity, New Zealand citizenship, age and work hours. Removing the ethnic differences in another three factors would increase the gap between European and Pasifika.

4.3.3.3. European–Asian gap

Prioritised Asian completed second-year courses at a rate of 2.21 percentage points below prioritised European. It can be seen from the results given in the third column of Table 4.10 that If Asian were given the same measured factors as European, we estimated that this would eliminate 0.77 percentage points or 34.84% of the ethnic gap in course completion. This result is statistically significant at a 1% level. The proportion of total explained ethnic gap decrease by about half in the second year compared to that in the first year.

The most important explanatory factors are ordered as part-time/New Zealand citizenship (both 0.42 percentage points or 19.00%), school deciles (0.24 percentage points or 10.86%), female representation (0.22 percentage points or 9.95%), other school factors/other parental factors (both 0.16 percentage points or 7.24%) and school qualifications (0.06 percentage points or 2.71%). Eliminating the differences in university distribution, parental education, prior activity, age and cohort year of 2011 between European and Asian would widen the ethnic gap by 0.91 percentage points, or 41.18%.

When using all ethnicities, the European-Asian gap in second-year course completion became 1.89 percentage points. Eliminating the differences in all the measurable factors between European and Asian, this could explain 0.76 percentage points or 40.21% of this ethnic gap in the course completion outcome (see the sixth column in Table 4.10). The primary factors to close the ethnic gap are part-time enrolment, New Zealand citizenship, female representation, school decile scores, other school factors, other parent-related factors and school qualifications. The factors widening the gap are university distribution, parents' highest level of qualification, prior activity, age and cohort year of 2011.

4.4. Qualification completion by age 24

4.4.1. Probit regression analysis

For qualification completion analysis, the sample is further restricted to the 2010 cohort and those enrolled in a three-year bachelor's qualification. Based on these students, qualification completion is defined as successfully completing a bachelor's qualification by age 24 (or the year 2019). The qualification being completed was not necessarily the same qualification as originally enrolled in due to the possibility that some students may change qualifications or switch universities during their studies. This qualification completion analysis used the same explanatory variables that were included in the previous two analyses, except for age and enrolment nature (e.g., part-time) being excluded (see Table 3.1 for the variables used).

Table A.4.11 in the Appendix provides the full set of results from Probit Maximum Likelihood estimations of qualification completion, with students being identified with both prioritised and all ethnicities. It is important to note that using the alternative ethnic designations does not make the regression results qualitatively different. The reported Pseudo R^2 statistic of 0.0926 suggests that the predictive power of the explanatory variables in this qualification completion analysis is stronger than the variables in the first-year course completion analysis, but weaker than the ones in the second-year course completion analysis. Female students on average are nearly 12 percentage points more likely than their male counterparts to complete a three-year bachelor's degree by age 24. This effect is statistically significant at a 1% level.

Compared to the omitted group of decile 6 schools, studying in deciles 1 to 4 schools is associated with a negative and statistically significant impact on the probability of completing a bachelor's qualification. The strongest negative effect found is that coming from a decile 1 school lower qualification completion by 21.07 percentage points, which is statistically significant at a 1% level. It was previously found that higher school decile scores (from 7 onwards) increased the probability of enrolling in university study, while participating in schools in the top two deciles (9 and 10) positively affected both first- and second-year course completion. However, the current regression results show that coming from deciles 7 to 10 schools does not statistically significantly lift the probability of qualification completion. These results together highlight that attending higher decile schools does not necessarily increase the likelihood of completing a three-year qualification, but coming from schools with lower decile scores statistically significantly reduces the probability of achieving this university outcome.

The type of school ownership has no statistically significant effects on the propensity to complete a three-year bachelor's programme. However, attending a single-gender school reduce the probability of the qualification completion by 1.73 percentage points, compared to studying in a coeducation school. This finding is different from zero at a 10% level. Attending schools located in minor urban areas increase the likelihood of degree completion by 3.23 percentage points, relative to studying in main urban area schools. This result is statistically significant at a 5% level. Coming from secondary urban or rural area schools appears to have no statistically significant impact on the probability of completing a three-year degree.

Similar to the results found in the previous analysis, the highest school qualification has persistently no statistically significant impact on the probability of qualification completion. Relative to never switching school, changing school twice or three times lowers the completion probability by about 8.5 percentage points. This effect is more than double (19.36 percentage points) for moving school 4 times or more. All these results are different from zero at a 1% level. Once again, truancy, suspension and/or unjustified absence have no statistically significant association with the completion outcome.

In terms of university factors, in comparison to being high school students, taking part in any other of the activities before enrolling in a qualification negatively and statistically significantly affect the likelihood of degree completion. Being employed have the strongest negative effect (nearly 29 percentage points) on the qualification completion outcome, while the weakest negative effect (11.30 percentage points) coming from being overseas.

Students coming from different universities tend to perform differently in terms of qualification completion. This finding is statistically significant at a 1% level, based on the Chi-square test. Having New Zealand citizenship increases the likelihood of qualification completion by 3.17 percentage points. This effect is different from zero at a 5% level.

Living in a single-parent household at the time of the 2013 Census lead to a 2.52 percentage-point reduction in the probability of graduating with a bachelor's degree. This result is statistically significant at a 10% level. Having one parent with a level 4 certificate decrease the likelihood of qualification completion by 4.19 percentage points, relative to having parents without an educational qualification, which is different from zero at a 10% level. This is the only statistically significant effect of parents' highest educational qualifications on qualification completion, although it was found to be an important explanatory factor in the analyses of the earlier three university outcomes. As for household income, living in a household with a $1 to

$30,000 income statistically significantly lower the likelihood of completing a three-year bachelor's degree by 4.13 percentage points, compared to coming from a household earning $30,001 to $50,000. The other household income ranges are found to have no statistically significant impact on the probability of degree completion. Lastly, parents' benefit and crime histories do not matter for the probability of completing a three-year bachelor programme.

Table 4.11. Key results on ethnicity from the Probit regression analysis for qualification completion by age 24

	Prioritised ethnicity			All ethnicities		
Variables	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Ethnicities						
Māori	-0.3940***	0.0392	-0.1414***	-0.3586***	0.0389	-0.1288***
Pasifika	-0.5998***	0.0507	-0.2153***	-0.5687***	0.0484	-0.2042***
Asian	-0.0009	0.0326	-0.0003	0.0195	0.0321	0.0070
MELAA	-0.2559***	0.0714	-0.0919***	-0.2477***	0.0712	-0.0889***
Other	-0.0528	0.1092	-0.0190	-0.0514	0.1065	-0.0185
Number of observations (n)	15,843			16,722		
Likelihood function	-9952.2624			-9952.4969		
Pseudo R^2	0.0926			0.0925		

This table only reports the results of identifying with ethnic minority groups for the university outcome. The full set of regression results are provided in Table A.4.11 in the Appendix, along with the full set of covariates that are held constant in this regression analysis.

* denotes statistical significance at a 10% level
** denotes statistical significance at a 5% level
*** denotes statistical significance at a 1% level

The key marginal effects of identifying with ethnic minority groups using both prioritised ethnicity and all ethnicities are given in Table 4.11. Māori students were 19.15 (using prioritised ethnicity) or 18.45 (using all ethnicities) percentage points less likely than European to complete a three-year qualification. The estimated marginal effect coefficients for Māori indicate that, if all other measurable factors are held constant, this ethnic difference would decrease to 14.14 percentage points when using prioritised ethnicity (a 26.16% decline) or to 12.88 percentage points when using all ethnicities (a 30.19% decline). The residual ethnic differences are both statistically significant at a 1% level. These numbers suggest that more than two-thirds of these overall gaps in degree completion between Māori and European remain, even after controlling for a wide array of background factors.

It was found that the qualification completion rate for Pasifika was 27.25 percentage points lower than that for European when using prioritised ethnicity. The difference was slightly lower at 27.23 percentage points, when using all ethnicities. The remaining difference declined to 21.53 (using prioritised ethnicity) or 20.42 (using all ethnicities) percentage points, closing 20.99% or 25.01% of the difference in qualification completion, when the considered categories of factors held constant. The differences that remained unexplained are different from zero at a 1% level.

Asian students had an average qualification completion rate that was 1.77 percentage points below European when using prioritised ethnicity. The difference became 1.21 percentage points when focusing on the students with multiple ethnicities. The remaining ethnic difference reduce to 0.03 percentage points when using prioritised ethnicity, which means that 98.31% of the difference could be eliminated if all the measurable factors were kept constant. More importantly, keeping everything else constant reverses the ethnic difference when using all ethnicities, which means that Asian are 0.70 percentage points more likely than their European peers to finish a qualification. Both these results in ethnic differences, however, are statistically insignificant at the conventional levels.

4.4.2. Impact of employment hours on qualification completion

Our earlier results showed that working during the academic months was associated with lower first-year course pass rates but had no statistically significant impact on second-year course completion performance. Compared to the course completions in the first and second years, qualification completion is an indicator that reflects long-run academic performance. Given the differences between the academic outcomes, we were now interested to know how employment affects qualification completion. The workload is measured as average weekly hours of work during the academic months of all the years enrolled at university to age 24 since the start of the degree qualification was considered in the participation analysis. As in the previous analyses, we used five dummy variables to identify students who working zero hours, low (between zero and less than 10 hours), medium (between 10 and less than 20 hours), high (20 or more hours) and those with unknown hours. Students zero working hours were set as the reference group.

The numbers working zero, low, medium, high and an unknown number of hours are comprised of 10.28, 5.72, 2.59, 14.47 and 66.96%, respectively, of the full sample. The

regression results in Table A.4.11 in the Appendix show that students working a low number of hours tend to be 7.20 percentage points more likely to complete a bachelor's qualification than those who had no job commitments. This effect is statistically significant at a 1% level. Working either a medium or a high number of hours have no measurable effect on qualification completion. These findings may indicate that working a small number of hours does not heavily crowd out the time spent on study during the university years; meanwhile it allows students to apply what they are being taught into classes in practice (e.g., Geel & Backes-Gellner, 2012; Quirk, Keith & Quirk, 2001; Scott-Clayton & Minaya, 2016). For example, in some of the universities in New Zealand, cooperative education provides final-year bachelor students opportunities to integrate the academic knowledge gained in university with the skills obtained from experiences in the workplace, which eventually improves the probability of graduating. One could also say that degree completion has a reverse causal effect on working a low number of hours. In other words, those successfully complete a qualification 'on time' are usually the ones who did not choose to work heavily while studying.

We reran the regression by dropping those with missing information on work hours. The proportion of the subsample working zero, low, medium and high hours is 31.14, 17.32, 7.80 and 43.75%, respectively. We find almost no difference between the results by using this subsample and the findings based on the full sample of students. Specifically, spending low hours at work increase the probability of completing a qualification by 7.19 percentage points, compared with working zero hours. This effect is also statistically significant at a 1% level. Neither working medium hours nor working high hours has a statistically significant impact on the degree completion outcome. These regression results are reported in Table A.4.13 in the Appendix.

If the analysis was replicated using only female students, we have a subsample of students that can be broken down into 9.95% working zero hours, 5.70% working a low number of hours, 2.51% working medium hours and 14.32% working a high number of hours, while 67.56% had an unknown number of working hours. The highly statistically significant positive effect (8.07 percentage points) of working a low number of hours on qualification completion continues to exist. Additionally, the results show that working a medium number of hours is 6.30 percentage points more likely for students to complete a three-year qualification than those not holding a part-time job. This effect is statistically significant at a 10% level. No statistically significant association is found between working high hours and qualification completion.

When restricting the analysis to male students, the subsample includes 10.76% of those working zero hours, 5.80% working low hours, 2.67% working a medium number of hours and 14.67% who worked a high number of hours, while 66.15% had an unknown number of hours worked. The results for male students are similar to the full sample results. Compared to working zero hours, spending a low number of hours in employment improve the probability of graduating by 6.12 percentage points, and this effect is statistically significant at a 5% level. Taking on medium and high employment hours both have no impact on the final stage of the outcome at university. It should be noted that work hours have different effects on qualification completion between female and male university students. In particular, working low hours have stronger and more statistically significant effect on female students' degree completion.

Table 4.12. Proportions of ethnic subsamples based on average employment hours during university years to age 24: prioritised ethnicity

	Sample proportion
Māori students	
Employment hours	
Zero	0.0894
Low (>0 and 10 hours)	0.0532
Medium (≥10 and <20 hours)	0.0213
High (≥20 hours)	0.1383
Unknown	0.6979
Pasifika students	
Employment hours	
Zero	0.1066
Low (>0 and 10 hours)	0.0439
Medium (≥10 and <20 hours)	0.0282
High (≥20 hours)	0.1630
Unknown	0.6583
Asian students	
Employment hours	
Zero	0.1027
Low (>0 and 10 hours)	0.0590
Medium (≥10 and <20 hours)	0.0240
High (≥20 hours)	0.1454
Unknown	0.6689
European students	
Employment hours	
Zero	0.1050

Low (>0 and 10 hours)	0.0575
Medium (≥10 and <20 hours)	0.0264
High (≥20 hours)	0.1432
Unknown	0.6682

This table only reports the proportions of the ethnic groups. The full set of proportions of all the subsamples are provided in Table A.4.12 in the Appendix.

The next analysis was repeated among the ethnic groups. Of the subsample containing both Māori and Pasifika students, 9.49% had zero, 5.06% had low, 2.53% had medium, 14.81% had high working hours, while 68.23% had unknown working hours. Surprisingly, we find that work hours have no effects on qualification completion for this particular subsample students (see Tables A.4.12 and A.4.13 in the Appendix). Also, such statistically insignificant correlations persist even when we separately looked at Māori (8.94% working zero hours, 5.32% working a low number of hours, 2.13% working a medium number of hours and 13.83% working high hours, with 69.79% having missing information), and Pasifika (10.66% who worked zero hours, 4.39% who worked a low number of hours, 2.82% who worked a medium number of hours and 16.30% who worked high hours, with 65.83% who had missing information) subgroups (see Tables 4.12 and 4.13). These results can also be interpreted as that working low hours has no detrimental or helpful impacts on the qualification completion of Māori and Pasifika ethnic groups.

When we concentrated on non-Māori and non-Pasifika groups (including Asian, European, MELLA and Other ethnicities), the subsample is comprised of 10.42% working zero hours, 5.83% working a low number of hours, 2.61% working a medium number of hours and 14.41% working high hours, while 66.67% had no data on work hours. Similar to the full sample results, spending a low number of hours in part-time employment statistically significantly increase the qualification completion probability by 7.40 percentage points, relative to having zero work hours, while working a medium or a high number of hours have no statistically significant impacts on achieving this university outcome.

If the analysis is confined to Asian or European students. The proportion having zero, low, medium, high and unknown work hours for the Asian subgroup is 10.27, 5.90, 2.40, 14.54 and 66.89%, respectively, while the equivalent numbers for the European subsample are 10.50, 5.75, 2.64, 14.32 and 66.82%. We find that working a low number of hours positively affected the degree completion propensities for both Asian and European, but the effect is stronger (10.31 percentage points) and less statistically significant (a 5% level) for Asian, while less strong (6.83 percentage points) and more statistically significant (a 1% level) for European.

There is no empirical evidence that working a medium or high number of hours has a measurable impact on qualification completion for Asian and European. One conclusion on the analysis results related to the ethnic groups is that the effect of holding a part-time job while studying on degree completion varies between ethnicities.

Lastly, we ran the regression by using the full sample and the redefined variables for work hours. The new thresholds for low, medium and high work hours respectively were between zero and less than 15 hours, between 15 and less than 30 hours, and 30 hours or more. The proportions of the sample having zero, low, medium, high hours and missing work hours are 10.28, 7.12, 3.03, 12.59 and 66.96%, respectively. Compared to the students not engaged in part-time employment, those working low hours are 6.25 percentage points more likely to graduate. This effect is statistically significant at a 1% level. The positive effect reduces to 4.78 percentage points for working a medium number of hours and is different from zero at a 10%. The impact of working high hours on degree completion again is not statistically significant at the conventional test levels. Putting together the results based on the full samples while applying the alternative measures to the work hours variables, suggests that although working no more than 30 hours a week during the university years has positive effect on qualification completion, the magnitude and statistical significance of the effect declines as the work hours increase within the time limit.

Table 4.13. Key results on the ethnicity from the Probit regression analysis for employment hours on qualification completion

	Prioritised ethnicity			All ethnicities		
	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Māori students						
Employment hours						
Low (>0 and 10 hours)	0.2529	0.1969	0.0840	-	-	-
Medium (≥10 and <20 hours)	0.3481	0.2735	0.1156	-	-	-
High (≥20 hours)	-0.0125	0.1570	-0.0042	-	-	-
Unknown	0.0614	0.1295	0.0204	-	-	-
Pasifika students						
Employment hours						
Low (>0 and 10 hours)	0.0666	0.2570	0.0198	0.1364	0.2522	0.0400
Medium (≥10 and <20 hours)	0.2209	0.3039	0.0655	0.1893	0.3005	0.0556
High (≥20 hours)	0.0209	0.1869	0.0062	0.1265	0.1800	0.0371
Unknown	-0.0467	0.1545	-0.0139	0.0254	0.1494	0.0075
Asian students						

Employment hours						
Low (>0 and 10 hours)	0.2817**	0.1307	0.1031**	0.3023**	0.1301	0.1107**
Medium (≥10 and <20 hours)	-0.0157	0.1765	-0.0057	-0.0045	0.1753	-0.0017
High (≥20 hours)	0.0134	0.1022	0.0049	0.0136	0.1013	0.0050
Unknown	0.0112	0.0837	0.0041	0.0121	0.0831	0.0044
European students						
Employment hours						
Low (>0 and 10 hours)	0.1899***	0.0672	0.0683***	0.1870***	0.0653	0.0674***
Medium (≥10 and <20 hours)	0.0605	0.0898	0.0218	0.0829	0.0868	0.0299
High (≥20 hours)	0.0587	0.0523	0.0211	0.0391	0.0507	0.0141
Unknown	0.0487	0.0427	0.0175	0.0309	0.0415	0.0111

This table only reports the results of the ethnic subsamples. The full set of regression results for all the subsamples are provided in Table A.4.13 in the Appendix.

* denotes statistical significance at a 10% level
** denotes statistical significance at a 5% level
*** denotes statistical significance at a 1% level

Contrasting the degree completion analysis results to those in the course completion analyses, working a low number of hours reduces first-year course completion rates and has no measurable impact on the course pass rate in the second year, but it increases the likelihood of completing a degree qualification. These findings together suggest that part-time work is associated with students' continued enrolment at university.

4.4.3. Decomposition analysis

The whole set of nonlinear (Fairlie) decomposition results in qualification completion between the main ethnic minority groups (Māori, Pasifika and Asian) and European, using both prioritised and all ethnicities, are summarised in Table A.4.14 located in the Appendix. As with the participation analysis, this decomposition also used the estimated coefficients from the pooled Probit regression on each pair of related comparison groups (i.e., European and Māori, European and Pasifika, and European and Asian). The results are also based on conducting 100 replications of the decomposition procedure.

Table 4.14. Key results on ethnicity from the nonlinear decomposition of qualification completion by age 24

	Prioritised ethnicity			All ethnicities		
	European vs Māori	European vs Pasifika	European vs Asian	European vs Māori	European vs Pasifika	European vs Asian
Total ethnic difference	0.1915	0.2725	0.0177	0.1845	0.2723	0.0121
Total explained differences	0.0452*** (0.0051) [23.60%]	0.0676*** (0.0106) [24.81%]	0.0146** (0.0070) [82.49%]	0.0462*** (0.0050) [25.04%]	0.0737*** (0.0097) [27.07%]	0.0144** (0.0068) [119.01%]
Unexplained differences	0.1463 [76.40%]	0.2049 [75.19%]	0.0031 [17.51%]	0.1383 [74.96%]	0.1986 [72.93%]	-
Number of observations (n)	11,640	11,184	12,972	12,372	12,003	13,755

This table only reports the overall decomposition results in the university outcome. The full set of decomposition results are provided in Table A.4.14 in the Appendix.

* denotes statistical significance at a 10% level
** denotes statistical significance at a 5% level
*** denotes statistical significance at a 1% level

4.4.3.1. European–Māori gap

It was found earlier that prioritised Māori, on average, were 19.15 percentage points less likely to complete a three-year degree than prioritised European. The contributions of each individual and category of factors in the overall difference in degree completion are given in the first column of Table A.4.14. It has been stated that single-parent household composition and parents' benefit and crime histories are grouped as other parental factors. The primary results reported in Table 4.14 indicate that if we gave Māori the same observed factors as European we could explain 4.52 percentage points or 23.60% of the gap in qualification completion between the ethnic groups. This result is statistically significant at a 1% level. The finding also suggests that Māori students are still 14.63 percentage points less likely to graduate from a three-year degree than comparable European, even after removing the differences in the factors included in the regression.

The most important factors explaining the European-Māori gap in completing a three-year degree are, in order, school deciles (2.45 percentage points or 12.79%), prior activity (1.60 percentage points or 8.36%), other school factors (0.77 percentage points or 4.02%), the highest secondary school qualification (0.69 percentage points or 3.60%) and parents' highest level of qualification (0.46 percentage points or 2.40%). Factors that widen the ethnic gap are

female representation, other parental factors, university distribution and New Zealand citizenship status. In particular, removing the ethnic differences in these four factors together would increase the qualification completion gap between European and Māori by 1.45 percentage points, or 7.57% (see Table A.4.14 in the Appendix).

The degree completion gap between European and Māori was 18.45 percentage points, when using all ethnicities. The results in the fourth column of Table 4.14 show that the total explaining European-Māori gap in degree completion increased to 4.62 percentage points or 25.04%. The primary factors that explain this ethnic gap are school decile scores, prior activity, other school factors, school qualifications, parental education attainment and employment hours. The gap would become wider if we eliminated the ethnic differences in female representation, other parental factors, New Zealand citizenship status and university distribution.

4.4.3.2. European–Pasifika gap

The qualification completion rate for prioritised Pasifika was 27.25 percentage points below the rate of prioritised European. It can be seen in the second column of Table 4.14 that eliminating the ethnic differences in all observable factors accounts for 6.76 percentage points of the European-Pasifika gap. This suggests that if we gave Pasifika students the same measured covariates as European students, we could close 24.81% of the ethnic gap. This total explained difference is statistically different from zero at a 1% level. It can be seen that accounting for the differences in the covariates in the regressions can explain similar proportions of the European-Māori and European-Pasifika gaps.

The primary factors contributing to the ethnic gap are ordered as school deciles (5.46 percentage points or 20.04%), school qualifications (1.50 percentage points or 5.50%), other school factors (0.47 percentage points or 1.72%), prior activity (0.21 percentage points or 0.77%), parental education (0.10 percentage points or 0.37%), New Zealand citizenship (0.09 percentage points or 0.33%), employment hours (0.03 percentage points or 0.11%) and other parental factors (0.02 percentage points or 0.07%). The negative explaining factors are female representation and university distribution. In particular, if Pasifika and European had the same female representation and university distribution, this would widen the ethnic gap in degree completion by 1.12 percentage points, or 4.11%.

Using all ethnicities, the European-Pasifika gap in completing a degree was 27.23 percentage points. The overall explainable gap in qualification completion between European and Pasifika

increases to 7.37 percentage points or 27.07% (see fifth column of Table 4.14). The most important factors narrowing the gap are school decile scores, school qualifications, other school factors, prior activity, other parental-related factors, New Zealand citizenship, hours spent in employment and parents' highest qualifications. Female representation and university distribution are the factors that increase the ethnic gap.

4.4.3.3. European–Asian gap

We previously found that Asian students were 1.77 percentage points less likely than their European peers to complete a three-year degree using prioritised ethnicity. The key decomposition results are given in the third column of Table 4.14. If Asian students had the same measurable factors as those for European students, this could eliminate 1.46 percentage points or 82.49% of the gap in degree completion between the ethnic groups. This finding is different from zero at a 5% level.

Different to Māori and Pasifika, the factor explaining the largest proportion of the gap in qualification completion between European and Asian is New Zealand citizenship status (0.80 percentage points or 45.20%). The second most important factor is other parental factors (0.76 percentage points or 42.94%), followed by female representation /school decile scores (both 0.66 percentage points or 37.29%), university distribution (0.47 percentage points or 26.55%) and school qualifications (0.33 percentage points or 18.64%). Eliminating the ethnic differences in the remaining four factors including prior activity, parental education, other school factors and employment hours, would together dramatically widen the gap by 2.22 percentage points, or 125.42%.

The European-Asian gap in qualification completion was 1.21 percentage points when using all ethnicities. The main decomposition results for this ethnic gap can be found in the sixth column of Table 4.14. It is highly important to note that the total ethnic gap in qualification completion is overexplained (1.44 percentage points or 119.01%). This result indicates that if we gave the same observable factors to Asian and European, the former group would be 0.23 percentage points more likely than the latter one to complete a three-year degree qualification. The most important factors contributing to the qualification completion gap are, in order, other parental factors, New Zealand citizenship status, school deciles, female representation, university distribution, school qualifications and other school factors. Factors increasing the gap are activity engaged in before enrolling in a degree qualification and parental education.

4.5. Conclusions

The gap between European and Māori was 22.30 percentage points in university participation by age 21, 13.14 percentage points in first-year course completion, 12.43 percentage points in second-year course completion and 19.15 percentage points in degree completion by age 24. It can be seen that European-Māori gaps in the academic outcomes are relatively bigger in the beginning and at the end of the university journey.

Our regression results suggest that holding constant other personal characteristics, school qualifications, school and parental factors could eliminate more than three quarters of the European-Māori in university participation. For those individuals who enrolled in university, no more than one half of this ethnic gap in (first and second year) course completions could be eliminated, once other personal characteristics, school qualifications, school, university and parental factors were held constant. Finally, only about one quarter of this ethnic gap in qualification completion could be closed, ceteris paribus. These results suggest that the proportions of these European-Māori gaps could be eliminated by holding the observed factors constant decline as we move from university participation to qualification completion.

The European-Pasifika gaps in university participation, first-year course completion, second-year course completion and degree completion were 13.48, 26.07, 24.85 and 27.25 percentage points, respectively. The numbers indicate that the achievement gaps between European and Pasifika students become wider from participation to degree completion.

Keeping constant the measurable factors could eliminate almost the entire European-Pasifika gap in university participation. Less than half of this ethnic gap in course completions could be eliminated once the relevant factors were controlled for. Holding other observed characteristics constant, we could close no more than one quarter of the gap in degree completion between the ethnic groups. Similarly to Māori, the percentages of explained European-Pasifika gaps in the university outcomes decreased over time.

The gap in university participation between European and Asian was –19.33 percentage points. European-Asian gaps in first-year course completion, second-year course completion and degree completion respectively are 2.01, 2.21 and 1.77 percentage points. As shown by the numbers, the Asian participation rate is higher than that of European (negative gap), and the positive gaps grow narrower from course completions to degree completion.

Holding constant the measurable factors could reduce the European-Asian gap in university participation by 70%. Less than 70% and more than 30% of these ethnic gaps in first and second year course completions, respectively, could be eliminated if adjusting for the observed covariates. In the end, approximately the whole gap in degree completion between these ethnic groups could be explained when keeping other factors constant. Unlike Māori and Pasifika, higher proportions of these European-Asian gaps could be eliminated when moving from participation to degree completion.

A prevailing concern is that employment while studying may lead to detrimental effects on academic achievement. Our regression results show that the hours spent in employment while studying had negative effects on course pass rates in the first year. However, the effect does not become stronger with the number of hours worked. For example, working either low and high hours statistically significantly reduce first-year course completion rates by 1.74 and 0.79 percentage points, respectively, while working a medium number of hours have no measurable effect on the course completion outcome. We find that these employment effects also vary by gender, enrolment nature, and ethnicity.

Surprisingly, our results indicate that holding a job while studying has no statistically significant effect on course pass rates in the second year. We, do however, find that working a low number of hours have a positive effect on degree completion. For example, spending a low number of hours at work statistically significantly increase the probability of completing a degree qualification by 7.20 percentage points. Working a medium or high number of hours are both associated with an statistically insignificant effects on degree completion. The employment effects on degree completion also vary by gender and ethnicity. These results together reveal that part-time employment may allow students to continue studying at university and to apply their academic learning into practice (e.g., cooperative education), which consequently increases the probability of degree completion.

This analysis of the impact of employment while studying on university outcomes has potential limitations. First, we approximated work hours, which may have been under or over-estimated. Second, we included a broad range of factors in our regressions, but we had no access to other potentially important variables, such as course letter grades, study attitudes and time management skills. A detailed discussion about research limitations is provided in Section 5.4. Given the inconsistent results, further work should try to isolate the one-way causal effects of employment while studying on academic outcomes at university.

Chapter 5. Discussion and conclusion

5.1. Introduction

This thesis has provided empirical evidence on the factors contributing to differences in a variety of academic outcomes at university between the main ethnic minority groups (Māori, Pasifika and Asian) and European in New Zealand using a combination of administrative and survey data.

Reviews of the relevant international and New Zealand literature was provided in Chapter 2. Students from ethnic minority groups, on average, have poorer academic outcomes at university compared to students from the dominant ethnic group in many countries. For example, in the US, African American and Hispanic students were less likely to participate in and graduate from universities than White students (e.g., Clotfelter, Ladd & Vigdor, 2015; Fletcher & Tienda, 2009). Findings from UK studies have shown that although Black and Caribbean students did not always have a lower university participation rate than equivalent White students, they were still less likely to complete their qualifications (e.g., Jackson, 2012; Rodgers, 2013).

In the context of New Zealand, Māori and Pasifika university students had worse outcomes in terms of university participation, GPA achievement and degree completion compared to their European counterparts (e.g., Cao & Maloney, 2018; Meehan, Pacheco & Pushon, 2019; Scott & Smart , 2005). It was found that employment while studying had different effects on educational achievement. This may be because the effects of working while studying on university performance are subject to the types of academic outcomes, the analysis methods used, the country of analysis, job characteristics and student characteristics (Neyt et al., 2019). This thesis extends the literature, especially Cao & Maloney (2018), by examining ethnic differences in academic outcomes at different stages of the university journey by using linked longitudinal data.

Chapter Three discussed the data and methodologies used in this project. Anonymous unit-level data stored in the Integrated Data Infrastructure (IDI) were used for this thesis. The IDI is a large database that is maintained and operated by Statistics New Zealand. The IDI holds data from government agencies, surveys conducted by Statistics New Zealand and non-government organisations. The data used in this thesis came from five sources: the Ministry of Education, Inland Revenue, the Ministry of Justice, the Ministry of Social Development and

the 2013 Census. These data can be linked by using a personal unique identifier variable (snz_uid), created by Statistics New Zealand.

The use of the IDI data allowed us to look at four university outcomes, including enrolling in a bachelor's qualification by age 21, first- and second-year course completions, and completing a bachelor's qualification by the age of 24, while controlling for a wider range of factors, such as individual demographic characteristics, school characteristics, parent-related backgrounds and university factors. The initial study sample included nearly 180,000 individuals who turned 15 years old in the years 2010 to 2012 while they were studying in a New Zealand school. The sample was split into three cohorts, based on the year that the students turned 15 years old. Different cohorts of students were used at different stages of the analysis due to data unavailability. In terms of methodology, Probit, Fractional Probit and formal decomposition techniques were utilised to answer the research questions mentioned in the introduction.

Chapter 4 reported our findings, which are summarised in the next section (Section 5.2) of this chapter. Implications, research limitations and possible further research directions are then discussed, before the conclusions are presented for this thesis.

5.2. Overall findings

University participation and first-year course completion analyses were based on all three age-15 cohorts (2010 to 2012). Second-year course completion analysis was restricted to the 2010 and 2011 cohorts and to this year at university. The sample for degree completion analysis was further restricted to the 2010 cohort and those enrolled in a three-year bachelor's programme. These sample restrictions were because of data unavailability.

Due to the focus on ethnic differences in university performance, ethnicity was the key independent variable for this thesis. We measured student ethnicity using both ethnicity prioritisation rules and all self-reported ethnicities. It is important to note that our analysis results are only slightly sensitive to the ways of measuring ethnicity.

Based on prioritised ethnicity, the sample breakdown, in descending order, was European (57.05%), Māori (20.97%), Asian (10.30%), Pasifika (9.14%), Middle Eastern/Latin American/African (1.76%) and Other (0.79%) (see footnote 9 for the sample breakdown when using all ethnicities). It is worth mentioning that our sample contained a higher proportion of Pasifika and Asian students, and about the same proportion of Māori students, compared to

the sample used by Meehan, Pacheco & Pushon (2019) taken from cohorts born between 1990 and 1994.

The mean university participation rate for European was 35.38%, whereas the rate for Māori was only 13.08%, which left a gap of 22.30 percentage points in the academic outcome between the ethnic groups. Similar gaps in first- and second-year course completions dropped to 13.14 (European with 83.71%, Māori with 70.57%) and 12.43 percentage points (European with 87.43%, Māori with 75.00%), respectively. The European-Māori gap in degree completion was slightly higher at 19.15 percentage points (European with 56.09%, Māori with 36.94%).

Our decomposition results showed that eliminating the differences in the measured covariates between European and Māori could narrow this ethnic gap by more than 85% in university participation. We could explain less than one third of these ethnic gaps in course completions and only about one quarter of the gap in degree completion. This suggests that increasing proportions of these Māori-European gaps are unexplained as we move to later university outcomes.

The most important factors that explain the lower university participation rate of Māori relative to European were highest school qualifications, school deciles and parents' highest level of education. The factors in explaining the ethnic gaps in course completions were school deciles, parental educational attainment, part-time enrolment, school qualifications. The factors in explaining the degree completion gap were school deciles, prior activity, other school factors and school qualifications.

The European-Pasifika gap increased from 13.48 percentage points (35.38% for European, 21.90% for Pasifika) in university participation to 26.07 percentage points (83.71% for European, 57.64% for Pasifika) in first-year course completion, and 24.85 percentage points (87.43% for European, 62.58% for Pasifika) in second-year course completion, and finally 27.25 percentage points (56.09% for European, 28.84% for Pasifika) in degree completion. Compared to Pasifika, Māori were less likely to enrol in university but more likely to pass courses in the first two years and to complete their degrees. Giving Pasifika the same observed factors as European, we could explain nearly the whole of the gap in university participation between European and Pasifika students. The total explained gaps in course completions and degree completion decreased to approximately one third and about one quarter, respectively. Our ability to explain the gaps between Māori or Pasifika and European erodes as we move to later academic outcomes at university.

The primary contributors to the European-Pasifika gap in university participation were school qualifications, school deciles and parental educational attainment. The contributors to the European-Pasifika gap in course completions were school deciles, part-time enrolment, parental education attainment and school qualifications. The factors that contribute to the European-Pasifika gap in degree completion were school deciles, the highest school qualification, other school factors and prior activity.

The gap in university enrolment between European and Asian was –19.33 percentage points (35.38% for European, 54.71% for Asian). This ethnic gap in first-year course completion was 2.01 percentage points (83.71% for European, 81.70% for Asian). The gap increased slightly to 2.21 percentage points (87.43% for European, 85.22% for Asian) in second-year course completion, but dropped to 1.77 percentage points (56.09% for European, 54.32% for Asian) in qualification completion. If Asian and European students were assigned the same characteristics, this could eliminate about 65% of this ethnic gap in university participation. The total explained gap decreased to more than 60% in first-year course completion and less than 35% in second-year course completion, but increased to more than 80% in qualification completion. If using all ethnicities, we could overexplain the gap in qualification completion between the ethnic groups.

The primary factors contributing to the European-Asian gap in university participation were highest secondary school qualifications, other school factors and parental education. The factors contributing to this ethnic gap in first-year course completion were New Zealand citizenship, other school factors, female representation/school deciles and enrolment status. The factors explaining this ethnic gap in second-year course completion were New Zealand citizenship status/part-time enrolment, school deciles, female representation and other parental/school factors. New Zealand citizenship status, other parental factors, female representation/school deciles and university distribution were seen as the main contributors to this ethnic gap in degree completion.

Our results generally are in line with the findings of Meehan, Pacheco & Pushon (2019), who found that eliminating ethnic differences in individual characteristics, socioeconomic status, school factors and parents' highest qualifications between Māori, Pasifika or Asian relative to European can explain approximately 87, over 100 and about 47% of these ethnic gaps, respectively, in enrolling in a bachelor's degree. In this thesis, we could explain about one quarter of the gap in degree completion rates between Māori or Pasifika and European. The proportion of explained gaps, however, was smaller relative to the similar ones found in Meehan, Pacheco & Pushon (2017) who used first-year course pass rates as one of the

independent variables while excluded all school academic performance from their analysis, which artificially assigned a greater expandatory power to the first-year outcome, which eventually biases the explained gaps upwards. For example, those achieve similar academic outcome in the first year tend to equally likely to complete their qualification.

5.3. Implications

Overall, the findings of this thesis provide a deeper understating of the factors that contribute to the differences in university academic performance between the primary ethnic minority groups and European. Our analytical results confirmed the relevant literature about ethnic differences in university success. Māori and Pasifika students had much lower levels of academic achievement than did European, from the beginning to the end of the university journey. The findings of this thesis may be helpful to policymakers and university administrators who intend to reduce ethnic differences in university-level education.

Our decomposition results revealed that most of the underachievement in university participation by Māori and Pasifika compared to European students could be eliminated if they had the same observed characteristics, while the main explanatory variables are the highest school qualification obtained, the decile score of the school last attended and parents' highest level of education attainment. These findings are consistent with Meehan, Pacheco & Pushon (2019), who highlighted the importance of socioeconomic status, school performance and parental education in determining the probability of undertaking bachelor-level study. Hence, our findings suggest that efforts to help Māori and Pasifika students should be mainly focused on improving their academic performance in school and the quality of schools in the lowest deciles. There also needs to an awareness that parental education has a substantial effect on the students' probability of participating at university. This is an important intergenerational effect. Although there is not much that can be done to change parental educational attainment for the current generation, improving educational achievement for the current generation can change future generations' participation rates at university.

It is important to recall that Cao & Maloney (2018) found that if Māori or Pasifika had the same individual characteristics, school factors and university enrolment patterns, this would reduce similar ethnic gaps in first-year course completion rates and GPAs by no more than one quarter. In the current thesis, we found school decile, parental education and enrolment status to be the most import factors in explaining the gaps between Māori/Pasifika and European in first- and second-year course completion, while school decile, university factors and other school factors are the key factors contributing to the ethnic gaps in qualification completion.

However, eliminating differences even in a wider range of observed factors (compared to the factors used in Cao & Maloney (2018)), we could still only explain no more than one-third of the Māori/Pasifika-European gaps in first- and second-year course completions, and degree completion. These findings suggest that the residual ethnic gaps are probably the result of discrimination or lack of other important explanatory variables.

Our analysis showed that Asian were more likely to commence a bachelor's qualification at university than European students, but that they lagged behind European after enrolment. Similarly to Māori and Pasifika, school qualifications and parental education are the main factors driving the difference in university participation between European and Asian students. School deciles tended to be a less important explanatory variable for the European-Asian gap. Again, the previous study, Cao and Maloney saw that removing differences in the observable factors reduced about half of the gap in first-year course completion rates and one third of the gap in GPA between European and Asian. The current analyses showed that if Asian had the same measurable factors as European, this would explain more than half of the ethnic gaps in first-year course and degree completion, and more than one third of the gap in second-year course completion. Being a New Zealand citizen was found to be the key factor that contributed to the European-Asian gaps in these post-participation university outcomes. For Asian students, holding New Zealand citizenship may imply that they grew up in the same cultural environment as European students and, more importantly, that they are fluent in English. As mentioned by Juhong & Maloney (2006), having English as a second language greatly affected Asian students' academic performance in university. Thus, the findings of the current thesis indicate that interventions intending to close the European-Asian gap in course and qualification completions should aim at removing English language barrier for Asian university learners.

5.4. Research limitations

This thesis has demonstrated the strength of using data from a large research database to examine the ethnic differences in a variety of academic outcomes in university education at the national level. Our findings add useful new insights into the literature on this topic. However, there a number of limitations that should be borne in mind. Some are related to data and others are connected with methodology. Of these data limitations, even though we could control for a broader range of variables, the data lacked some important factors. For example, we knew whether or not a course had been successfully completed, but no information was available on the course letter grade. Other relevant variables that could have been used in our analysis if they were available include motivation for attending university, study attitude and

peer effects, etc. It is likely that more of the substantial unexplained differences in course and qualification completions between Māori/Pasifika and European could have been eliminated if the mentioned unavailable variables were controlled for. As mentioned in Section 4.2.2, the hours spent on wage and salary work were not given in the tax data which forced us to roughly calculate the hours worked by using the available information and assuming the students were paid minimum wage. As a result, the true work hours of some students might be either under- or overestimated.

Second, some of the variables used in our analysis had incomplete data. For example, information on school qualifications was unavailable for slightly more than 20% of the students used in the participation analysis. Only about 72% of the students' parent(s) could be identified due to restricted access to the 2006 Census data. However, we chose not to drop the observations with variables having missing data. This is because the missing values on the variables may not be randomly distributed across students, thus simply dropping these observations could reduce sample size and bias the estimated results. Third, variables such as household income and last school deciles were measured at one time point, which was different from when we actually wanted them to be measured. The ideal measure for such variables would be the average values over a reasonable interval of time. For example, school deciles should be measured by the mean deciles of all high schools attended; similarly, the average earnings of household members between the ages six to sixteen of the student(s) should be used to measure household income. Nevertheless, with the data currently accessible, there is not much we could do to improve the measurement of these variables.

In terms of methodological limitations, we acknowledge that our study sample was confined only to bachelor students at universities, although other types of tertiary organizations, such as institutes of technology, polytechnics and wānanga also offer bachelor-level education. The concentration on universities is because they are the preeminent source of post-school qualifications and the estimated returns on university qualifications are generally found to be substantially greater than those from other tertiary institutions. It is unlikely that our findings can be generalised to students from other tertiary education providers in New Zealand.

Second, it should also be noted that the qualification completion analysis was restricted to students who enrolled in a three-year bachelor's programme, given that school data were unavailable for before 2008 and tertiary data after 2019 were not accessible. One could argue that students who undertook a bachelor's programme requiring longer years to complete may not have altered their university performance much by ethnicity because students admitted into those programmes who usually are more academically prepared, regardless of ethnicity.

One could also argue that the ethnic differences could be even larger in programmes such as law, medicine and engineering. Therefore, it is unsure whether our findings provide reliable information about understanding ethnic differences in qualification completion to all bachelor-level students enrolling in different programmes. Despite the existence of these limitations, this thesis would have been impossible without using the data in the IDI.

5.5. Recommendations for further research

There are several potential areas for future research that are beyond the scope of the current project. The first of these would be to include course letter grades (GPAs) in analysis. To do this, it would require integrating data on course GPAs provided by universities with the IDI. This process, however, could take months or years to complete. If this has been done, GPA could be used as a dependent variable. For example, apart from course completion, GPA would be another outcome variable to be looked at during first- and second-year university study. A main benefit of doing this would allow us to verify our expectation that working 20 hours or more a week while studying has a stronger detrimental impact on university performance, because course grades are believed to be better indicators of overall academic performance. However, it has been argued that course GPAs sometimes do not provide a reliable measure of academic achievement that students gain at university (Johnson, 2008).

Second, in addition to students at universities, the study sample could be expanded to cover those undertaking a bachelor's qualification at other types of tertiary education institutions (institutes of technology, polytechnics and wānanga). There are different student bodies at universities compared to non-university institutions. Specifically, non-university institutions usually admit relatively more students from ethnic minority groups and disadvantaged socioeconomic backgrounds; these students are on average less academically ready for bachelor-level study. Thus, considering students from both university and non-university organisations could provide us with the whole picture of ethnic gaps in tertiary education.

Third, if we would have access to later data some of the issues mentioned regarding data unavailability will have been solved. We have been told that the Inland Revenue Department has started collecting data on the work hours of taxpayers since the end of 2019. We believe that once the actual information related to employment hours can be accessed, the potential measurement errors on the self-estimated hours will be avoided. As a consequence, our analysis will yield more reliable results about the impact of term-time employment on university performance. As already mentioned, information on parental backgrounds was unavailable for about 30% of the students and the analysis on degree completion was based on only those

enrolling in a three-year degree qualification. The proportion of students with missing parental information could reduce to within 10% or less, and those undertaking a degree requiring a longer completion time would be included in the analysis, if we had tertiary data in the IDI available up to the year 2023 or later. In this way, the results could undoubtedly be generalised to the entire population of bachelor students at university education in New Zealand or in other similar OECD countries (e.g., Australia).

5.6. Conclusion

Ethnic differences in university success have received a great deal of public attention. In this thesis, we found that Māori and Pasifika lagged behind their European counterparts in terms of participation, first- and second-year course completion rates, and qualification completion in bachelor-level study across all universities in New Zealand. Compared with European, Asian tended to have lower course completion rates and be less likely to complete a three-year degree although they attended university at a higher rate. Decomposition results highlighted to what extent the factors can explain the underperformance of the ethnic minority groups relative to European. These findings have important implications for closing the ethnic gaps in the academic outcomes in university education. Overall, our findings suggest for earlier interventions to improve high school academic performance and quality of school in the lowest deciles for Māori and Pasifika students, by doing so enhance their university performance. The decile of the school last attended and pre-university achievement appear to be relatively less important for Asian students. To support this group of students, actions should focus on removing the English language barrier. In terms of the unexplained ethnic differences, future analysis is needed to identify whether the remaining ethnic differences are the result of discrimination.

References

Ahlburg, D. A., McCall, B. P., & Na, I. G. (2002). Time to dropout from college: A hazard model with endogenous waiting. *Unpublished paper. Industrial Relations Center, University of Minnesota*.

Allen, J., & Robbins, S. B. (2008). Prediction of college major persistence based on vocational interests, academic preparation, and first-year academic performance. *Research in Higher Education*, *49*(1), 62-79.

Arulampalam, W., Naylor, R. A., & Smith, J. (2012). Am I missing something? The effects of absence from class on student performance. *Economics of Education review*, *31*(4), 363-375.

Becker, G. S. (1993). *Human capital: A theoretical and empirical analysis, with special reference to education,* 3rd edition. University of Chicago press.

Beffy, M., Fougère, D., & Maurel, A. (2010). *The effect of part-time work on post-secondary educational attainment: New evidence from French data* (No. 5069). IZA Discussion Papers.

Benseman, J., Coxon, E., Anderson, H., & Anae, M. (2006). Retaining non-traditional students: Lessons learnt from Pasifika students in New Zealand. *Higher Education Research and Development, 25*(2), 147–162.

Bijsmans, P., & Schakel, A. H. (2018). The impact of attendance on first-year study success in problem-based learning. *Higher Education*, *76*(5), 865-881.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, 436-455.

Boliver, V. (2013). How fair is access to more prestigious UK universities?. *The British Journal of Sociology*, *64*(2), 344-364.

Boven, N., Exeter, D., Sporle, A., & Shackleton, N. (2020). The implications of different ethnicity categorisation methods for understanding outcomes and developing policy in New Zealand. *Kōtuitui: New Zealand Journal of Social Sciences Online*, *15*(1), 123-139.

Boyd, S., Chalmers, A., & Kumekawa, E. (2001). *Beyond school: Final year school students' experiences of the transition to tertiary study or employment.* Wellington: New Zealand Council for Educational Research.

Bozick, R. (2007). Making it through the first year of college: The role of students' economic resources, employment, and living arrangements. *Sociology of Education*, *80*(3), 261-285.

Brahm, T., Jenert, T., & Wagner, D. (2017). The crucial first year: a longitudinal study of students' motivational development at a Swiss Business School. *Higher Education*, *73*(3), 459-478.

Brawley, S. (1993). 'No" White Policy" in NZ': Fact and Fiction in New Zealand's Asian Immigration Record, 1946-1978. *New Zealand Journal of History*, *27*(1), 16-36.

Buscha, F., Maurel, A., Page, L., & Speckesser, S. (2012). The effect of employment while in high school on educational attainment: A conditional difference-in-differences approach. *Oxford Bulletin of Economics and Statistics*, *74*(3), 380-396.

Callender, C. (2008). The impact of term-time employment on higher education students' academic attainment and achievement. *Journal of Education Policy*, *23*(4), 359-377.

Cappellari, L., & Lucifora, C. (2009). The "Bologna Process" and college enrollment decisions. *Labour Economics*, *16*(6), 638-647.

Cao, Z., & Maloney, T. (2018). Decomposing ethnic differences in university academic achievement in New Zealand. *Higher Education*, *75*(4), 565-587.

Cingano, F., & Cipollone, P. (2007). University drop-out. the case of Italy, Temi di discussione (Economic working papers) 626, Bank of Italy. *Economic Research Department*.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2015). *Public universities, equal opportunity, and the legacy of Jim Crow: Evidence from North Carolina* (No. w21577). National Bureau of Economic Research.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederick D. Weinfeld, and Robert L. York. 1966. *Equality of Educational Opportunity.* Washington: U.S. Department of Health, Education, and Welfare, Office of Education.

Connor, H., Tyers, C., Modood, T., & Hillage, J. (2004). Why the difference? A closer look at higher education minority ethnic students and graduates. *Institute for Employment Studies Research Report*, *552*.

Cormack, D., & Robson, C. (2011). *Classification and output of multiple ethnicities: considerations for monitoring Māori health*. Te Rōpū Rangahau Hauora a Eru Pōmare.

Crawford, R. (2016). New Zealand Productivity Commission Research Note 2016/1: History of tertiary education reforms in New Zealand. Retrieved 1 August 2019 from https://www.productivity.govt.nz/assets/Documents/history-of-tertiary-educaion/65759a16ed/History-of-tertiary-education-reforms.pdf.

Credé, M., Roch, S. G., & Kieszczynka, U. M. (2010). Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research*, *80*(2), 272-295.

Darling-Hammond, L., LaFors, J., & Snyder, J. (2001). Educating teachers for California's future. *Teacher Education Quarterly*, 9-55.

Darolia, R. (2014). Working (and studying) day and night: Heterogeneous effects of working on the academic performance of full-time and part-time students. *Economics of Education Review*, *38*, 38-50.

DeAngelo, L., Franke, R., Hurtado, S., Pryor, J. H., & Tran, S. (2011). Completing college: Assessing graduation rates at four-year institutions. Los Angeles: Higher Education Research Institute, UCLA.

Di Pietro, G. (2004). The determinants of university dropout in Italy: a bivariate probability model with sample selection. *Applied Economics Letters*, *11*(3), 187-191.

Donnelly, B. (2003). Inquiry into decile funding in New Zealand State and integrated schools. *Report of the Education and Science Committee*, 16.

Earle, D. (2007). *Te Whai i Ngā Taumata Atakura: Supporting Maori Achievement in Bachelors Degrees*. Tertiary Sector Performance Analysis and Reporting, Ministry of Education.

Fairlie, R. W. (1999). The absence of the African-American owned business: An analysis of the dynamics of self-employment. *Journal of Labour Economics*, *17*(1), 80-108.

Fairlie, R. W. (2005). An extension of the Blinder-Oaxaca decomposition technique to logit and probit models. *Journal of Economic and Social Measurement*, *30*(4), 305-316.

Fergusson, D. M., & Woodward, L. J. (2000). Family socioeconomic status at birth and rates of university participation. *New Zealand Journal of Educational Studies*, *35*(1), 25-36.

Ferrão, M., & Almeida, L. (2019). Differential effect of university entrance score on first-year students' academic performance in Portugal. *Assessment & Evaluation in Higher Education*, *44*(4), 610-622.

Fletcher, J. M., & Tienda, M. (2009). High school classmates and college success. *Sociology of Education*, *82*(4), 287-314.

Fletcher, J., & Tienda, M. (2010). Race and ethnic differences in college achievement: Does high school attended matter?. *The Annals of the American Academy of Political and Social Science*, *627*(1), 144-166.

Geel, R., & Backes-Gellner, U. (2012). Earning while learning: When and how student employment is beneficial. *Labour*, *26*(3), 313-340.

Geisinger, B. N., & Raman, D. R. (2013). Why they leave: Understanding student attrition from engineering majors. *International Journal of Engineering Education*, *29*(4), 914.

Guiffrida, D. A., Lynch, M. F., Wall, A. F., & Abel, D. S. (2013). Do reasons for attending college affect academic outcomes?: A test of a motivational model from a self-determination theory perspective. *Journal of College Student Development*, *54*(2), 121-139.

Haultain, S. A. (2009). Deciding to enter tertiary education and taking on debt: A longitudinal perspective. PhD Thesis, University of Canterbury.

Heinesen, E. (2010). Estimating Class-size Effects using Within-school Variation in Subject-specific Classes. *The Economic Journal*, *120*(545), 737-760.

Hoare, A., & Johnston, R. (2011). Widening participation through admissions policy–a British case study of school and university performance. *Studies in Higher Education*, *36*(1), 21-41.

Hout, M. (2012). Social and economic returns to college education in the United States. *Annual Eeview of Sociology*, *38*, 379-400.

Howe, K. R. (2003). *The quest for origins: who first discovered and settled the Pacific islands?*. University of Hawaii Press.

Jackson, M. (2012). Bold choices: How ethnic inequalities in educational attainment are suppressed. *Oxford Review of Education*, *38*(2), 189-208.

Jia, P., & Maloney, T. (2015). Using predictive modelling to identify students at risk of poor university outcomes. *Higher Education*, *70*(1), 127-149.

Johnes, G., & McNabb, R. (2004). Never give up on the good times: student attrition in the UK. *Oxford Bulletin of Economics and Statistics*, *66*(1), 23-47.

Johnson, M. (2008). Grading in competence-based qualifications–is it desirable and how might it affect validity?. *Journal of Further and Higher Education*, *32*(2), 175-184.

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

Kalenkoski, C. M., & Pabilonia, S. W. (2012). Time to work or time to play: The effect of student employment on homework, sleep, and screen time. *Labour Economics*, *19*(2), 211-221.

Keith, P. M., Byerly, C., Floerchinger, H., Pence, E., & Thornberg, E. (2006). Deficit and resilience perspectives on performance and campus comfort of adult students. *College Student Journal*, *40*(3), 546-557.

Kennett, D. J., Reed, M. J., & Stuart, A. S. (2013). The impact of reasons for attending university on academic resourcefulness and adjustment. *Active Learning in Higher Education*, *14*(2), 123-133.

Light, A., & Strayer, W. (2002). From Bakke to Hopwood: Does race affect college attendance and completion?. *Review of Economics and Statistics*, *84*(1), 34-44.

Lorah, J., & Ndum, E. (2013). Trends in Achievement Gaps in First-Year College Courses for Racial/Ethnic, Income, and Gender Subgroups: A 12-Year Study. ACT Research Report Series 2013 (8). *ACT, Inc.*

Marie, D., Fergusson, D. M., & Boden, J. M. (2008). Educational achievement in Maori: The roles of cultural identity and social disadvantage. *Australian Journal of Education*, *52*(2), 183-196.

Massey, D. S. (2006). Social background and academic performance differentials: White and minority students at selective colleges. *American Law and Economics Review*, *8*(2), 390-409.

McCall, B., Smith, J., & Wunsch, C. (2016). Government-sponsored vocational education for adults. *Handbook of the Economics of Education* (Vol. 5, pp. 479-652). Elsevier.

Meehan, L., Pacheco, G., & Pushon, Z. (2017). Explaining ethnic disparities in bachelor's qualifications: Participation, retention and completion in NZ. *New Zealand Productivity Commission Working Paper 2017*, *1*.

Meehan, L., Pacheco, G., & Pushon, Z. (2019). Explaining ethnic disparities in bachelor's degree participation: evidence from NZ. *Studies in Higher Education*, *44*(7), 1130-1152.

Ministry of Education. (2001). Profile & Trends 2000: New Zealand's Tertiary Education Sector. Retrieved 1 August 2018 from https://www.educationcounts.govt.nz/__data/assets/pdf_file/0003/9246/PT-2000.pdf.

Ministry of Education. (2002). Profile & Trends 2001: New Zealand's Tertiary Education Sector. Retrieved 1 August 2018 from https://www.educationcounts.govt.nz/__data/assets/pdf_file/0004/9247/PT-2001.pdf.

Ministry of Education. (2006). Profile & Trends 2005: New Zealand's Tertiary Education Sector. Retrieved 1 August 2018 from https://www.educationcounts.govt.nz/publications/series/2531/profile_and_trends_2005.

Ministry of Education. (2006). OECD thematic review of tertiary education: New Zealand country background report. Retrieved 1 August 2018 from https://www.oecd.org/newzealand/36441052.pdf.

Ministry of Education. (2013). Ka Hikitia–Accelerating success 2013–2017: The Māori education strategy. Retrieved 31 August 2019 from https://www.education.govt.nz/assets/Documents/Ministry/Strategies-and-policies/Ka-Hikitia/KaHikitiaAcceleratingSuccessEnglish.pdf.

Ministry of Education. (2013). Pasifika Education Plan 2013-2017. Retrieved 31 August 2019 from https://www.education.govt.nz/assets/Documents/Ministry/Strategies-and-policies/PasifikaEdPlan2013To2017V2.pdf.

Ministry of Education. (2014). Ethnic Group Codes. Retrieved on 29 August 2020 from https://www.educationcounts.govt.nz/__data/assets/pdf_file/0019/9514/Ethnic-Group-Codes-Nov2014.pdf.

Ministry of Education. (2014). Tertiary education strategy 2014–2019. Retrieved 31 August 2019 from https://www.education.govt.nz/assets/Documents/Further-education/Tertiary-Education-Strategy.pdf.

Ministry of Education, (2014). What young graduates earn when they leave study. Retrieved 21 February 2021 from https://www.educationcounts.govt.nz/__data/assets/pdf_file/0010/146548/What-young-graduates-earn-when-they-leave-study.pdf.

Ministry of Education. (2017). Profile & Trends 2016: New Zealand's Annual Tertiary Education Enrolments [Part 1 of 6]. Retrieved 1 August 2018 from https://www.educationcounts.govt.nz/__data/assets/pdf_file/0010/182296/2016-pt-provider-based-enrolments-part1b.pdf.

Ministry of Justice. (2019). Convictions in the general population. Retrieved on 3 March 2021 from https://www.justice.govt.nz/assets/Documents/Publications/Factsheet-proportion-ever-convicted.pdf.

Modood, T. (2006). Ethnicity, Muslims and higher education entry in Britain. *Teaching in Higher Education*, *11*(2), 247-250.

Mutuku, S., & Kiilu, M. (2016). Influence of demographic factors on academic performance among primary teacher trainees-a case study of machakos teachers college. *International Journal of Educational Studies*, *3*(01), 7-11.

Neumark, D. (1988). Employers' discriminatory behavior and the estimation of wage discrimination. *Journal of Human Resources*, 279-295.

Neyt, B., Omey, E., Verhaest, D., & Baert, S. (2019). Does student work really affect educational outcomes? A review of the literature. *Journal of Economic Surveys*, *33*(3), 896-921.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 693-709.

Oaxaca, R. L., & Ransom, M. R. (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics*, *61*(1), 5-21.

Olani, A. (2009). Predicting first year university students' academic success. *Electronic Journal of Research in Educational Psychology*, *7*(3), 1053-1072.

Osaikhiuwu, O. C. (2014). Institutional factors affecting the academic performance of public administration students in a Nigerian University. *Public Administration Research*, *3*(2), 171-177.

Parent, D. (2006). Work while in high school in Canada: its labour market and educational attainment effects. *Canadian Journal of Economics/Revue Canadienne D'économique*, *39*(4), 1125-1150.

Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Applied Econometrics*, *11*(6), 619-632.

Pampel, F. C., Krueger, P. M., & Denney, J. T. (2010). Socioeconomic disparities in health behaviours. *Annual review of sociology*, *36*, 349-370.

Parker, F. W. (1901). Principles of correlation. *The Course of Study*, *1*(6), 506-509.

Powdthavee, N., & Vignoles, A. (2009). The socio-economic gap in university dropout. *The BE Journal of Economic Analysis & policy*, *9*(1).

Quirk, K. J., Keith, T. Z., & Quirk, J. T. (2001). Employment during high school and student achievement: Longitudinal analysis of national data. *The Journal of Educational Research*, *95*(1), 4-10.

Reid, G., Bycroft, C., & Gleisner, F. (2016). Comparison of ethnicity information in administrative data and the census. *Wellington, New Zealand: Statistics New Zealand*.

Richardson, J. J., Kemp, S., Malinen, S., & Haultain, S. A. (2013). The academic achievement of students in a New Zealand university: Does it pay to work?. *Journal of Further and Higher Education*, *37*(6), 864-882.

Richardson, J. T. (2012). Face-to-face versus online tuition: Preference, performance and pass rates in white and ethnic minority students. *British Journal of Educational Technology*, *43*(1), 17-27.

Richardson, J. T., & Woodley, A. (2003). Another look at the role of age, gender and subject as predictors of academic attainment in higher education. *Studies in Higher Education*, *28*(4), 475-493.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353.

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*(5), 535-550.

Russell, L. J. (1945). The principle of causality. In *Proceedings of the Aristotelian Society* (Vol. 46, pp. 105-126). Aristotelian Society, Wiley.

Schneider, B., & Saw, G. (2016). Racial and ethnic gaps in postsecondary aspirations and enrolment. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *2*(5), 58-82.

Scott, D. (2005a). Retention, completion and progression in tertiary education in New Zealand. *Journal of Higher Education Policy and Management*, *27*(1), 3-17.

Scott, D. (2005b). What factors make a difference to getting a degree in New Zealand?. Retrieved 31 August 2019 from https://www.educationcounts.govt.nz/__data/assets/pdf_file/0007/9547/modelling-completion-factors-report-sfs.pdf.

Scott, D. (2008). *How does achievement at school affect achievement in tertiary education*. Tertiary Sector Performance Analysis and Reporting, Strategy and System Performance, Ministry of Education.

Scott, D., & Smart, W. (2005). What factors make a difference to getting a degree in New Zealand? Education Counts. *Tertiary Sector Performance Analysis & Reporting, Ministry of Education, New Zealand*.

Scott-Clayton, J., & Minaya, V. (2016). Should student employment be subsidized? Conditional counterfactuals and the outcomes of work-study participation. *Economics of Education Review*, *52*, 1-18.

Shulruf, B., Hattie, J., & Tumen, S. (2008). Individual and school factors affecting students' participation and success in higher education. *Higher education*, *56*(5), 613-632.

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

Smith, J., & Naylor, R. (2005). Schooling effects on subsequent university performance: evidence for the UK university population. *Economics of Education Review*, *24*(5), 549-562.

Sopoaga, F., Zaharic, T., Kokaua, J., Ekeroma, A. J., Murray, G., & van der Meer, J. (2013). Pacific students undertaking the first year of health sciences at the University of Otago, and factors associated with academic performance. *NZ Med J*, *126*, 96-108.

Spence, M. (1973). Job Market Signalling. *The Quarterly Journal of Economics*, *87*(3), 355-374.

Staff, J., & Mortimer, J. T. (2007). Educational and work strategies from adolescence to early adulthood: Consequences for educational attainment. *Social Forces*, *85*(3), 1169-1194.

Stegers-Jager, K. M., Steyerberg, E. W., Cohen-Schotanus, J., & Themmen, A. P. (2012). Ethnic disparities in undergraduate pre-clinical and clinical performance. *Medical Education*, *46*(6), 575-585.

Stockfelt, S. (2018). Ethnic variation in higher education participation amongst males in the UK: the mediating effects of attitudes and prior attainment. *Studies in Higher Education*, *43*(11), 1895-1911.

Strathdee, R., & Engler, R. (2012). Who is missing from higher education in New Zealand?. *British Educational Research Journal*, *38*(3), 497-514.

Tertiary Education Commission. (2016). Introducing Tū Māia e te Akonga. Retrieved 31 August 2019 from http://www.tec.govt.%20nz/assets/Publications-and-others/Tu-Maia-e-te-%20Akonga-brochure.pdf.

Tertiary Education Commission. (2017). Pasifika operational strategy 2017-2020. Retrieved 31 August 2019 from https://tec.govt.nz/assets/Publications-and-others/7d0971ac24/Pasifika-Operational-Strategy-2017-2020.pdf.

Tertiary Education Commission. (2018). Māori and Pasifika achievement is a learner success issue. Retrieved 31 August 2019 from https://www.tec.govt.nz/news-and-consultations/archived-news/maori-and-pasifika-achievement-learner-success-issue/.

Tickell, G., & Smyrnios, K. X. (2005). Predictors of tertiary accounting students' academic performance: a comparison of year 12-to-university students with TAFE-to-university students. *Journal of Higher Education Policy and Management*, *27*(2), 239-259.

Ting, S. M., & Bryant, A. (2001). The impact of acculturation and psychosocial variables on academic performance of Native American and Caucasian college freshmen. *Journal of College Admission*, 22-28.

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

Tumen, S., Shulruf, B., & Hattie, J. (2008). Student pathways at the university: Patterns and predictors of completion. *Studies in Higher Education*, *33*(3), 233-252.

Tym, C., McMillion, R., Barone, S., & Webster, J. (2004). First-Generation College Students: A Literature Review. *TG (Texas Guaranteed Student Loan Corporation)*.

Universities New Zealand. (2018). *Achieving parity for Mäori and Pasifika: The university sector view* (No. 1). UNZ Parity Discussion Paper.

Urban, M., Jones, E., Smith, G., Evans, C., Maclachlan, M., & Karmel, T. (1999). Completions: Undergraduate academic outcomes for 1992 commencing students. *Higher Education Division, Department of Education, Training and Youth Affairs, Occasional Paper Series*, (99-G).

US Department of Labour. (2020). College enrolment and work activity of 2019 high school graduates. Retrieved 19 March 2021 from https://www.bls.gov/news.release/archives/hsgec_04282020.htm.

van Herpen, S. G., Meeuwisse, M., Hofman, W. A., Severiens, S. E., & Arends, L. R. (2017). Early predictors of first-year academic success at university: pre-university effort, pre-university self-efficacy, and pre-university reasons for attending university. *Educational Research and Evaluation*, *23*(1-2), 52-72.

Wenz, M., & Yu, W. C. (2010). Term-time employment and the academic performance of undergraduates. *Journal of Education Finance*, 358-373.

Zaitseva, E., Milsom, C., & Stewart, M. (2013). Connecting the dots: Using concept maps for interpreting student satisfaction. *Quality in Higher Education*, *19*(2), 225-247.

Zorlu, A. (2013). Ethnic disparities in higher education. *IZA Journal of Migration*, *2*(1), 1-21.

Statistical Appendix

Table A.3.3. Ethnic differences in the university outcomes: two different ways of measuring ethnicity[a]

Ethnicities	Participation		First-year course completion		Second-year course completion		Qualification completion	
	Prioritized ethnicity[b]	All Ethnicities[c]	Prioritized ethnicity	All ethnicities	Prioritized ethnicity	All ethnicities	Prioritized ethnicity	All ethnicities
Māori	0.1308 n=37,536 (20.97%)	0.1308 n=37,536 (19.57%)	0.7057 n=4,821 (8.57%)	0.7057 n=4821 (8.13%)	0.7500 n=2,400 (7.57%)	0.7500 n=2,400 (7.20%)	0.3694 n=1,413 (8.92%)	0.3694 n=1,413 (8.45%)
Pasifika	0.2190 n=16,359 (9.14%)	0.2096 n=18,393 (9.59%)	0.5764 n=3,555 (6.32%)	0.5796 n=3825 (6.45%)	0.6258 n=1,836 (5.79%)	0.6271 n=1,971 (5.91%)	0.2884 n=957 (6.04%)	0.2816 n=1,044 (6.24%)
Asian	0.5471 n=18,435 (10.30%)	0.5404 n=18,996 (9.91%)	0.8170 n=10,068 (17.91%)	0.8152 n=10,245 (17.27%)	0.8522 n=5,967 (18.81%)	0.8513 n=6,054 (18.16%)	0.5432 n=2,745 (17.33%)	0.5418 n=2,796 (16.72%)
MELAA	0.4252 n=3,147 (1.76%)	0.4219 n=3,228 (1.68%)	0.7466 n=1,338 (2.38%)	0.7506 n=1,359 (2.29%)	0.7801 n=723 (2.28%)	0.7846 n=738 (2.21%)	0.4370 n=357 (2.25%)	0.4417 n=360 (2.15%)
European	0.3538 n=102,135 (57.05%)	0.3457 n=112,029 (58.42%)	0.8371 n=35,970 (63.98%)	0.8330 n=38,547 (65.00%)	0.8743 n=20,502 (64.63%)	0.8702 n=21,870 (65.59%)	0.5609 n=10,227 (64.55%)	0.5539 n=10,959 (65.54%)

Other	0.3333 n=1,422 (0.79%)	0.3239 n=1,584 (0.83%)	0.7962 n=471 (0.84%)	0.8000 n=510 (0.86%)	0.8367 n=294 (0.93%)	0.8641 n=309 (0.93%)	0.5208 n=144 (0.91%)	0.5098 n=153 (0.91%)
Overall	0.3157 n=179,034 (100%)	0.3110 n=191,766 (100%)	0.8032 n=56,223 (100%)	0.8011 n=59,307 (100%)	0.8439 n=31,722 (100%)	0.8417 n=33,342 (100%)	0.5209 n=15,843 (100%)	0.5167 n=16,722 (100%)
Ethnic differences								
European – Māori	0.2230	0.2149	0.1314	0.1273	0.1243	0.1202	0.1915	0.1845
European – Pasifika	0.1348	0.1361	0.2607	0.2534	0.2485	0.2431	0.2725	0.2723
European – Asian	-0.1933	-0.1947	0.0201	0.0178	0.0221	0.0189	0.0177	0.0121

[a] The values reported in this table were based on the numbers of observations that have been randomly rounded to base 3 or supressed if the original observation count is fewer than 6 due to the confidentiality requirements from Statistics New Zealand.

[b] The top value in each cell is the university outcome, the number of student observations (n) being ethnically defined in a particular way is shown in the middle, while the bottom value in bracket describes the share of the number of student observations in relation to the overall observations under the ethnic identification.

[c] Students can appear in different ethnic groups if they self-reported more than one ethnicity. A student is allowed to maximum self-report up to three different ethnicities.

Table A.4.1. Probit regression results for participation in a bachelor's degree with a New Zealand university by age 21

Variables	Prioritised ethnicity			All ethnicities		
	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Constant	-2.7938***	0.0750	-	-2.7934***	0.0750	-
Female	0.3107***	0.0086	0.0550***	0.3107***	0.0086	0.0550***
Cohort years						
Cohort 2011	0.0741***	0.0107	0.0131***	0.0742***	0.0107	0.0131***
Cohort 2012	0.0702***	0.0113	0.0124***	0.0701***	0.0113	0.0124***
Ethnicities						
Māori	-0.2845***	0.0136	-0.0503***	-0.2863***	0.0134	-0.0507***
Pasifika	-0.0311*	0.0185	-0.0055*	-0.0414**	0.0174	-0.0073**
Asian	0.3309***	0.0147	0.0586***	0.3277***	0.0144	0.0580***
MELAA	0.1530***	0.0304	0.0271***	0.1411***	0.0300	0.0250***
Other	0.0369	0.0464	0.0065	0.0506	0.0446	0.0090
School deciles						
Decile 1	-0.4371***	0.0299	-0.0774***	-0.4334***	0.0298	-0.0767***
Decile 2	-0.2766***	0.0246	-0.0490***	-0.2754***	0.0246	-0.0487***
Decile 3	-0.1283***	0.0227	-0.0227***	-0.1270***	0.0227	-0.0225***
Decile 4	-0.0386*	0.0200	-0.0068*	-0.0381*	0.0200	-0.0067*
Decile 5	-0.0570***	0.0194	-0.0101***	-0.0562***	0.0194	-0.0099***
Decile 7	0.0462**	0.0188	0.0082**	0.0469**	0.0188	0.0083**
Decile 8	0.1322***	0.0172	0.0234***	0.1324***	0.0172	0.0234***
Decile 9	0.1661***	0.0177	0.0294***	0.1663***	0.0177	0.0294***

Decile 10	0.2958***	0.0185	0.0524***	0.2963***	0.0185	0.0524***
Decile unknown	0.4000***	0.0807	0.0708***	0.4008***	0.0807	0.0709***
School authority						
State integrated	0.0552***	0.0130	0.0098***	0.0554***	0.0130	0.0098***
Private	0.2000***	0.0213	0.0354***	0.1995***	0.0213	0.0353***
Unknown	-0.0936	0.1598	-0.0166	-0.0933	0.1598	-0.0165
School gender						
Single sex	0.1976***	0.0106	0.0350***	0.1978***	0.0106	0.0350***
Unknown	0.8678***	0.1561	0.1536***	0.8668***	0.1561	0.1534***
School area						
Minor Urban Area	-0.0155	0.0165	-0.0027	-0.0156	0.0165	-0.0028
Secondary Urban Area	-0.0404**	0.0180	-0.0072**	-0.0409**	0.0180	-0.0072**
Rural Area	-0.0070	0.0341	-0.0012	-0.0070	0.0341	-0.0012
Unknown	0.2576***	0.0780	0.0456***	0.2576***	0.0779	0.0456***
School region						
Northland	-0.1766***	0.0283	-0.0313***	-0.1778***	0.0283	-0.0315***
Waikato	-0.1975***	0.0172	-0.0349***	-0.1982***	0.0172	-0.0351***
Bay of Plenty	-0.2579***	0.0197	-0.0456***	-0.2585***	0.0197	-0.0457***
Gisborne	-0.2196***	0.0463	-0.0389***	-0.2202***	0.0463	-0.0390***
Hawkes Bay	-0.2264***	0.0240	-0.0401***	-0.2272***	0.0240	-0.0402***
Taranaki	-0.2066***	0.0300	-0.0366***	-0.2076***	0.0300	-0.0367***
Manawatu-Whanganui	-0.2089***	0.0216	-0.0370***	-0.2097***	0.0217	-0.0371***
Wellington	-0.2036***	0.0153	-0.0360***	-0.2036***	0.0153	-0.0360***
West Coast	-0.3259***	0.0615	-0.0577***	-0.3276***	0.0615	-0.0580***
Canterbury	-0.2387***	0.0156	-0.0422***	-0.2394***	0.0156	-0.0424***

Otago	-0.2179***	0.0232	-0.0386***	-0.2188***	0.0231	-0.0387***
Southland	-0.2470***	0.0327	-0.0437***	-0.2480***	0.0327	-0.0439***
Tasman	-0.3630***	0.0483	-0.0643***	-0.3635***	0.0483	-0.0643***
Nelson	-0.2787***	0.0389	-0.0493***	-0.2794***	0.0389	-0.0495***
Marlborough	-0.4253***	0.0512	-0.0753***	-0.4265***	0.0512	-0.0755***
Unknown	-1.5157***	0.1117	-0.2683***	-1.5165***	0.1117	-0.2684***
Highest school qualification						
NCEA Level 1	0.6217***	0.0584	0.1100***	0.6218***	0.0584	0.1100***
NCEA Level 2	1.1620***	0.0535	0.2056***	1.1617***	0.0535	0.2056***
NCEA Level 3	2.9807***	0.0526	0.5275***	2.9805***	0.0526	0.5274***
Overseas qualification	2.5933***	0.0571	0.4590***	2.5938***	0.0571	0.4590***
Other qualification	1.8640***	0.0533	0.3299***	1.8638***	0.0533	0.3298***
Unknown qualification	-0.6235***	0.0705	-0.1104***	-0.6233***	0.0705	-0.1103***
Switching school						
1	0.0434***	0.0157	0.0077***	0.0437***	0.0157	0.0077***
2	-0.0081	0.0172	-0.0014	-0.0081	0.0172	-0.0014
3	-0.0661***	0.0210	-0.0117***	-0.0659***	0.0210	-0.0117***
4 or more	-0.2517***	0.0263	-0.0445***	-0.2516***	0.0263	-0.0445***
Truancy, suspension and/or unjustified absence	-0.0356	0.0335	-0.0063	-0.0348	0.0335	-0.0062
Single-parent household	-0.0269**	0.0134	-0.0048**	-0.0276**	0.0134	-0.0049**
Parents highest qualification						
Level 1 to 3 Certificate	0.1827***	0.0209	0.0323***	0.1817***	0.0209	0.0322***
Level 4 Certificate	0.1552***	0.0234	0.0275***	0.1535***	0.0234	0.0272***
Level 5 to 6 Diploma	0.2861***	0.0231	0.0506***	0.2845***	0.0231	0.0504***
Bachelor’s degree and Level 7 Qualification	0.4085***	0.0220	0.0723***	0.4072***	0.0220	0.0721***

Postgraduate	0.4918***	0.0241	0.0870***	0.4904***	0.0241	0.0868***
Overseas School Qualification	0.3048***	0.0234	0.0539***	0.3033***	0.0234	0.0537***
Unknown qualification	0.0666*	0.0356	0.0118*	0.0669*	0.0356	0.0118*
Household income						
Zero	0.1420	0.0949	0.0251	0.1449	0.0949	0.0256
1 to 30,000	0.0145	0.0242	0.0026	0.0147	0.0242	0.0026
50,001 to 70,000	0.0028	0.0218	0.0005	0.0027	0.0218	0.0005
70,001 to 100,000	0.0166	0.0204	0.0029	0.0164	0.0204	0.0029
100,001 or more	0.0907***	0.0194	0.0160***	0.0901***	0.0194	0.0159***
Unknown	-0.0168	0.0209	-0.0030	-0.0168	0.0209	-0.0030
Parents' benefit and charges known	-0.1255***	0.0448	-0.0222***	-0.1238***	0.0448	-0.0219***
Parents' proportions days living on benefit	-0.0199	0.0288	-0.0035	-0.0199	0.0288	-0.0035
Parents' crime charges	-0.0015	0.0012	-0.0003	-0.0015	0.0012	-0.0003
Number of observations (n)	179,034			191,766		
Likelihood function	-56617.074			-56612.861		
Pseudo R^2	0.4929			0.4929		

* denotes statistical significance at a 10% level; ** denotes statistical significance at a 5% level; *** denotes statistical significance at a 1% level.

Table A.4.2. Nonlinear decomposition of participation in a bachelor's degree with a New Zealand university by age 21

	Prioritised ethnicity			All ethnicities		
	European vs Māori	European vs Pasifika	European vs Asian	European vs Māori	European vs Pasifika	European vs Asian
Total ethnic difference	0.2230	0.1348	-0.1933	0.2149	0.1361	-0.1947
Explained ethnic difference[a]						
Female	-0.0001*** (0.0000) [-0.04%]	-0.0016*** (0.0001) [-1.19%]	0.0017*** (0.0001) [0.88%]	-0.0001*** (0.0000) [-0.05%]	-0.0014*** (0.0001) [-1.03%]	0.0016*** (0.0001) [0.82%]
Cohort years	-0.0001*** (0.0000) [-0.04%]	-0.0003*** (0.0000) [-0.22%]	-0.0002*** (0.0000) [-0.10%]	-0.0001*** (0.0000) [-0.05%]	-0.0002*** (0.0000) [-0.15%]	-0.0002*** (0.0000) [-0.10%]
School deciles	0.0309*** (0.0014) [13.86%]	0.0396*** (0.0019) [29.38%]	-0.0046*** (0.0005) [-2.38%]	0.0308*** (0.0014) [14.33%]	0.0386*** (0.0018) [28.36%]	-0.0055*** (0.0005) [-2.82%]
School qualification	0.1280*** (0.0008) [57.40%]	0.0831*** (0.0006) [61.65%]	-0.0866*** (0.0007) [-44.78%]	0.1287*** (0.0007) [59.89%]	0.0827*** (0.0005) [60.76%]	-0.088*** (0.0007) [-45.20%]
Other school factors	0.0157*** (0.0007) [7.04%]	-0.0077*** (0.0009) [-5.71%]	-0.0274*** (0.0014) [-14.17%]	0.0161*** (0.0007) [7.49%]	-0.0065*** (0.0008) [-4.78%]	-0.0269*** (0.0013) [-13.82%]
Parental education	0.0166*** (0.0011) [7.44%]	0.0157*** (0.0010) [11.65%]	-0.0099*** (0.0013) [-5.12%]	0.0169*** (0.0010) [7.86%]	0.0152*** (0.0009) [11.17%]	-0.0102*** (0.0012) [-5.24%]
Household income	0.0047*** (0.0006) [2.11%]	0.0045*** (0.0008) [3.34%]	0.0025*** (0.0006) [1.29%]	0.0045*** (0.0006) [2.09%]	0.0043*** (0.0007) [3.16%]	0.0024*** (0.0005) [1.23%]

Other parental factors	-0.0046*** (0.0009) [-2.06%]	-0.0023*** (0.0005) [-1.71%]	0.0003 (0.0002) [0.16%]	-0.0047*** (0.0009) [-2.19%]	-0.0022*** (0.0005) [-1.62%]	0.0004* (0.0002) [0.21%]
Total explained differences	0.1911*** (0.0009) [85.70%]	0.1310*** (0.0020) [97.18%]	-0.1242*** (0.0020) [64.22%]	0.1921*** (0.0009) [89.39%]	0.1305*** (0.0017) [95.89%]	-0.1264*** (0.0019) [64.92%]
Unexplained differences	0.0319 [14.30%]	0.0038 [2.82%]	-0.0692 [35.78%]	0.0228 [10.61%]	0.0056 [4.11%]	-0.0683 [35.08%]
Number of observations (n)	139,671	118,494	120,567	149,565	130,422	131,025

All decompositions used the estimated coefficients from pooled Probit regressions of relevant ethnic groups (European and Māori, European and Pasifika and European and Asian). A dummy variable was created to indicate the each of the ethnic minority groups used in the regression. The decomposition results reported in this table are based on 100 replications.

[a] Standard errors associated with these individual or groups of variables are listed in parentheses ('( )') below these estimated effects. Percentage changes in the total ethnic differences associated with these individual or groups of variables are shown in square brackets ('[ ]').

* denotes statistical significance at a 10% level; ** denotes statistical significance at a 5% level; *** denotes statistical significance at a 1% level.

Table A.4.3. Probit regression results for first-year course completion

Variables	Prioritised ethnicity			All ethnicities		
	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Constant	0.5077***	0.1882	-	0.5079***	0.1879	-
Female	0.2172***	0.0094	0.0550***	0.2171***	0.0094	0.0550***
Cohort years						
Cohort 2011	0.0357***	0.0118	0.0091***	0.0347***	0.0118	0.0088***
Cohort 2012	0.0556***	0.0126	0.0141***	0.0545***	0.0126	0.0138***
Ethnicities						
Māori	-0.3143***	0.0164	-0.0796***	-0.2743***	0.0162	-0.0695***
Pasifika	-0.5654***	0.0194	-0.1432***	-0.5303***	0.0185	-0.1344***
Asian	-0.0255*	0.0144	-0.0065*	-0.0100	0.0141	-0.0025
MELAA	-0.2209***	0.0294	-0.0560***	-0.2043***	0.0291	-0.0518***
Other	-0.0896*	0.0489	-0.0227*	-0.0587	0.0471	-0.0149
School deciles						
Decile 1	-0.4871***	0.0356	-0.1234***	-0.5014***	0.0356	-0.1270***
Decile 2	-0.2489***	0.0296	-0.0631***	-0.2548***	0.0296	-0.0646***
Decile 3	-0.2013***	0.0260	-0.0510***	-0.2072***	0.0260	-0.0525***
Decile 4	-0.0866***	0.0233	-0.0219***	-0.0897***	0.0233	-0.0227***
Decile 5	-0.0539**	0.0219	-0.0136**	-0.0545**	0.0219	-0.0138**
Decile 7	0.0268	0.0204	0.0068	0.0266	0.0204	0.0067
Decile 8	0.0070	0.0193	0.0018	0.0072	0.0193	0.0018
Decile 9	0.0443**	0.0191	0.0112**	0.0449**	0.0191	0.0114**
Decile 10	0.0457**	0.0196	0.0116**	0.0470**	0.0196	0.0119**

Decile unknown	-0.1261	0.1026	-0.0320	-0.1287	0.1026	-0.0326
School authority						
State integrated	-0.0418***	0.0132	-0.0106***	-0.0424***	0.0132	-0.0107***
Private	-0.0320	0.0201	-0.0081	-0.0304	0.0201	-0.0077
Unknown	0.2933	0.2331	0.0743	0.3070	0.2330	0.0778
School gender						
Single sex	0.0041	0.0109	0.0010	0.0045	0.0109	0.0011
Unknown	-0.0798	0.2178	-0.0202	-0.0930	0.2178	-0.0236
School area						
Minor Urban Area	0.0833***	0.0199	0.0211***	0.0853***	0.0199	0.0216***
Secondary Urban Area	0.0325	0.0210	0.0082	0.0328	0.0209	0.0083
Rural Area	0.0107	0.0412	0.0027	0.0116	0.0412	0.0029
Unknown	0.2668***	0.0928	0.0676***	0.2687***	0.0930	0.0681***
Highest school qualification						
NCEA Level 1	0.0722	0.1877	0.0183	0.0613	0.1875	0.0155
NCEA Level 2	0.0555	0.1783	0.0141	0.0422	0.1780	0.0107
NCEA Level 3	0.4057**	0.1772	0.1028**	0.3938**	0.1769	0.0998**
Overseas qualification	0.6515***	0.1786	0.1650***	0.6385***	0.1783	0.1618***
Other qualification	0.2202	0.1776	0.0558	0.2064	0.1772	0.0523
Unknown qualification	0.5444**	0.2386	0.1379**	0.5280**	0.2380	0.1338**
Switching school						
1	-0.0162	0.0163	-0.0041	-0.0159	0.0163	-0.0040
2	-0.0628***	0.0185	-0.0159***	-0.0618***	0.0185	-0.0157***
3	-0.1436***	0.0233	-0.0364***	-0.1420***	0.0233	-0.0360***
4 or more	-0.3114***	0.0329	-0.0789***	-0.3100***	0.0329	-0.0785***

Truancy, suspension and/or unjustified absence	0.0305	0.0369	0.0077	0.0294	0.0369	0.0074
Age						
Under 18	0.4548***	0.1063	0.1152***	0.4554***	0.1066	0.1154***
19	0.0146	0.0113	0.0037	0.0140	0.0113	0.0036
20	0.0250	0.0161	0.0063	0.0227	0.0161	0.0057
21	-0.0360	0.0249	-0.0091	-0.0343	0.0250	-0.0087
22 or older	0.0708	0.0623	0.0179	0.0659	0.0622	0.0167
Employment hours						
Low (>0 and <10) hours	-0.0688**	0.0294	-0.0174**	-0.0674**	0.0294	-0.0171**
Medium (≥10 and <20) hours	0.0001	0.0352	0.0000	-0.0005	0.0352	-0.0001
High (≥20) hours	-0.0313*	0.0185	-0.0079*	-0.0309*	0.0186	-0.0078*
Unknown	-0.0299**	0.0149	-0.0076**	-0.0292**	0.0149	-0.0074**
Prior activity						
Non-employed or beneficiary	-0.1871***	0.0408	-0.0474***	-0.1852***	0.0408	-0.0469***
employed	-0.1388***	0.0205	-0.0352***	-0.1350***	0.0205	-0.0342***
Other type student	-0.1729***	0.0278	-0.0438***	-0.1736***	0.0278	-0.0440***
House person or retired	-0.2325	0.1538	-0.0589	-0.2282	0.1529	-0.0578
Overseas	0.1703***	0.0343	0.0431***	0.1756***	0.0344	0.0445***
University						
University B	-0.0008	0.0203	-0.0002	-0.0005	0.0203	-0.0001
University C	-0.0048	0.0188	-0.0012	-0.0029	0.0188	-0.0007
University D	0.0135	0.0153	0.0034	0.0142	0.0153	0.0036
University E	-0.0795***	0.0191	-0.0201***	-0.0772***	0.0191	-0.0196***
University F	0.0975***	0.0379	0.0247***	0.1023***	0.0379	0.0259***
University G	-0.0294*	0.0151	-0.0075*	-0.0280*	0.0151	-0.0071*

University H	0.1490***	0.0161	0.0377***	0.1504***	0.0161	0.0381***
New Zealand citizenship	0.1421***	0.0162	0.0360***	0.1454***	0.0162	0.0368***
Part-time	-0.8657***	0.0179	-0.2193***	-0.8680***	0.0179	-0.2199***
Single-parent household	-0.0780***	0.0148	-0.0198***	-0.0776***	0.0147	-0.0197***
Parents highest qualification						
Level 1 to 3 Certificate	0.0486*	0.0253	0.0123**	0.0527**	0.0253	0.0133**
Level 4 Certificate	0.0391	0.0282	0.0099	0.0434	0.0282	0.0110
Level 5 to 6 Diploma	0.1227***	0.0270	0.0311***	0.1272***	0.0270	0.0322***
Bachelor's degree and Level 7 Qualification	0.1824***	0.0254	0.0462***	0.1858***	0.0254	0.0471***
Postgraduate	0.2659***	0.0270	0.0674***	0.2682***	0.0270	0.0680***
Overseas School Qualification	0.0758***	0.0260	0.0192***	0.0745***	0.0261	0.0189***
Unknown qualification	-0.0347	0.0411	-0.0088	-0.0369	0.0411	-0.0094
Household income (in NZD)						
Zero	-0.0440	0.0948	-0.0111	-0.0435	0.0954	-0.0110
1 to 30,000	-0.0174	0.0267	-0.0044	-0.0179	0.0267	-0.0045
50,001 to 70,000	-0.0309	0.0243	-0.0078	-0.0292	0.0242	-0.0074
70,001 to 100,000	-0.0179	0.0225	-0.0045	-0.0156	0.0225	-0.0040
100,001 or more	-0.0167	0.0213	-0.0042	-0.0133	0.0212	-0.0034
Unknown	-0.0773***	0.0235	-0.0196***	-0.0756***	0.0235	-0.0192***
Parents' benefit and charges known	-0.1450***	0.0493	-0.0367***	-0.1484***	0.0493	-0.0376***
Parents' proportions days living on benefit	0.0113	0.0308	0.0029	0.0104	0.0308	0.0026
Parents' crime charges	-0.0004	0.0019	-0.0001	-0.0003	0.0019	-0.0001
Number of observations (n)	56,223			59,307		
Likelihood function	-25595.507			-25602.173		
Pseudo R^2	0.0820			0.0818		

* denotes statistical significance at a 10% level; ** denotes statistical significance at a 5% level; *** denotes statistical significance at a 1% level.

Table A.4.4. Proportions of student subsamples based on first-year employment hours: prioritised ethnicity

	Sample proportion[a]
Original thresholds of hours	
Employment hours	
Zero	0.1093
Low (>0 and 10 hours)	0.0308
Medium (≥10 and <20 hours)	0.0189
High (≥20 hours)	0.1368
Unknown	0.7043
Only Known hours	
Employment hours	
Zero	0.3695
Low (>0 and 10 hours)	0.1043
Medium (≥10 and <20 hours)	0.0639
High (≥20 hours)	0.4625
Full-time students	
Employment hours	
Zero	0.1096
Low (>0 and 10 hours)	0.0309
Medium (≥10 and <20 hours)	0.0188
High (≥20 hours)	0.1370
Unknown	0.7037
Female students	
Employment hours	
Zero	0.1084
Low (>0 and 10 hours)	0.0302
Medium (≥10 and <20 hours)	0.0186
High (≥20 hours)	0.1367
Unknown	0.7061
Male students	
Employment hours	
Zero	0.1106
Low (>0 and 10 hours)	0.0317
Medium (≥10 and <20 hours)	0.0193
High (≥20 hours)	0.1367
Unknown	0.7018
Māori and Pasifika students	
Employment hours	

Zero	0.1146
Low (>0 and 10 hours)	0.0279
Medium (≥10 and <20 hours)	0.0197
High (≥20 hours)	0.1425
Unknown	0.6953
Māori students	
Employment hours	
Zero	0.1095
Low (>0 and 10 hours)	0.0299
Medium (≥10 and <20 hours)	0.0193
High (≥20 hours)	0.1400
Unknown	0.7019
Pasifika students	
Employment hours	
Zero	0.1214
Low (>0 and 10 hours)	0.0253
Medium (≥10 and <20 hours)	0.0202
High (≥20 hours)	0.1459
Unknown	0.6863
Non-Māori and non-Pasifika students	
Employment hours	
Zero	0.1084
Low (>0 and 10 hours)	0.0313
Medium (≥10 and <20 hours)	0.0187
High (≥20 hours)	0.1358
Unknown	0.7059
Asian students	
Employment hours	
Zero	0.1043
Low (>0 and 10 hours)	0.0316
Medium (≥10 and <20 hours)	0.0176
High (≥20 hours)	0.1302
Unknown	0.7163
European students	
Employment hours	
Zero	0.1104
Low (>0 and 10 hours)	0.0310
Medium (≥10 and <20 hours)	0.0190
High (≥20 hours)	0.1367

Unknown	0.7029
New thresholds of hours	
Employment hours	
Zero	0.1093
Low (>0 and <15 hours)	0.0408
Medium (≥15 and <30 hours)	0.0236
High (≥30 hours)	0.1220
Unknown	0.7043

[a] The values reported in this table were produced based on the number of observations and have been randomly rounded to base 3 or supressed if the original observation count is fewer than 6 due to the confidentiality requirements from Statistics New Zealand.

Table A.4.5. Probit regression results for employment hours on first-year course completion by student subsamples

	Prioritised ethnicity			All ethnicities		
	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Only Known hours						
Employment hours						
Low (>0 and 10 hours)	-0.0691**	0.0294	-0.0174**	-0.0675**	0.0294	-0.0170**
Medium (≥10 and <20 hours)	0.0009	0.0352	0.0002	0.0006	0.0352	0.0001
High (≥20 hours)	-0.0307*	0.0186	-0.0077*	-0.0304	0.0186	-0.0076
Full-time students						
Employment hours						
Low (>0 and 10 hours)	-0.0698**	0.0303	-0.0171**	-0.0686**	0.0303	-0.0168**
Medium (≥10 and <20 hours)	0.0106	0.0367	0.0026	0.0099	0.0367	0.0024
High (≥20 hours)	-0.0325*	0.0192	-0.0080*	-0.0322*	0.0193	-0.0079*
Unknown	-0.0299*	0.0154	-0.0073*	-0.0294*	0.0155	-0.0072*
Female students						
Employment hours						
Low (>0 and 10 hours)	-0.0786**	0.0396	-0.0184**	-0.0775**	0.0395	-0.0182**
Medium (≥10 and <20 hours)	0.0153	0.0470	0.0036	0.0133	0.0470	0.0031
High (≥20 hours)	-0.0226	0.0249	-0.0053	-0.0222	0.0250	-0.0052
Unknown	-0.0161	0.0202	-0.0038	-0.0152	0.0202	-0.0036
Male students						
Employment hours						
Low (>0 and 10 hours)	-0.0627	0.0437	-0.0175	-0.0612	0.0437	-0.0171
Medium (≥10 and <20 hours)	-0.0237	0.0535	-0.0066	-0.0227	0.0535	-0.0063

High (≥20 hours)	-0.0468*	0.0277	-0.0131*	-0.0467*	0.0277	-0.0130*
Unknown	-0.0495**	0.0220	-0.0138**	-0.0491**	0.0220	-0.0137**
Māori and Pasifika students						
Employment hours						
Low (>0 and 10 hours)	-0.0663	0.0737	-0.0225	-0.0638	0.0734	-0.0217
Medium (≥10 and <20 hours)	0.1436*	0.0836	0.0488*	0.1415*	0.0838	0.0480*
High (≥20 hours)	-0.0145	0.0422	-0.0049	-0.0134	0.0423	-0.0046
Unknown	-0.0187	0.0344	-0.0063	-0.0175	0.0345	-0.0060
Māori students						
Employment hours						
Low (>0 and 10 hours)	-0.0514	0.0965	-0.0165	-	-	-
Medium (≥10 and <20 hours)	0.0898	0.1167	0.0288	-	-	-
High (≥20 hours)	-0.0380	0.0588	-0.0122	-	-	-
Unknown	-0.0525	0.0479	-0.0169	-	-	-
Pasifika students						
Employment hours						
Low (>0 and 10 hours)	-0.0802	0.1179	-0.0290	-0.1530	0.1129	-0.0552
Medium (≥10 and <20 hours)	0.2092*	0.1205	0.0755*	0.1989*	0.1167	0.0717*
High (≥20 hours)	0.0099	0.0614	0.0036	-0.0150	0.0595	-0.0054
Unknown	0.0140	0.0498	0.0051	-0.0139	0.0483	-0.0050
Non-Māori and non-Pasifika students						
Employment hours						
Low (>0 and 10 hours)	-0.0692**	0.0322	-0.0164**	-0.0692**	0.0322	-0.0164**
Medium (≥10 and <20 hours)	-0.0314	0.0384	-0.0075	-0.0313	0.0384	-0.0074
High (≥20 hours)	-0.0369*	0.0207	-0.0088*	-0.0369*	0.0207	-0.0088*

Unknown	-0.0324**	0.0165	-0.0077**	-0.0324**	0.0165	-0.0077**
Asian students						
Employment hours						
Low (>0 and 10 hours)	-0.0308	0.0667	-0.0076	-0.0316	0.0663	-0.0078
Medium (≥10 and <20 hours)	0.0442	0.0874	0.0108	0.0636	0.0868	0.0157
High (≥20 hours)	-0.0501	0.0432	-0.0123	-0.0519	0.0427	-0.0128
Unknown	-0.0387	0.0347	-0.0095	-0.0387	0.0343	-0.0095
European students						
Employment hours						
Low (>0 and 10 hours)	-0.0866**	0.0378	-0.0201**	-0.0829**	0.0364	-0.0196**
Medium (≥10 and <20 hours)	-0.0730*	0.0438	-0.0170*	-0.0763*	0.0424	-0.0180*
High (≥20 hours)	-0.0464*	0.0241	-0.0108*	-0.0466**	0.0232	-0.0110**
Unknown	-0.0414**	0.0192	-0.0096**	-0.0423**	0.0186	-0.0100**
New thresholds of hours						
Employment hours						
Low (>0 and <15 hours)	-0.0520**	0.0264	-0.0132**	-0.0509*	0.0264	-0.0129*
Medium (≥15 and <30 hours)	0.0050	0.0330	0.0013	0.0050	0.0330	0.0013
High (≥30 hours)	-0.0359*	0.0190	-0.0091*	-0.0356*	0.0190	-0.0090*
Unknown	-0.0299**	0.0149	-0.0076**	-0.0292**	0.0149	-0.0074**

* denotes statistical significance at a 10% level; ** denotes statistical significance at a 5% level; *** denotes statistical significance at a 1% level.

Table A.4.6. Linear decomposition of first-year course completion

	Prioritised ethnicity			All ethnicities		
	European vs Māori	European vs Pasifika	European vs Asian	European vs Māori	European vs Pasifika	European vs Asian
Total ethnic difference	0.1314	0.2607	0.0201	0.1273	0.2534	0.0178
Explained ethnic difference[a]						
Female	-0.0027***	-0.0040***	0.0025***	-0.0026***	-0.0038***	0.0025***
	(0.0005)	(0.0005)	(0.0003)	(0.0004)	(0.0005)	(0.0003)
	[-2.05%]	[-1.53%]	[12.44%]	[-2.04%]	[-1.50%]	[14.04%]
Cohort years	0.0001	-0.0003**	-0.0001*	0.0001	-0.0002*	-0.0001**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
	[0.08%]	[-0.12%]	[-0.50%]	[0.08%]	[-0.08%]	[-0.56%]
School deciles	0.0199***	0.0466***	0.0025***	0.0200***	0.0456***	0.0030***
	(0.0018)	(0.0033)	(0.0006)	(0.0018)	(0.0031)	(0.0006)
	[15.14%]	[17.87%]	[12.44%]	[15.71%]	[18.00%]	[16.85%]
School qualification	0.0056***	0.0079***	-0.0011**	0.0055***	0.0080***	-0.0010**
	(0.0006)	(0.0007)	(0.0005)	(0.0006)	(0.0007)	(0.0005)
	[4.26%]	[3.03%]	[-5.47%]	[4.32%]	[3.16%]	[-5.62%]
Other school factors	0.0033***	0.0037***	0.0033***	0.0031***	0.0042***	0.0034***
	(0.0008)	(0.0009)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
	[2.51%]	[1.42%]	[16.42%]	[2.44%]	[1.66%]	[19.10%]
Age	0.0003	0.0003	-0.0005***	0.0003	0.0004	-0.0005***
	(0.0002)	(0.0003)	(0.00010	(0.0002)	(0.0003)	(0.0001)
	[0.23%]	[0.12%]	[-2.49%]	[0.24%]	[0.16%]	[-2.81%]
Employment hours	0.0000	-0.0001	0.0001	0.0000	-0.0002	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
	[0.00%]	[-0.04%]	[0.50%]	[0.00%]	[-0.08%]	[0.56%]
New Zealand citizenship	-0.0018***	0.0018***	0.0072***	-0.0017***	0.0015***	0.0068***
	(0.0003)	(0.0004)	(0.0009)	(0.0003)	(0.0003)	(0.0009)
	[-1.37%]	[0.69%]	[35.82%]	[-1.34%]	[0.59%]	[38.20%]

Part-time	0.0074***	0.0188***	0.0013*	0.0073***	0.0178***	0.0011
	(0.0011)	(0.0016)	(0.0008)	(0.0011)	(0.0016)	(0.0008)
	[5.63%]	[7.21%]	[6.47%]	[5.73%]	[7.02%]	[6.18%]
University	-0.0024***	-0.0072***	-0.0022*	-0.0022***	-0.0074***	-0.0023*
	(0.0006)	(0.0013)	(0.0012)	(0.0006)	(0.0012)	(0.0012)
	[-1.83%]	[-2.76%]	[-10.95%]	[-1.73%]	[-2.92%]	[-12.92%]
Prior activity	0.0022***	0.0016***	-0.0015***	0.0022***	0.0018***	-0.0014***
	(0.0004)	(0.0005)	(0.0004)	(0.0004)	(0.0005)	(0.0004)
	[1.67%]	[0.61%]	[-7.46%]	[1.73%]	[0.71%]	[-7.87%]
Parental education	0.0083***	0.0128***	0.0009	0.0080***	0.0118***	0.0003
	(0.0016)	(0.0015)	(0.0014	(0.0016)	(0.0014)	(0.0013)
	[6.32%]	[4.91%]	[4.48%]	[6.28%]	[4.66%]	[1.69%]
Other parental factors	0.0001	0.0010	0.0002	0.0004	0.0018*	0.0003
	(0.0015)	(0.0011)	(0.0009)	(0.0014)	(0.0010)	(0.0009)
	[0.08%]	[0.38%]	[1.00%]	[0.31%]	[0.71%]	[1.69%]
Total explained differences	0.0403***	0.0829***	0.0126***	0.0404***	0.0813***	0.0122***
	(0.0025)	(0.0042)	(0.0024)	(0.0024)	(0.0040)	(0.0023)
	[30.67%]	[31.80%]	[62.69%]	[31.74%]	[32.08%]	[68.54%]
Unexplained differences	0.0911***	0.1778***	0.0075**	0.0869***	0.1721***	0.0056
	(0.0051)	(0.0070)	(0.0036)	(0.0051)	(0.0067)	(0.0036)
	[69.33%]	[68.20%]	[37.31%]	[68.26%]	[67.92%]	[31.46%]
Number of observations (n)	40,791	39,525	46,038	43,368	42,372	48,792

All decompositions used the estimated coefficients from pooled OLS regressions of the relevant ethnic groups (European and Māori, European and Pasifika and European and Asian). A dummy variable was created to indicate the each of the ethnic minority groups used in the regression.

[a] Standard errors associated with these individual or groups of variables are listed in parentheses ('( )') below these estimated effects. Percentage changes in the total ethnic differences associated with these individual or groups of variables are shown in square brackets ('[ ]').

* denotes statistical significance at a 10% level; ** denotes statistical significance at a 5% level; *** denotes statistical significance at a 1% level.

Table A.4.7. Probit regression results for second-year course completion

Variables	Prioritised ethnicity			All ethnicities		
	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Constant	1.0859***	0.2677	-	1.0514***	0.2676	-
Female	0.2780***	0.0130	0.0584***	0.2778***	0.0130	0.0584***
Cohort year						
Cohort 2011	0.0317**	0.0140	0.0066**	0.0305**	0.0140	0.0064**
Ethnicities						
Māori	-0.3734***	0.0237	-0.0784***	-0.3279***	0.0237	-0.0689***
Pasifika	-0.6053***	0.0269	-0.1271***	-0.5599***	0.0258	-0.1176***
Asian	-0.0716***	0.0197	-0.0150***	-0.0470**	0.0194	-0.0099**
MELAA	-0.3110***	0.0404	-0.0653***	-0.2889***	0.0400	-0.0607***
Other	-0.0644	0.0674	-0.0135	-0.0168	0.0658	-0.0035
School deciles						
Decile 1	-0.3978***	0.0522	-0.0835***	-0.4146***	0.0520	-0.0871***
Decile 2	-0.1922***	0.0417	-0.0404***	-0.2002***	0.0416	-0.0421***
Decile 3	-0.1434***	0.0358	-0.0301***	-0.1490***	0.0358	-0.0313***
Decile 4	-0.0874***	0.0321	-0.0183***	-0.0909***	0.0321	-0.0191***
Decile 5	-0.0459	0.0300	-0.0096	-0.0490	0.0300	-0.0103
Decile 7	0.0458	0.0283	0.0096	0.0461	0.0283	0.0097
Decile 8	0.0230	0.0266	0.0048	0.0236	0.0266	0.0050
Decile 9	0.0616**	0.0262	0.0129**	0.0629**	0.0262	0.0132**
Decile 10	0.0710***	0.0268	0.0149***	0.0719***	0.0268	0.0151***
Decile unknown	0.3486*	0.1898	0.0732*	0.3302*	0.1908	0.0694*

School authority						
State integrated	-0.0092	0.0189	-0.0019	-0.0084	0.0189	-0.0018
Private	0.0199	0.0284	0.0042	0.0238	0.0285	0.0050
Unknown	0.3660	0.3099	0.0768	0.3867	0.3109	0.0812
School gender						
Single sex	-0.0177	0.0150	-0.0037	-0.0173	0.0150	-0.0036
Unknown	-0.4723	0.2989	-0.0992	-0.4800	0.3014	-0.1008
School area						
Minor Urban Area	0.0350	0.0283	0.0074	0.0394	0.0282	0.0083
Secondary Urban Area	-0.0044	0.0299	-0.0009	-0.0022	0.0299	-0.0005
Rural Area	0.0935	0.0622	0.0196	0.0989	0.0621	0.0208
Unknown	0.1441	0.1487	0.0303	0.1502	0.1518	0.0316
Highest school qualification						
NCEA Level 1	-0.3812	0.2574	-0.0800	-0.3634	0.2577	-0.0763
NCEA Level 2	-0.2720	0.2415	-0.0571	-0.2656	0.2415	-0.0558
NCEA Level 3	0.0136	0.2400	0.0028	0.0232	0.2399	0.0049
Overseas qualification	0.0886	0.2416	0.0186	0.0973	0.2415	0.0204
Other qualification	-0.1512	0.2406	-0.0317	-0.1429	0.2406	-0.0300
Unknown qualification	-0.1058	0.3640	-0.0222	-0.1081	0.3633	-0.0227
Switching school						
1	-0.0202	0.0207	-0.0042	-0.0196	0.0207	-0.0041
2	-0.1036***	0.0245	-0.0217***	-0.1011***	0.0245	-0.0212***
3	-0.1540***	0.0340	-0.0323***	-0.1523***	0.0340	-0.0320***
4 or more	-0.3148***	0.0510	-0.0661***	-0.3085***	0.0511	-0.0648***
Truancy, suspension and/or unjustified absence	0.0211	0.0509	0.0044	0.0193	0.0509	0.0041

Age						
Under 18	0.2035	0.4342	0.0427	0.1972	0.4343	0.0414
19	0.0013	0.0688	0.0003	0.0034	0.0689	0.0007
20	0.0644	0.0683	0.0135	0.0649	0.0684	0.0136
21	0.0694	0.0695	0.0146	0.0672	0.0697	0.0141
22 or older	-0.0926	0.0762	-0.0194	-0.0874	0.0764	-0.0184
Employment hours						
Low (>0 and <10) hours	0.0377	0.0399	0.0079	0.0359	0.0399	0.0075
Medium (≥10 and <20) hours	0.0653	0.0513	0.0137	0.0634	0.0512	0.0133
High (≥20) hours	0.0046	0.0252	0.0010	0.0053	0.0253	0.0011
Unknown	0.0213	0.0203	0.0045	0.0217	0.0203	0.0046
Prior activity						
Non-employed or beneficiary	-0.0968*	0.0562	-0.0203*	-0.0917*	0.0565	-0.0193*
employed	-0.1151***	0.0294	-0.0242***	-0.1072***	0.0294	-0.0225***
Other type student	-0.1868***	0.0379	-0.0392***	-0.1871***	0.0379	-0.0393***
House person or retired	-0.0048	0.1850	-0.0010	-0.0156	0.1847	-0.0033
Overseas	0.0434	0.0462	0.0091	0.0508	0.0463	0.0107
University						
University B	-0.0849***	0.0282	-0.0178***	-0.0829***	0.0282	-0.0174***
University C	-0.1029***	0.0254	-0.0216***	-0.0984***	0.0253	-0.0207***
University D	-0.0068	0.0214	-0.0014	-0.0052	0.0214	-0.0011
University E	-0.1341***	0.0278	-0.0282***	-0.1294***	0.0278	-0.0272***
University F	0.2435***	0.0605	0.0511***	0.2541***	0.0604	0.0534***
University G	-0.1300***	0.0217	-0.0273***	-0.1263***	0.0217	-0.0265***
University H	0.0569***	0.0218	0.0120***	0.0566***	0.0218	0.0119***

New Zealand citizenship	0.0987***	0.0220	0.0207***	0.1042***	0.0220	0.0219***
Part-time	-1.0370***	0.0202	-0.2178***	-1.0383***	0.0202	-0.2181***
Single-parent household	-0.0645***	0.0207	-0.0135***	-0.0621***	0.0208	-0.0130***
Parents highest qualification						
Level 1 to 3 Certificate	0.0456	0.0357	0.0096	0.0518	0.0357	0.0109
Level 4 Certificate	0.0486	0.0400	0.0102	0.0585	0.0399	0.0123
Level 5 to 6 Diploma	0.1199***	0.0378	0.0252***	0.1279***	0.0377	0.0269***
Bachelor's degree and Level 7 Qualification	0.1700***	0.0359	0.0357***	0.1763***	0.0358	0.0370***
Postgraduate	0.1789***	0.0378	0.0376***	0.1842***	0.0377	0.0387***
Overseas School Qualification	0.1129***	0.0367	0.0237***	0.1135***	0.0367	0.0238***
Unknown qualification	0.0698	0.0656	0.0147	0.0717	0.0654	0.0151
Household income (in NZD)						
Zero	0.2090	0.1354	0.0439	0.2145	0.1376	0.0450
1 to 30,000	-0.0015	0.0379	-0.0003	-0.0044	0.0379	-0.0009
50,001 to 70,000	-0.0212	0.0339	-0.0044	-0.0195	0.0338	-0.0041
70,001 to 100,000	0.0314	0.0317	0.0066	0.0336	0.0318	0.0071
100,001 or more	0.0201	0.0298	0.0042	0.0252	0.0298	0.0053
Unknown	-0.0663**	0.0330	-0.0139**	-0.0658**	0.0331	-0.0138**
Parents' benefit and charges known	-0.1495*	0.0771	-0.0314*	-0.1517**	0.0770	-0.0319**
Parents' proportions days living on benefit	0.0579	0.0463	0.0122	0.0551	0.0463	0.0116
Parents' crime charges	0.0020	0.0027	0.0004	0.0020	0.0027	0.0004
Number of observations (n)	31,722			33,342		
Likelihood function	-12115.468			-12121.923		
Pseudo R^2	0.1184			0.1179		

* denotes statistical significance at a 10% level; ** denotes statistical significance at a 5% level; *** denotes statistical significance at a 1% level.

Table A.4.8. Proportions of student subsamples based on second-year employment hours: prioritised ethnicity

	Sample proportion[a]
Original thresholds of hours	
Employment hours	
Zero	0.1120
Low (>0 and 10 hours)	0.0321
Medium (≥10 and <20 hours)	0.0183
High (≥20 hours)	0.1424
Unknown	0.6954
Only Known hours	
Employment hours	
Zero	0.3676
Low (>0 and 10 hours)	0.1052
Medium (≥10 and <20 hours)	0.0599
High (≥20 hours)	0.4672
Full-time students	
Employment hours	
Zero	0.1122
Low (>0 and 10 hours)	0.0318
Medium (≥10 and <20 hours)	0.0182
High (≥20 hours)	0.1427
Unknown	0.6950
Female students	
Employment hours	
Zero	0.1088
Low (>0 and 10 hours)	0.0322
Medium (≥10 and <20 hours)	0.0175
High (≥20 hours)	0.1434
Unknown	0.6983
Male students	
Employment hours	
Zero	0.1168
Low (>0 and 10 hours)	0.0319
Medium (≥10 and <20 hours)	0.0193
High (≥20 hours)	0.1409
Unknown	0.6914
Māori and Pasifika students	
Employment hours	

Zero	0.1140
Low (>0 and 10 hours)	0.0312
Medium (≥10 and <20 hours)	0.0156
High (≥20 hours)	0.1508
Unknown	0.6891
Māori students	
Employment hours	
Zero	0.1113
Low (>0 and 10 hours)	0.0325
Medium (≥10 and <20 hours)	0.0113
High (≥20 hours)	0.1400
Unknown	0.7050
Pasifika students	
Employment hours	
Zero	0.1175
Low (>0 and 10 hours)	0.0294
Medium (≥10 and <20 hours)	0.0196
High (≥20 hours)	0.1648
Unknown	0.6672
Non-Māori and non-Pasifika students	
Employment hours	
Zero	0.1117
Low (>0 and 10 hours)	0.0322
Medium (≥10 and <20 hours)	0.0187
High (≥20 hours)	0.1411
Unknown	0.6964
Asian students	
Employment hours	
Zero	0.1111
Low (>0 and 10 hours)	0.0322
Medium (≥10 and <20 hours)	0.0176
High (≥20 hours)	0.1388
Unknown	0.7009
European students	
Employment hours	
Zero	0.1131
Low (>0 and 10 hours)	0.0322
Medium (≥10 and <20 hours)	0.0190
High (≥20 hours)	0.1414

Unknown	0.6943
New thresholds of hours	
Employment hours	
Zero	0.1120
Low (>0 and <15 hours)	0.0419
Medium (≥15 and <30 hours)	0.0245
High (≥30 hours)	0.1263
Unknown	0.6954

[a] The values reported in this table were produced based on the number of observations and have been randomly rounded to base 3 or supressed if the original observation count is fewer than 6 due to the confidentiality requirements from Statistics New Zealand.

Table A.4.9. Probit regression results for employment hours on second-year course completion by student subsamples

	Prioritised ethnicity			All ethnicities		
	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Only Known hours						
Employment hours						
Low (>0 and 10 hours)	0.0371	0.0400	0.0078	0.0349	0.0399	0.0073
Medium (≥10 and <20 hours)	0.0672	0.0515	0.0141	0.0654	0.0514	0.0137
High (≥20 hours)	0.0079	0.0254	0.0017	0.0088	0.0254	0.0019
Full-time students						
Employment hours						
Low (>0 and 10 hours)	0.0649	0.0436	0.0125	0.0629	0.0436	0.0121
Medium (≥10 and <20 hours)	0.0810	0.0548	0.0156	0.0795	0.0547	0.0153
High (≥20 hours)	0.0111	0.0269	0.0021	0.0118	0.0269	0.0023
Unknown	0.0236	0.0216	0.0045	0.0244	0.0216	0.0047
Female students						
Employment hours						
Low (>0 and 10 hours)	0.0807	0.0557	0.0148	0.0772	0.0556	0.0141
Medium (≥10 and <20 hours)	0.0702	0.0736	0.0128	0.0709	0.0736	0.0130
High (≥20 hours)	0.0018	0.0347	0.0003	0.0041	0.0347	0.0007
Unknown	0.0450	0.0280	0.0082	0.0463	0.0280	0.0085
Male students						
Employment hours						
Low (>0 and 10 hours)	-0.0094	0.0581	-0.0023	-0.0099	0.0581	-0.0024
Medium (≥10 and <20 hours)	0.0509	0.0708	0.0126	0.0484	0.0708	0.0120

High (≥20 hours)	0.0122	0.0366	0.0030	0.0115	0.0367	0.0028
Unknown	-0.0037	0.0293	-0.0009	-0.0039	0.0294	-0.0010
Māori and Pasifika students						
Employment hours						
Low (>0 and 10 hours)	0.1023	0.0945	0.0313	0.0999	0.0945	0.0306
Medium (≥10 and <20 hours)	0.0209	0.1360	0.0064	0.0147	0.1360	0.0045
High (≥20 hours)	0.0017	0.0597	0.0005	0.0045	0.0599	0.0014
Unknown	0.0397	0.0484	0.0121	0.0411	0.0486	0.0126
Māori students						
Employment hours						
Low (>0 and 10 hours)	0.1273	0.1228	0.0361	-	-	-
Medium (≥10 and <20 hours)	0.2169	0.1927	0.0614	-	-	-
High (≥20 hours)	0.0147	0.0874	0.0042	-	-	-
Unknown	-0.0185	0.0685	-0.0053	-	-	-
Pasifika students						
Employment hours						
Low (>0 and 10 hours)	0.0207	0.1490	0.0068	-0.0311	0.1434	-0.0103
Medium (≥10 and <20 hours)	-0.0971	0.1909	-0.0320	-0.1420	0.1908	-0.0468
High (≥20 hours)	-0.0119	0.0837	-0.0039	-0.0377	0.0819	-0.0124
Unknown	0.1026	0.0699	0.0338	0.0563	0.0680	0.0186
Non-Māori and non-Pasifika students						
Employment hours						
Low (>0 and 10 hours)	0.0231	0.0440	0.0045	0.0395	0.0430	0.0078
Medium (≥10 and <20 hours)	0.0707	0.0551	0.0137	0.0846	0.0541	0.0167
High (≥20 hours)	0.0030	0.0279	0.0006	0.0000	0.0273	0.0000

Unknown	0.0165	0.0224	0.0032	0.0228	0.0219	0.0045
Asian students						
Employment hours						
Low (>0 and 10 hours)	0.0567	0.0951	0.0119	0.0516	0.0942	0.0108
Medium (≥10 and <20 hours)	0.1031	0.1061	0.0216	0.1103	0.1058	0.0231
High (≥20 hours)	0.0445	0.0558	0.0093	0.0360	0.0553	0.0075
Unknown	0.0481	0.0455	0.0101	0.0446	0.0451	0.0093
European students						
Employment hours						
Low (>0 and 10 hours)	0.0456	0.0514	0.0085	0.0454	0.0496	0.0087
Medium (≥10 and <20 hours)	0.0554	0.0654	0.0104	0.0752	0.0634	0.0144
High (≥20 hours)	0.0004	0.0331	0.0001	-0.0107	0.0318	-0.0020
Unknown	0.0109	0.0262	0.0020	0.0165	0.0253	0.0032
New thresholds of hours						
Employment hours						
Low (>0 and <15 hours)	0.0434	0.0360	0.0091	0.0416	0.0360	0.0087
Medium (≥15 and <30 hours)	-0.0035	0.0458	-0.0007	-0.0034	0.0457	-0.0007
High (≥30 hours)	0.0104	0.0259	0.0022	0.0111	0.0260	0.0023
Unknown	0.0213	0.0203	0.0045	0.0217	0.0203	0.0046

* denotes statistical significance at a 10% level; ** denotes statistical significance at a 5% level; *** denotes statistical significance at a 1% level.

Table A.4.10. Linear decomposition of second-year course completion

	Prioritised ethnicity			All ethnicities		
	European vs Māori	European vs Pasifika	European vs Asian	European vs Māori	European vs Pasifika	European vs Asian
Total ethnic difference	0.1243	0.2485	0.0221	0.1202	0.2431	0.0189
Explained ethnic difference[a]						
Female	-0.0029*** (0.0007) [-2.33%]	-0.0048*** (0.0007) [-1.93%]	0.0022*** (0.0005) [9.95%]	-0.0029*** (0.0007) [-2.41%]	-0.0046*** (0.0007) [-1.89%]	0.0023*** (0.0005) [12.17%]
Cohort year	0.0001 (0.0001) [0.08%]	-0.0001 (0.0001) [-0.04%]	-0.0001 (0.0001) [-0.45%]	0.0001 (0.0001) [0.08%]	-0.0001 (0.0001) [-0.04%]	-0.0001 (0.0001) [-0.53%]
School deciles	0.0131*** (0.0020) [10.54%]	0.0333*** (0.0041) [13.40%]	0.0024*** (0.0007) [10.86%]	0.0131*** (0.0019) [10.90%]	0.0329*** (0.0038) [13.53%]	0.0022*** (0.0007) [11.64%]
School qualification	0.0031*** (0.0006) [2.49%]	0.0047*** (0.0008) [1.89%]	0.0006 (0.0005) [2.71%]	0.0031*** (0.0006) [2.58%]	0.0052*** (0.0008) [2.14%]	0.0007 (0.0005) [3.70%]
Other school factors	0.0021** (0.0009) [1.69%]	0.0009 (0.0009) [0.36%]	0.0016* (0.0009) [7.24%]	0.0020** (0.0009) [1.66%]	0.0015* (0.0009) [0.62%]	0.0017** (0.0009) [8.99%]
Age	0.0008** (0.0004) [0.64%]	0.0001 (0.0003) [0.04%]	-0.0002 (0.0001) [-0.90%]	0.0009*** (0.0003) [0.75%]	0.0004 (0.0004) [0.16%]	-0.0003* (0.0001) [-1.59%]

Employment hours	0.0001 (0.0001) [0.08%]	0.0001 (0.0001) [0.04%]	0.0000 (0.0000) [0.00%]	0.0001 (0.0001) [0.08%]	0.0001 (0.0001) [0.04%]	0.0000 (0.0000) [0.00%]
New Zealand citizenship	-0.0009** (0.0004) [-0.72%]	0.0009** (0.0004) [0.36%]	0.0042*** (0.0010) [19.00%]	-0.0008** (0.0004) [-0.67%]	0.0007** (0.0003) [0.29%]	0.0039*** (0.0010) [20.63%]
Part-time	0.0108*** (0.0022) [8.69%]	0.0436*** (0.0034) [17.55%]	0.0042*** (0.0013) [19.00%]	0.0108*** (0.0022) [8.99%]	0.0415*** (0.0032) [17.07%]	0.0041*** (0.0013) [21.69%]
University	-0.0017*** (0.0006) [-1.37%]	-0.0070*** (0.0014) [-2.82%]	-0.0054*** (0.0014) [-24.43%]	-0.0015** (0.0006) [-1.25%]	-0.0071*** (0.0013) [-2.92%]	-0.0052*** (0.0013) [-27.51%]
Prior activity	0.0010*** (0.0004) [0.80%]	0.0007 (0.0006) [0.28%]	-0.0015*** (0.0004) [-6.79%]	0.0010*** (0.0003) [0.83%]	0.0008 (0.0005) [0.33%]	-0.0014*** (0.0004) [-7.41%]
Parental education	0.0057*** (0.0018) [4.59%]	0.0075*** (0.0016) [3.02%]	-0.0019 (0.0016) [-8.60%]	0.0052*** (0.0017) [4.33%]	0.0074*** (0.0015) [3.04%]	-0.0020 (0.0016) [-10.58%]
Other parental factors	-0.0002 (0.0017) [-0.16%]	0.0027** (0.0011) [1.09%]	0.0016 (0.0012) [7.24%]	0.0001 (0.0016) [0.08%]	0.0026** (0.0010) [1.07%]	0.0017 (0.0012) [8.99%]
Total explained differences	0.0311*** (0.0033) [25.02%]	0.0826*** (0.0059) [33.24%]	0.0077*** (0.0029) [34.84%]	0.0312*** (0.0032) [25.96%]	0.0813*** (0.0055) [33.44%]	0.0076*** (0.0028) [40.21%]

Unexplained differences	0.0932*** (0.0066) [74.98%]	0.1659*** (0.0087) [66.76%]	0.0144*** (0.0042) [65.16%]	0.0890*** (0.0066) [74.04%]	0.1618*** (0.0083) [66.56%]	0.0113*** (0.0041) [59.79%]
Number of observations (n)	22,902	22,338	26,469	24,270	23,841	27,924

All decompositions used the estimated coefficients from pooled OLS regressions of relevant ethnic groups (European and Māori, European and Pasifika and European and Asian). A dummy variable was created to indicate the each of the ethnic minority groups used in the regression.

[a] Standard errors associated with these individual or groups of variables are listed in parentheses ('( )') below these estimated effects. Percentage changes in the total ethnic differences associated with these individual or groups of variables are shown in square brackets ('[ ]').

* denotes statistical significance at a 10% level; ** denotes statistical significance at a 5% level; *** denotes statistical significance at a 1% level.

Table A.4.11. Probit regression results for qualification completion by age 24

Variables	Prioritised ethnicity			All ethnicities		
	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Constant	-0.2169	0.4423	-	-0.2252	0.4435	-
Female	0.3286***	0.0215	0.1180***	0.3293***	0.0215	0.1182***
Ethnicities						
Māori	-0.3940***	0.0392	-0.1414***	-0.3586***	0.0389	-0.1288***
Pasifika	-0.5998***	0.0507	-0.2153***	-0.5687***	0.0484	-0.2042***
Asian	-0.0009	0.0326	-0.0003	0.0195	0.0321	0.0070
MELAA	-0.2559***	0.0714	-0.0919***	-0.2477***	0.0712	-0.0889***
Other	-0.0528	0.1092	-0.0190	-0.0514	0.1065	-0.0185
School deciles						
Decile 1	-0.5869***	0.0990	-0.2107***	-0.5955***	0.0989	-0.2138***
Decile 2	-0.2756***	0.0704	-0.0989***	-0.2824***	0.0704	-0.1014***
Decile 3	-0.3220***	0.0613	-0.1156***	-0.3270***	0.0612	-0.1174***
Decile 4	-0.1157**	0.0535	-0.0416**	-0.1165**	0.0535	-0.0418**
Decile 5	-0.0720	0.0491	-0.0259	-0.0750	0.0491	-0.0269
Decile 7	-0.0514	0.0458	-0.0185	-0.0521	0.0458	-0.0187
Decile 8	-0.0687	0.0431	-0.0247	-0.0692	0.0431	-0.0248
Decile 9	0.0065	0.0425	0.0023	0.0065	0.0425	0.0023
Decile 10	-0.0198	0.0435	-0.0071	-0.0200	0.0435	-0.0072
Decile unknown	0.3095	0.2350	0.1111	0.3075	0.2349	0.1104
School authority						

State integrated	-0.0318	0.0316	-0.0114	-0.0326	0.0315	-0.0117
Private	-0.0108	0.0457	-0.0039	-0.0075	0.0457	-0.0027
Unknown	0.1395	0.5219	0.0501	0.1505	0.5220	0.0540
School gender						
Single sex	-0.0481*	0.0248	-0.0173*	-0.0465*	0.0248	-0.0167*
Unknown	-0.2386	0.5071	-0.0857	-0.2416	0.5072	-0.0867
School area						
Minor Urban Area	0.0899**	0.0443	0.0323**	0.0918**	0.0443	0.0330**
Secondary Urban Area	-0.0127	0.0460	-0.0046	-0.0106	0.0460	-0.0038
Rural Area	0.0045	0.0952	0.0016	0.0068	0.0953	0.0024
Unknown	-0.2641	0.2022	-0.0948	-0.2686	0.2021	-0.0964
Highest school qualification						
NCEA Level 1	0.0179	0.4368	0.0064	0.0066	0.4381	0.0024
NCEA Level 2	-0.1450	0.4158	-0.0521	-0.1527	0.4172	-0.0548
NCEA Level 3	0.5333	0.4125	0.1915	0.5253	0.4139	0.1886
Overseas qualification	0.4532	0.4153	0.1627	0.4434	0.4166	0.1592
Other qualification	0.0826	0.4139	0.0297	0.0729	0.4152	0.0262
Unknown qualification	0.2129	0.6183	0.0764	0.2016	0.6181	0.0724
Switching school						
1	-0.0134	0.0308	-0.0048	-0.0141	0.0308	-0.0051
2	-0.2355***	0.0409	-0.0845***	-0.2337***	0.0409	-0.0839***
3	-0.2365***	0.0703	-0.0849***	-0.2325***	0.0703	-0.0835***
4 or more	-0.5392***	0.1133	-0.1936***	-0.5326***	0.1137	-0.1912***
Truancy, suspension and/or unjustified absence	0.0341	0.0812	0.0122	0.0312	0.0812	0.0112
Employment hours						

Low (>0 and <10) hours	0.2006***	0.0542	0.0720***	0.1987***	0.0542	0.0714***
Medium (≥10 and <20) hours	0.1080	0.0725	0.0388	0.1062	0.0725	0.0381
High (≥20) hours	0.0476	0.0423	0.0171	0.0470	0.0423	0.0169
Unknown	0.0331	0.0347	0.0119	0.0330	0.0347	0.0119
Prior activity						
Non-employed or beneficiary	-0.8054***	0.0961	-0.2892***	-0.8007***	0.0961	-0.2875***
Employed	-0.8004***	0.0448	-0.2873***	-0.7957***	0.0448	-0.2857***
Other type student	-0.6597***	0.0613	-0.2368***	-0.6609***	0.0613	-0.2373***
House person or retired	-0.7858**	0.3962	-0.2821**	-0.7770**	0.3983	-0.2790**
Overseas	-0.3148***	0.0657	-0.1130***	-0.3136***	0.0657	-0.1126***
University						
University B	0.0653	0.0506	0.0234	0.0670	0.0507	0.0241
University C	-0.1482***	0.0430	-0.0532***	-0.1451***	0.0430	-0.0521***
University D	0.2104***	0.0345	0.0755***	0.2115***	0.0345	0.0759***
University E	0.0978**	0.0454	0.0351**	0.1001**	0.0454	0.0360**
University F	0.0972	0.0884	0.0349	0.1002	0.0884	0.0360
University G	-0.0115	0.0338	-0.0041	-0.0098	0.0338	-0.0035
University H	0.1339***	0.0347	0.0481***	0.1345***	0.0347	0.0483***
New Zealand citizenship	0.0884**	0.0369	0.0317**	0.0923**	0.0369	0.0332**
Single-parent household	-0.0703**	0.0353	-0.0252**	-0.0681*	0.0353	-0.0244*
Parents highest qualification						
Level 1 to 3 Certificate	-0.0542	0.0636	-0.0195	-0.0508	0.0636	-0.0182
Level 4 Certificate	-0.1166*	0.0702	-0.0419*	-0.1127*	0.0702	-0.0405*
Level 5 to 6 Diploma	0.0087	0.0667	0.0031	0.0143	0.0667	0.0051
Bachelor’s degree and Level 7 Qualification	-0.0431	0.0635	-0.0155	-0.0413	0.0635	-0.0148

Postgraduate	-0.0286	0.0668	-0.0103	-0.0268	0.0668	-0.0096
Overseas School Qualification	-0.0242	0.0661	-0.0087	-0.0270	0.0662	-0.0097
Unknown qualification	-0.1956*	0.1068	-0.0702*	-0.1966*	0.1068	-0.0706*
Household income (in NZD)						
Zero	-0.2890	0.2520	-0.1038	-0.2632	0.2512	-0.0945
1 to 30,000	-0.1151*	0.0663	-0.0413*	-0.1169*	0.0663	-0.0420*
50,001 to 70,000	0.0064	0.0582	0.0023	0.0079	0.0582	0.0028
70,001 to 100,000	0.0265	0.0538	0.0095	0.0304	0.0538	0.0109
100,001 or more	0.0002	0.0509	0.0001	0.0060	0.0509	0.0022
Unknown	0.0262	0.0579	0.0094	0.0286	0.0579	0.0103
Parents' benefit and charges known	-0.2542**	0.1229	-0.0912**	-0.2547**	0.1229	-0.0915**
Parents' proportions days living on benefit	-0.0173	0.0733	-0.0062	-0.0203	0.0733	-0.0073
Parents' crime charges	0.0049	0.0046	0.0018	0.0049	0.0046	0.0018
Number of observations (n)	15,843			16,722		
Likelihood function	-9952.2624			-9952.4969		
Pseudo R^2	0.0926			0.0925		

* denotes statistical significance at a 10% level; ** denotes statistical significance at a 5% level; *** denotes statistical significance at a 1% level.

Table A.4.12. Proportions of student subsamples based on average employment hours during university years to age 24: prioritised ethnicity

	Sample proportion[a]
Original thresholds of hours	
Employment hours	
Zero	0.1028
Low (>0 and 10 hours)	0.0572
Medium (≥10 and <20 hours)	0.0259
High (≥20 hours)	0.1447
Unknown	0.6696
Only Known hours	
Employment hours	
Zero	0.3114
Low (>0 and 10 hours)	0.1732
Medium (≥10 and <20 hours)	0.0780
High (≥20 hours)	0.4375
Female students	
Employment hours	
Zero	0.0995
Low (>0 and 10 hours)	0.0570
Medium (≥10 and <20 hours)	0.0251
High (≥20 hours)	0.1432
Unknown	0.6756
Male students	
Employment hours	
Zero	0.1076
Low (>0 and 10 hours)	0.0580
Medium (≥10 and <20 hours)	0.0267
High (≥20 hours)	0.1467
Unknown	0.6615
Māori and Pasifika students	
Employment hours	
Zero	0.0949
Low (>0 and 10 hours)	0.0506
Medium (≥10 and <20 hours)	0.0253
High (≥20 hours)	0.1481
Unknown	0.6823
Māori students	
Employment hours	

Zero	0.0894
Low (>0 and 10 hours)	0.0532
Medium (≥10 and <20 hours)	0.0213
High (≥20 hours)	0.1383
Unknown	0.6979
Pasifika students	
Employment hours	
Zero	0.1066
Low (>0 and 10 hours)	0.0439
Medium (≥10 and <20 hours)	0.0282
High (≥20 hours)	0.1630
Unknown	0.6583
Non-Māori and non-Pasifika students	
Employment hours	
Zero	0.1042
Low (>0 and 10 hours)	0.0583
Medium (≥10 and <20 hours)	0.0261
High (≥20 hours)	0.1441
Unknown	0.6676
Asian students	
Employment hours	
Zero	0.1027
Low (>0 and 10 hours)	0.0590
Medium (≥10 and <20 hours)	0.0240
High (≥20 hours)	0.1454
Unknown	0.6689
European students	
Employment hours	
Zero	0.1050
Low (>0 and 10 hours)	0.0575
Medium (≥10 and <20 hours)	0.0264
High (≥20 hours)	0.1432
Unknown	0.6682
New thresholds of hours	
Employment hours	
Zero	0.1028
Low (>0 and <15 hours)	0.0712
Medium (≥15 and <30 hours)	0.0303
High (≥30 hours)	0.1259

Unknown	0.6696

[a] The values reported in this table were produced based on the number of observations and have been randomly rounded to base 3 or supressed if the original observation count is fewer than 6 due to the confidentiality requirements from Statistics New Zealand.

Table A.4.13. Probit regression results for employment hours on qualification completion by student subsamples

	Prioritised ethnicity			All ethnicities		
	Estimated coefficients	Standard errors	Marginal effects	Estimated coefficients	Standard errors	Marginal effects
Only Known hours						
Employment hours						
Low (>0 and 10 hours)	0.2014***	0.0546	0.0719***	0.1995***	0.0546	0.0712***
Medium (≥10 and <20 hours)	0.1037	0.0729	0.0370	0.1022	0.0729	0.0365
High (≥20 hours)	0.0482	0.0426	0.0172	0.0476	0.0426	0.0170
Female students						
Employment hours						
Low (>0 and 10 hours)	0.2299***	0.0723	0.0807***	0.2267***	0.0723	0.0796***
Medium (≥10 and <20 hours)	0.1794*	0.0977	0.0630*	0.1781*	0.0977	0.0625*
High (≥20 hours)	0.0609	0.0562	0.0214	0.0597	0.0562	0.0210
Unknown	0.0802*	0.0462	0.0282*	0.0801*	0.0462	0.0281*
Male students						
Employment hours						
Low (>0 and 10 hours)	0.1671**	0.0825	0.0612**	0.1666**	0.0825	0.0610**
Medium (≥10 and <20 hours)	-0.0035	0.1092	-0.0013	-0.0052	0.1092	-0.0019
High (≥20 hours)	0.0379	0.0646	0.0139	0.0379	0.0646	0.0139
Unknown	-0.0232	0.0529	-0.0085	-0.0230	0.0529	-0.0084
Māori and Pasifika students						
Employment hours						
Low (>0 and 10 hours)	0.2093	0.1531	0.0672	0.2010	0.1533	0.0644
Medium (≥10 and <20 hours)	0.2775	0.1983	0.0890	0.2707	0.1985	0.0868

High (≥20 hours)	-0.0013	0.1178	-0.0004	-0.0032	0.1179	-0.0010
Unknown	0.0200	0.0977	0.0064	0.0185	0.0978	0.0059
Māori students						
Employment hours						
Low (>0 and 10 hours)	0.2529	0.1969	0.0840	-	-	-
Medium (≥10 and <20 hours)	0.3481	0.2735	0.1156	-	-	-
High (≥20 hours)	-0.0125	0.1570	-0.0042	-	-	-
Unknown	0.0614	0.1295	0.0204	-	-	-
Pasifika students						
Employment hours						
Low (>0 and 10 hours)	0.0666	0.2570	0.0198	0.1364	0.2522	0.0400
Medium (≥10 and <20 hours)	0.2209	0.3039	0.0655	0.1893	0.3005	0.0556
High (≥20 hours)	0.0209	0.1869	0.0062	0.1265	0.1800	0.0371
Unknown	-0.0467	0.1545	-0.0139	0.0254	0.1494	0.0075
Non-Māori and non-Pasifika students						
Employment hours						
Low (>0 and 10 hours)	0.2039***	0.0582	0.0740***	0.2025***	0.0572	0.0736***
Medium (≥10 and <20 hours)	0.0822	0.0781	0.0298	0.0945	0.0766	0.0343
High (≥20 hours)	0.0533	0.0455	0.0193	0.0423	0.0447	0.0154
Unknown	0.0381	0.0372	0.0138	0.0259	0.0366	0.0094
Asian students						
Employment hours						
Low (>0 and 10 hours)	0.2817**	0.1307	0.1031**	0.3023**	0.1301	0.1107**
Medium (≥10 and <20 hours)	-0.0157	0.1765	-0.0057	-0.0045	0.1753	-0.0017
High (≥20 hours)	0.0134	0.1022	0.0049	0.0136	0.1013	0.0050

Unknown	0.0112	0.0837	0.0041	0.0121	0.0831	0.0044
European students						
Employment hours						
Low (>0 and 10 hours)	0.1899***	0.0672	0.0683***	0.1870***	0.0653	0.0674***
Medium (≥10 and <20 hours)	0.0605	0.0898	0.0218	0.0829	0.0868	0.0299
High (≥20 hours)	0.0587	0.0523	0.0211	0.0391	0.0507	0.0141
Unknown	0.0487	0.0427	0.0175	0.0309	0.0415	0.0111
New thresholds of hours						
Employment hours						
Low (>0 and <15 hours)	0.1741***	0.0507	0.0625***	0.1726***	0.0507	0.0620***
Medium (≥15 and <30 hours)	0.1331*	0.0682	0.0478*	0.1324*	0.0682	0.0475*
High (≥30 hours)	0.0377	0.0435	0.0135	0.0367	0.0436	0.0132
Unknown	0.0331	0.0347	0.0119	0.0330	0.0347	0.0119

* denotes statistical significance at a 10% level; ** denotes statistical significance at a 5% level; *** denotes statistical significance at a 1% level

Table A.4.14. Nonlinear decomposition of qualification completion by age 24

	Prioritised ethnicity			All ethnicities		
	European vs Māori	European vs Pasifika	European vs Asian	European vs Māori	European vs Pasifika	European vs Asian
Total ethnic difference	0.1915	0.2725	0.0177	0.1845	0.2723	0.0121
Explained ethnic difference[a]						
Female	-0.0060*** (0.0005) [-3.13%]	-0.0091*** (0.0007) [-3.34%]	0.0066*** (0.0005) [37.29%]	-0.0060*** (0.0005) [-3.25%]	-0.0095*** (0.0007) [-3.49%]	0.0066*** (0.0005) [54.55%]
School deciles	0.0245*** (0.0047) [12.79%]	0.0546*** (0.0083) [20.04%]	0.0066*** (0.0020) [37.29%]	0.0246*** (0.0046) [13.33%]	0.0544*** (0.0076) [19.98%]	0.0068*** (0.0019) [56.20%]
School qualification	0.0069*** (0.0010) [3.60%]	0.0150*** (0.0014) [5.50%]	0.0033*** (0.0011) [18.64%]	0.0064*** (0.0009) [3.47%]	0.0156*** (0.0013) [5.73%]	0.0028** (0.0011) [23.14%]
Other school factors	0.0077*** (0.0022) [4.02%]	0.0047* (0.0025) [1.72%]	-0.0008 (0.0023) [-4.52%]	0.0078*** (0.0022) [4.23%]	0.0061*** (0.0023) [2.24%]	0.0002 (0.0023) [1.65%]
Employment hours	0.0000 (0.0004) [0.00%]	0.0003 (0.0004) [0.11%]	-0.0002 (0.0002) [-1.13%]	0.0001 (0.0004) [0.05%]	0.0006 (0.0004) [0.22%]	0.0000 (0.0002) [0.00%]
New Zealand citizenship	-0.0017 (0.0011) [-0.89%]	0.0009 (0.0007) [0.33%]	0.0080*** (0.0028) [45.20%]	-0.0015 (0.0011) [-0.81%]	0.0008 (0.0006) [0.29%]	0.0073*** (0.0028) [60.33%]

University	-0.0019 (0.0018) [-0.99%]	-0.0021 (0.0037) [-0.77%]	0.0047 (0.0037) [26.55%]	-0.0011 (0.0017) [-0.60%]	-0.0016 (0.0034) [-0.59%]	0.0046 (0.0036) [38.02%]
Prior activity	0.0160*** (0.0010) [8.36%]	0.0021 (0.0013) [0.77%]	-0.0149*** (0.0009) [-84.18%]	0.0164*** (0.0010) [8.89%]	0.0054*** (0.0013) [1.98%]	-0.0144*** (0.0008) [-119.01%]
Parental education	0.0046 (0.0037) [2.40%]	0.0010 (0.0039) [0.37%]	-0.0063 (0.0045) [-35.59%]	0.0046 (0.0037) [2.49%]	0.0005 (0.0036) [0.18%]	-0.0069 (0.0045) [-57.02%]
Other parental factors	-0.0049 (0.0034) [-2.56%]	0.0002 (0.0034) [0.07%]	0.0076** (0.0034) [42.94%]	-0.0051 (0.0035) [-2.76%]	0.0014 (0.0031) [0.51%]	0.0074** (0.0033) [61.16%]
Total explained differences	0.0452*** (0.0051) [23.60%]	0.0676*** (0.0106) [24.81%]	0.0146** (0.0070) [82.49%]	0.0462*** (0.0050) [25.04%]	0.0737*** (0.0097) [27.07%]	0.0144** (0.0068) [119.01%]
Unexplained differences	0.1463 [76.40%]	0.2049 [75.19%]	0.0031 [17.51%]	0.1383 [74.96%]	0.1986 [72.93%]	-
Number of observations (n)	11,640	11,184	12,972	12,372	12,003	13,755

All decompositions used the estimated coefficients from pooled Probit regressions of relevant ethnic groups (European and Māori, European and Pasifika and European and Asian). A dummy variable was created to indicate the each of the ethnic minority groups used in the regression. The decomposition results reported in this table are based on 100 replications.

[a] Standard errors associated with these individual or groups of variables are listed in parentheses ('( )') below these estimated effects. Percentage changes in the total ethnic differences associated with these individual or groups of variables are shown in square brackets ('[ ]').

* denotes statistical significance at a 10% level; ** denotes statistical significance at a 5% level; *** denotes statistical significance at a 1% level.