School of Economics

Working Paper Series

Explaining ethnic disparities in bachelor's degree participation: Evidence from NZ

Lisa Meehan, Gail Pacheco and Zoe Pushon

Working Paper – 17/03

Explaining ethnic disparities in bachelor's degree participation: Evidence from NZ

Lisa Meehan[a,1], Gail Pacheco[b] and Zoe Pushon[c,2]

[a] New Zealand Productivity Commission, PO Box 8036, Wellington 6143, New Zealand.

[b] Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand.

[c] New Zealand Productivity Commission, PO Box 8036, Wellington 6143, New Zealand.

Corresponding author: Gail Pacheco, gail.pacheco@aut.ac.nz, +64 9 921 9999 ext 5708

Abstract

There are substantial ethnic gaps in higher education in NZ, despite more than a decade of considerable policy effort aimed at this concern. This study uses newly linked administrative data to examine the underachievement of Māori and Pasifika relative to Europeans. We follow a population cohort born between 1990 and 1994 from school through to young adulthood to assess the relative contributions of prior academic performance, socioeconomic status and parental education to these gaps.

Controlling for the relevant covariates narrows the Māori-European gap, and eliminates the Pasifika-European gap in bachelor's degree participation rates. Utilising Fairlie decompositions, we find that school performance is by far the largest contributor to the ethnic gaps. Low socioeconomic status and parental education are also pertinent, but less important. Our results suggest that ethnic-based policies aimed at encouraging participation are likely to have a limited effect if used in isolation, and signal the need for policy interventions earlier in the education system.

Keywords: higher education; ethnicity; bachelor's degree; decomposition

JEL classification: I21, I24

[1] Presently at OECD.

[2] Present address: Castalia Strategic Advisors, 1747 Pennsylvania Avenue NW, Washington DC 20006.

Disclaimer

Access to the data presented was managed by Statistics NZ under strict microdata access protocols and in accordance with the security and confidentiality provisions of the Statistic Act 1975. These findings are not Official Statistics. The opinions, findings, recommendations, and conclusions expressed are those of the authors, not Statistics NZ, the NZ Productivity Commission or AUT.

Acknowledgements

We are grateful to several organisations for providing us with helpful comments and assistance with the data. This includes the Ministry of Education (Ralf Engler and David Earle), the Tertiary Education Commission (Anne Broadbent and Sergei Koudrin), the Treasury (Sarah Tumen), Motu (Isabelle Sin) and Statistics NZ's microdata team. Thank you to Tim Maloney and Isabelle Sin for providing valuable review comments. We also thank the Productivity Commission's 'New models of tertiary education' inquiry team for sponsoring this research and providing useful feedback. We also received helpful feedback from the participants of several workshops. Any errors or omissions remain the responsibility of the authors.

Funding: This work was supported by the NZ Productivity Commission.

1 Introduction

Globally, the number of people participating in higher education continues to rise, with about one third of those in the school leaver age group now participating in post-school education (based on recent UNESCO data – see Marginson, 2016). Despite this growth, disparities in participation by ethnicity are still prevalent, and have been documented by many studies, including Clotfelter, Ladd, and Vigdor (2015) for the US, Crawford and Greaves (2015) for the UK, and Earle (2008) for NZ.

In NZ, the under-representation of Māori and Pasifika ethnic groups at higher levels of post-school education – specifically degree-level study – is of particular policy interest.[3] This is despite the fact that Māori and Pasifika have similar levels of participation as Europeans in post-school education more generally. In 2015 for example, total enrolment rates for post-school education among 18 and 19 year olds were similar for Māori (42.8%) and Europeans (42.5%) and somewhat higher for Pasifika (48.8%). However, Māori and Pasifika had much lower rates of enrolment in bachelor's qualifications - 10.3% and 14.3% respectively compared with 20.4% for Europeans (Ministry of Education, 2016).

There have been numerous policy initiatives in NZ aimed at closing the ethnic gaps in education. For example, boosting the achievement of Māori and Pasifika is currently one of six key government priorities relating to higher education (Ministry of Education, 2014). Crawford (2016), however, laments the fact that despite NZ's long history of policies aimed at increasing participation of Māori and Pasifika, these groups remain under-represented at bachelor's-level study and above.

With the advent of newly linked administrative data in NZ, the time is now ripe to empirically investigate the key factors associated with ethnic disparities in bachelor's qualifications and provide new insights into this old question. To this end, we make use of linked data from a number of NZ government agencies on a population cohort of young people born between 1990 and 1994. These data allow us to track the academic progress of the cohort through high school and on to higher education (until the end of 2014), and examine the likelihood of participating in degree-level study while controlling for the majority of explanatory variables suggested by past studies.

Our main research questions are: *How do degree-level participation rates differ by ethnicity in NZ? To what extent does controlling for relevant individual, school, household and parental characteristics close the ethnic gaps in participation rates; and is this adjustment similar for both Māori and Pasifika, in terms of their gap with Europeans? In explaining the differences, what is the relative importance of the three key factors suggested by the literature - ie, socioeconomic status, school performance, and parents' educational attainment?* While there are a myriad of variables that have been found to be linked with higher education outcomes, these three stand out from the crowd and, to our knowledge, there are no studies that have had adequate data to include all three factors at a population level.

Given the binary nature of the participation outcome, the research questions detailed above are primarily tackled via the use of maximum likelihood estimation of a probit regression model. To address the last question, we also use Fairlie non-linear decomposition models to assess the relative proportions of total ethnic differences that can be explained by the various individual, school, household and parental characteristics.

These questions have potentially important policy implications. For example, if Māori and Pasifika have lower participation rates even after taking into account differences in school performance and other observable characteristics, then policies aimed at encouraging their enrolment in bachelor's degrees may help in lifting their participation. If we find that lower socioeconomic status is largely to blame for the ethnic disparities, then perhaps greater financial assistance for priority groups would aid

[3] This under-representation is part of a wider trend in NZ of generally worse outcomes in a broad range of areas (such as income, health and housing) for Māori and Pasifika in comparison with other ethnic groups (particularly Europeans).

in closing the socioeconomic gap, and subsequently, the education gap. Further, if disparities in degree-level study largely reflect disparities that emerge earlier in the education system, then policy efforts may be best directed towards closing the gaps in school achievement.

The rest of this paper is organised as follows: Section 2 briefly discusses the relevant prior research; Section 3 provides an overview of the NZ education system; Section 4 describes the linked administrative data used and provides descriptive statistics by ethnicity for both the outcome and explanatory variables; Section 5 presents a brief description of the methods used in this research; Section 6 discusses the core results; Section 7 presents a number of variants to our base specifications; and Section 8 concludes and outlines potential future directions for this research agenda.

2 Prior research

Notable contributions in this area include: Cameron and Heckman (2001); Clotfelter et al. (2015); Black and Sufi (2002); Chowdry et al. (2013); and Reardon et al. (2012) - all of which investigate reasons for ethnic differences in enrolment. For instance, Clotfelter et al. (2015) finds that socioeconomic status is an important factor in explaining the likelihood of entering a four-year university in North Carolina. Similarly, Cameron and Heckman (2001) find that differences in parental income make a significant contribution to the ethnic disparity.

Maani (2006) and Strathdee and Engler (2012) are two NZ studies that look at participation in higher education. Maani (2006) uses a longitudinal cohort of all children born in one NZ city (Christchurch) in 1977, and finds that the likelihood of university enrolment at age 18 increases with parental income and education, holding IQ and high school performance at their means. Strathdee and Engler (2012) restrict their sample to those who qualified for university entrance and find that prior school achievement is a strong predictor of the propensity to enrol in higher education.

In general, prior literature has identified three pieces of information that are crucial to explaining variation in the outcome of interest: prior academic performance, parents' educational attainment, and socioeconomic status. The first of these is highlighted by Chowdry et al. (2013) who show that poor achievement in schools is a strong predictor of participation rates in higher education in the UK. The authors conclude by emphasizing the need for policy intervention early in the education system.

The second factor that appears to be of immense value in explaining higher education outcomes is the level of parents' qualifications. Its importance is illustrated by Clotfelter et al. (2015), who find that once parental education is added to the analysis, the predicted propensity for blacks to participate in degree study is higher than their white counterparts. This analysis focussed on data from the North Carolina System in the US, and to our knowledge no study has been able to include parents' education level in a population-wide analysis of higher education outcomes, possibly due to difficulties in obtaining the necessary information.

The third major factor often present in prior literature is socioeconomic status. Given that students who drop out of higher education often cite financial difficulties as an important factor in their decision (Yorke, 1998), it is little wonder that socioeconomic status often comes to the forefront. It is important to note, however, that socioeconomic status tends to be measured via a number of proxies, based on the particular data constraints facing researchers. For instance, it has been captured by family income (eg, Black & Sufi, 2002), deprivation index for neighbourhood meshblocks (eg, Chowdry et al., 2013) and even by parental occupation (eg, Fergusson & Woodward, 2000). The way socioeconomic status is measured is important, as some measures may capture short-term financial and credit constraints, whereas others may reflect a long-term perspective on household resources. The latter is of more relevance to explaining variations in higher-education participation (Cameron & Heckman, 2001).

In terms of available information on socioeconomic status in NZ, school decile has often been used as a proxy (see Earle, 2007; Engler, 2010; Juhong & Maloney, 2006). The school decile system in NZ is used to allocate school funding. School deciles are based on the catchment area of a school. A decile rating of 1 is assigned to the 10% of schools with the highest proportion of students from low socioeconomic communities. However, it is widely acknowledged that school decile may not be the best source of socioeconomic information, as the catchment area for schools (particularly high schools) can be quite large and encompass a range of different communities.

As will be evident in the data section, the use of linked administrative data in this study allows us to replace decile with a more refined measure of socioeconomic status - the deprivation index. This index is calculated for each meshblock in NZ and is based on nine variables from the Census, reflecting eight dimensions of deprivation.[4] We use information on the address each individual in our population cohort lived at most when aged 15 and 16, and find the associated deprivation score based on their location.

It is important to note at this point that while there has been reasonable consensus in the literature regarding the importance of the three factors outlined above; their relative importance is still subject to debate. For instance, Chowdry et al. (2013) and Crawford (2014) argue that the prior achievement gap at UK high schools is more important than the socioeconomic gap for enrolment and retention. On the other hand, Rothstein (2009) argues that the school the student is sourced from is even more important than the prior performance of the individual student.

Our study aims to make three contributions to the literature. First, the use of newly linked administrative data permits a number of data-related advancements. By making use of eight different life-course datasets collected for the NZ population, we use information on the majority of the covariates suggested by the literature. For example, we have linked information on parents' qualifications, past academic performance at school, characteristics of the school attended, distance to the nearest tertiary education campus, and so on. The data also allows us to use a better measure of socioeconomic status that is based on the individual's address (when aged 15-16) rather than the decile of the school attended. Second, our analysis is undertaken at the population level. The vast majority of relevant studies have relied on survey data, or information from one or a few colleges/universities, or limited the sample to only those eligible to apply for admission. Instead, we take a broader approach, looking at life-course information to assess the relevant determinants at a population level. Finally, we employ multiple methods to investigate the factors, or set of factors, most relevant in explaining participation outcomes. In particular, in addition to the standard probabilistic frameworks that provide marginal effects of different covariates, we also use Fairlie decompositions to quantify the separate contributions of differences in observable characteristics. To date, this decomposition method has been under-utilised in understanding the relative importance of different factors in explaining ethnic gaps in higher education[5].

[4] A meshblock is the smallest geographic unit used by Statistics NZ. See Section 4 for more details.

[5] An exception to this is recent work by Ciao and Maloney (2017), which utilises Fairlie decompositions to examine ethnic differences in university grade outcomes and course completion rates.

3 Overview of the NZ qualifications system

The NZ Qualifications Authority (NZQA) is a government agency that manages the NZ Qualifications system. This includes administering the National Certificate of Educational Achievement (NCEA) – the most prevalent form of high school qualification in NZ.[6] NCEA is designed to measure a student's performance against standards of achievement or competency.

Each NCEA subject (such as mathematics or English) is made up of a number of standards (which can be thought of as sub-sections of that subject). Each standard has a defined credit value (usually between 3 to 6 credits) and an associated level (1, 2, or 3). Typically, a high school student attempts Level 1 standards when aged 15 or 16 in Year 11; Level 2 standards in Year 12, and Level 3 standards in Year 13 (the last year of high school).

To attain an NCEA Level 1 qualification a student must accumulate 80 credits at any level, of which 10 credits must be achieved for literacy and 10 credits for numeracy. For a NCEA Level 2 qualification, a student must attain 60 credits at Level 2 or above and 20 credits at any other level; and for a NCEA Level 3 qualification, a student must gain 60 credits at Level 3 and 20 credits at Level 2 or above.

Regardless of NCEA level, students are graded on a four-point scale for each achievement standard undertaken: 'not achieved', 'achieved', 'merit', and 'excellence'. For an NCEA qualification to be endorsed with merit (excellence) a student must gain 50 credits at merit (excellence) for that level of qualification or above.

In our analysis, we use information on whether an individual attained achieved; merit or excellence in their NCEA Level 1 qualification as an indicator for academic performance at school. This is the best source of such information that is available across the majority of our population. This is because school is compulsory in NZ from 6 to 16 years of age, and therefore almost all individuals in our population cohort have attempted NCEA Level 1.

Post-school education in NZ includes both higher and vocational education. The tertiary education system comprises 10 levels that range from a Level 1 certificate to a Level 10 doctoral degree. Levels 1-3 are broadly comparable to NCEA (ie, senior high school) qualifications, Levels 46 generally cover trades, technical and business qualifications, and Levels 7-10 generally covers degrees, graduate- and postgraduate qualifications.

Bachelor's degrees are studied predominantly at universities, but can also be undertaken at polytechnics and wānanga.[7] A typical bachelor's degree takes three years of full-time study to complete.

4 Data

This paper uses individual-level data from Statistics NZ's Integrated Data Infrastructure (IDI), which is a large database linking microdata about individuals and households in NZ. The IDI includes administrative data from a long list of government agencies, Statistics NZ surveys, and information from non-government organisations. This study relies mostly on the administrative data sources in order to focus on a population cohort. All data within the IDI are in confidentialised form, but can be

[6] For more information on the NCEA qualification see http://www.nzqa.govt.nz/qualifications-standards/qualifications/ncea/

[7] NZ has eight, predominantly government-funded, universities, 18 polytechnics and three wānanga. A wānanga is a tertiary education institution that provides programmes in a Māori cultural context.

linked via an individual's unique identifier (snz_uid), which permits a longitudinal perspective in our quest to better understand determinants of participation in bachelor's qualifications in NZ.[8]

Based on information from the IDI, we define the participation outcome variable as enrolling in a bachelor's qualification before the age of 20.[9] This age restriction is primarily driven by the fact that those entering tertiary education at or after the age of 20 are generally not required to meet the same entry criteria as those enrolling before the age of 20.[10]

We focus on the following population of interest:

- Individuals born between 1 July 1990 and 30 June 1994, divided into four July-to-June year cohorts.[11] Throughout this paper, we refer to cohorts by their end year.
- Individuals enrolled at a NZ high school for at least one day, who lived in NZ for at least 300 days in each of the calendar years in which they turned 15 and 16 and achieved at least one standard in their 15th or 16th year. These restrictions ensure we are focusing on individuals who undertook their final year of compulsory schooling in NZ.[12] These restriction mean that we are largely studying the output of the NZ education system, and excluding most foreign students, for whom NCEA qualifications are a noisy measure of school attainment;
- Individuals who lived in NZ for at least 200 days in at least one of the two years from their 18th to 20th birthdays. This restriction means our target sample does not include individuals who were primarily based overseas when making decisions about participating in a bachelor's qualification in NZ before the age of 20.

A key feature of the data is the number of life-course factors that we can control for in our analysis. These factors include: measures of socioeconomic status (when the individual was aged 15-16), characteristics of the high school the individual attended, their academic performance, distance to the nearest tertiary-provider campus, parents' educational achievement, and a host of other covariates detailed in Table 1. This table provides definitions of all variables (both dependent and explanatory) and cites the IDI source of each variable.

Table 2 subsequently provides descriptives of all variables from Table 1, and does so for the aggregate sample, as well as the ethnic sub-samples. There are five broad ethnic categories – Māori (accounting for 21% of the population), Pasifika (8%), Asian (7%), Other (2%) and European (62%).[13] Many administrative data sources in the IDI contain information on ethnicity. However, an individual's ethnic identity is not static and can change over time and by data source. For consistency, we therefore use ethnicity as denoted in the school enrolment data, and to overcome the issue of multiple ethnic responses we assign individuals to one ethnicity using a prioritised classification (noting that less than 7% of our target population identified with multiple ethnicities).[14] The order of prioritisation is: Māori, Pasifika, Asian, Other, European. That is, if an individual ever records an association with Māori (in the school

[8] Throughout this paper, observation counts are randomly rounded to base 3 in accordance with Statistics NZ confidentiality requirements.

[9] Participation equates to being enrolled in a programme greater than 0.03 EFTS, which is a week of study.

[10] Students entering a bachelor's qualification at a NZ university before the age of 20 are generally required to have University Entrance or an equivalent qualification. However, universities offer special admission to those aged 20 or over who do not meet this criterion. See http://www.nzqa.govt.nz/qualifications-standards/awards/university-entrance/ and http://www.universitiesnz.ac.nzstudying-in-nz/domestic#Discretionary%20Entrance.

[11] The July-June period roughly matches a school-intake cohort.

[12] School is compulsory in NZ until the age of 16. It is also important to note that some NZ students study towards international qualifications (such as International Baccalaureate or Cambridge exams) instead of undertaking NCEA qualifications. These individuals will appear in the data on school enrolments, but the data do not allow us to identify students who undertake international qualifications. We therefore exclude students who attended schools that primarily offer international qualifications.

[13] Due to the small nature of the 'Other' ethnic sub-group, we do not include descriptive statistics and results for this category.

[14] The implication of using prioritised ethnicity is that multiple ethnic responses are not accounted for, thereby concealing diversity within, and overlap between, ethnic groups.

enrolment data) then, by prioritisation, they are recorded as Māori regardless of the other identified ethnicities. If an individual records an association with Pasifika (but not Māori) then they are recorded as Pasifika; and so on.

As shown in Table 2, participation in bachelor's study is lowest among Māori. Just over 15% of Māori in our population cohort participate; while the comparable figure is close to 40% for Europeans; and just over 70% for Asians. Pasifika participation is a little higher (relative to Māori) at about 22%. These descriptive statistics are consistent with the NZ literature (Earle, 2007; Strathdee & Engler, 2012) where it is highlighted that Māori and Pasifika are under-represented at degree-level study.

We now turn to factors that may explain bachelor's degree participation, and in particular the lower participation among Māori and Pasifika. First, we expect lower socioeconomic status individuals to have lower participation rates because, for example, they have greater family financial constraints. Much of the prior NZ literature has had to rely on school decile as a proxy for socioeconomic status. Based on address information available in the IDI, we are able to use the deprivation index instead, which is a more refined measure than school decile.[15] The ethnic patterns of deprivation are as expected with Māori and Pasifika experiencing the highest levels of deprivation (average index scores of 7.18 and 8.12 respectively) and Europeans having the lowest levels (average index score of 4.62).

We expect school performance to be another useful explanatory variable. As explained in Section 3, to measure school performance we use information on the individual's NCEA Level 1 results, which are the earliest school results that are available in our data.[16] This is measured on a four-point scale: 'not achieved', 'achieved', 'merit', and 'excellence'. This is the best source of information on academic performance at a population level. Using the earliest school results that are available is likely to involve less endogeneity bias than using higher levels of NCEA results qualification (Levels 2 and 3), as students who are already planning to continue on to tertiary education are much more likely to attempt higher levels than those who are not.

There are particularly wide ethnic gaps in school achievement at NCEA Level 1 (Table 2). Between 18% and 19% of Māori and Pasifika attain NCEA Level 1 with merit or excellence, whereas the comparable figure for Europeans is more than double (44%), and for Asians is more than triple (62%). Over 30% of Māori and 21% of Pasifika in our population cohort do not attain NCEA Level 1, compared with about 12% for Europeans and just 5% for Asians. While reasons for these substantial differences at age 15-16 are not the focus of this paper, it is certainly a worrying signal of the potential source of ethnic variations in participation in bachelor's study in NZ.

As a measure of student disengagement, we also include information on the number of school notifications (that is, suspensions, stand-downs and serious truancies) an individual received during the years of their 15th and 16th birthdays. As indicated in Table 2, Europeans received an average of 0.14 notifications; while Māori received more than twice as many, at 0.40.

In terms of individual characteristics, the ethnic sub-groups are fairly evenly divided by gender. However, it is worth noting that females are more likely to enrol in bachelor's study - around two-thirds of students participating for each ethnicity are female (not shown in Table 2).

---

[15] The deprivation index is constructed for each meshblock in NZ. A meshblock is the smallest geographic unit used by Statistics NZ. The median size of a meshblock in 2006 was approximately 87 people and around 35 households. In 2014 the median size of a high-school catchment was 402 households.

[16] Studies often use childhood cognitive test results as these are considered to be less influenced by environmental factors and therefore a better reflection of innate ability. For example, Maani (2006) uses IQ scores at age 8 and Chowdry et al. (2013) uses national assessed standardised achievement test scores at age 11. We do not have measures of childhood ability in our data, but use the earliest available measure (NCEA Level 1).

To measure migrant status, we use information on whether the individual has a NZ birth record. Only 3% of Māori and 11% of Europeans are migrants, compared with 23% for Pasifika and 65% for Asians.[17]

Table 2 also shows the number of times the average individual changed schools during their 15th and 16th years. Frequent school switches are usually associated with low levels of socioeconomic status, in conjunction with a lower likelihood of stable housing arrangements. As the number of school switches rise, there is a greater likelihood that these disruptions lead to disengagement from the education system. We find, on average, that European students switch schools only 0.09 times during their 15th and 16th years, while Māori switch on average 0.23 times, and Pasifika 0.13 times.

Existing literature, such as Black et al. (2015), suggests that the characteristics of the school attended may also be important, even after accounting for an individual's academic performance at school. Therefore, we also include a set of explanatory variables on school characteristics. These include whether the high school the individual attended was single sex or co-educational; and whether the school authority was state, state-integrated or private. State schools are funded by the central government. State-integrated schools are former private schools that have been integrated into the state education system. They are also funded by the government, but retain their special character or religious affiliation. Private schools are mostly privately funded, although they receive a little government funding.

Around 29% of the population cohort attended single-sex schools, and it is noticeable that this figure was lower for Māori, at just over 22%. In terms of school authority, about 85% of our cohort went to state schools, a figure that rises to close to 98% if including state-integrated schools. This high proportion of students attending state schools is expected for two reasons. First, only a small proportion of NZ students attend private schools. Second, our population excludes students attending schools that predominantly offer non-NCEA qualifications such as International Baccalaureate (IB) and Cambridge Examinations, and private schools are more likely to offer these qualifications.[18]

School characteristics also vary by ethnicity. Māori students are more likely to attend state schools, and less likely to attend state-integrated or private schools than Europeans. Pasifika are less likely to attend a state school and more likely to attend a state-integrated school than Europeans.

Another factor potentially relevant to investigating ethnic disparities in bachelor's participation, is distance to the nearest provider. This is based on the Euclidean distance between the individual's home address at age 15-16 and the nearest tertiary-provider campus that offers bachelor's qualifications. Existing literature suggests that long distances may pose a barrier to tertiary education, particularly for lower-income households or certain ethnic groups, for financial or cultural reasons (eg, Gibbons & Vignoles, 2012). Reflecting the higher concentration of Pasifika and Asians living in urban areas, Pasifika and Asians tend to live the closest to educational providers, with an average distance of just 5km; while Europeans and Māori live further away with distances of 17.6km and 15km respectively (Table 2).

The final set of variables detailed in Table 2 relate to parent's educational attainment. Statistics NZ has recently added information to the IDI that allows links to be made between individuals and their parents.[19] We find that approximately 91% of our population of interest have identified parents via this added information. Using the parent's unique identifier allows us to source information on the parents'

[17] As the indigenous population, the percentage of Māori migrants may seem high. However, this may be because the NZ population is quite internationally mobile, with one of the largest diaspora in the world (in per capita terms) (Conway, 2016). In particular, NZ and Australia have an open labour market and a substantial number of Māori live in Australia.

[18] Indeed, state and state-integrated schools must offer NCEA, but can choose to offer other qualifications in addition to NCEA. Private schools are not obliged to offer NCEA.

[19] Statistics NZ have collated information from several sources to identify parent-child links in the IDI, including birth records, Ministry of Social Development data relating to social welfare payments and family tax credits, Census 2013, and several household surveys.

highest qualification from the 2013 Census.[20] Few studies in NZ have looked at the impact of parental education on students' higher education outcomes, despite Māori and Pasifika being more likely to have parents without formal qualifications (Ministry of Social Development, 2008). This family background information is vital for better understanding participation at university. Numerous past studies (Kane, 1994; Li, 2007; and Finnie, Wismer & Mueller, 2015) have found a strong positive correlation between parents' educational attainment and that of their children. This correlation may be due to family expectations or a demonstration effect. Close to 30% of the Pasifika and Māori sub-groups in our cohort have parents with no school qualifications; while the comparable numbers for Europeans and Asians are about 13% and 16%. Likewise, a smaller percentage of Māori and Pasifika individuals have parents with bachelor's or postgraduate degrees relative to Europeans and Asians.

[20] We can obtain information on parental qualifications from the Census 2013 for 72% of our population cohort. While this sub-population is not random, it is broadly similar to our total population. The main difference is that this sub-population is less likely to include migrants.

Table 1 Variable definitions and sources

Ethnic sub-groups: Māori	Dummy variable: 1 if Māori; 0 otherwise	MoE School
Pasifika	Dummy variable: 1 if Pasifika; 0 otherwise	
Asian	Dummy variable: 1 if Asian; 0 otherwise	
European	Dummy variable: 1 if European; 0 otherwise	
Other	Dummy variable: 1 if Other ethnicities; 0 otherwise	
Dependent variable		
Participation	Dummy variable: 1 if enrolled in a bachelor's qualification prior to the age of 20; 0 otherwise	MoE Tertiary
Explanatory variables		
Cohort year: 1991	Dummy variable: 1 if born between 1st July 1990 and 30th June 1991; 0 otherwise	
1992	Dummy variable: 1 if born between 1st July 1991 and 30th June 1992; 0 otherwise	SNZ Spine
1993	Dummy variable: 1 if born between 1st July 1992 and 30th June 1993; 0 otherwise	
1994	Dummy variable: 1 if born between 1st July 1993 and 30th June 1994; 0 otherwise	
Individual characteristics		
Male	Dummy variable: 1 if male; 0 otherwise	SNZ Spine
Switching school	Number of times switched schools when aged 15-16	MoE School
Migrant	Dummy variable: 1 if not born in NZ; 0 otherwise	DIA Births
Socioeconomic status		
NZ Deprivation index	Categorical variable: Ordinal scale ranging from 1 to 10. Deprivation index for census meshblock individual lived in for the most number of days when aged 15-16, with a higher number reflecting a more deprived meshblock.	SNZ Address, NZDep2006
School characteristics		
Single-sex high school	Dummy variable: 1 if single-sex school; 0 otherwise	MoE School
School authority (3)	Dummy variable: 1 if state school; 0 otherwise	
	Dummy variable: 1 if state integrated school; 0 otherwise	
	Dummy variable: 1 if private school or other type of school; 0 otherwise	
School performance and engagement		
Academic performance (4)	Dummy variable: 1 if not attained NCEA Level 1; 0 otherwise	MoE School
	Dummy variable: 1 if attained NCEA Level 1 with 'Achieved'; 0 otherwise	
	Dummy variable: 1 if attained NCEA Level 1 with 'Merit'; 0 otherwise	
	Dummy variable: 1 if attained NCEA Level 1 with 'Excellence'; 0 otherwise	
School notifications	Number of school notifications (suspensions, stand-downs and serious truancies)	MoE School
Other variables		
Distance	Euclidean distance between home when aged 15-16 to nearest tertiary delivery site that offers a bachelor's degree (Based on the address where the individual spent the most days during the calendar years of their 15th and 16th birthdays).	SNZ Address, MoE Tertiary and TEC
Parents' education (5)		
No qualification	Dummy variable: 1 if parents have no qualifications	
School	Dummy variable: 1 if highest educational attainment of parents is school qualification; 0 otherwise	Census 2013
Post-school	Dummy variable: 1 if highest educational attainment of parents is a post-school qualification; 0 otherwise	
Bachelor's	Dummy variable: 1 if highest educational attainment of parents is a bachelor's qualification; 0 otherwise	
Postgraduate	Dummy variable: 1 if highest educational attainment of parents is a postgraduate qualification; 0 otherwise	

Notes: MOE = Ministry of Education; SNZ = Statistics NZ; DIA = Department of Internal Affairs; TEC = Tertiary Education Commission; NZDep2006 = The deprivation index for each meshblock based on Census 2006 data (see Salmond, Crampton and Atkinson 2007). Information on the location of delivery sites was provided to the authors by the Tertiary Education Commission.

Table 2 Variable definitions and sources

	Total	European	Māori	Pasifika	Asian
Ethnic breakdown for full sample (%)	100.00	62.17	20.77	8.39	6.97
Dependent variable: Participation (%)	35.19	39.19	15.55	21.57	70.67
Explanatory variables					
Cohort year: Cohort 1991 (%)	23.99	24.72	22.33	23.09	23.94
Cohort 1992 (%)	24.98	25.06	25.40	24.97	23.39
Cohort 1993 (%)	25.47	25.18	26.04	26.13	25.78
Cohort 1994 (%)	25.57	25.04	26.23	25.81	26.89
Individual characteristics					
Male (%)	49.76	50.18	49.54	49.13	47.45
Switching school (number of times aged 15-16)	0.12	0.09	0.23	0.13	0.07
Migrant (%)	15.15	10.59	2.95	23.21	65.00
Socioeconomic information: NZ Deprivation index	5.52	4.62	7.18	8.12	5.50
School characteristics					
Single sex high school (%)	28.78	30.96	22.55	27.47	29.42
School authority: State (%)	85.18	83.45	92.04	80.98	85.95
State Integrated (%)	12.60	13.47	7.41	18.61	11.95
Private or Other (%)	2.22	3.08	0.55	0.42	2.10
School performance and engagement					
NCEA Level 1 attained with: Not achieved (%)	15.87	11.65	30.31	21.28	5.39
Achieved	46.20	43.98	52.92	60.21	32.32
Merit	30.91	36.25	15.84	17.06	43.60
Excellence	7.02	8.13	2.95	1.44	18.69
Number of school notifications	0.20	0.14	0.40	0.26	0.06
Other variable: Distance to nearest delivery site (km)	15.08	17.69	15.07	4.97	5.74
Sample size	189,364	117,730	39,330	15,895	13,203
Parent's education					
No qualification	17.90	12.57	28.92	29.80	15.70
School	32.50	32.84	29.03	39.29	32.79
Post-school	29.78	33.54	27.82	19.85	16.53
Bachelor's	13.41	14.06	10.17	8.18	23.37
Postgraduate	6.41	7.00	4.06	2.88	11.62
Sample size	137,145	83,787	30,432	11,496	9,390

5 Methodology

We use maximum-likelihood probit analysis to examine the effects of the various covariates (outlined in Section 4) on our participation outcome variable. The general probit model is described as follows:

$$Y_i^* = X_i'\beta + u_i \quad (1)$$

Y_i^* is the latent variable related to participation. We observe Y_i equal to 1 if $Y_i^* > 0$ and equal to 0 if $Y_i^* \leq 0$. X_i is a vector of individual, family and school characteristics. Assuming that the error term, u_i, is normally distributed, the probit model can be described as:

$$Prob(P_i = 1) = F(X_i'\beta) \quad (2)$$

where $F(\cdot)$ is the cumulative normal distribution function.

To allow for parameter heterogeneity, we run separate probit regressions for each ethnic sub-group. Additionally, we also perform two iterations within each ethnic sub-group; one based on the full population sample; and one based on the sub-sample for which there is information on parents' educational attainment. The linkage rate for including this information reduces the sample to 72% of the size of the population cohort.

We also extend our empirical analysis with the use of Fairlie (1999, 2005) decompositions. This method assesses the extent to which our observable information explains the ethnic gap in the outcome variable. This method extends the Blinder-Oaxaca technique to non-linear models (Blinder, 1973; Oaxaca, 1973). Several previous studies (such as Chowdry et al., 2013 and Vignoles & Powdthavee, 2009) assess the relative importance of observable factors by sequentially adding variables. However, we prefer the Fairlie method as the sequential method may be sensitive to the order in which variables are added.

Using the results from the probit model, the mean difference in participation probabilities for Europeans and Māori (for example) is decomposed as follows:

$$\bar{P}_i^E - \bar{P}_i^M = \left[\sum_{i=1}^{N^E} \frac{F\left(X_i^E \widehat{\beta^E}\right)}{N^E} - \sum_{i=1}^{N^M} \frac{F\left(X_i^M \widehat{\beta^E}\right)}{N^M}\right] + \left[\sum_{i=1}^{N^M} \frac{F(X_i^M \widehat{\beta^E})}{N^M} - \sum_{i=1}^{N^M} \frac{F(X_i^M \widehat{\beta^M})}{N^M}\right] \quad (3)$$

where superscripts *E* and *M* identify coefficients and values from the European and Māori sub-populations respectively. $\bar{P}_i$ is the average probability of participation, *N* is the size of the sub-population and $\hat{\beta}$ is the coefficient from the probit regressions in (2).

The first bracket on the right-hand side represents the 'explained' portion of the ethnic gap in the outcome variable, based on the differences in the distribution of measured variables (*X*) for European and Māori individuals. The second bracket is more difficult to interpret and equates to the 'unexplained' portion of the gap.

A well-known issue in implementing this decomposition method is whether the estimated coefficients used to weight the explained component of the decomposition should be those relating to Europeans or Māori, or estimated from a pooled regression. If there are large differences in the estimated coefficients between Europeans and Māori then these different approaches can yield quite different results. We use the estimated coefficients from the regression for Europeans since a narrowing of the ethnicity gap is likely to be due to Māori converging towards Europeans.[21] As the Fairlie method requires equally-sized sub-populations of European and Māori individuals, we match a random sub-sample of European individuals to the smaller sub-population of Māori individuals based on their predicted probabilities of participation. Results may vary based on the characteristics of the random sub-sample of European individuals used in the matching, so we report average estimates of repeated random sub-samples. It is also worth noting that the results could be sensitive to the ordering of variables, and we therefore apply random ordering, as suggested by Fairlie (2016).

[21] We also test our results against those using estimated coefficients from a pooled regression that includes a dummy variable for the ethnic minority group (to account for the issue of the inappropriate transfer of some of the unexplained part of the gap into the explained component that arises with the use of the pooled method - see Jann, 2008; Fortin, 2006; and Fairlie, 2005).

Finally, while the example above relates to decomposing the participation gap between European and Māori; we also repeat this analysis for European and Pasifika, and for European and Asian.

6 Results

6.1 Marginal effects

Table 3 provides the marginal effects from separate probit estimations of the participation outcome variable for each of the four ethnic sub-groups: European, Māori, Pasifika and Asian. Two specifications are reported - column (A) is based on the full population cohort while column (B) is based on the sample that has information on parents' education level. We initially focus our interpretation efforts on column (A) for each of the ethnic groups.

The two main covariates of interest in column (A) are socioeconomic status and school performance. Holding all other covariates at their means, as expected, growing up in a higher socioeconomic environment is associated with a higher probability of participating in a bachelor's qualification for all ethnicities. For example, Europeans who lived in a meshblock with the highest socioeconomic status (ie, deprivation index of 1) when they were aged 15-16 were 13.5 percentage points more likely to participate in bachelor's study compared with those that lived in a meshblock with the lowest socioeconomic status (ie, deprivation index of 10). The estimated marginal effect decreases as socioeconomic status decreases, relative to the lowest socioeconomic group (ie, deprivation index of 10). The trend is similar for other ethnicities, but the marginal effects are of a smaller magnitude.

The other main covariate of interest for the column (A) model is prior academic achievement. There is clearly a rising likelihood of participating in a bachelor's qualification as we compare achieving NCEA Level 1, with attaining it with merit, or excellence, relative to those who did not achieve NCEA Level 1. However, the magnitude of these differences varies across ethnic groups. For instance, for Asians, the probability of participating in a bachelor's qualification increases just over two-fold when comparing achieving NCEA Level 1 to attaining excellence in that qualification. This compares to nearly a seven-fold increase for European and Pasifika, and more than a ten-fold increase for Māori. These results signal that it is not enough for Māori to just attain an NCEA Level 1 qualification, it appears imperative to have that qualification endorsed with either merit or excellence if we wish to improve their propensity for participating in bachelor's qualifications.

Looking to the other covariates, we find that the probability of participating in a bachelor's qualification appears to decrease for the 1994 cohort (where the 1991 cohort is the excluded category). This is the case for all ethnic groups except for Pasifika. This may be because the effect of the global financial crisis was more acute for the earlier cohorts, resulting in a greater propensity to participate in further education.

As Table 3 illustrates, males are less likely to participate in a bachelor's qualification relative to females. Migrants have a higher probability of participating for the European, Māori and Pasifika sub-groups, but the effect is statistically insignificant for the Asian sub-group.

In terms of changing schools - for Europeans and Pasifika, switching schools more often decreases the probability of participating, but interestingly, there is no effect for Māori and Asians. Each additional school change when aged 15-16 reduces the likelihood of participating in bachelor's study by between 2 to 3 percentage points for Europeans and Pasifika.

Attending a single-sex high school appears to increase the probability of participating for all ethnic groups. It is unclear why this might be. While it is well documented that girls do better in single-sex schools, this effect remains even after controlling for an individual's school achievement. It may be that single-sex schools have a stronger tendency to direct students towards higher education resulting in a positive peer effect. Or if single-

sex schools are generally considered more academic, there may be a selection effect whereby those who expect, or whose family expect them, to pursue higher education are more likely to attend single-sex schools.[22]

Other school characteristics appear to make a statistically significant difference to the probability to participate – albeit, the level of significance varies for the different ethnicities. There is a positive marginal effect for Māori and Pasifika in attending a private school over a state school (5.9 and 10 percentage points respectively), while private school attendance is associated with a reduced probability of participation for Europeans of 4.4 percentage points. Interpreting the value of private schools per se from these results should be done so cautiously, as schools that provide IB and/or Cambridge exams were removed from our target population, and private schools are more likely to offer these qualifications. Hence, these results are based on a sub-set of the total private schools in NZ.

Table 3 also shows a strong negative association between school notifications – a proxy for disengagement at school – and the likelihood of bachelor's study, with all ethnicities appearing to be negatively affected. One additional notification reduces the probability of participation for Māori by 3.7 percentage points and 6.4 percentage points for Pasifika. The marginal effects for Europeans and Asians are larger at 9.4 and 7.7 percentage points.

The final variable included in the column (A) specifications is distance to the nearest campus that offers bachelor's qualifications. Consistent with evidence for England showing that distance has little or no impact on the decision to participate (Gibbons & Vignoles, 2012), we find that the magnitude of the marginal effects for distance are very small in size (while still being statistically significant for most ethnicities). Table 3 shows that for each additional 10km an individual lives further away from a delivery site, the marginal effect on the propensity to participate in bachelor's study decreases by at most 1.1 percentage points (Model B for Asians). Given that the descriptives in Table 2 shows that individuals live on average, between 5km and 18km away from a campus, it appears that distance plays an extremely minor role in determining the likelihood of undertaking a bachelor's qualification.

Parents' educational attainment is included in the specifications in column (B). These results indicate that the likelihood of an individual undertaking a bachelor's qualification increases with the level of their parents' education. For example, for Pasifika, having parents with at least school qualifications increases the likelihood of participating in a bachelor's degree by 1.9 percentage points, compared with the excluded group of parents with no qualifications. This impact rises to 15.1 percentage points if the parents' highest qualification is a postgraduate one. It is also worth noting that the addition of parent's educational attainment in the model does mildly dampen the role socioeconomic status and school performance.

These results are relatively easy to summarise – consistent with the extant literature, the three factors of importance are socioeconomic status, prior performance in school, and parents' educational attainment. What is most interesting from these findings is the relative contributions of each of these factors – which indicate that prior performance in school plays the largest role, by far. Section 6.2 will explore the extent to which ethnic differences in the observed factors contribute to the ethnic gap in participation, and which factors make the largest contribution to the explained gap.

[22] Usually, students attending state schools in NZ attend their local school. However, there is overlap in some school catchments, so students may live within more than one school catchment. In addition, if a school has extra places after all local students have enrolled, it can offer these places to students outside the catchment, with placements decided via a ballot.

Table 3 Marginal effects from maximum likelihood probit analysis of participation in a bachelor's qualification

	European		Māori		Pasifika		Asian	
	(A)	(B)	(A)	(B)	(A)	(B)	(A)	(B)
Cohort year: - Cohort 1992	-0.004	-0.009*	-0.003	-0.001	0.007	0.009	-0.004	-0.016
- Cohort 1993	-0.011**	-0.023***	0.003	0.002	0.006	0.008	0.004	-0.015
- Cohort 1994	-0.035***	-0.044***	-0.008**	-0.007	0.002	-0.007	-0.023*	-0.042***
Individual characteristics								
Male	-0.060***	-0.070***	-0.027***	-0.029***	-0.067***	-0.071***	-0.005	-0.022**
Migrant	0.050***	0.048***	0.019**	0.018*	0.026***	0.041***	0.011	0.013
Switching school	-0.031***	-0.027***	0.001	-0.003	-0.025**	-0.031**	-0.009	-0.021
Socioeconomic status: Deprivation index								
= 1	0.135***	0.103***	0.078***	0.051***	0.058**	0.050*	0.090***	0.057**
= 2	0.109***	0.071***	0.051***	0.032***	0.064***	0.050*	0.096***	0.066***
= 3	0.096***	0.066***	0.046***	0.027***	0.061***	0.031	0.114***	0.097***
= 4	0.074***	0.044***	0.039***	0.034***	0.022	0.014	0.113***	0.088*
= 5	0.076***	0.047***	0.025***	0.013**	0.029*	0.022	0.093***	0.073***
= 6	0.052***	0.024**	0.026***	0.018***	0.035***	0.020	0.073***	0.042*
= 7	0.044***	0.028**	0.027***	0.019***	0.008	0.001	0.086***	0.059**
= 8	0.031***	0.019*	0.013***	0.011**	0.019**	0.020*	0.081***	0.068***
= 9	0.029***	0.015	0.014***	0.010**	0.004	-0.003	0.043***	0.024
School characteristics								
Single-sex high school	0.057***	0.045***	0.016***	0.013***	0.029***	0.030***	0.038***	0.024**
School authority: State Integrated	0.065***	0.050***	0.026***	0.018***	0.028***	0.034***	0.019	0.028*
Private or Other	-0.044***	-0.088***	0.059**	0.047*	0.10*	0.069	-0.057	0.003
School performance and engagement								
NCEA Level 1 : Achieved	0.129***	0.122***	0.068***	0.064***	0.125***	0.135***	0.367***	0.381***
Merit	0.657***	0.614***	0.496***	0.462***	0.590***	0.596***	0.821***	0.805***
Excellence	0.875***	0.838***	0.729***	0.667***	0.830***	0.828***	0.883***	0.874***
Number of school notifications	-0.094***	-0.081***	-0.037***	-0.031***	-0.064***	-0.076***	-0.077***	-0.077***
Distance to nearest provider (10 kms)	-0.008***	-0.009***	-0.002***	-0.002***	-0.004	-0.007*	-0.011***	-0.014***
Parents' highest qualification:								
- School	-	0.049***	-	0.023***		0.019**	-	0.048***
- Post-school	-	0.064***	-	0.029***		0.050***	-	0.012
- Bachelor's	-	0.174***	-	0.068***		0.137***	-	0.113***
- Postgraduate	-	0.220***	-	0.117***		0.151***	-	0.152***
Pseudo R^2	0.368	0.367	0.355	0.346	0.298	0.311	0.320	0.342
N	117,732	83,787	39,330	30,432	15,897	11,496	13,203	9,390

Note: Reference groups = Cohort 1991, male, NZ born, deprivation index=10, co-ed school, state school authority, not attained NCEA, and parents with no qualifications. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

6.2 Decomposing ethnic differences

We now take a more detailed look at how much of the ethnic gap in can be explained by differences in observed characteristics across ethnicities, and how much of a contribution each characteristic makes to the overall explained gap. Table 4 provides results of Fairlie decompositions of the participation gap. We compare Māori, Pasifika and Asian ethnic groups with Europeans using the relevant coefficients from the regressions relating to Europeans in specification (B) (ie, models including parental education).[23] We group the contributions of related factors, such as the cohort year, school characteristics, NCEA level 1 outcomes and parents' educational attainment.

Taking the gap in participation rates between Māori and Europeans as an example, these decompositions allow us to explore two questions. First, if we gave Māori the same observed characteristics as Europeans, to what extent would the ethnic gap in participation rates close? That is, how much of the gap is explained by the factors included in our model, and how is due to unobserved factors, which could be culturally specific? Second, which of the factors in our model contribute the most to this explained gap? For instance, how much of a contribution do differences in school performance between Māori and Europeans make relative to the contribution of differences in socioeconomic status?

While the actual participation rate for Europeans is about 34.75%, it is only 14.91% for Māori - a 19.84 percentage point difference (see second row of Table 4). The actual participation rate for Pasifika is 23.31% - a 11.45 percentage point difference compared with Europeans. Asians have a higher participation rate than Europeans of 71.73%, a -36.97 percentage point difference. [24]

Differences in the distribution of characteristics across our ethnic sub-groups reflect the explained proportion of the ethnic gaps. We find that all individual, household, school and other characteristics collectively explain 86.69% of the Māori-European gap. The observed characteristics over-explains the Pasifika-European gap, and less than half of the higher participation rate of Asians relative to Europeans is explained by differences in observed characteristics.

This tells us that Asians, who have a participation rate that is 36.97 percentage points higher than Europeans, would experience a decline of 17.32 percentage points (less than half the raw gap) if given the same characteristics of the European sub-group. If Pasifika had the same characteristics of the European population, this would raise their participation rate by 14.42 percentage points, to a level that would be above their European counterparts. For Māori, if given the same observed characteristics as Europeans, this would raise their participation rate in bachelor's qualifications by 17.20 percentage points. That is, it would greatly reduce, but not entirely eliminate the Māori-European participation gap.

Regardless of ethnicity, the largest contributor of the participation gaps is school performance. This factor contributes to more than three-quarters of the total 'explained'. More specifically, if we gave the same NCEA level 1 results of Europeans to Māori and Pasifika, this would raise their participation rates by 13.39 and 11.47 percentage points respectively; and drop the Asian participation rate by 13.58 percentage points.

The second and third most important factors depend on the ethnic comparison. For Māori and Pasifika, the second most important factor is socioeconomic status, followed closely by parent's highest qualification. That is, the lower average socioeconomic status and lower level of parental education

[23] For the participation decompositions, as a sensitivity test, we also used estimated coefficients from a pooled regression that included a dummy variable for the minority ethnic group. This made little difference to the results.

[24] Note that the raw participation, retention and completion rates reported in Table 4 are slightly different to the descriptive statistics given in Table 2. This is because Table 2 descriptive statistics are based on the whole population, while Table 4 uses the sub-sample for which parents' education was available (see Section 4).

among Māori and Pasifika go some way to explaining the lower participation rates for these ethnic groups. For Asians, socioeconomic status is not as important in explaining the gap, but being a migrant is a significant contributor. That is, migrants have a higher likelihood of participation, so the greater share of migrants in the Asian population accounts for some of the higher participation for this ethnic group.

Table 4 Decomposition of ethnic differences in bachelor's degree participation

	Māori vs European	Pasifika vs European	Asian vs European
Ethnic difference[25] (% point: European – Ethnic minority)	34.75 – 14.91	34.75 - 23.31	34.75 - 71.73
Total difference (% point)	**19.84**	**11.45**	**-36.97**
Cohort year	0.03***	0.03***	0.05***
Individual characteristics			
Male	0.03***	-0.10***	-0.19***
Switching school	0.13***	0.03***	-0.04***
Migrant	0.18***	-0.36***	-1.82***
Socioeconomic information			
Deprivation index	1.45***	2.11***	0.61***
School characteristics	0.23***	-0.35***	-0.04***
School performance and engagement			
NCEA Level 1	13.39***	11.47***	-13.58***
Course pass rate	-	-	
Number of school notifications	0.53***	0.30***	-0.24***
Other variable			
Distance to nearest delivery site	-0.02**	-0.53***	-0.70***
Parents' highest qualification	1.28***	1.83***	-1.28***
Total difference explained (% point)	17.20	14.42	-17.32
Proportion of gap explained	86.69%	125.94%	46.85%
Sample size	114,216	95,280	93,174

In summary, for Māori in particular, there are clearly drivers of the participation gap that are not observed in the data, and these may be culturally specific to this ethnic group. Similarly, there may be specific factors that are not observed and relate to why Asians are over-represented in the participation statistics, relative to Europeans.

The story for Pasifika is somewhat different. Our model over-explains the Pasifika-European gap in participation – that is, if Pasifika had the same characteristics as the European population, this would raise their participation rates to a level above their European counterparts. Of course raising participation rates is only one element of successful outcomes in higher education. Beyond participation, it is important to also understand the factors related to retention and completion of bachelor's degrees. For instance, Clotfelter et al. (2015), found that black students were more likely to

[25] Based on the sub-sample of the population cohort that has linked information regarding parents' educational attainment.

enrol in college relative to white students with similar backgrounds, but black students who attended college were less likely to complete a degree.

7 Sensitivity analysis

We now undertake a number of variants of our main models as checks on the robustness of our results. Focussing on the specifications that include parents' education (Model B of Table 3), we add school fixed effects, use alternative measures of school performance, alternative ethnicity classifications, and trial adding mother's education and father's education to our model as separate variables. Result tables for the marginal effects and Fairlie decompositions relating to these model variants are provided in Appendix A.

7.1 School fixed effects

The potential link between high school characteristics and higher educational outcomes is acknowledged in the literature. For example, Fletcher and Tienda (2010) finds that the racial gap in achievement across four Texas universities disappeared after controlling for high-school fixed effects. Using administrative data from a NZ university, Shulruf, Hattie and Tumen (2008) finds that the location of the school and characteristics of the student body have a significant effect on first-year GPA. While some school characteristics are included in our main regressions, there may still be remaining unobserved differences across schools. Therefore, as the first check for robustness, we follow recent work by Chowdry et al. (2013), who include school fixed effects in their analysis of participation rates in higher education in the UK. This involves replacing all school characteristic variables in Model B of Table 3 (ie, those related to gender of the high school and type of school authority) with a dummy variable for each school in the data set – approximately 500 schools in total.[26]

In general, across all ethnicities, the inclusion of school fixed effects makes little difference to the marginal effects for most of the variables. The main difference is that the marginal effects for deprivation decrease with the inclusion of school fixed effects. For example, for Europeans, in Model B of Table 3, an individual with a deprivation level of 1 was 10.3 percentage points more likely to participate than an individual with a deprivation level of 10. The inclusion of school fixed effects lowers this marginal affect to 7.9 percentage points (Table A.1). This most likely reflects that schools in NZ generally draw their students from the local area, resulting in an association between the school attended and deprivation status. However, the overall conclusions of those with lower deprivation levels being more likely to participate in bachelor's study still holds, and the marginal effects for Europeans are still larger than for other ethnicities.

Turning to the Fairlie decompositions, the inclusion of school fixed effects slightly increases the role played by schools, relative to the results shown in Table 4. For instance, school characteristics explained 0.23 percentage points of the Māori-European gap in Table 4, whereas school fixed effects explain 1.68 percentage points of the gap in Table A.3. At the same time, the inclusion of school fixed effects has lowered the proportion of the gaps explained by differences in NCEA Level 1. For example, in Table 4, NCEA Level 1 explained 13.39 percentage points of the Māori-European gap, whereas it contributes 10.14 percentage points in the model with school fixed effects. This suggests that at least a small part of the ethnic differences in participation that are attributed to NCEA results in Model B may reflect school-level effects, which could for instance include peer effects.

[26] A few schools with very small numbers of students with no variation in outcomes were excluded from this analysis.

7.2 Measuring school performance

As described in Section 3, NCEA results are graded on a four point scale for each achievement standard undertaken: 'not achieved', 'achieved', 'merit' and 'excellence'. For an NCEA level to be endorsed with merit or excellence, an individual requires a particular number of credits with merit or excellence. This can be attained for the level they are currently studying or can be back credited if attained at higher levels. It is therefore possible, for example, for someone to 'achieve' NCEA Level 1 in their first attempt, but in their next year, achieve sufficient NCEA Level 2 standards to have the NCEA Level 1 qualification upgraded to 'merit'.

While analysis in this study has utilised information that includes this back crediting, it is a useful sensitivity check to see if the results change if we focus on only first-attempt outcomes for NCEA level 1 attainment. As shown in Table A.1 removing the back crediting generally increases the marginal effects of school performance for all ethnicities. This is to be expected since the inclusion of back-credited results moderates the differences in student achievement. For example, for Māori students, relative to those who did not attain, the marginal effect of achieving NCEA Level 1 in Table 3 (Model B) was 6.4% for 'achieve', 46.2% for 'merit' and 66.7% for 'excellence', but the removal of back crediting increases these to 10.6% for 'achieve', 51.6% for 'merit' and 75.8% for 'excellence' (see Table A.1).

Turning to the Fairlie results, the removal of back crediting does not change the overall trends. In particular, differences in NCEA results still explain the largest proportion of the ethnic gaps in participation.

7.3 Ethnicity classification

This study has relied on prioritised ethnic identities, such that the order of priority is Māori, Pasifika, Asian, Other, and then European. We avoided allowing individuals to have multiple ethnic identities in order to not double count individuals across ethnic groups. An alternative method of assigning ethnicity is to drill down the comparisons to those that identify solely as Māori, Pasifika, Asian or European for that matter. This is another way of investigating how sensitive the results may be to different ways of assigning ethnic identity from the IDI.

The rates of participation in bachelor's degrees for those who identify only as Māori are lower than those who were classified as Māori using prioritised ethnicity (13.33 versus 14.91). A similar pattern emerges for Pasifika (21.76% for sole Pasifika versus 23.31% using the prioritised method) and Asian (70.67% versus 71.73%). This results in somewhat larger participation gaps for Māori, Pasifika and Asians relative to Europeans.

However, the contributions of each of the variables to the ethnic gaps in participation do not change much. The main difference is that NCEA Level 1 makes a larger contribution, reflecting a larger NCEA Level 1 performance gap between Europeans and sole Māori (and sole Pasifika) relative to Māori-European (and Pasifika-European) gap under the prioritised ethnicity classification. But once again, the differences are not particularly large (13.39 versus 14.37 percentage points for the Māori-European gap, and 11.47 versus 12.59 percentage points for the Pasifika-European gap) – see results provided in Table A.2.

7.4 Mother's versus father's education

A number of studies that investigate children's educational attainment make use of parents' education levels, such as the parents' years of schooling or qualification attained (see Datcher, 1982; Hill & Duncan, 1987). The majority of these studies have found that mother's education is more strongly associated with the child's educational outcome than the father's education level (Haveman & Wolfe, 1995). For instance, Marks (2008) used data from 30 countries to examine the influence of a number of family background characteristics on children's educational outcomes (with respect to literacy and numeracy performance) and found that the impact of mother's education was usually greater or similar to that of the father's. One notable exception is Gang and Zimmerman (2000), which finds, for Germany, that father's education is a more important influence than mother's education.

Based on this international evidence, our final sensitivity test involves including separate variables for mother's and father's education to explore the hypothesis that mothers' education is more strongly associated with the participation outcome than father's education. The results of this exploration are illustrated in Tables A.2 and A.4.

In terms of the marginal effects on an individual's propensity to participate in a bachelor's qualification, mother's and father's education levels seem to be of similar magnitude (see Table A.2). For the Fairlie decompositions (see Table A.4), mother's highest qualification makes a larger contribution to the Māori-European and Pasifika-European participation gaps than father's highest qualification, although the difference is not great (0.77 percentage points for mother's qualification and 0.66 percentage points for father's qualifications in both cases). For the Asian-European gap, the father's highest qualification makes a larger contribution to higher participation rates among Asians (-0.72 percentage points for father's qualification versus -0.53 for mother's qualifications). In general, these results signal that any differing role played by mother's versus father's education with respect to their child's educational choices is minor, with the evidence pointing to mother's education being slightly more relevant for Māori, and Pasifika, and father's education being of marginally greater importance for Asians.

8 Conclusions

There are substantial ethnic disparities in the rates of bachelor's degree participation in NZ. Compared with Europeans, Māori and Pasifika have much lower participation rates, while Asians have much higher rates.

We explore the drivers of these ethnic differences using newly linked administrative data. This allows us to follow a population cohort and control for the majority of factors (based on individual, school and parental characteristics) suggested by existing literature. This includes the three drivers that stand out (from past research) as being particularly important: socioeconomic status, school performance and parental education. In addition, we decompose the ethnic education gaps to measure the degree to which they are explained by each of these observable factors. To our knowledge, this is the first study that is able to control for all three main factors at a population level and which uses decompositions to examine ethnic differences in bachelor's degree participation.

Our results that are consistent with the extant literature – the three sets of variables with the largest marginal effects were socioeconomic status, prior performance in school, and parent's educational attainment. A key finding from these results was in terms of the relative contributions of each factor, which indicated that prior performance in school played a substantially greater role than the other two factors.

Both the probit models and the Fairlie decompositions highlighted the critical importance of doing well in the first set of assessments that students face at high school (NCEA Level 1). The vast majority of the explained ethnic gaps were due to differences in school performance (at age 15-16), with socioeconomic status and parental education also playing important, but much smaller, roles. Given the substantial ethnic divide apparent in individuals' early high school outcomes, further work should focus on school achievement and how this feeds into achievement at higher levels of education. This analysis could be undertaken with the use of transition models, which would not only allow us to investigate the drivers of high school outcomes (at all NCEA levels), but would also address potential selection and identification issues.

In the decomposition analysis, we find that our covariates explain just over 86% of the Māori-European gap in participation, indicating the presence of unobservable factors driving a portion of the ethnic gap. In contrast, the Pasifika-European gap in participation is entirely explained by differences in observed characteristics. Indeed, if Pasifika were given the same characteristics as their European counterparts, Pasifika would have higher rates of participation than Europeans.

While not the primary focus of this study, it is useful to note that differences in characteristics do not explain much of the higher rates of participation among Asians relative to Europeans. In fact, less than half the Asian-European gap could be explained. This suggests that unobservable factors, such as cultural values and attitudes, may be particularly relevant in explaining the educational overachievement of Asians in bachelor's study in NZ.

Overall, our results suggest that ethnic-based policies aimed at encouraging entrance to bachelor's degrees are likely to have a limited effect if used in isolation. Rather, our findings highlight the need for policy intervention earlier in the education system to help lift the NCEA performance of Māori and Pasifika, and in doing so improve the likelihood of their participation in higher education qualifications, such as bachelor's degrees.

References

Black, S. E., & Sufi, A. (2002). *Who goes to college? Differential enrollment by race and family background.* NBER Working Paper Series, no. 9310. Cambridge, MA, United States: National Bureau of Economic Research.

Black, S. E., Lincove, J., Cullinane, J., & Veron, R. (2015). Can you leave high school behind? *Economics of Education Review*, 46, 52-63.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources, 8*(4), 436-455.

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

Chowdry, H., Crawford, C., Dearden, L., Goodman, A., & Vignoles, A. (2013). Widening participation in higher education: Analysis using linked administrative data. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 176*(2), 431-457.

Ciao, Z., & Maloney, T. (2017). *Decomposing ethnic differences in university academic achievement in New Zealand*. AUT Working Paper. Auckland: Auckland University of Technology.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2015). *Public universities, equal opportunity, and the legacy of Jim Crow: Evidence from North Carolina*. NBER Working Paper Series, no. 21577. Cambridge, MA, United States: National Bureau of Economic Research.

Conway, P. (2016). *Achieving New Zealand's productivity potential.* NZPC Research Paper, no. 2016/1. Wellington, New Zealand: New Zealand Productivity Commission.

Crawford, C. (2014). *Socio-economic differences in university outcomes in the UK: Drop-out, degree completion and degree class*. IFS Working Paper, no. 14/31. London, United Kingdom: Institute for Fiscal Studies.

Crawford, C., & Greaves, E. (2015). *Socio-economic, ethnic and gender differences in HE participation.* Research Paper, no. 186. London, United Kingdom: Department for Business Innovation and Skills.

Crawford, R. (2016). *History of tertiary education reforms in New Zealand.* NZPC Research Note, no. 2016/1. Wellington, New Zealand: New Zealand Productivity Commission.

Datcher, L. (1982). Effects of community and family background on achievement. *The Review of Economics and Statistics, 64*(1), 32-41.

Earle, D. (2007). *Te Whai i Ngā Taumata Atakura: Supporting Māori achievement in bachelors degrees*. Wellington, New Zealand: Ministry of Education.

Earle, D. (2008). Hei titiro anō i te whāinga Māori achievement in bachelors degrees revisited. Wellington, New Zealand: Ministry of Education.

Engler, R. (2010). School leavers' progression to bachelors-level study. Wellington, New Zealand: 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 Labor 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.

Fairlie, R. W. (2016). *Addressing path dependence and incorporating sample weights in the nonlinear Blinder-Oaxaca Decomposition technique for logit, probit and other nonlinear models*. Santa Cruz, CA, United States: Department of Economics, Univeristy of California.

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.

Finnie, R., Wismer, A., & Mueller, R. E. (2015). Access and barriers to postsecondary education: Evidence from the Youth in Transition Survey. *Canadian Journal of Higher Education, 45*(2), 229-262.

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.

Fortin, N. M. (2006). *Greed, altruism, and the gender wage gap.* Canada: Department of Economics, University of British Columbia.

Gang, I. N., & Zimmerman, K. F. (2000). Is child like parent? Educational attainment and ethnic origin. *The Journal of Human Resources, 35*(3), 550-569.

Gibbons, S., & Vignoles, A. (2012). Geography, choice and participation in higher education in England. *Regional Science and Urban Economics, 42*(1), 98-113.

Haveman, R. & Wolfe, B. (1995). The determinants of children's attainment: A review of methods and findings. *Journal of Economic Literature, 33*(4), 1829-1878.

Hill, M. S. & Duncan, G. J. (1987). Parental family income and the socioeconomic attainment of children. *Social Science Research, 16*, 39-73.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal, 8*(4): 453-479.

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

Kane, T. J. (1994). College entry by blacks since 1970: The role of college costs, family background, and the returns to education. *Journal of Political Economy, 102*(5), 878-911.

Li, W. (2007). Family background, financial constraints and higher education attendance in China. *Economics of Education Review, 26*(6), 724-734.

Maani, S. (1999). *Private and public returns to investments in secondary and higher education in New Zealand over time: 1981-1996*. Treasury Working paper, no. 99/02. Wellington, New Zealand: New Zealand Treasury.

Maani, S., & Maloney, T. (2004). *Returns to post-school qualifications: New evidence based on the HLFS Income Supplement (1997-2002)*. Wellington, New Zealand: Department of Labour.

Maani, S. (2006). Parental income and the choice of participation in university, polytechnic or employment at age 18: A longitudinal study. *Research on Economic Inequality*, *13*, 217-248.

Marginson, S. (2016). The worldwide trend to high participation higher education: Dynamics of social stratification in inclusive systems. *Higher Education*, *72*(4), 1-22.

Marks, G. N. (2008). Are father's or mother's socioeconomic characteristics more important influences on student performance? Recent international evidence. *Social Indicators research, 85*(2), 293-309.

Ministry of Education. (2014). *Tertiary Education Strategy 2014-2019*. Wellington, New Zealand: Ministry of Education and Ministry of Business, Employment and Innovation.

Ministry of Education. (2016). *Participation Rates.* [Data file]. Wellington, New Zealand: Education counts.

Ministry of Social Development. (2008). *Children and young people: Indicators of wellbeing in New Zealand 2008*. Wellington, New Zealand: MSD.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review, 14*(3), 693-709.

Reardon, S. F., Baker, R., & Klasik, D. (2012*). Race, income, and enrollment patterns in highly selective colleges, 1982-2004*. Working Paper, no. 526. Stanford, CA, United States: Center for Education Policy Analysis.

Rothstein, J. (2009). *SAT scores, high schools, and collegiate performance predictions.* Paper presented at the annual meeting of the National Council on Measurement in Education, Montreal, CA.

Salmond, C., Crampton, P., & Atkinson, A. (2007). *NZDep2006 Index of Deprivation*. Dunedin, New Zealand: University of Otago.

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.

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

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

Yorke, M. (1998). Undergraduate non-completion in England: some implications for the higher education system and its institutions. *Tertiary Education and Management*, *4*(1), 59-70.

Appendix A: Sensitivity analysis tables

Table A.1 Marginal effects from probit analysis of participation in a bachelor's qualification: School fixed effects and first NCEA attempt

	European		Māori		Pasifika		Asian	
	School FE	First NCEA	School FE	First NCEA	School FE	First NCEA	School FE	First NCEA
Cohort year: - Cohort 1992	-0.007	-0.010*	0.0008	-0.002	0.010	0.013	-0.013	-0.018
- Cohort 1993	-0.022***	-0.019***	0.0037	0.004	0.010	0.019*	-0.016	-0.007
- Cohort 1994	-0.044***	-0.014***	-0.006	0.001	-0.004	0.009	-0.037**	-0.016
Individual characteristics								
Male	-0.064***	-0.091***	-0.030***	-0.039***	-0.055***	-0.083***	-0.033***	-0.041***
Migrant	0.037***	0.054***	0.012	0.029***	0.044***	0.033***	0.016	0.006
Switching school	-0.027***	-0.046***	-0.003	-0.010**	-0.030**	-0.051***	-0.018	-0.032*
Socioeconomic status: Deprivation index = 1	0.079***	0.131***	0.042***	0.070***	0.044	0.063**	0.037	0.059**
= 2	0.055***	0.097***	0.025***	0.047***	0.042	0.059**	0.045	0.066***
= 3	0.059***	0.093***	0.025***	0.041***	0.017	0.038*	0.086***	0.106***
= 4	0.035***	0.067***	0.032***	0.047***	0.007	0.052**	0.074**	0.088***
= 5	0.047***	0.061***	0.014**	0.023***	0.017	0.030	0.071**	0.082***
= 6	0.027**	0.038***	0.018***	0.028***	0.020	0.023	0.043	0.046*
= 7	0.032***	0.039***	0.019***	0.025***	-0.0004	-0.0004	0.060*	0.060***
= 8	0.023**	0.028***	0.008	0.015***	0.026**	0.024**	0.079***	0.068***
= 9	0.020*	0.022**	0.008*	0.013***	-0.003	-0.0002	0.024	0.018
School characteristics								
Single-sex high school	-	0.049***	-	0.015***	-	0.046***	-	0.032***
School authority: State Integrated	-	0.069***	-	0.030***	-	0.038***	-	0.034**
Private or Other	-	-0.069***	-	0.081**	-	0.107	-	0.017
School performance and engagement								
NCEA Level 1 : Achieved	0.119***	0.220***	0.061***	0.106***	0.131***	0.199***	0.402***	0.515***
Merit	0.615***	0.687***	0.464***	0.516***	0.610***	0.661***	0.817***	0.833***
Excellence	0.850***	0.865***	0.673***	0.758***	0.836***	0.885***	0.881***	0.873***
Number of school notifications	-0.083***	-0.117***	-0.029***	-0.045***	-0.080***	-0.103***	-0.067***	-0.094***
Distance to nearest provider (10 kms)	-0.006***	-0.010***	-0.002*	-0.003***	-0.009	-0.012***	-0.008	-0.014***
Parents' highest qualification: - School	0.050***	0.070***	0.021***	0.030***	0.017*	0.032***	0.030	0.066***
- Post-school	0.062***	0.088***	0.027***	0.036***	0.048***	0.065***	-0.000	0.024
- Bachelor's	0.157***	0.212***	0.062***	0.090***	0.150***	0.162***	0.086***	0.135***
- Postgraduate	0.193***	0.268***	0.105***	0.155***	0.138***	0.187***	0.124***	0.184***
School fixed effects	Yes	No	Yes	No	Yes	No	Yes	No
Pseudo R^2	0.385	0.303	0.367	0.277	0.329	0.230	0.364	0.286
N	84,426	83,787	30,966	30,432	11,298	11,496	8,958	9,390

Note: 'School FE' replaces the characteristics of the school attended with a dummy for the school attended. 'First NCEA' uses the first NCEA Level 1 result obtained rather than the highest NCEA Level 1 result ever obtained. Reference groups = Cohort 1991, male, NZ born, deprivation index=10, co-ed school, state school authority, not attained NCEA, and parents with no qualifications. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table A.2 Marginal effects from probit analysis of participation in a bachelor's qualification: Single ethnicity and separating parents' education

	European		Māori		Pasifika		Asian	
	Single ethnicity	Mother/Father	Single ethnicity	Mother/Father	Single ethnicity	Mother/Father	Single ethnicity	Mother/Father
Cohort year: - Cohort 1992	-0.009*	-0.002	0.002	-0.002	0.008	0.06	-0.018	-0.006
- Cohort 1993	-0.023***	-0.008*	0.003	0.003	0.011	0.005	-0.017	0.001
- Cohort 1994	-0.044***	-0.029***	-0.003	-0.008**	0.000	-0.000	-0.041***	-0.027**
Individual characteristics								
Male	-0.070***	-0.067***	-0.030***	-0.028***	-0.068***	-0.070***	-0.023**	-0.010
Migrant	0.048***	0.028***	0.006	0.016**	0.031***	0.026***	-0.013	0.001
Switching school	-0.027***	-0.034***	-0.002	-0.001	-0.037***	-0.027***	-0.015	-0.007
Socioeconomic status: Deprivation index = 1	0.103***	0.110***	0.041***	0.066***	0.039	0.036	0.053*	0.078***
= 2	0.071***	0.087***	0.025***	0.042***	0.041	0.047**	0.070***	0.087***
= 3	0.066***	0.078***	0.019**	0.040***	0.034	0.043**	0.100***	0.105***
= 4	0.044***	0.058***	0.027***	0.035***	0.030	0.011	0.093***	0.102***
= 5	0.047***	0.061***	0.009	0.020***	0.024	0.018	0.084***	0.083***
= 6	0.024**	0.041***	0.011*	0.023***	0.015	0.027**	0.059**	0.059***
= 7	0.028**	0.036***	0.014**	0.023***	-0.009	-0.000	0.065***	0.081***
= 8	0.019*	0.025**	0.003	0.011**	0.021*	0.014	0.065***	0.073***
= 9	0.015	0.026**	0.008	0.013***	-0.006	0.0007	0.025	0.039*
School characteristics								
Single-sex high school	0.045***	0.052***	0.013***	0.014***	0.027***	0.027***	0.023*	0.036***
School authority: State Integrated	0.050***	0.060***	0.023***	0.023***	0.032***	0.030***	0.027	0.015
Private or Other	-0.088***	-0.047***	0.061*	0.056**	0.090	0.087*	-0.016	-0.051
School performance and engagement								
NCEA Level 1 : Achieved	0.122***	0.130***	0.059***	0.066***	0.129***	0.123***	0.391***	0.375***
Merit	0.614***	0.642***	0.437***	0.473***	0.590***	0.575***	0.804***	0.816***
Excellence	0.838***	0.859***	0.641***	0.700***	0.823***	0.819***	0.866***	0.877***
Number of school notifications	-0.081***	-0.091***	-0.030***	-0.035***	-0.066***	-0.063***	-0.060***	-0.076***
Distance to nearest provider (10 kms)	-0.009***	-0.007***	-0.002**	-0.002***	-0.009**	-0.003	-0.017***	-0.009***
Parents' highest qualification: - School	0.049***	-	0.019***	-	0.021**	-	0.055***	-
- Post-school	0.064***	-	0.023***	-	0.043***	-	0.037**	-
- Bachelor's	0.174***	-	0.057***	-	0.129***	-	0.134***	-
- Postgraduate	0.220***	-	0.104***	-	0.142***	-	0.166***	-
Mother's highest qualification: - School	-	0.050***	-	0.018***	-	0.020**	-	0.021
- Post-school	-	0.076***	-	0.025***	-	0.043***	-	-0.019
- Bachelor's	-	0.154***	-	0.056***	-	0.107***	-	0.073***
- Postgraduate	-	0.192***	-	0.094***	-	0.127***	-	0.124***
- Missing	-	0.109***	-	0.024***	-	0.012	-	0.010
Father's highest qualification: - School	-	0.049***	-	0.029***	-	0.017	-	0.031*
- Post-school	-	0.050***	-	0.019***	-	0.047***	-	0.005
- Bachelor's	-	0.177***	-	0.070***	-	0.136***	-	0.113***
- Postgraduate	-	0.193***	-	0.097***	-	0.124***	-	0.131***
- Missing	-	0.079***	-	0.017***	-	-0.019**	-	-0.013
Pseudo R^2	0.367	0.376	0.341	0.363	0.302	0.308	0.344	0.330
N	83,787	117,729	22,245	39,330	9,834	15,897	8,643	13,203

Note: 'Single ethnicity' uses individuals who identify with one ethnicity only. 'Mother/Father' separates parents' education into mother's and father's education. Reference groups = Cohort 1991, male, NZ born, deprivation index=10, co-ed school, state school authority, not attained NCEA, and parents with no qualifications. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table A.3 Decomposition of ethnic differences in participation: School fixed effects and first NCEA attempt

	Māori vs European		Pasifika vs European		Asian vs European	
	School FE	**First NCEA**	**School FE**	**First NCEA**	**School FE**	**First NCEA**
Ethnic difference[1] (% point: European – Ethnic minority)	34.56 – 14.68	34.75 – 14.91	34.56 – 23.19	34.75 – 23.31	34.56 – 71.64	34.75 – 71.73
Total difference (% point)	**19.88**	**19.84**	**11.37**	**11.45**	**-37.08**	**-36.97**
Cohort year	0.01**	0.01***	0.02***	0.02***	0.01**	0.01*
Individual characteristics						
Male	0.06***	0.05***	-0.07**	-0.12***	-0.28***	-0.28***
Switching school	0.25***	0.27***	0.05***	0.06***	-0.09***	-0.09***
Migrant	0.16***	0.23***	-0.32***	-0.46***	-1.52***	-2.14***
Socioeconomic information						
Deprivation index	1.55***	2.19***	2.33***	3.17***	0.52***	0.79***
School characteristics	-	0.44***	-	-0.41***	-	-0.03***
School Fixed Effects	1.68***	-	-1.60*	-	-3.59***	-
School performance and engagement						
NCEA Level 1	10.14***	10.32***	8.96***	9.32***	-11.20***	-11.10***
Number of school notifications	0.91***	0.89***	0.55***	0.52***	-0.43***	-0.39***
Other variable						
Distance to nearest delivery site	-0.01	-0.03***	-0.38***	-0.70***	-0.46***	-0.84***
Parents' highest qualification	1.63***	1.82***	2.26***	2.52***	-1.23***	-1.50***
Total difference explained (% point)	16.39	16.20	11.83	13.93	-18.26	-15.59
Proportion of gap explained	82.45%	81.66%	103.98%	121.66%	49.25%	42.16%
Sample size	115,947	114,216	96,174	95,280	93,984	93,174

Note: Several variables are grouped into the following clusters of information: Cohort year includes cohort 1992, 1993 and 1994; School characteristics include single sex school, state integrated authority and private school; NCEA Level 1 includes attaining this level with achievement, merit and excellence; and Parents' highest qualification includes school, post-school, bachelor's and postgraduate.

Note: 'School FE' replaces the characteristics of the school attended with a dummy for the school attended. 'First NCEA' uses the first NCEA Level 1 result obtained rather than the highest NCEA Level 1 result ever obtained.

[1] Based on the sub-sample of the population cohort that has linked information regarding parents' educational attainment.

Table A.4 Decomposition of ethnic differences in participation: Single ethnicity and separating parents' education

	Māori vs European		Pasifika vs European		Asian vs European	
	Single ethnicity	Mother/Father	Single ethnicity	Mother/Father	Single ethnicity	Mother/Father
Ethnic difference[2] (% point: European – Ethnic minority)	34.75 –13.33	39.19 – 15.55	34.75 – 21.76	39.19 – 21.57	34.75 – 72.91	39.19 – 70.67
Total difference (% point)	**21.42**	**23.64**	**12.99**	**17.62**	**-38.16**	**-31.48**
Cohort year	0.03***	0.02***	0.031***	0.02***	0.05***	0.03***
Individual characteristics						
Male	0.08***	0.05***	-0.08***	-0.02***	-0.17***	-0.17***
Switching school	0.15***	0.18***	0.03***	0.06***	-0.04***	-0.03***
Migrant	0.18***	0.13***	-0.40***	-0.20***	-1.94***	-0.98***
Socioeconomic information						
Deprivation index	1.58***	1.60***	2.24***	2.38***	0.65***	0.67***
School characteristics	0.25***	0.45***	-0.42***	-0.16***	-0.02	0.05***
School performance and engagement						
NCEA Level 1	14.37***	16.30***	12.59***	15.39***	-13.90***	-11.59***
Number of school notifications	0.57***	0.63***	0.32***	0.40***	-0.25***	-0.19***
Other variable						
Distance to nearest delivery site	0.02**	-0.06***	-0.55***	-0.45***	-0.72***	-0.55***
Parents' highest qualification	1.40***	-	2.02***	-	-1.25***	-
Mother's highest qualification	-	0.77***	-	0.77***	-	-0.53***
Father's highest qualification	-	0.66***	-	0.66***	-	-0.72***
Total difference explained (% point)	18.65	20.76	15.78	18.81	-17.58	-13.98
Proportion of gap explained	87.07%	87.79%	121.49%	106.78%	46.08%	44.42%
Sample size	106,032	157,062	93,618	133,623	92,427	130,935

Note: Several variables are grouped into the following clusters of information: Cohort year includes cohort 1992, 1993 and 1994; School characteristics include single sex school, state integrated authority and private school; NCEA Level 1 includes attaining this level with achievement, merit and excellence; and Mother's (Father's) highest qualification includes school, post-school, bachelor's and postgraduate.

Note: 'Single ethnicity' uses individuals who identify with one ethnicity only. 'Mother/Father' separates parents' education into mother's and father's education.

[2] 'Single ethnicity' column results are based on those with parents' education information. 'Mother/Father education' column results are based on all individuals.