Background paper 2
Contributors to education achievement — Indigenous and non-Indigenous primary school students

Key points

- Data on all schools and Year 3 and 5 students in 2013 and 2014 were used to investigate reading and numeracy achievement for Indigenous and non-Indigenous primary school students.
- Most of the variation in achievement is attributable to student characteristics rather than school characteristics. This suggests that policies targeting the needs of individual students are likely to be important in improving student achievement.
- A greater proportion of variation in achievement is attributable to schools for Indigenous students than non-Indigenous students, suggesting that school characteristics (including social and peer influences) are relatively more important for Indigenous students than non-Indigenous students. A preliminary analysis suggests that this result is highly influenced by very remote schools.
- Consistent with past studies, most of the variation attributable to schools is explained by characteristics observed in the data, while little of that attributable to students can be explained. Unobserved student characteristics that might contribute to achievement include cognitive abilities and student attitudes.
- Indigenous students have lower levels of achievement even after characteristics such as language background and socioeconomic status (SES) are taken into account.
- Both Indigenous and non-Indigenous students perform less well when the proportion of Indigenous students in a school is higher, even after accounting for other student and school characteristics.
- For both Indigenous and non-Indigenous students, attending a school where a large proportion of students have a language background other than English does not have an influence on student achievement.
- For Indigenous students, the most important observed characteristics in explaining achievement are their own SES and that of their school community, along with the school attendance rate and the school's share of Indigenous students. The region in which they attend school is not as important as other characteristics.
- For non-Indigenous students, student and school SES are the most important factors, accounting for nearly three-quarters of the variation in achievement that can be explained by observed factors.
- About half of the gap in average achievement between Indigenous and non-Indigenous students is attributed to differences in the observed characteristics of schools and students in each group, while half is due to differences in how their characteristics are associated with achievement.
- Students at some schools do considerably better (or worse) than might be expected. This could be partly be due to unobserved school characteristics. However, there is little overlap in the high-achieving schools for Indigenous students and those for non-Indigenous students. Unobserved school characteristics that work well for one group of students may not necessarily work as well for the other, which highlights the importance of taking Indigenous status into account in schools.
- Further insight into why certain schools, particularly those with Indigenous students, perform well could be gained by examining high-performing schools to see what unobserved characteristics set them apart. These insights could be used to inform policy and lift Indigenous performance in other schools.

About this paper

Improving Indigenous education achievement is a key government objective, as part of a wider goal of reducing the disadvantage experienced by Indigenous Australians.[1] In 2008, the Council of Australian Governments set a target to halve the gap in reading, writing and numeracy achievements for Indigenous children within ten years (COAG 2011). Despite these commitments to reducing the gap, there has been little evidence of improvement (PC 2015). This raises the question of what more could be done to improve Indigenous education achievement.

This paper uses a novel dataset to shed light on the contributors to reading and numeracy achievement for Indigenous primary school students. The relationships between Indigenous students' achievement and a range of school and student characteristics are tested. A key question is whether the approaches that benefit non-Indigenous students also work for Indigenous students. To examine this, the analysis distinguishes between these two groups.

The paper seeks to address the following questions:

- What are the relative contributions of school and student characteristics to students' reading and numeracy achievement? In other words, how much of the variation in achievement between students is accounted for by the schools that they attend and how much by differences in students' characteristics? (section 5)
- What is the relationship between achievement and each school and student characteristic observed in the data when other characteristics are taken into account? In other words, what is the average association (measured in reading or numeracy test score points) between achievement and each characteristic? (section 6)
- Which of the observed characteristics are most important in explaining reading and numeracy achievement? That is, how much of the variation in achievement between students is accounted for by each characteristic observed in the data? (section 7)
- How much of the gap in average test scores between Indigenous and non-Indigenous students is explained by differences in their observed characteristics and how much by differences in the relationships between those characteristics and achievement? (section 8)
- Are there some schools where Indigenous students do better than would be expected, given their observed characteristics and those of the schools that they attend? Do non-Indigenous students at these schools also do relatively well? (section 9)

[1] The term 'Indigenous Australians' is used throughout this background paper to refer to Australia's first peoples. While this is convenient for drafting, it does not reflect the diversity of Indigenous peoples, communities and nations.

Although there are several Australian studies that look into school and student contributors to achievement, few specifically consider Indigenous students, and none have examined Indigenous students at the primary school level. Furthermore, most of the literature has used national survey data or administrative records from state education departments. The Commission is unaware of any research into the contributors to student achievement that uses national administrative data covering all students and schools.

The research in this paper adds to the literature by examining reading and numeracy achievement using national data (compiled by the Australian Curriculum, Assessment and Reporting Authority (ACARA)) that cover all Year 3 and Year 5 students across Australia in 2013 and 2014. The data include students' test scores from the National Assessment Program — Literacy and Numeracy (NAPLAN). This dataset has only recently become accessible to researchers and provides a rich source of de-identified information on students and schools. The analysis pools data from 2013 and 2014 to increase the size of the dataset and the likelihood that significant relationships will be identified, if they exist, particularly for Indigenous students. The results in this paper focus on reading and numeracy achievement for Year 5 students only. Similar conclusions are drawn from analyses for Year 3 students and for other NAPLAN test domains. For brevity, these results are not reported in this paper, but are available in annex B.

A further point is that there has been little publicly available research involving the evaluation of schools that perform better than expected for Indigenous or non-Indigenous students (chapter 4). While the analysis in this study is not able to evaluate schools, it does attempt to identify schools that are performing better than would be expected given their characteristics observed in the data and the characteristics of the students attending those schools. Schools such as these could be examined to see whether there are unobserved characteristics contributing to the higher achievement of their students, and if these characteristics could be replicated across other schools.

This paper conveys the key findings from the analysis. A conceptual framework for the analysis is first described (section 1). Then, brief and relatively simple descriptions of the statistical techniques (section 2) and underlying data (section 4) are included, and results are presented graphically as far as is practicable. More detailed and technical descriptions of the work are presented in annex A. The paper also presents some findings from the relevant Australian literature (section 3).

1 Framework of the analysis

Conceptual framework

Many different factors can affect the academic achievement of a student (Hanushek 1986; Hattie 2003; Zubrick et al. 2000). These factors can be grouped into ecological categories based on the contexts in which they affect achievement. For example, student achievement can be affected by:

- social influences, such as the characteristics of regions in which students live
- school-specific influences, such as principals, school policies and school culture
- influences of school peers, such as peer attitudes towards education
- teacher influences, such as teacher experience and expectations
- family influences, such as the highest education level and occupation of parents
- student-specific influences, such as gender, language background and cognitive ability.

The factors within these ecological contexts can be further categorised as ‘policy factors’ or ‘environmental factors’. ‘Policy factors’ are within the control of education policy makers, while ‘environmental factors’ are not within the control of policy makers but might be of interest in their own right.

For example, at a broad state or national level, policy can be modified through changes to school curricula or funding for particular education programs. At a more narrow level, principals, school councils, and individual teachers can influence the academic achievement of students, for example, by changing school policies, teaching methods or teacher attitudes and expectations.

Environmental factors are generally beyond the control of education institutions and policy makers but also have an important influence on academic achievement. Broad environmental factors can include aspects of the social contexts in which students live, such as the neighbourhood or school community. For example, the number of employment opportunities in a region or the proportion of students with a language background other than English at a school could have an impact on achievement, but cannot be directly influenced by education policy makers. Environmental factors can also include those specific to each student, such as the student’s family situation, upbringing and own innate abilities and learning capacity. For example, parental attitudes towards education, and a student’s health, could play a role in the academic achievement of a student.

Some social, school-specific and teacher influences can be modified through policy, while most peer, family, and student-specific influences are environmental. Nevertheless, policy could still be implemented based on the knowledge of how environmental factors influence achievement. For example, if students with a non-English language background tend to have lower achievement, policy that is aimed at improving achievement specifically for that group of students could be introduced.

A summary of how specific factors are expected to be related to student achievement, according to the literature, is presented in table 1. Further findings from the literature are presented in section 3.

Operationalising the conceptual framework

In operationalising the conceptual framework, the ecological categories of factors that influence achievement are classified into two groups relating to school-level and student-level characteristics. School-level characteristics include school-specific characteristics mentioned above, as well as social influences and peer influences that are the same for all students within a school. Student-level characteristics include characteristics of the student and of their parents and family. As for teacher influences, these are reflected at the school level to the extent that they are the same for all students within a school, and at a student level to the extent that their influences differ between students within a school. For simplicity, from here on, school-level characteristics and student-level characteristics are referred to as 'school characteristics' and 'student characteristics' respectively in this paper. The various factors that influence achievement can be brought together in a formal model that describes how each contributes to achievement (annex A). Both policy factors and environmental factors at school and student levels should be taken into account in order to produce unbiased estimates of their relationships with achievement. For the characteristics that are examined in this study, table 1 describes how and why each is expected to be related to student achievement in the present analysis.

As described in table 1, the expected relationships for some characteristics in the analysis are because of the omission of potentially important factors. Not all the factors that could affect achievement are analysed in this study due to a lack of data (section 4). Therefore, a distinction is made between characteristics that are observed in the ACARA data and those that are unobserved. Examples of characteristics that might have an influence on achievement, classified into those that are observed in the ACARA data and those that are unobserved, are illustrated in figure 1. Some of the unobserved data exist at a national level but are not available for the present study because they are not linked with the school and student data provided by ACARA. It is noted that there are many potential influences on achievement and the examples in figure 1 are not exhaustive. Furthermore, some unobserved characteristics that influence achievement may never be able to be identified or observed.

Table 1 Expected relationships between observed characteristics and achievement

Characteristic	*Expected relationship*[a]	*Reason for relationship*	*Source(s)*
Student characteristics			
Age	Uncertain	• Some students who entered school at a younger age may have done so because they had higher ability than their peers. • Delayed school entry or grade repetition may be associated with lower cognitive maturity. • Older students who have not been held back a grade may achieve higher scores because they are more developmentally mature (cognitively and emotionally).	Fertig and Kluve (2005); Grissom (2004)
Male	Negative (larger in reading, smaller in numeracy)	• Girls may be more motivated to perform well in class than boys. • Boys may be more prone to disruptive and inattentive classroom behaviours. • Girls may experience higher levels of maths anxiety than boys, which is related to poorer maths performance.	Devine et al. (2012); Fergusson and Horwood (1997); Pomerantz, Altermatt and Saxon (2002); Voyer and Voyer (2014)
Language background other than English (LBOTE)	Negative for Indigenous, uncertain for non-Indigenous	• Students with a LBOTE may face a disadvantage because the test is delivered in English and may be biased towards mainstream culture. • Parents with a LBOTE may have more difficulties assisting their children's schooling. • Cultural differences in parental expectations and parenting styles may mean that students from some non-English backgrounds perform better than those from English-speaking backgrounds.	Abedi (2003); Marks, McMillan and Hillman (2001); Nous Group (2011)
Parental education level	Positive	• More highly educated parents may have higher educational expectations of their children. • More highly educated parents may be more likely to provide a home environment that helps to stimulate achievement.	Chevalier (2004); Davis-Kean (2005); Plug and Vijverberg (2003)
Parental occupation	Positive	• Parents in more highly skilled occupations may earn more and be more able to financially provide their children with resources so that they can reach their full schooling potential. • Parents who are employed may act as role models, encouraging their children to also work hard. • Parents who are employed may have richer social networks. Students may benefit from the information learned through these social networks.	Chevalier (2004); Morgan and Sørensen (1999); Plug and Vijverberg (2003); Schulz (2005); Wiese and Freund (2011)
Student mobility	Negative	• Moving schools may cause disruptions in schooling and create gaps in learning through higher rates of absence. • Mobility can have negative impacts on student engagement, attendance and behaviour. • Highly mobile students may be experiencing problems in family or personal situations, which can affect their school performance.	Boon (2011); Lu and Rickard (2016)

(continued next page)

Table 1 (continued)

Characteristic	*Expected relationship*[a]	*Reason for relationship*	*Source(s)*
School characteristics			
School sector (government, Independent or Catholic)	None	• Australian studies generally suggest that there are no significant differences in test results between different school sectors after student background characteristics (such as socioeconomic status) are controlled for.	Cobbold (2015)
Primary and secondary schools combined	Positive	• Combined schools may have more opportunities to offer cross-age tutoring, and allow students to participate in programs for students of different ages. • Combined schools may give primary school students access to staff expertise across all years and across different subject areas.	Victorian Department of Education and Training (2016)
Student–teacher ratio	Uncertain	• Smaller class sizes mean that teachers can potentially focus more time on the individual learning needs of each student, as opposed to the class as a whole. • The effect of class size depends on whether or how teachers actually change their teaching strategy for different sized classes. • Smaller class sizes may mean that lower quality teachers are hired to staff the extra classes.	Ehrenberg et al. (2001); Hoxby (2000b)
Non-teaching staff–student ratio[b]	Uncertain	• The role of teaching aides (and other support staff) is to help students who may need additional assistance to achieve learning outcomes. • Schools that have more students with learning difficulties are likely to employ more support staff. Given the lack of a disability indicator in this analysis, the influence of learning difficulties on achievement could be partially captured in the estimated influence of non-teaching staff.	Bourke (2008)
School enrolments	Negative	• Smaller schools may enable closer relationships to be developed between students and teachers and make teachers assume more responsibility for student learning. • Larger schools may increase specialisation among teachers, as well as increase opportunities for students to develop social relationships. • Empirical studies generally find negative relationships between school size and achievement.	Howley (1996); Kuziemko (2006); Lee and Loeb (2000); Leithwood and Jantzi (2007)
School finances	Uncertain	• Increasing expenditure per student may give students access to more or better school resources, such as higher quality teachers and technology. • Schools may not effectively use extra funds to improve the learning environment.	Elliot (1998); Greenwald, Hedges and Laine (1996); Hanushek (1997)

(continued next page)

Table 1 (continued)

Characteristic	*Expected relationship*[a]	*Reason for relationship*	*Source(s)*
School attendance rate	Positive	• The attendance rate of a school may reflect the ability of a school to engage its students. • Low school attendance may affect regular attenders because teachers must accommodate chronic absentees in the same class. • School-level attendance may be a proxy for student-level attendance. Students with low attendance rates may be more likely to have behavioural issues or disruptive family environments, which could have an impact on their achievement.	Daraganova, Mullan and Edwards (2014); Hancock et al. (2013); Rothman (2001)
Academic orientation (school test participation rate)	Uncertain	• High rates of NAPLAN participation within a school may reflect a more academically-oriented culture. Students at these schools may be encouraged to perform better by their parents or teachers. • The influence of attending an academically-oriented school may depend on whether parents and teachers use encouraging behaviours (which could increase achievement) or pressuring behaviours (which could lower achievement).	Gemici, Lim and Karmel (2013); Rogers et al. (2009)
Peer demographics (Indigenous, LBOTE)	Uncertain	• Teachers may have lower expectations of Indigenous students and students with a LBOTE. Higher proportions of these students may create a culture in which students are expected to perform badly. • Students may perform better when they are surrounded by peers with similar characteristics because teachers may be able to more easily accommodate their learning styles.	Hoxby and Weingarth (2005) Hoxby (2000a)
Peer SES (parental education level and occupation, school fees)[c]	Positive	• Average SES may affect the school's disciplinary climate and teachers' teaching styles. • High-SES schools may benefit from greater parental support. For example, richer parents may purchase learning resources that spread across the school. • Peer pressure and peer competition in high-SES schools may encourage students to work harder.	Van Ewijk and Sleegers (2010); Hoxby (2000a); Perry and McConney (2010)
State	Mixed	• There are differences in curriculums, schooling systems and programs administered across different states, which could affect achievement. • There may be unobserved differences in student cohorts across states. For example, there may be cultural differences between Indigenous students in Western Australia and those in Victoria.	Australian Government (2015); Zubrick et al. (2006)
Remoteness	Negative	• Schools in more remote areas may find it harder to recruit experienced teachers, and may experience higher teacher turnover. • Teachers in more remote schools may have fewer opportunities for professional development. • Schools in more remote areas may have more limited access to educational facilities and resources. • The relatively low educational requirements of jobs in more remote communities may limit students' educational aspirations.	Lamb, Glover and Walstab (2014); Nous Group (2011)

(continued next page)

Table 1 (continued)

Characteristic	*Expected relationship*[a]	• *Reason for relationship*	*Source(s)*
Other			
Year	Uncertain	• NAPLAN scores may not be perfectly comparable across years due to potential errors in the process of equating scores across years. • There may be unobserved differences in student cohorts across years that could affect achievement.	ACARA (2014b); annex A

a Expected relationships are specific to the current analysis, based on the assumption of linear relationships. For some characteristics, expected relationships may be affected by omitted factors, as described in the 'reason for relationship' column. **b** Non-teaching staff include teachers' aides and assistants, specialist support staff (such as counsellors), administrative and clerical staff, building operations, general maintenance and other service staff (ACARA nd). **c** A school fees variable (standardised by school sector) and interactions between fees and school sector are included in the analysis as a proxy for the parental income of the school community (annex A).

Figure 1 Examples of characteristics influencing student achievement[a,b]

	Observed in the dataset	Unobserved – data exist but not included in dataset	Unobserved – data do not exist
Social	• Remoteness • State	• Local unemployment rate	• Libraries and educational facilities
School	• School sector • Number of enrolments • Staff numbers • Attendance rate • Finances	• Average satisfaction of teachers • Teacher and principal turnover • Principal characteristics	• School policies • School culture • Educational resources • Extracurricular activities
Peers	• % Indigenous students • % LBOTE students • % parents by education / occupation category	• Health • School satisfaction	• Cognitive abilities • Attitudes • Aspirations
Teacher		• Age • Gender • LBOTE • Experience • Qualifications	• Teaching style • Attitudes
Family	• Parental education • Parental occupation	• Parent LBOTE	• Parent engagement • Home learning activities
Student	• Age • Gender • LBOTE	• Health and disability • Attendance	• Cognitive abilities • Attitudes • Aspirations

Grouping of characteristics in the statistical analysis:
School-level Student-level Unobserved

[a] Unobserved characteristics are categorised according to whether they exist at a national level. Unobserved data that exist can include data that are believed to be held in administrative records. [b] School-level characteristics are those that are the same for all students within a school, while student-level characteristics are those that differ between students within a school.

2 Statistical techniques

Three different statistical techniques are used in the analysis presented in this paper. A summary of each technique is presented in this section, with details available in annex A.

First, a regression technique called multilevel modelling is used to determine how much of the variation in students' NAPLAN scores is attributable to their characteristics (both observed and unobserved) and how much is attributable to the characteristics of the schools that they attend (both observed and unobserved). That is, the total variation in student achievement is divided into parts explained by student characteristics (student-level variation) and school characteristics (school-level variation). Student-level variation takes into account the variation in individual student's scores within each school. School-level variation measures how much variation there is in average NAPLAN scores across schools.

The proportion of total variation attributable to schools reflects the extent to which students within a school are similar to each other. Students within a school might be similar because they face the same policies, such as school programs and teaching methods, set by school decision makers. Students could also resemble each other because they have similar environmental characteristics, for example, if they face similar levels of disadvantage or have parents with similar levels of education.

For both school-level and student-level variation, the proportion of variation explained by observed characteristics and the proportion of variation that is unexplained (and attributable to unobserved characteristics) are also identified. This is done by examining how much of each of school- and student-level variation remains unexplained after observed characteristics are added to the model.

Multilevel modelling also permits analysis of relationships between each characteristic observed in the ACARA data and students' NAPLAN scores, taking other characteristics into account (box 1).

The technique also identifies 'school effects', which describe the additional influence of each school on student achievement, after all observed characteristics have been controlled for. These effects are interpreted in the literature as capturing the impact of unobserved school characteristics on student achievement. An analysis of school effects is used to distinguish schools where students do considerably better or worse than their own observed characteristics, and those of the schools that they attend, would suggest.

Fixed effects modelling is an alternative approach to multilevel modelling and was considered for these parts of the analysis but, for a number of reasons discussed in annex A, multilevel modelling was preferred.

Second, a technique called dominance analysis is used to identify how much different sets of observed school and student characteristics contribute to explaining the variation in student achievement.

Box 1 Interpreting results in a multiple regression analysis

Multiple regression analysis, including the multilevel modelling using in this paper, identifies relationships between multiple characteristics and an outcome variable of interest. This type of analysis produces a 'regression coefficient' for each characteristic, which describes how a change in the characteristic is associated with a change in the outcome variable — on average. For example, thinking about the characteristic 'age' (measured in years) in the analysis of student test scores, the coefficient on age is an estimate of the average difference in scores associated with a one year increase in age. The 'on average' nature of the coefficient is important. Say the coefficient on 'age' is 2. This does not mean that a one year increase in age is associated with a 2-point increase in test score for all students, just that that is the association on average.

'Holding other characteristics constant'

For any characteristic, the results of a multiple regression analysis tend to be different from those from an analysis that looks only at the relationship between that characteristic and the outcome. This is because there are often other characteristics associated with both the outcome and the characteristic of interest. These other associations are not taken into account in a simple analysis. For example, parental education is associated with both Indigenous background and reading achievement. Indigenous students are more likely than non-Indigenous students to have parents with low levels of education (background paper 1). And students whose parents have lower levels of education tend to do less well on tests. In a simple analysis it is unclear whether the relatively low test scores of Indigenous students reflect their parents' education or something else related to Indigenous background — the unique relationships cannot be isolated.

Multiple regression analysis aims to identify the unique relationship between each characteristic and the outcome. The estimated coefficient for a particular characteristic is interpreted 'holding other characteristics constant'. For example, the relationship between Indigenous background and test scores is interpreted assuming that other characteristics for which measures are available in the data, such as parental education, have been taken into account.

Interpreting coefficients on categorical characteristics

Some characteristics represent a number of sub-groups, or categories. Gender is an example, with categories of male and female. When characteristics of this type are included in a regression analysis, a coefficient is produced for each of the characteristic's categories except for one 'default category'. The coefficients on each category are interpreted relative to that default category. For example, the default category for mother's highest education level is 'Year 9 or below'. The coefficient on the 'Year 12' category is then interpreted as the average difference in test scores between students with mothers whose highest education level is Year 12 and students with mothers whose highest education level is Year 9 or below, holding other characteristics constant.

Statistical significance

Regression analysis also produces statistics that enable the researcher to infer whether the relationship between a characteristic and the outcome variable is 'statistically significant' or not. A relationship is considered significant if there is a small probability of observing such a relationship due to chance alone when a true relationship does not actually exist. The column charts illustrating relationships between a given characteristic and test score (in section 6 of this paper) show vertical lines on each column that represent 95 per cent 'confidence intervals'. If the confidence interval does not cross zero on the vertical axis, then the relationship is statistically significant at a 5 per cent level of significance — the probability of observing such a relationship when no true relationship exists is less than 5 per cent.

Finally, a technique called Blinder-Oaxaca decomposition is used to split the gap in average test scores between Indigenous and non-Indigenous students into three parts. The technique attributes part of the gap to differences in the average characteristics of Indigenous and non-Indigenous students, another part to differences in the relationships between those characteristics and achievement and a remaining part to differences in characteristics and relationships that can occur at the same time. This provides insight into whether average test scores between Indigenous and non-Indigenous students differ mainly because they have different observed characteristics or because they have different relationships between those characteristics and achievement.

3 Existing evidence

Australian evidence on education outcomes

A number of Australian studies have examined the relative contributions of school and student characteristics to education outcomes using multilevel modelling techniques. They have generally used national survey data or state-specific data and focused on secondary school education outcomes, including tertiary entrance scores, literacy and numeracy achievement and retention (table 2). These studies consistently found that school characteristics (both observed and unobserved) accounted for a relatively small share of the variation in education outcomes (typically about 20 per cent) compared with student characteristics (80 per cent).[2] In addition, although the variation attributable to schools was mostly explained by characteristics for which data were available in those studies, much of the variation attributable to students remained unexplained and was due to characteristics that were not available.

Conclusions about which observed factors contributed most to the variation in an outcome under study varied slightly across studies. However, the fact that different studies analyse different outcomes and included different explanatory factors means that results are not necessarily comparable.

In terms of school-level variation, several studies found that, out of all observed factors, differences in the socioeconomic status (SES) of both students and their school communities made the largest contribution to the variation in outcomes attributable to schools (Lamb et al. 2004; Lokan, Greenwood and Cresswell 2008; Nous Group 2011).[3] Marks, McMillan and Hillman (2001) found that over a third of the total variation between schools was explained by their students' prior achievement and SES, when no other characteristics were controlled for. In contrast, Gemici, Lim and Karmel (2013) found that

2 Studies of education outcomes in other countries have found similar results, though actual percentages may differ (for example, Lokan, Greenwood and Cresswell (2008)).

3 Because each school has a different student profile (that is, compositional differences in the demographic and socioeconomic characteristics of students), differences in students' characteristics can explain some of the variation in achievement attributed to schools.

school sector was the largest predictor of school-level variation in tertiary entrance scores while school SES was insignificant, holding student characteristics (including prior achievement) constant. However, for the probability of enrolling in university, the most important school characteristic was found to be the proportion of students with a language background other than English (which was positively associated with university enrolment), followed by school sector and school SES. Marks (2010), in analysing tertiary entrance scores, found school SES to be insignificant after student SES and other student characteristics were taken into account.

Although SES appears to explain a lot of the variation in education outcomes in several studies, particularly of the variation attributable to schools, it is important to bear in mind that most of the variation is still at the student level, which is largely unexplained.

Table 2 **Examples of multilevel education modelling in Australia**

Author(s)	*Dataset*[a]	*Outcome variable*	*Variation (%)*			
			Total school-level variation	School-level variation explained[b]	Student-level variation explained[b]	Total variation explained[b]
Gemici, Lim & Karmel (2013)	PISA–LSAY 2006	Tertiary entrance score	20	64[c]	38[c]	43[c]
		Probability of attending university	25	94[c]	**na**	**na**
Lamb (2015)	NAPLAN (NSW) 2009, 2013	Year 3 numeracy	17–19	73–77	7–8	18–21[c]
		Year 5 numeracy	21	68–73	6–8	18–22[c]
Lamb et al. (2004)	TIMSS 1996	Junior secondary school mathematics achievement	24	88	23	39[c]
	VCAA (VIC) 2000	Year 12 study score	30	89	43	57[c]
	LSAY 1995	Tertiary entrance score	45	37	10	22[c]
		Student retention	12	72	13	20[c]
Lokan, Greenwood & Cresswell (2008)	PISA 2000	Reading achievement	17	80	26	35[c]
Marks, McMillan & Hillman (2001)	LSAY 1995	Tertiary entrance score	22	**na**	**na**	**na**
Marks (2010)	PISA 2003	Tertiary entrance score	25	**na**	**na**	**na**
Nous Group (2011)	PISA 2009	Reading, mathematics and science achievement	25–27	84–88[c]	27–33[c]	43–47[c]
Rothman & McMillan (2003)	LSAY 1995, 1998	Reading and mathematics achievement	14–18	58–66[c]	10–11[c]	17–20[c]

[a] LSAY — Longitudinal Surveys of Australian Youth; PISA — Programme for International Student Assessment; TIMSS — Third International Mathematics and Science Survey; VCAA — Victorian Curriculum and Accreditation Authority. [b] 'Explained' variation is the variation that can be accounted for by characteristics observed in the data underlying the analysis. [c] Commission calculation based on reported results because estimates of the percentages of explained variation were not reported. **na** Not available.

Characteristics available in survey data

Studies based on survey data often contain information about student and school characteristics that is not available in datasets derived from administrative records. Reported results, however, suggest that these additional characteristics tend to have a small influence on achievement relative to other characteristics, and much of the variation in student outcomes remains unexplained. That is, although characteristics that are available in survey data may be significantly related to student achievement, they do not necessarily explain a lot of the variation in student achievement.

Student educational characteristics

Lokan, Greenwood and Cresswell (2008) found that time spent on homework and home educational resources only accounted for about 4 per cent of total variation in reading achievement, based on their reported results. In contrast, school and student SES-related characteristics (including parental education and wealth) explained about 16 per cent of total variation. Enjoyment of reading was a relatively important factor in explaining variation attributable to students however, accounting for about 7 per cent of total variation, including nearly a third of the proportion of all student variation that was explained. This result could partly be due to reverse causality — in addition to reading enjoyment having a positive effect on reading scores, students who achieve high reading scores could enjoy reading more.

Student psychological characteristics

Marks, McMillan and Hillman (2001) found that a student's self-concept of ability explained an additional 5 per cent of total variation in individual tertiary entrance scores after controlling for year 9 achievement, SES, demographic factors, school sector and remoteness (which, as a group, explained 30 per cent of total variation).

Teacher characteristics

Evidence on the relationship between teacher qualifications and education outcomes is mixed, with some studies finding positive relationships (Nous Group 2011) and others finding no relationship (Gemici, Lim and Karmel 2013; Lamb and Fullarton 2001). Teacher morale and disciplinary climate have also been found to explain less than 1 per cent of the total variation in reading achievement (Lokan, Greenwood and Cresswell 2008). Teacher efficacy, as measured by students' perceptions of their teachers, has been found to have a positive relationship with tertiary entrance scores (Marks 2010).

Evidence on unobserved school effects

A number of studies have used multilevel models to examine differences in 'school effects', or school 'value added', that is, the additional influence of a school on achievement after other observed characteristics are controlled for. These school effects provide an indication of how unobserved characteristics that differ between schools, such as teacher quality, school culture and school ethos, influence achievement. For example, Gemici, Lim and Karmel (2013) presented the distribution of school effects on tertiary entrance scores, and concluded that unobserved school characteristics have a sizable impact. Nous Group (2011) examined school effects on reading, mathematics and science achievement by school sector. They found that government, Independent and Catholic sectors have broadly the same distributions of school effects after controlling for other school and student characteristics. Comparing government schools with all non-government schools, their findings suggested that there was more variation in school effects within the government sector, with some schools having a particularly large influence on achievement and others having a much smaller influence.

Evidence on Indigenous education outcomes

Indigenous background has been analysed to an extent in previous research, but few studies examine Indigenous students specifically to see how contributors to their education outcomes differ to those for non-Indigenous students. Most studies only included an indicator for Indigenous background within regression analysis for all students. These studies found that Indigenous students had lower education outcomes than non-Indigenous students (for example, Biddle and Cameron (2012), Hancock et al. (2013), Marks, McMillan and Hillman (2001) and Nous Group (2011)).

Studies based on cross-tabulations show that there are differences between Indigenous and non-Indigenous students in factors that could influence achievement, such as SES, home educational resources, attendance and student engagement (De Bortoli and Cresswell 2004; De Bortoli and Thompson 2010).

A number of the characteristics that differ between Indigenous and non-Indigenous students were found to be significantly related to achievement in reading, maths and science tests for 15-year old Indigenous students after controlling for other observed factors in a regression analysis (De Bortoli and Thompson 2010). For example, home educational resources, engagement in reading and academic self-concept were found to be related to Indigenous students' reading literacy. However, there were relatively few Indigenous students included in this analysis — relationships between other observed characteristics and achievement could exist but might not have been able to be identified in the analysis.

A few recent studies separate the gap in average achievement between Indigenous and non-Indigenous students into a share that is attributable to differences in the observed characteristics of Indigenous and non-Indigenous students and the schools they attend, and a share attributable to differences in the relationships between those characteristics and achievement (that is, differences in regression coefficients). For example, in a study of 15-year olds using school and student data from the Programme for International Student Assessment (PISA) 2009, differences between Indigenous and non-Indigenous students in the observed characteristics of schools and students explained 44 per cent of the total gap in mean reading test scores, and differences in relationships explained 37 per cent (Song, Perry and McConney 2014).[4] Another report analysing PISA and LSAY data from 2006 and 2009 found that between 50 and 63 per cent of the gap in mean reading scores was explained by differences in characteristics and between 38 and 50 per cent was attributed to differences in relationships or other factors, depending on the method of analysis (Mahuteau et al. 2015).

The above findings suggest that if Indigenous students had the same characteristics as non-Indigenous students, then the gap in average achievement would be smaller than if the relationships between those characteristics and achievement for Indigenous students were the same as for non-Indigenous students. However, these studies only focused on Indigenous students at secondary school. The decomposition analysis in this study complements those above by examining whether the same applies to Indigenous students in primary school.

4 Data description

As noted above, the analyses presented in this paper are based on data compiled by ACARA for Year 3 and Year 5 students in Australia in 2013 and 2014. The data are de-identified — that is, they do not contain identifying information such as the names or addresses of schools or students. The variables provided include:

- students' NAPLAN test scores in the knowledge domains of reading, writing, numeracy and language conventions (spelling, grammar and punctuation)
- students' characteristics, including age, gender, Indigenous background, language background other than English (LBOTE), parental education, parental occupation and whether they attended the same school in Years 3 and 5 (for Year 5 students only)
- school characteristics, including school sector (government, Independent or Catholic), state[5], remoteness, staff numbers, student numbers, proportions of students who are Indigenous or who have a LBOTE, school finances and attendance rates.

In total, the data include over 250 000 students per year and year level (with Indigenous students comprising about 5 per cent) and over 7000 schools. A profile of both Indigenous and non-Indigenous students is presented in background paper (BP) 1.

4 It is not clear from this report what the remaining 19 per cent is attributable to.

5 The term 'states' is used throughout this background paper to refer to states and territories.

The analyses in this paper exclude some students because they had missing data or because they were the only student in their school who sat the relevant NAPLAN test. About 19 per cent of Indigenous students (corresponding to about a third of all schools with Indigenous students) and about 8 per cent of non-Indigenous students (corresponding to about 7 per cent of schools) were excluded. Furthermore, some summary variables were constructed from the available data. Details of the excluded students and the constructed dataset are provided in annex A. A table of descriptive statistics for the samples used in the analysis is also available in annex A.

One of the main benefits of using the ACARA data is that they essentially cover all schools and all students who sat the NAPLAN tests (notwithstanding some exclusions). This means that the results are less susceptible to the bias that can arise when analysis is based on a sample of the population. The risk with a sample is that the members are systematically different from the population that they represent. The number of Indigenous students in the dataset is also large, meaning that Indigenous students can be analysed separately from non-Indigenous students. In many datasets, the small number of Indigenous students rules out analysis of this type.

However, the data and analyses are not without limitations and these need to be borne in mind when considering the results presented in the paper. First, the results are conditional on students sitting the relevant NAPLAN test. Students who do not participate in the test appear to have characteristics that are systematically different from those of students who do participate (BP 1). The results, therefore, cannot be generalised to all students but have to be interpreted as applying only to students with characteristics like those of the students who sat the relevant test. There may also be differences in student cohorts from year to year, which would reduce the extent to which the results can be generalised to students in a different cohort. Furthermore, questions on NAPLAN tests can favour some particular groups of students more than others (annex A). To the extent that scores reflect characteristics other than student ability, results could be biased.

Another downside is that the ACARA data do not contain information on some potentially important student characteristics related to achievement, such as health, cognitive and non-cognitive functioning, learning disabilities, home educational resources and student attendance (figure 1). The data also lack information about some school-related characteristics such as teacher quality and school educational resources (but do include data on attendance rates at a school level). As described above, some of these characteristics have been analysed in studies using survey data, but not all studies report how much of the variation in achievement was explained by these characteristics. Where shares of explained variation were reported, these characteristics were found to explain relatively little of the total variation.

The fact that these characteristics are not included within the ACARA data raises the possibility that the relationships that are estimated between achievement and observed characteristics might be biased. This will be the case if there is an association between omitted and observed characteristics (annex A).

One further point is that it was anticipated that the ACARA data would support separate analyses by state and remoteness region for Indigenous students because the number of students in the data is relatively large. However, multilevel models draw on information about students within each school. Many schools have a very small Indigenous enrolment (and the extent to which this is the case only came to light when ACARA data on students and schools were linked). Tests of the effect of this data characteristic on multilevel modelling results led to a conclusion that the ACARA data could support national level analysis but not state or remoteness level analysis (annex A).

Despite these limitations, the ACARA dataset is a very useful source of information and sheds some light on Indigenous and non-Indigenous primary school outcomes. The results and conclusions in the following sections are for Year 5 students. Year 5 students were chosen as the main focus because they enable conclusions to be made about student mobility, as measured by whether the student attended the same school when they were in Year 3 and Year 5. Comprehensive tables of these results can be found in annex A.

5 How much of the variation in student achievement is attributable to schools?

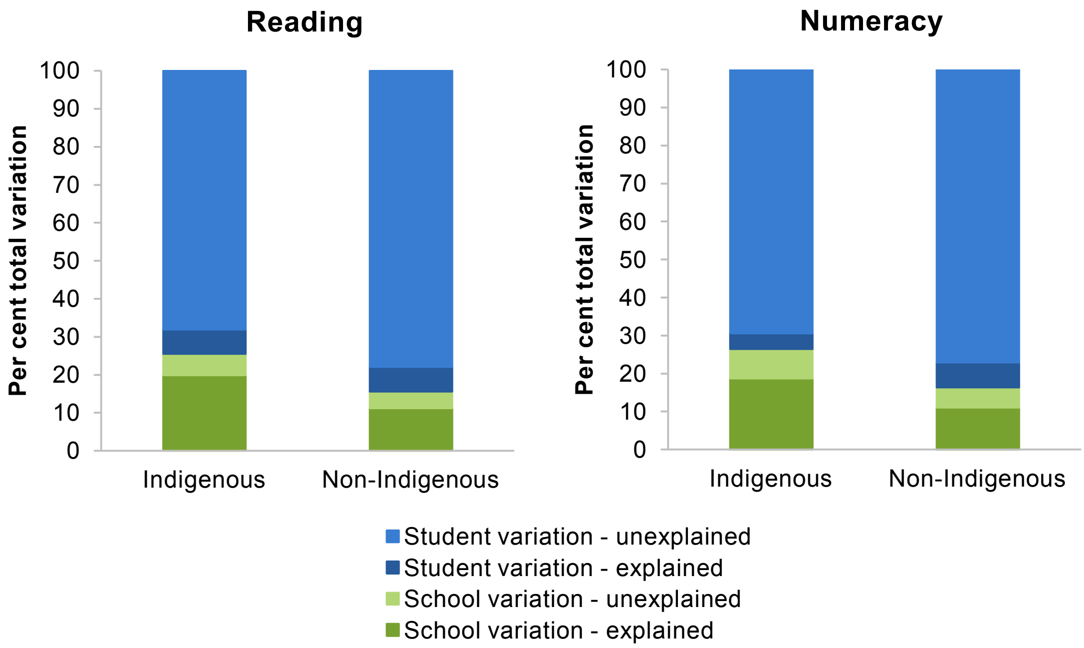
Consistent with past studies, the majority of the variation in students' reading and numeracy scores is attributable to students' characteristics rather than the characteristics of the schools that they attend (figure 2). This suggests that policies aimed at targeting the needs of individual students are likely to be important in improving student achievement. An individualised approach to improving student achievement could involve, for example, teachers assessing what each student knows, adapting their teaching methods to the needs of the individual student, and then evaluating the impact on student achievement (chapter 4).

Also consistent with previous research, a large proportion (about 70 per cent) of the variation attributable to schools is estimated to be explained by characteristics that are observed in the data, while very little (less than 10 per cent) of the variation attributable to student characteristics is estimated to be explained by observed characteristics. This means that most of the variation in achievement is unexplained. There could be a vast range of factors that contribute to the unexplained components of both student-level and school-level variation, not all of which can be identified or observed. Some of the unexplained variation could be due to student characteristics that are not available in the data, for example, cognitive abilities and attitudes. At the school level, the unexplained variation partly captures unobserved characteristics like school culture and teacher quality (to the extent that they are constant across all students at the school), but these explain relatively little of the overall variation in NAPLAN test scores.[6]

[6] Some unobserved characteristics, such as school culture, could be partially captured in the explained variation if they are related to observed characteristics. It is not possible to test whether this is the case.

Figure 2 **School characteristics explain a relatively small proportion of the overall variation in students' test scores**

Reading and numeracy, by Indigenous status (Year 5, 2013 and 2014 pooled)



Source: Commission estimates based on ACARA data (unpublished).

The total percentage of variation explained by observed characteristics, as represented by the R^2 of the multilevel model, is estimated to be 31 per cent for Indigenous students in reading (annex A). Including the variation explained by unobserved school characteristics, this only increases to 36 per cent.

In examining the results for Indigenous and non-Indigenous students separately, a notable difference is that a greater proportion of the total variation in students' test scores is attributable to schools for Indigenous students than for non-Indigenous students — 26 per cent compared with 16 per cent for both reading and numeracy. Therefore, even though student characteristics matter more to achievement than the characteristics of the schools that they attend, school characteristics matter relatively more for Indigenous students than for non-Indigenous students.

That said, a preliminary analysis indicates that the result for Indigenous students is strongly influenced by very remote schools (annex A). When schools in very remote regions are excluded from the analysis, the share of the variation in Indigenous students' test scores that is attributable to the characteristics of the schools that they attend drops to about 17 per cent — very similar to the share for all non-Indigenous students. Among schools in very remote areas it would appear to be considerably higher. Such a difference between levels of remoteness is not observed for non-Indigenous students because few non-Indigenous students attend very remote schools relative to metropolitan or provincial schools. Excluding remote schools from the analysis of non-Indigenous students led to a drop of less than one per cent in the proportion of variation attributable to schools.

The idea that the school-level variation may be larger for Indigenous students in very remote schools is probably not surprising. The analysis suggests that Indigenous students attending very remote schools are more alike than those attending schools in other remoteness areas and that there are larger differences between very remote schools in their effects on students' achievement. This could be for a number of reasons.

One reason is that it is likely that Indigenous students within the same very remote communities have similar observed and unobserved characteristics — they may be from smaller communities in which people have very similar family characteristics, customs and habits, and face the same community resources.

In addition to having more similar social and family influences, another reason is that very remote schools may be smaller, so students attending the same school in a very remote area may be more likely to have been taught by the same teachers. Therefore, the influence of teachers may be more likely to be captured in school-level variation than student-level variation.

Overall, the characteristics of some very remote communities and teachers at very remote schools may be beneficial for student achievement, while others may make it more difficult to perform well at school. This school-level homogeneity may be less likely to be present for students in metropolitan areas, where communities tend to be larger and may encompass a wide range of socioeconomic and family circumstances, customs and cultures. Students attending the same school in metropolitan areas may also have been taught by different teachers, because schools tend to be larger.

6 How do observed characteristics relate to achievement?

The results presented in this section illustrate how specific observed characteristics are related to student achievement. For each characteristic, the analysis examines whether there is a statistically significant relationship with achievement, and if so, it discusses the direction and strength of the relationship in terms of NAPLAN score points. In this paper, the strength of the relationship (or size of the coefficient) between a given characteristic and achievement is interpreted relative to the standard deviation of test scores for the sample (box 2).

Box 2 **Interpreting the size of a coefficient**

Whether a coefficient that describes the relationship between a characteristic and an outcome is considered to be large or small is subjective. The interpretation is guided by the context of the analysis.

In this study, a large coefficient between a characteristic and achievement, in terms of test scores, is considered to be the equivalent of about one quarter of a standard deviation. Assuming a normally distributed sample, an increase in test score of a quarter of a standard deviation for a student who was at the 50th percentile of scores would bring them up 10 percentage points to the 60th percentile.

A small coefficient is considered to be less than one tenth of a standard deviation. Again assuming a normally distributed sample, an increase in test score of one tenth of a standard deviation would bring a student who was at the 50th percentile of scores up to the 54th percentile.

The following table shows the size of small and large coefficients in terms of score points for Year 5 Indigenous and non-Indigenous students in reading and numeracy, based on the standard deviation of test scores for students included in the sample.

	Indigenous		*Non-Indigenous*	
	Small	Large	Small	Large
Reading	8.0	20.0	7.5	18.8
Numeracy	6.9	17.3	7.5	18.7

How do Indigenous students perform relative to non-Indigenous students?

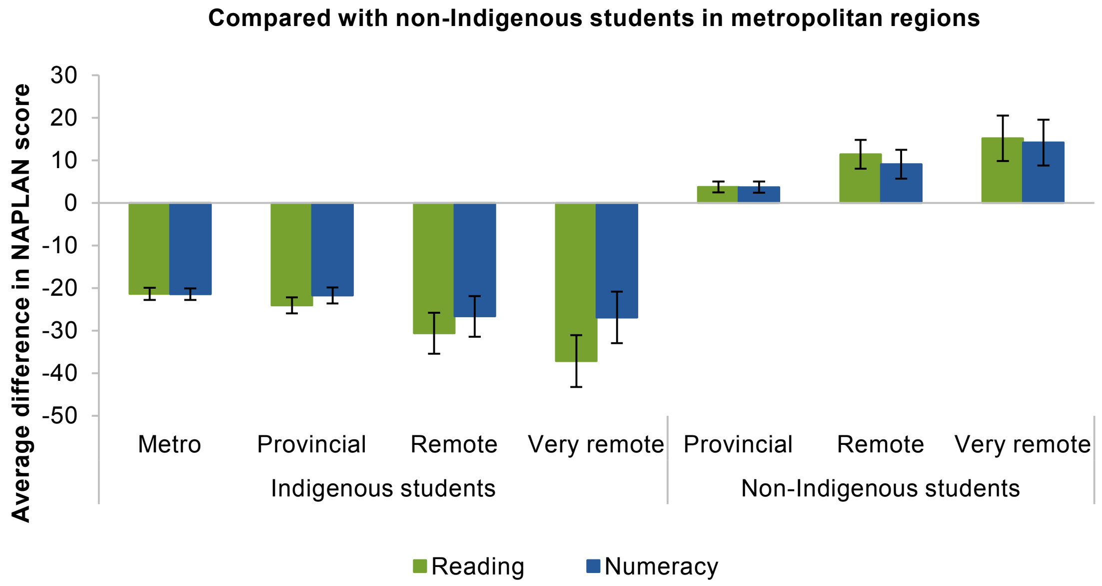
Data for Indigenous and non-Indigenous students were pooled to examine how Indigeneity is associated with NAPLAN test scores. These regression analyses (using multilevel models) took into account other student characteristics as well as a range of school characteristics, described in table 1.

In the analysis using Year 5 data, Indigenous students in a metropolitan area and with an English-speaking background score an average of 21 points less on the NAPLAN reading and numeracy tests than their non-Indigenous peers, other observed characteristics held constant (figure 3). That is, a statistically significant difference in achievement between Indigenous and non-Indigenous students remains after other characteristics like parental education and occupation are taken into account. This finding reflects unobserved characteristics related to being Indigenous and could include characteristics such as discrimination, expectations and attitudes towards education.

Indigenous students also achieve lower scores than non-Indigenous students in metropolitan areas the more remote they are (figure 3). Indigenous students in very remote areas are expected to score 37 points less than non-Indigenous students in metropolitan areas. In contrast, the results suggest that non-Indigenous students in more remote areas perform better than those in metropolitan areas, after other characteristics are taken into account. Further discussion on this point is presented below.

Figure 3 **Indigenous students in more remote areas perform less well, relative to their non-Indigenous peers[a,b,c]**

Reading and numeracy (Year 5, 2013 and 2014 pooled)



[a] Regression coefficients on Indigenous status and on Indigenous × remoteness interaction terms have been summed together where relevant to produce these estimates. [b] Vertical lines represent 95 per cent confidence intervals (box 1). [c] Relationships for categorical variables should be interpreted relative to the default category of non-Indigenous students in metropolitan regions (box 1).

Source: Commission estimates based on ACARA data (unpublished).

How are student characteristics related to achievement for Indigenous students and non-Indigenous students?

For this part of the analysis, regressions were performed for Indigenous and non-Indigenous students separately to see how observed characteristics relate to NAPLAN scores for each group. In many cases, the interpretations of results refer to studies cited in table 1.

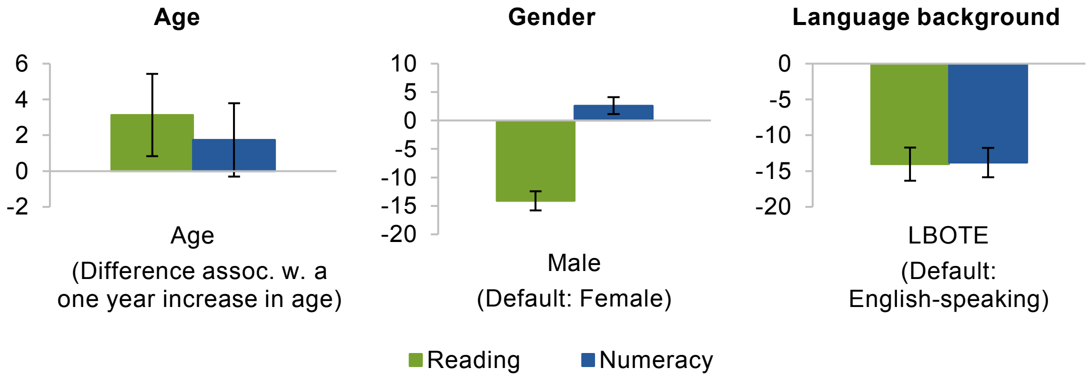
With respect to student demographic characteristics, age, gender and language background are all found to be related to NAPLAN reading and numeracy scores for both Indigenous and non-Indigenous students (figures 4 and 5). Results are generally similar for reading and numeracy, but there are a few areas of difference in the influences of gender and language background.

Both Indigenous and non-Indigenous boys tend to achieve lower average scores than girls in reading and higher scores than girls in numeracy. This may be because girls experience more anxiety over mathematical subjects (table 1).

Indigenous and non-Indigenous students with a LBOTE both achieve lower scores in reading, compared with students with an English language background. Indigenous students with a LBOTE also perform less well in numeracy, but non-Indigenous students with a LBOTE perform better than those with an English-speaking background, on average. It may be that non-Indigenous students with a LBOTE predominantly come from backgrounds with high parental expectations (table 1), and that difficulties associated with achieving high scores in subjects that rely on a good understanding of the English language, such as reading, are not as salient in numeracy.

Figure 4 **Indigenous students' demographic characteristics are related to achievement[a,b]**

Average difference in NAPLAN score, reading and numeracy (Indigenous, Year 5, 2013 and 2014 pooled)



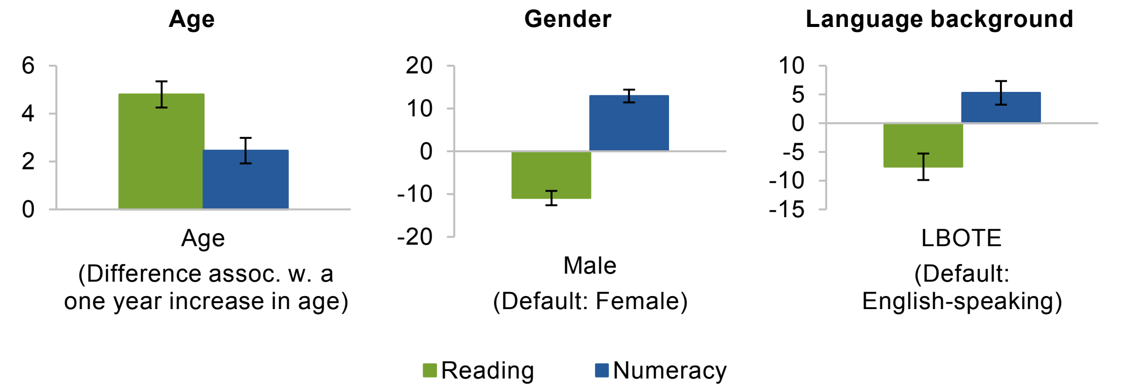
[a] Vertical lines represent 95 per cent confidence intervals (box 1). [b] Relationships for categorical variables should be interpreted relative to the default category (box 1).

Source: Commission estimates based on ACARA data (unpublished).

Student SES, as reflected by parental education and occupation, is also significantly related to NAPLAN scores. As estimated relationships tend to be stronger for mother's education than father's education and for father's occupation than mother's occupation, these characteristics were chosen to be depicted in figures 6 and 7.

Figure 5 **Non-Indigenous students' demographic characteristics are also related to achievement[a,b]**

Average difference in NAPLAN score, reading and numeracy (non-Indigenous, Year 5, 2013 and 2014 pooled)



[a] Vertical lines represent 95 per cent confidence intervals (box 1). [b] Relationships for categorical variables should be interpreted relative to the default category (box 1).

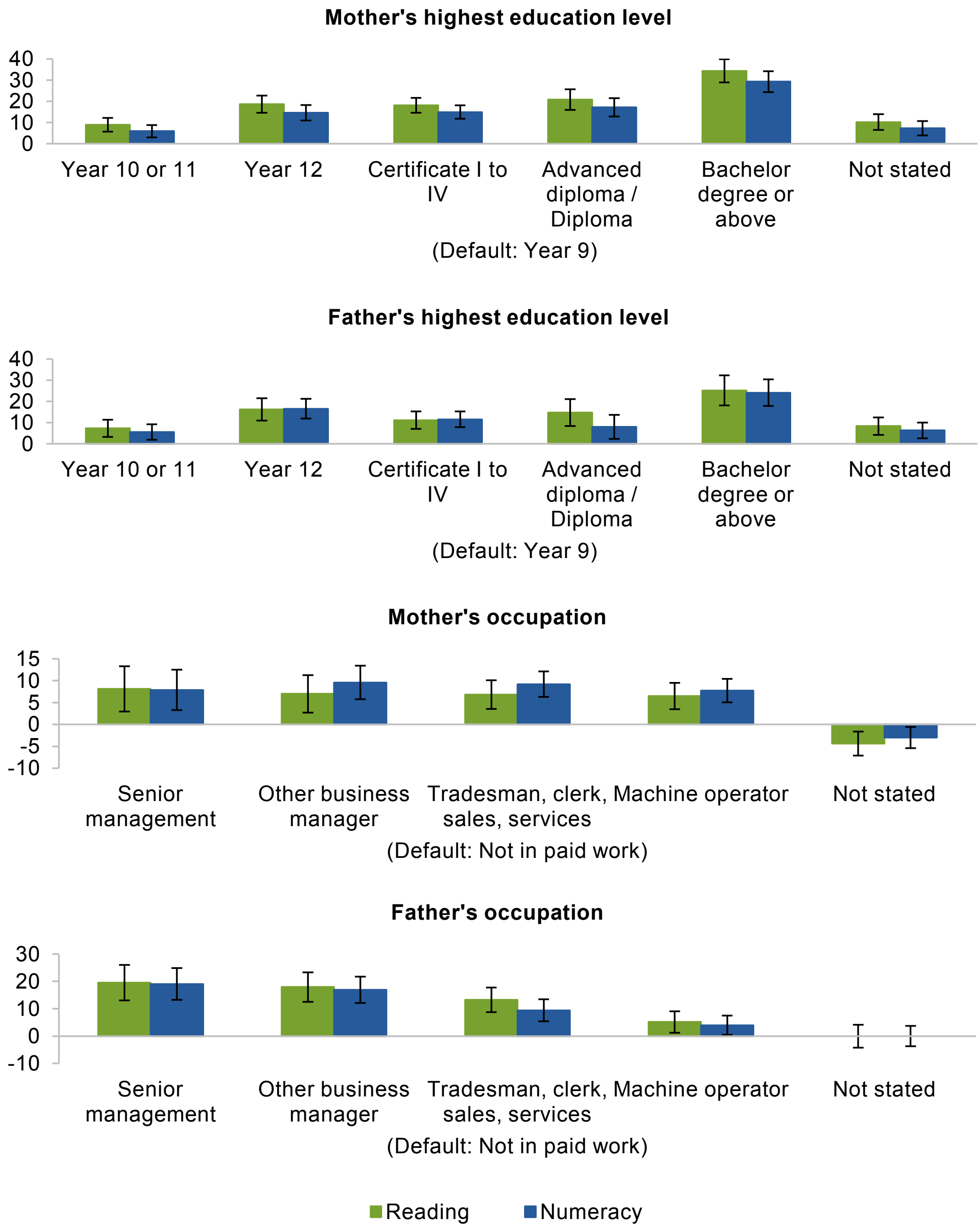
Source: Commission estimates based on ACARA data (unpublished).

In general, students perform better in reading and numeracy the higher their parents' level of education. For example, on average, Indigenous and non-Indigenous children with a mother whose highest level of education is Year 12 have reading test scores 19 points higher than children with a mother whose highest level of education is Year 9 or below, when other characteristics are taken into account. These results could, for example, reflect parental attitudes towards education (table 1). They might also reflect a parents' ability to help their child with their study, both directly (for example, through help with homework) and indirectly (for example, through reading to children).

Children whose parents are employed also tend to perform better than students with parents who are not in paid work, and the estimated relationships between parental occupation and achievement are larger for more highly skilled occupations. These results might reflect the impact of employment on the availability of resources that are relevant to education in a child's home (for example, books and good nutrition) (table 1). Parents in more highly skilled occupations might earn more and be better able to financially provide for their children's education. They may also act as positive role models and encourage their children to perform better (table 1). The observation that estimated relationships tend to be larger for fathers' occupations than mothers' occupations could reflect the fact that some mothers who are not in paid work have skills that would enable them to acquire paid work, but choose to stay at home to care for their children.

Figure 6 **Indigenous students achieve higher scores on average if their parents are more highly educated and employed[a,b]**

Average difference in NAPLAN score, reading and numeracy (Indigenous, Year 5, 2013 and 2014 pooled)



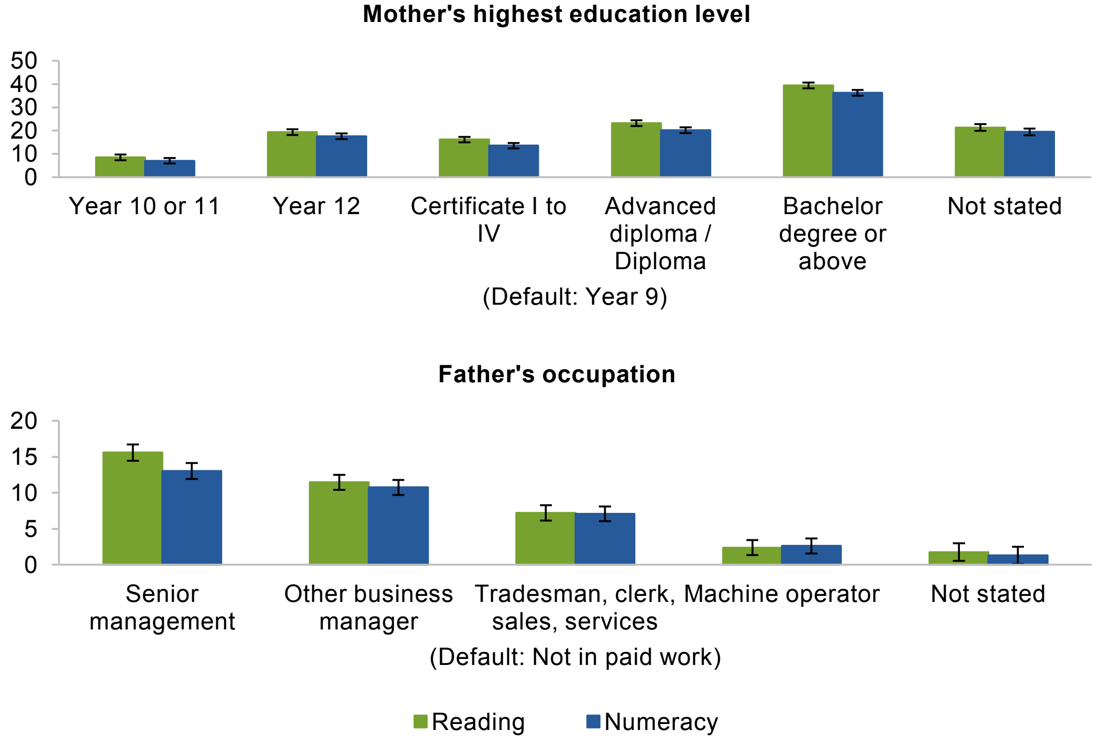
[a] Vertical lines represent 95 per cent confidence intervals (box 1). [b] Relationships for categorical variables should be interpreted relative to the default category (box 1).

Source: Commission estimates based on ACARA data (unpublished).

The analysis suggests that Year 5 students who attended the same school in Year 3 perform better than those who changed schools, but the size of the association is small — for reading, there is about an 8-point difference in scores for Indigenous students, and a 6-point difference for non-Indigenous students. This suggests that student mobility, or characteristics associated with mobility, can have a negative but small association with student achievement. However, the measure of student mobility used was a simple indicator of whether the student attended the same school in Years 3 and 5, and the number of school changes was not observed. More sophisticated measures of student mobility that can capture the frequency of moves may find that mobility has larger influences on achievement. More detailed data on student mobility exist in state records but were not available for this study. Further research in this area would be worth pursuing.

Figure 7 **Non-Indigenous students also perform better on average if their parents are more highly educated and employed[a,b]**

Average difference in NAPLAN score, reading and numeracy (non-Indigenous, Year 5, 2013 and 2014 pooled)



[a] Vertical lines represent 95 per cent confidence intervals (box 1). [b] Relationships for categorical variables should be interpreted relative to the default category (box 1).

Source: Commission estimates based on ACARA data (unpublished).

How are school characteristics related to achievement for Indigenous students and non-Indigenous students?

A range of school characteristics and their relationships with NAPLAN scores were also examined. The relationships for a number of characteristics of interest — some peer and school-related characteristics (percentage of Indigenous students, percentage of students with a LBOTE, school attendance rate) and broader social characteristics (state and remoteness) — are presented in figure 8 for Indigenous students and figure 9 for non-Indigenous students.

Other school characteristics were also included in the analysis, some of which are related to achievement. For example, Indigenous and non-Indigenous students tend to perform slightly better in reading in Independent schools than in government schools, but there is no significant difference in numeracy. Even though there are some significant relationships, school sector and other characteristics did not make a large contribution to explaining the overall variation in student achievement. Coefficients for these other characteristics are presented in annex A and are not discussed in detail here.

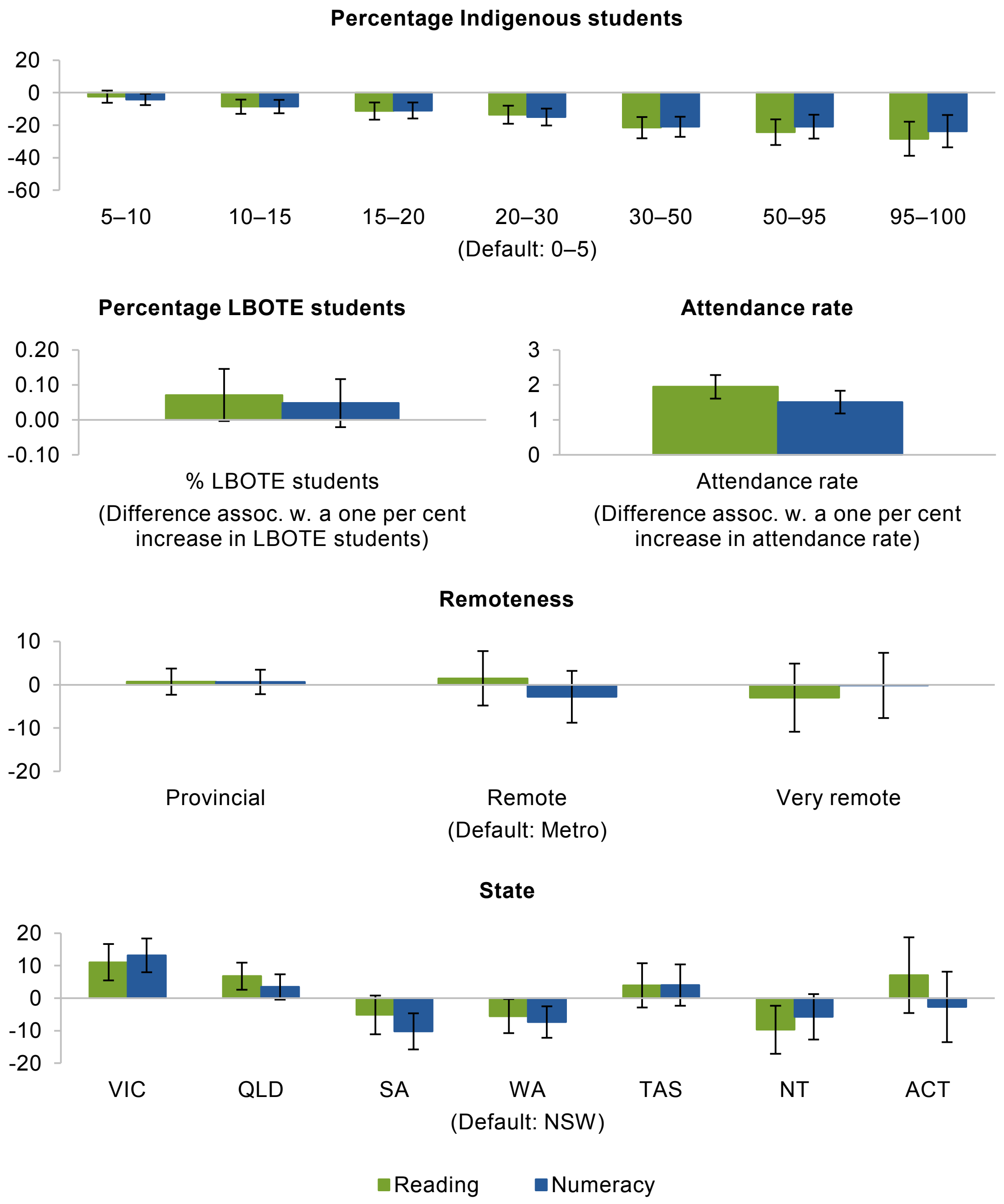
The results suggest that the percentage of Indigenous students at a school is negatively associated with NAPLAN reading and numeracy scores for both Indigenous and non-Indigenous students. Indigenous students at schools where 95 to 100 per cent of students are Indigenous are expected to score 28 points less in reading compared with Indigenous students at schools with up to 5 per cent Indigenous students, controlling for other observed characteristics. Non-Indigenous students at schools with more than 95 per cent Indigenous students are estimated to score, on average, 77 points less than non-Indigenous students at schools with up to 5 per cent Indigenous students.

Overall, this suggests that there may be wider disadvantage associated with the proportion of Indigenous students in a school that influences achievement for both Indigenous and non-Indigenous students, in addition to the influence of Indigenous background at a student level. There may also be other factors at play. For example, if teachers tend to have lower expectations of Indigenous students (consciously or subconsciously), there may be a greater culture of low expectations for students the larger the proportion of Indigenous students at a school (table 1). These results also suggest that if teachers do tend to adapt their teaching styles to suit the learning styles of the majority of their students (table 1), it does not appear to be enough to offset the overall negative relationship between the proportion of Indigenous students at a school and achievement on average.

The percentage of students with a LBOTE is not significantly related to NAPLAN scores for both Indigenous and non-Indigenous students after controlling for other characteristics. Therefore, even though having a LBOTE tends to be negatively related to scores at the student level, it does not have a significant effect at the school level. That is, there are no peer effects associated with attending a school where a large proportion of students have a LBOTE.

Figure 8 Indigenous students' achievement is related to the characteristics of their school[a,b]

Average difference in NAPLAN score, reading and numeracy (Indigenous, Year 5, 2013 and 2014 pooled)

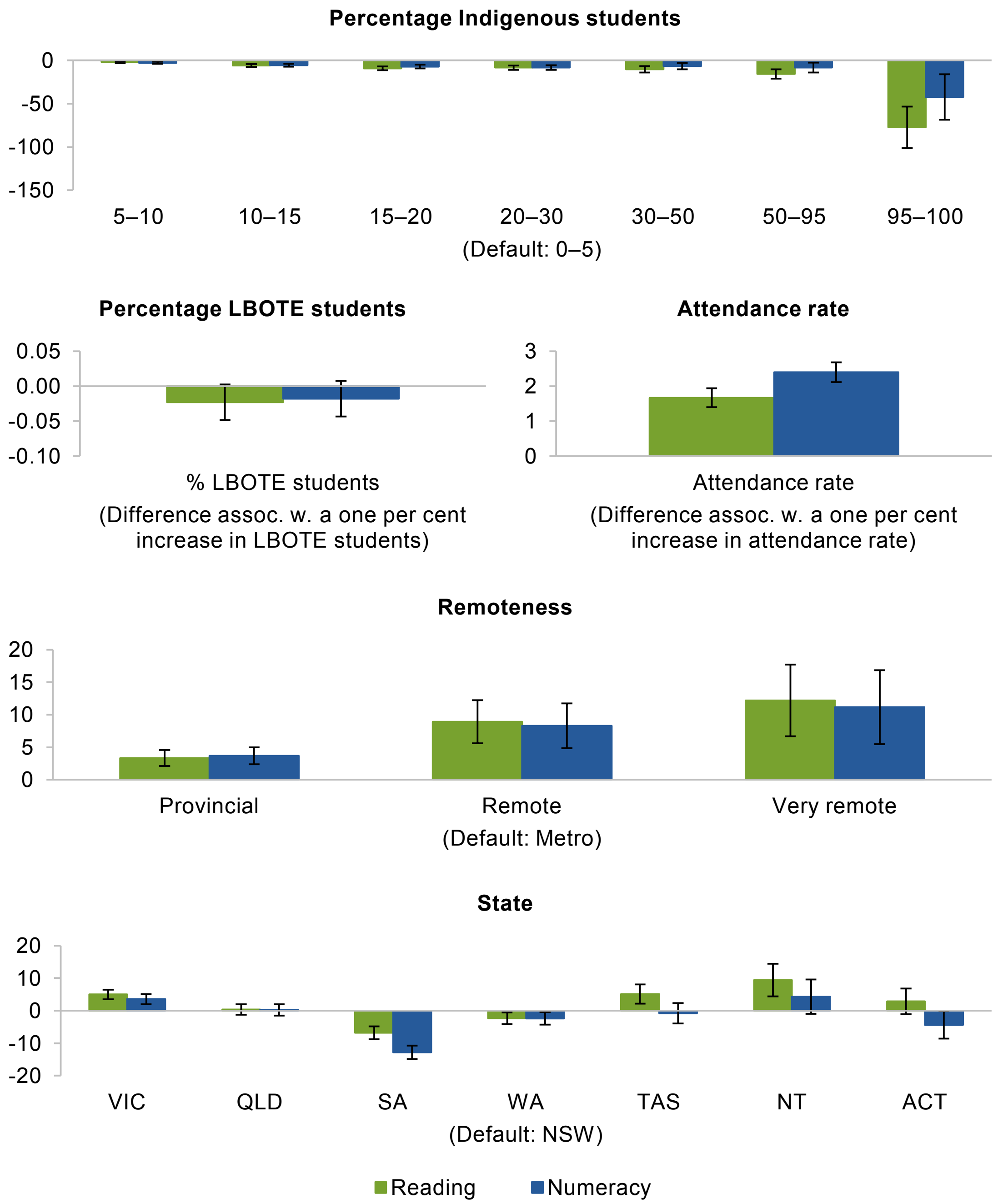


[a] Vertical lines represent 95 per cent confidence intervals (box 1). [b] Relationships for categorical variables should be interpreted relative to the default category (box 1).

Source: Commission estimates based on ACARA data (unpublished).

Figure 9 **Non-Indigenous students' achievement is also related to the characteristics of their school[a,b]**

Average difference in NAPLAN score, reading and numeracy (non-Indigenous, Year 5, 2013 and 2014 pooled)



[a] Vertical lines represent 95 per cent confidence intervals (box 1). [b] Relationships for categorical variables should be interpreted relative to the default category (box 1).

Source: Commission estimates based on ACARA data (unpublished).

The attendance rate for a school is positively related to NAPLAN reading and numeracy for Indigenous and non-Indigenous students — a 1 per cent increase in the attendance rate of a school is associated with about a 2-point increase in score on average. In other words, students in a school with an attendance rate of 95 per cent are expected to have a test score that is about 10 points higher, other characteristics equal, than those in schools with an attendance rate of 90 per cent. It may be, for example, that higher attendance rates enable teachers to teach more effectively, or that having better teachers, or some other unobserved characteristics of the school, induces a higher attendance rate (table 1). It may also be that the school attendance rate is a proxy for attendance at a student level — it is likely to be difficult for students to learn as well if they miss too much school.

As for broader community characteristics, school remoteness is not significant in explaining Indigenous student achievement after other characteristics are taken into account, despite there being large differences in the medians of NAPLAN scores by remoteness in a simple analysis (BP 1). This suggests that it is mainly observed characteristics that are associated with remoteness, such as the percentage of Indigenous students and school attendance rates, that explain variation in educational achievement, rather than other aspects of remoteness. Although preliminary results described in section 5 suggested that there was much more school-level variation in scores for Indigenous students in very remote schools, these results suggest that the average change in NAPLAN score associated with attending a very remote school is no different to attending a less remote school after other characteristics have been controlled for, even if there is greater variation around the average.

The results suggest that remoteness is associated with achievement for non-Indigenous students but positively rather than negatively. That is, non-Indigenous students in more remote areas are expected to perform better than those in less remote areas, after other observed characteristics are taken into account. It is not clear why this might be the case. One possibility is that non-Indigenous students in more remote areas could be more likely to have unobserved characteristics that have a positive impact on achievement. Despite the positive relationship, it is not very strong, according to the interpretations in box 2. It is also important to keep in mind that many very remote schools have observed characteristics that are associated with lower test scores, such as a high proportion of Indigenous students and relatively low attendance rates. Non-Indigenous students in very remote schools may not necessarily perform better than those in metropolitan schools in absolute terms, but they may perform better on average given these other characteristics.

Although remoteness is not significant for Indigenous students, reading and numeracy scores do differ depending on state. Indigenous students in Victoria and Queensland perform better on average in reading than those in New South Wales, while those in Western Australia and the Northern Territory perform less well. In numeracy, Indigenous students in Victoria achieve higher scores than those in New South Wales, and those in South Australia and Western Australia achieve lower scores. That said, although these relationships are significant, they are not particularly large. Students in other states do not perform significantly better or worse than those in New South Wales.

Some of these associations appear to differ for non-Indigenous students. For example, non-Indigenous students in Victoria, Tasmania and the Northern Territory achieve higher scores in reading compared with New South Wales, whereas those in South Australia and Western Australia achieve lower scores. In numeracy, the directions of these relationships are the same, except non-Indigenous students in Tasmania and Northern Territory do not perform better or worse than those in New South Wales.

7 Which observed characteristics are most important in explaining achievement?

While the previous section examined characteristics that are significantly related to achievement, that analysis does not shed light on how important those characteristics are to explaining achievement — that is, which characteristics make the largest contribution to explaining the variation in achievement. A strong average relationship does not necessarily mean that a characteristic makes a large contribution. This is because the characteristic may only represent a small proportion of students and thus only explains a small share of total variation. Furthermore, if the achievement of students with a characteristic varies a lot, the average relationship will not be a good description for many students with that characteristic, and therefore the characteristic will not explain a large proportion of total variation.

This section discusses the results of the dominance analysis to see which observed characteristics are most important for Indigenous and non-Indigenous students. Dominance analysis involves estimating the contribution of a particular characteristic to explained variation, by examining the change in explained variation that occurs when the characteristic is added to the model. Because of relationships between characteristics, the estimated contribution of a characteristic depends on which other characteristics have already been included in the model. In order to obtain an overall estimate of the contribution of a characteristic, multiple estimates of the contribution are generated with varying combinations of characteristics included in the model, and an average is taken.

This procedure is computationally demanding, so characteristics were grouped into ten sets to reduce the computational load. For example, parental education and occupation characteristics were grouped into a set relating to 'student SES'. Similarly, characteristics relating to the socioeconomic compositions of school communities were grouped into a set relating to 'school SES'.

Overall, as noted above, more of the variation in NAPLAN scores is explained by observed characteristics for Indigenous students than for non-Indigenous students — in reading, about 31 per cent of variation is explained for Indigenous students, compared with about 18 per cent for non-Indigenous students. However, as these R^2 estimates show, most of the variation still remains unexplained.

Of the sets of observed characteristics, the most important for both student groups is student SES (figure 10). However, there are differences between Indigenous and non-Indigenous students in the importance of SES relative to other observed characteristics — students' SES only accounts for about 5 per cent of total variation in reading scores for Indigenous students, but 9 per cent of total variation for non-Indigenous students. It is noted that estimates of the contribution of student SES that are generated when school SES is not taken into account incorporate the contributions of both sets of characteristics. Therefore, the average estimated contribution of student SES is overstated — it incorporates some of the contribution of school SES.

School SES also represents a large proportion of total variation for non-Indigenous students, ranking second with a contribution of 5 per cent. This means that SES-related characteristics account for nearly three-quarters of the variation in reading scores that is explained by observed characteristics for non-Indigenous students. This is consistent with other studies that find SES to be an important characteristic for students generally.

Although SES is important for Indigenous students, other characteristics also represent relatively large shares of explained variation. For Indigenous students, the attendance rate of a school is the second most important characteristic, contributing about 5 per cent of the total variation in reading scores. This suggests that initiatives that aim to improve attendance among Indigenous students could have a large effect on achievement, relative to initiatives that lead to changes in some of the other observed characteristics.

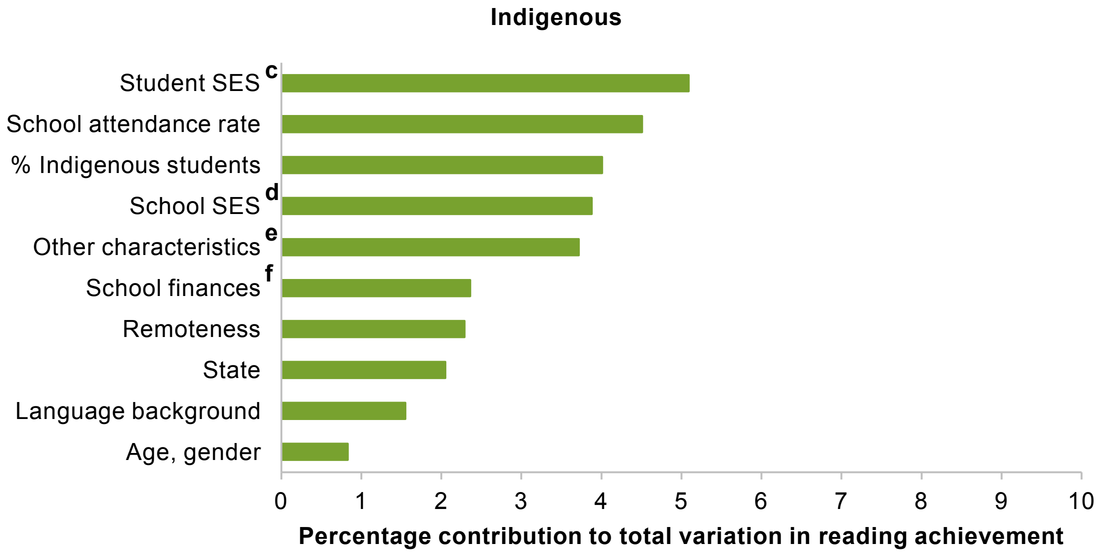
The result also points to the potential value to researchers of access to data on attendance at a student level. Given school attendance rates explain a relatively large share of the variation in Indigenous students' reading scores, this raises the question of whether the result reflects the influence of the attendance of individual students, or the attendance rate at the school level (table 1). These have different policy implications, as the former suggests that action to encourage individual children to attend school will have an influence, while the latter points to a different set of questions around why the school cannot achieve higher attendance rates. For example, if schools tend to have low attendance rates because teachers have difficulties engaging their students, then changes in teaching methods and curricula may help to improve attendance.

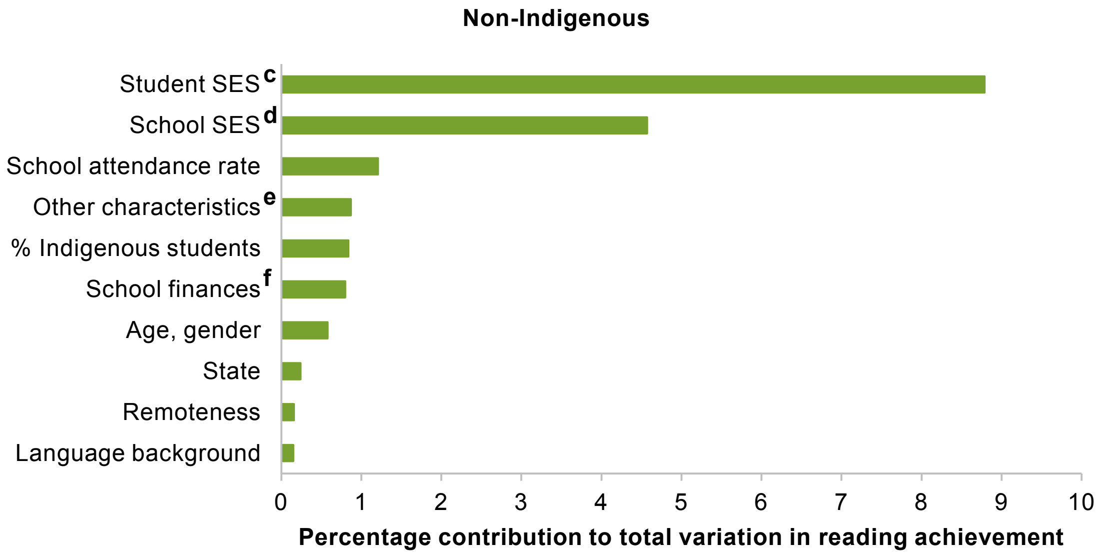
Attendance is closely followed by the percentage of Indigenous students at a school and school SES, which each contribute about 4 per cent to total variation, reflecting the importance of concentrated social disadvantage to educational achievement. These characteristics could have important impacts on achievement through teacher expectations, for example (table 1).

Attendance rates and the percentage of Indigenous students are also relatively important for non-indigenous students, ranking third and fifth respectively in the list of sets of characteristics, but they are much less important when compared with the large contribution of SES.

Figure 10 Socioeconomic status and attendance are important characteristics in explaining reading achievement[a,b]

Reading by Indigenous status (Year 5, 2013 and 2014 pooled)





[a] Results presented are general dominance statistics, which reflect the average contribution of a set of observed characteristics to the model's explained variation. [b] The relative importance of calendar year was not examined, but was included in all dominance analysis regressions. Calendar year explained 1 per cent or less of the variation in NAPLAN scores. [c] Student socioeconomic status: mother's and father's highest education level, mother's and father's occupation. [d] School socioeconomic status: percentage of mothers and fathers by highest education level, percentage of mothers and fathers by occupation, school fees and parent contributions per student (standardised by school sector) interacted with school sector. [e] Other characteristics: school sector, combined school indicator, average class size, non-teaching staff per student, number of enrolments, percentage LBOTE students, test participation rate, student mobility indicator. [f] School finances per student: recurrent funding (less school fees), capital income deductions, capital expenditure.

Source: Commission estimates based on ACARA data (unpublished).

The measures of broad social characteristics included in the analysis (state and remoteness) are comparatively unimportant in explaining variation in achievement, but are relatively more important for Indigenous students than non-Indigenous students. State and remoteness are estimated to contribute 2 per cent each to the total variation in reading scores for Indigenous students, and less than 1 per cent in total for non-Indigenous students. It is likely that SES-related characteristics capture much of the disadvantage associated with remoteness. The relatively small contribution of states suggests that differences in state curriculums are not very important. The larger contribution of state and remoteness to explained variation in the scores for Indigenous students compared with non-Indigenous students could reflect greater cultural differences between Indigenous communities in different locations, or differences in Indigenous-specific interventions between different states and remoteness areas.

The other sets of characteristics examined, including age and gender, language background, school finances and factors that individually explained relatively small proportions of variation, are not as important as SES, attendance and percentage Indigenous students in explaining reading achievement. These overall conclusions are generally similar for NAPLAN numeracy achievement.

8 Is the achievement gap mostly explained by differences in characteristics or differences in relationships?

The difference in mean NAPLAN scores between Indigenous and non-Indigenous students was analysed by performing Blinder-Oaxaca decompositions. As mentioned in section 2, this technique involves splitting the gap in mean scores into three components.

One part of the gap is attributed to differences in the average school and student characteristics observed for each group of students. This is known as the endowment effect. For example, if parental education is a characteristic included in the model, the technique identifies how much of the gap in mean NAPLAN scores is due to the difference between Indigenous and non-Indigenous students in their average levels of parental education. This is done for each observed characteristic analysed and the portions of the gap explained by differences in each characteristic are summed together. This sum is interpreted as the total portion of the gap that is explained by differences in average characteristics between Indigenous and non-Indigenous students.

Another portion of the gap is attributed to differences in the relationships between observed characteristics and achievement — the coefficient effect. This component takes into account the fact that the relationship between a particular characteristic (such as parental education) and achievement could be larger or smaller for Indigenous students than for non-Indigenous students on average. As an example, Indigenous students with a parent who has completed Year 12 might perform 10 points better on average than

Indigenous students with a parent who has completed Year 9, while for non-Indigenous students the relationship might be 20 points. These differences in relationships between each characteristic and achievement may contribute to the gap in average scores between Indigenous and non-Indigenous students. Differences in unobserved characteristics between Indigenous and non-Indigenous students are also captured in the coefficient effect.

A remaining part of the gap in average test scores is attributed to differences in characteristics and differences in relationships that can occur simultaneously — the interaction term. This term is an artefact of the decomposition method and cannot be interpreted intuitively.

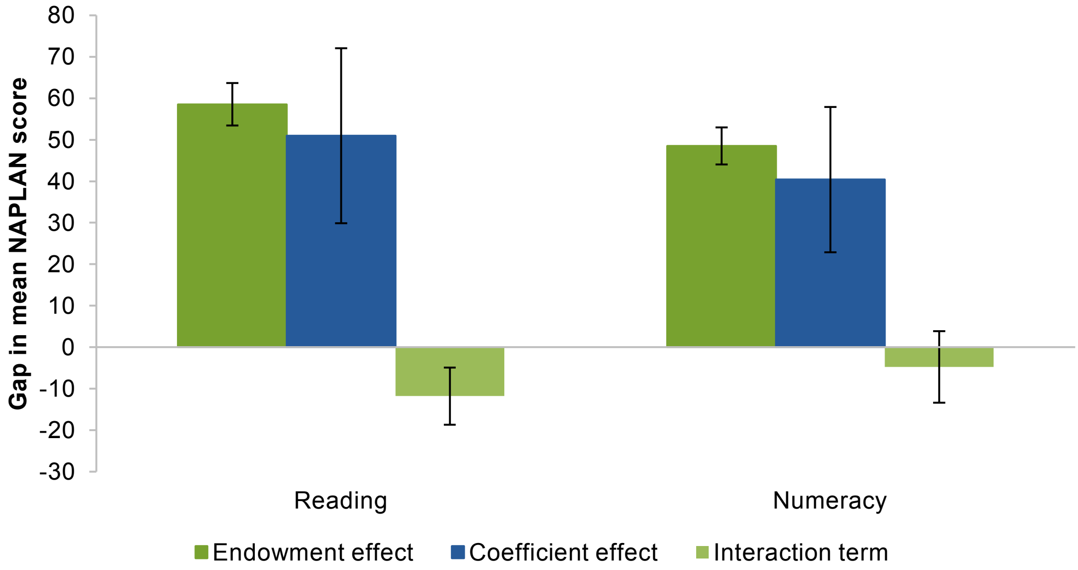
This section presents the results of decompositions for Year 5 students in 2014.

The total gap in mean reading scores between Indigenous and non-Indigenous students was 98 points. The results suggest that 59 points is due solely to differences in school and student characteristics between the two groups (the endowment effect) (figure 11). This indicates that, if the average Indigenous student had the same characteristics as the average non-Indigenous student, the gap in mean reading scores would be expected to more than halve. About 51 points of the gap is due solely to differences in the relationships between those characteristics and reading scores (the coefficient effect). This suggests that if Indigenous students faced the same relationships between characteristics and reading scores as non-Indigenous students, the gap would be expected to reduce by about half. Similar conclusions arise from the analysis of numeracy scores.

The coefficient effect results indicate that, to fully reduce the gap between Indigenous and non-Indigenous students, policy makers must address the reasons for the differences in relationships between characteristics and achievement. However, as mentioned in section 6, the relationships between particular observed characteristics and achievement are fairly similar between both Indigenous and non-Indigenous students. As a result, a large proportion of the coefficient effect may be due to the effect of unobserved characteristics (such as discrimination) that are related to being Indigenous.

Figure 11 **Differences in characteristics explain more of the average gap in achievement than differences in the relationships between characteristics and scores[a]**

Reading and numeracy (Year 5, 2014)



[a] Vertical lines represent 95 per cent confidence intervals (box 1).

Source: Commission estimates based on ACARA data (unpublished).

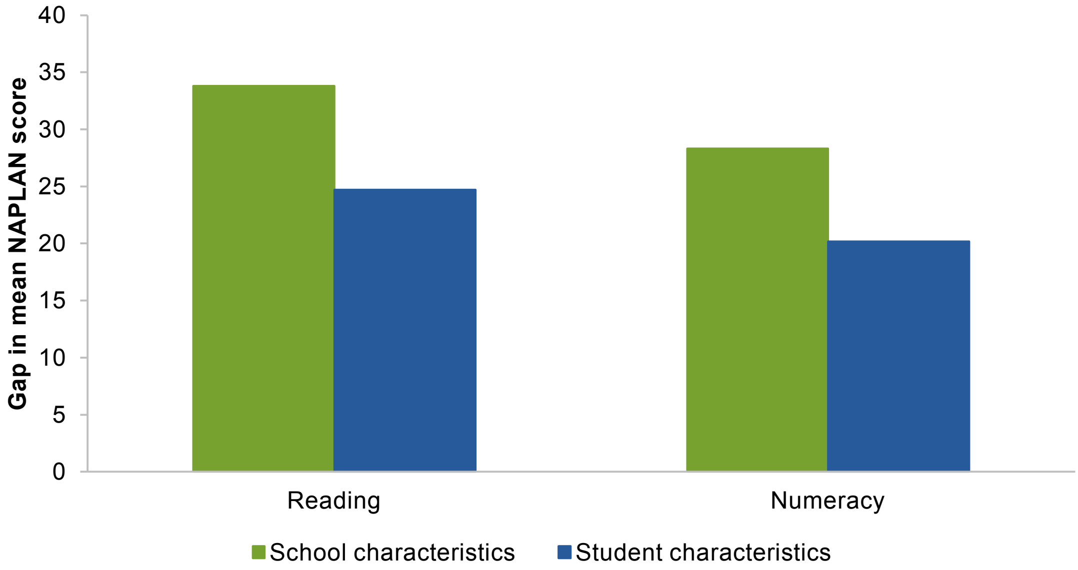
The interaction term, which accounts for differences in endowments and coefficients that occur at the same time, is relatively small for reading scores and is not statistically different from zero in the analysis of numeracy scores. As mentioned, this term cannot be interpreted intuitively.

How much of the gap is explained by school and student characteristics?

The endowment effect can be further separated into the effects of each observed characteristic. Overall, differences in school characteristics account for a greater share of the effect (about 58 per cent) than student characteristics (about 42 per cent) for both reading and numeracy tests (figure 12).

Figure 12 **Differences in school characteristics explain more of the average gap in achievement than differences in student characteristics**

Reading and numeracy (Year 5, 2014)



Source: Commission estimates based on ACARA data (unpublished).

The two characteristics that represent the largest proportions of the gap explained by observed school characteristics are the percentage of Indigenous students and the school attendance rate. These each account for about a third of the total endowment effect attributable to school characteristics. The results suggest that if Indigenous students were distributed in the same way as non-Indigenous students across schools according to the percentage of Indigenous students and school attendance rates, the gap in mean reading scores would be expected to close by 24 points — a quarter of the total gap in mean scores. While this scenario is likely infeasible in practice, it re-iterates findings from the previous section about the importance of attendance and the percentage of Indigenous students in a school. The results illustrate the potential value to education achievement of improving attendance and addressing characteristics that are likely related to the percentage of Indigenous students at a school (such as Indigenous disadvantage, expectations and attitudes towards education).

Among student characteristics, differences between Indigenous and non-Indigenous students in the highest level of education of mothers account for over a third of the endowment effect attributable to student characteristics for reading scores. Mother's education is followed closely by differences in father's occupation. Together, this suggests that having positive parent role models and breaking the cycle of educational disadvantage is important to closing the gap in average achievement levels for future generations of Indigenous children.

In theory, the coefficient effect can also be separated according to specific characteristics. In practice, the coefficient effect for each characteristic is highly dependent on how the characteristics are scaled or measured, though the overall estimate of the coefficient effect remains the same.[7] Given this sensitivity to scaling, detailed results of the coefficient effect are not discussed in this paper.

9 How well do schools perform relative to their predicted performance?

Given the observed characteristics of schools and students, students at some schools perform better than expected while students at other schools perform worse than expected. As noted above, researchers interpret ‘school effects’ as reflecting how much better or worse schools perform relative to their predicted performance. A school effect equal to zero indicates that the school is performing exactly as predicted by their observed characteristics. The school effect is not explained by observed characteristics and is at least partly attributed to a school’s unobserved characteristics. This could include important elements such as school culture and teacher quality, to the extent that they are constant within a school. School effects could also reflect social and community influences that are common across all students attending the same school, not just the characteristics of the school itself.

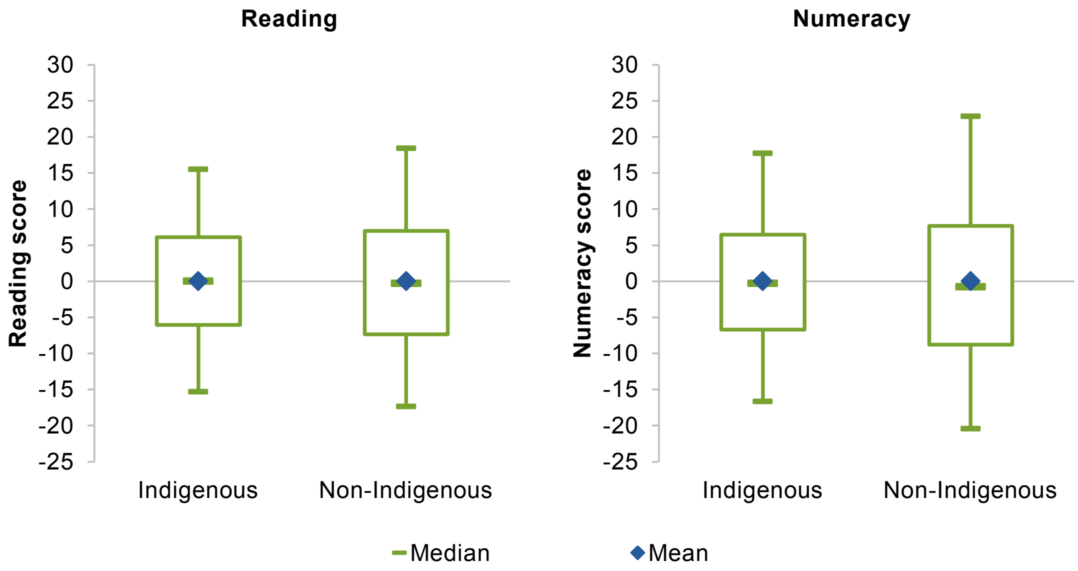
The spread of the distributions of school effects appear fairly similar across the schools with Indigenous students and the schools with non-Indigenous students that are examined in the analysis (figure 13).[8] Among schools with Indigenous students, attending a school that performs in the top five per cent in reading scores is associated with an increase in students’ scores by about 16 points or more. For non-Indigenous students, attending a school that performs in the top five per cent is associated with a similar increase of about 18 points or more. For the top five per cent of schools in numeracy, school effects range from at least 18 points and 23 points for schools with Indigenous students and schools with non-Indigenous students respectively.

7 The Blinder-Oaxaca decomposition analysis leads to characteristics that have large averages to have large coefficient effects. Transforming the scale of the characteristic (for example, by centring the variable to have a mean of zero) results in estimates that are more similar in size, but the direction of the effect becomes less intuitive (annex A).

8 As mentioned in annex A, a brief simulation study suggested that the inclusion of more interaction terms between factors in the modelled equations may result in lower variance in the estimated school effects.

Figure 13 **Some schools perform better or worse than expected, given their observed characteristics[a,b]**

Box and whisker plot of school effects, reading and numeracy by Indigenous status (Year 5, 2013 and 2014 pooled)



a The edges of the boxes represent school effects between the 25th and 75th percentiles. The 'whiskers' extending out from the boxes show the values for the 5th and 95th percentiles. b Estimated school effects are adjusted by the level of uncertainty in the estimates (annex A).

Source: Commission estimates based on ACARA data (unpublished).

Overall, these school effects are reasonably large (according to the interpretation of effect sizes in box 2). This suggests that unobserved school characteristics do play a role in student achievement. However, the reason why some schools outperform others after their observed characteristics are taken into account cannot be identified in this analysis.

The study also examined the consistency of schools identified in the top and bottom five per cent of the distribution for Indigenous and non-Indigenous students. For reading, only about 17 per cent of the top five per cent for non-Indigenous students were in the top five per cent for Indigenous students. For the bottom five per cent, the degree of overlap was 12 per cent. (Further details are provided in annex A.) The reason for the lack of consistency across Indigenous and non-Indigenous students could partly be because unobserved school characteristics that contribute to high achievement for non-Indigenous students may not necessarily work the same for Indigenous students, and vice versa. This highlights the importance of tailoring instruction to the needs of individual students (chapter 4).

What are the observed characteristics of achievement outlier schools for Indigenous students?

The following analysis defines achievement outlier schools as those that perform much better or much worse than expected, given their observed characteristics. Schools identified as achievement outliers can differ between analyses. The results indicate that some schools that perform within the top five per cent for a particular test domain in one year level are not necessarily within the top five per cent for that test domain in another year level. This is partly because of the relatively narrow spread of the distribution of school effects — a few points of difference between analyses across year levels could result in a school being within the top five per cent in one analysis but out of the top five per cent in another analysis. As a result, three categories of high-achieving outlier schools and low-achieving outlier schools are identified based on how consistently they perform across analyses.

- Category A high-achieving (low-achieving) schools perform within the top (bottom) five per cent in *both* reading and numeracy tests for both Year 3 and Year 5 students.
- Category B high-achieving (low-achieving) schools perform within the top (bottom) five per cent in *either* reading or numeracy tests for both Year 3 and Year 5 students.
- Category C high-achieving (low-achieving) schools perform within the top (bottom) five per cent for either Year 3 or Year 5 and within the top (bottom) ten per cent for the other year level, for either reading or numeracy tests.

Overall, there were a total of 90 schools that Indigenous students attend that were identified as high-achieving, and 70 identified as low-achieving.

While examining the observed characteristics of these schools does not provide any additional insights into why those schools perform better or worse than expected (because all observed characteristics have already been taken into account in generating the expectation), a brief profile of category A, B and C high-achieving and low-achieving outlier schools for Indigenous students is presented to provide some context for the results.

Schools that are identified as high-achieving or low-achieving for Indigenous students in reading are disproportionately schools in very remote areas, schools in the Northern Territory and schools that have a high Indigenous enrolment (table 3). That is, compared with the proportions of all schools with Indigenous students that are very remote, in the Northern Territory or have a very high Indigenous enrolment, there are higher proportions of schools that have those characteristics in the top or bottom five per cent. For example, very remote schools represent only about 6 per cent of all primary schools with Indigenous students. However, they represent about 34 per cent of the high-achieving schools and 21 per cent of the low-achieving schools. It is noted that the proportions are generally higher among category A and B schools than among category C schools. That is, schools that are more clearly among the high-achieving or low-achieving schools are more likely to be in very remote areas in the Northern Territory and have a high proportion of Indigenous students.

The observation may suggest that there is greater variation in school performance in very remote areas, in the Northern Territory and in schools with a high proportion of Indigenous students — some of these schools do much better than expected, while others do much less well, after their other characteristics are taken into account. As described in section 5, the unobserved characteristics of the families and communities that these students belong to may be more homogeneous, and the students are more likely to have been taught by the same teachers, and these characteristics may be conducive or disadvantageous to school achievement, leading to greater variation in school performance.

However, the result may also reflect the fact that the technique used to estimate school effects adjusts the estimates depending on how reliable they are (annex A). The technique places greater weight on school effects that are more reliably estimated and less weight on school effects that are less reliably estimated. All else equal, the less reliable the estimate, the smaller the estimated school effect. Estimates are less reliable if there are few students at the school included in the analysis. Conversely, estimates are more reliable if there are a large number of students at the school included in the analysis. Schools with a large number of Indigenous students included in the analysis are disproportionately those in very remote areas in the Northern Territory that have a high proportion of Indigenous enrolments. Therefore, the technique is likely to reduce the estimated effects of these schools by a smaller amount, relative to other schools that have few Indigenous students, leading to a greater proportion of these schools in both the high-achieving and low-achieving outlier groups.

An analysis of the high-achieving and low-achieving schools for non-Indigenous students produces a profile that is more similar to the profile for all schools. There is a higher proportion of schools where there are more non-Indigenous students (such as in metropolitan areas). This could indicate that there is greater variation in school effects among non-Indigenous students in metropolitan areas, or could again reflect the fact that estimates of school effects for schools with a large number of students in the analysis are more reliable.

Table 3 **Achievement outlier schools for Indigenous students are disproportionately in very remote areas in the Northern Territory with many Indigenous students**

Indigenous students only (2013 and 2014 pooled)

	Number of high-achieving outlier schools					*Number of low-achieving outlier schools*					*All schools*
Category	A	B	C	Total	Total %	A	B	C	Total	Total %	Total %
Total	13	33[a]	44	90		6	28[b]	36	70		
Remoteness											
Metro	3	7	14	24	26.7	0	1	15	16	22.9	50.7
Provincial	4	9	17	30	33.3	0	15	16	31	44.3	39.1
Remote	1	1	2	5	5.6	1	5	2	8	11.4	4.3
Very remote	5	15	11	31	34.4	5	7	3	15	21.4	6.0
State											
NSW	2	5	12	19	21.1	1	10	10	21	30.0	34.5
VIC	0	1	1	2	2.2	0	0	4	4	5.7	10.8
QLD	4	8	11	23	25.6	0	5	8	13	18.6	25.0
SA	0	1	3	4	4.4	2	0	1	3	4.3	6.6
WA	1	8	5	14	15.6	0	3	9	12	17.1	13.0
TAS	1	1	2	4	4.4	0	1	1	2	2.9	3.9
NT	5	9	9	23	25.6	3	9	3	15	21.4	4.5
ACT	0	0	1	1	1.1	0	0	0	0	0.0	1.6
% Indigenous students											
0–5	0	4	4	8	8.9	0	0	5	5	7.1	32.9
5–10	1	2	8	11	12.2	0	1	11	12	17.1	25.4
10–15	1	3	8	12	13.3	0	1	6	7	10.0	14.7
15–20	0	2	4	6	6.7	0	1	4	5	7.1	7.4
20–30	2	1	4	7	7.8	0	4	1	5	7.1	6.5
30–50	2	3	4	9	10.0	0	6	2	8	11.4	4.6
50–95	2	6	3	11	12.2	1	7	2	10	14.3	3.7
95–100	5	12	9	26	28.9	5	8	5	18	25.7	4.8

[a] Of the 33 category B high-achieving outlier schools, 21 were within the top five per cent in reading for both Year 3 and Year 5 students, while the remaining 12 were within the top five per cent in numeracy for both year levels. [b] 14 schools were within the bottom five per cent in reading for both Year 3 and Year 5 students, and the remaining 14 were within the bottom five per cent in numeracy for both year levels.

Source: Commission estimates based on ACARA data (unpublished).

Further analysis of high-achieving schools

Further analysis of high-achieving schools has the potential to inform policy and improve the performance of other schools. Additional insight into why certain schools perform well relative to other schools could be achieved by examining characteristics that are not available in the ACARA data, to see what sets them apart. An evaluation of this type could not be done for this study because of the de-identified nature of the data. It is also noted that different techniques could identify a different set of high-performing schools. Before investigating the high-performing schools in this analysis, it would be informative to crosscheck the rankings of schools with rankings obtained from alternative methods, such as those described in annex A. It may be worth examining the schools that perform consistently well across different methods of analyses, to see what could be contributing to their better-than-expected performance.

Schools could be outperforming similar schools for various reasons. Examples of unobserved characteristics that could potentially contribute to a school being high-achieving for Indigenous students include school programs that target the education of Indigenous students, the leadership of principals and the general attitudes and expectations of teachers at the school. Some of these unobserved characteristics may be replicable in other schools, and could be used to inform policy in order to lift Indigenous performance. For example, if a number of schools that outperform similar schools employ a particular teaching strategy, then that teaching strategy could be trialled in some other schools with an aim to improve student achievement. It is important to pilot such strategies to be able to correctly assess the effects of a policy initiative and see whether it leads to improvements in achievement, before considering whether to implement the initiative across schools more widely.

Annex A — Modelling NAPLAN test scores: data, research methodology and results

This annex details the conceptual model, data, statistical methods and results of the modelling of National Assessment Program — Literacy and Numeracy (NAPLAN) test scores in background paper (BP) 2. The motivation behind this analysis and the research questions examined in this study are described in BP 2. In brief, the goal is to shed light on the contributors to academic achievement for Indigenous Australian primary school students. The analysis also aims to identify whether there are schools that perform better than predicted given their observed characteristics, based on the estimation of 'school effects'. These results are used to provide insights into how Indigenous achievement might be improved.

Section A.1 of this annex describes the framework for understanding contributors to student achievement. This is formalised in the education production function, on which the analysis of NAPLAN scores is based. Section A.2 then describes the data obtained from the Australian Curriculum, Assessment and Reporting Authority (ACARA) and the processes undertaken in creating the final dataset used in the analysis. Descriptive statistics for students included in the modelling are also presented.

The remainder of the annex covers the statistical methods used in the analysis. Section A.3 compares the use of multilevel and fixed effects models in analysing how observed factors relate to NAPLAN scores, in partitioning variation in student achievement into school and student components, and in identifying school effects. Particular attention is paid to the relative advantages and disadvantages of each method in answering this study's research questions. Section A.4 then explains dominance analysis, which is used to evaluate the relative importance of observed factors in explaining NAPLAN scores. Finally, section A.5 describes the Blinder-Oaxaca decomposition method, which is used to decompose the gap in mean scores between Indigenous and non-Indigenous students. Only a subset of modelling results are presented in this annex, but further tables of modelling results are available in spreadsheet format in annex B.

A.1 The education production function

As described in BP 2, student achievement can be affected by factors that are within the control of policy makers, as well as environmental factors that are not within the control of policy makers. These factors can also be grouped into school-level and student-level factors.

The means by which various factors influence student achievement is formalised in the education production function. In this framework, student achievement (the output of the education process) is directly related to a set of inputs (Hanushek 1979, 1986). Education is also cumulative, that is, past inputs influence current levels of achievement. The major factors influencing achievement as described by Hanushek (1986) and recognised by Hattie (2003) include, for example:

- factors that can potentially be influenced by policy makers, such as school curricula, principals, teachers and school characteristics
- the influences of peers and family background (environmental factors)
- innate student ability or learning capacity (environmental factors).

Other conceptual models of child development acknowledge broader influences. For example, children's families, schools and community can be set within a wider social, economic, cultural and political context, which could also have influences on student achievement (Zubrick et al. 2000). These are also generally considered to be environmental rather than policy factors.

Both policy factors and environmental factors should be taken into account in a model in order to produce unbiased estimates of the different factors relationships with achievement.

A general form of the conceptual model, based on Hanushek (1979) and Todd and Wolpin (2003), and incorporating broader social factors, is as follows.

$$A_{ist} = f\left(C_{is}^{(t)}, S_{is}^{(t)}, P_{is}^{(t)}, T_{is}^{(t)}, F_{is}^{(t)}, I_{is}, \varepsilon_{ist}\right) \tag{1}$$

where:

- A_{ist} is the academic achievement of student i at school s and time t
- $C_{is}^{(t)}$ is a vector of broad social influences, cumulative to time t
- $S_{is}^{(t)}$ is a vector of school-related influences, cumulative to time t
- $P_{is}^{(t)}$ is a vector of peer influences, cumulative to time t
- $T_{is}^{(t)}$ is a vector of teacher influences, cumulative to time $tF_{is}^{(t)}$ is a vector of family background influences, cumulative to time t
- I_{is} is a vector of innate student learning capacity
- ε_{ist} allows for measurement error in test scores.

Estimating the education production function

The exact choice of explanatory factors, and hence the empirical specification, is usually guided by the availability of data (Hanushek 1986). That is also the case for this study. As data are only available on current factors in the ACARA data, a 'contemporaneous' modelling specification is estimated, which relates achievement to current factors (Todd and Wolpin 2003).[9] This specification assumes that only current factors influence current achievement (or that factors do not change over time) and that observed factors are not related to unobserved factors relating to achievement. While such assumptions are strong and unlikely to hold true, these models have often been estimated in the education literature as few alternatives exist when data are limited.

As noted in BP 2, the ACARA data do not include information on all the factors in the conceptual model. While some data are available on social, school, peer, family background and student influences in the ACARA dataset, no data are observed for teachers (BP 2, figure 1). The influence of teachers on achievement will be reflected at a school level, to the extent that it is constant across schools, and at a student level to the extent that it is specific to each student.

The lack of data on particular factors raises the issue of estimates potentially being biased because of variables omitted from the model. For example, there is a lack of data on teacher quality. If highly educated parents (a factor that is observed in the data) are better able to choose schools with high teacher quality, then the influence of teacher quality will be captured in the estimate of the influence of parental education, thus biasing the result. These issues are discussed further in section A.3.

Based on the available data, the following empirical specification was estimated, where examples of variables included in each vector of inputs is shown in BP 2, figure 1.

$$A_{is} = f(C_{is}, S_{is}, P_{is}, F_{is}, I_{is}, \varepsilon_{is}) \qquad (2)$$

Section A.3 describes the exact functional forms of the models used in this study.

A.2 Data description

The data provided by ACARA consist of de-identified school and student data for Year 3 and Year 5 students in 2013 and 2014. Separate data files were provided for student-level data and school-level data for each year. Student records were merged with their school's

9 An alternative specification is the value-added model, which analyses the influences of inputs while controlling for past achievement (Hanushek 1979). Although a measure of past achievement is available in the ACARA data for Year 5 students (in the form of their Year 3 NAPLAN scores), the value-added specification was not analysed in this study. Both contemporaneous and value-added specifications impose strong assumptions on the model, and the benefits of a value-added approach are uncertain when the potential for omitted variable bias is taken into account (Todd and Wolpin 2003).

records (with a match rate of nearly 100 per cent) to create a comprehensive dataset for analysis.

Data on NAPLAN scores and test participation were available for each student in all test domains: reading, writing, numeracy and language conventions (box A.1). Other student characteristics in the data included: age; gender; Indigenous status; language background other than English (LBOTE); parental education; parental occupation; and, for Year 5 students, an indicator for whether the student attended the same school when they sat the NAPLAN tests in Year 3 (a measure of student mobility).

At the school level, available variables included: school sector (government, Independent or Catholic); state; remoteness (metropolitan, provincial, remote or very remote); Index of Community Socio-Educational Advantage (ICSEA) value (which captures the educational advantage of a school's student population, as determined by students' family backgrounds, remoteness and the proportion of Indigenous students); teaching and non-teaching staff numbers; student numbers; proportions of female, Indigenous and LBOTE students; school finances (for 2013 only) and overall attendance rates for the school (for 2014 only).

Box A.1 **About NAPLAN**

The National Assessment Program — Literacy and Numeracy (NAPLAN) is a nationwide test for students in Years 3, 5, 7 and 9 that has been administered in May each year since 2008. NAPLAN assesses skills in four domains: reading, numeracy, writing and language conventions (spelling, grammar and punctuation).

NAPLAN scores are reported as scaled scores across all year levels. Scales were fixed based on the results of all students from the 2008 NAPLAN tests so that each test scale has a range of 0 to 1000, a mean of 500 and a standard deviation of 100 (Holmes-Smith 2012). The same scaled score represents the same level of achievement across all year levels and across all test years, with the exception of writing because of a change in the writing task from narrative to persuasive writing in 2011 (ACARA 2014c).

National minimum standards are based on scaled scores. The national minimum standard score for Year 3 students is 270, while that for Year 5 students is 374 (Holmes-Smith 2012).

The main purpose of NAPLAN is to provide information to governments, education authorities and schools to 'inform the development of strategies to improve the literacy and numeracy skills of students in all schools across Australia' (Australian Government 2013, p. 1). It is not designed to directly measure the level or progress of an individual student's knowledge or skills, and ought to be considered as a complement to teacher judgment and school-based assessment programs (VCAA 2013, p. 1).

For ease of reporting, the report and BP 2 focus on reading and numeracy results for Year 5 students. The analysis was performed on both Year 3 and Year 5 students, and results were similar, but Year 5 students were chosen as the main focus because they enable conclusions to be made about student mobility (for which an indicator is only available for Year 5 students). Reading was chosen because it is seen as a core skill that is important for all areas of learning (Song, Perry and McConney 2014), and numeracy was chosen to provide a measure of mathematical achievement. It is worth noting that scores across all tests are highly correlated, with correlations of about 0.8. Consequently, separate analyses for Year 3 students and for other test domains lead to similar overall results. The results of all analyses are presented in annex B.

Data modifications

Some modifications to the data were made to facilitate the analyses. These modifications consisted of the creation of new variables to use in the multivariate modelling, and amendments to correct for likely processing errors in the data. In addition, school variables that were missing only for one year were proxied by the value of the variable in the year for which it was available.

Variable creation

Highest level of education

Separate variables were provided on parents' school education and post-school education (including certificates, diplomas and degrees). These were amalgamated to form a highest education level variable for each parent.

In creating the variable, it is assumed that a Certificate I to IV qualification (including trade certificate) is a higher level of education than Year 12. The ABS classifies Certificate I and II qualifications as lower than Year 10 (ABS 2015), but the post-school education data on certificates could not be separated by qualification level.

In cases where a parent stated a level of school education but did not state their post-school education, preliminary regression results suggested that the relationship with student NAPLAN scores was similar to cases where a parent reported that they had the same level of school education and no post-school education. Therefore, it was assumed that all parents who did not state their post-school education had no post-school education. A separate category in the highest education variables was created for parents who did not report any school or post-school education.

School socioeconomic status

A number of variables were created to take into account the socioeconomic status (SES) of a school community, based on the notion that SES is a function of income, education and occupation.[10] For education and occupation, variables were created for the percentage of

[10] These variables were used instead of school ICSEA to control for SES because ICSEA takes into account additional factors such as remoteness and the percentage of Indigenous students (as described below).

parents of Year 3 and 5 students in a school community at each highest education level, and for the percentage of parents in each occupation category. Average school fees and parent contributions were standardised by school sector (such that each school sector has a mean of zero and a standard deviation of one).[11] This standardised fees variable, and interactions between this variable and school sector, were also included in the model to proxy income. The interactions between standardised fees and school sector take into account the fact that an independent school that has fees equivalent to the average fee level for independent schools may have students from families with higher incomes on average than a government school that has average government school fees. The creation of a composite school SES variable was considered but not pursued within the project's timeframe.

A variable containing a measure of SES in the school's geographic area (deciles of the Index of Relative Socioeconomic Advantage and Disadvantage (IRSAD)) was provided by ACARA at a later stage in the project. This could not be incorporated into the analysis within the timeframe of the project, but the IRSAD decile variable was regressed on the variables that were used to control for school SES in this study to examine how well they compared. The school SES variables in this study explained about 43 per cent of the variation in the IRSAD decile. The unexplained variation may be partly due to the range of variables included in the IRSAD, such as the number of bedrooms, rent and mortgage payments and number of cars (ABS 2014), that are not reflected in the variables that control for school SES in this study. Furthermore, the IRSAD decile variable reflects the SES of people and households in an area wider than the school community. In some cases, the SES of a school community may differ from the wider community.

Missing values

Missing values for the school-level proportions of female, Indigenous and LBOTE students were replaced with zero when they were reasonably expected to be zero based on the available data for student characteristics.

Proxy variables

School finance data were only available for 2013, but were applied to schools in 2014 to act as a proxy measure for school finances in 2014. Similarly, school attendance rates were only available for 2014, but were used as a proxy for 2013 attendance rates.

[11] This standardisation means that it is possible for the average of this variable across all students to be negative, depending on how students are distributed across schools and school sectors.

Data issues

NAPLAN reliability

Three types of error can affect the reliability of NAPLAN scores: measurement error, sampling error and equating error.

Measurement error

Measurement error reflects the ability of NAPLAN to measure the true values of constructs over time and across different groups of students. It can arise for a number of reasons: test scores may be affected by the types of questions asked, the stress associated with tests, teachers 'teaching to the test', and cultural and linguistic bias in test questions. These are described in detail below. While some of these issues can be controlled for through the inclusion of the right environmental factors in the model, data on these factors is not always available.

The limited number of questions in NAPLAN tests can mean that students are only asked about a subset of all the possible skills they could be expected to possess in a given domain of education (Australian Government 2010, 2014). As test questions change from year to year, so too do the precise skills that are tested. Consequently, a student's scores may reflect the questions asked rather than their true underlying ability in a given domain.[12]

Test conditions have been demonstrated to affect student performance through increased stress levels. For example, UK research shows that mandatory standardised testing can increase stress levels among students, and in doing so, reduce test performance and increase the margin of error in test results as students are likely to make errors due to nerves as well as lack of ability (Connor 2001). This has implications for this study to the extent that Indigenous students face different degrees of stress or respond differently to stress.

Critics of NAPLAN also suggest that test conditions are too inflexible for students with learning difficulties and disabilities (Australian Government 2014). To the extent that rates of disability differ between Indigenous and non-Indigenous students, this means there is a risk that differences in scores due to disability could be mistakenly attributed to Indigeneity.

Teachers, parents or private tutors 'teaching to the test' can contribute to measurement error. Evidence of NAPLAN-specific teaching or coaching has been provided by the Australian Education Union (2010, p. 3) and in media reports (Bagshaw 2015; Bita 2015; Morris 2013; Topsfield 2013). To the extent that certain groups of students are taught to

[12] Computer adaptive tests will be introduced from 2017 onwards, which will better tailor test items to the ability levels of individual students (ACARA 2015b). Computer adaptive tests select questions based on responses to prior questions and so require fewer questions than paper-based tests to obtain an equally accurate measure of student ability (thus reducing measurement error).

prepare for NAPLAN tests more than others, this could increase measurement error in NAPLAN results. These groups of students may perform better not because they have a higher level of competency in the subject area, but because they were better prepared for the tests.

Cultural and linguistic bias, particularly against Indigenous students and students with a refugee background, has been a persistent concern (Australian Government 2010, 2014). There is evidence that test questions sometimes assume knowledge that Indigenous students from remote communities (and students with refugee backgrounds) may not possess, such as about cinemas and newspaper deliveries (Wigglesworth, Simpson and Loakes 2011). Critics also argue that NAPLAN tests are an inappropriate instrument for measuring the skills of students that do not possess Standard Australian English (SAE) as a first language, because of different language-learning pathways (Wigglesworth, Simpson and Loakes 2011). Linguistic bias is an issue even for numeracy tests, especially when questions represent problems using words and require students to translate these words into mathematical representations (Harris et al. 2013).

Cultural and linguistic bias in the regression analysis is controlled for to an extent with the inclusion of variables for Indigeneity and LBOTE. However, fluency in SAE cannot be identified — some Indigenous students may be fluent while others may be learning SAE in conjunction with either a traditional Aboriginal language, a creole or a distinctive Aboriginal variety of English. This could mean that differences in scores attributed to remoteness may actually be due to differences in Indigenous language background.

Sampling error

Sampling error can also affect the reliability of NAPLAN scores. Sampling error refers to the difference between the true value of test scores and the estimated value of test scores caused by observing a sample rather than the population as a whole. This can affect the extent to which the results of this study can be generalised.

- At a student level, sampling error occurs because absences and withdrawals mean that NAPLAN is not a true census of students. Non-participation is not necessarily random. For example, Indigenous students are less likely to participate, particularly those in more remote areas.
- At a school level, sampling error occurs because the student cohort changes from year to year. This matters because the aim of NAPLAN is to make inferences about schools and education systems rather than about the cohort of students in any given year (ACARA 2015c).

Equating error

Equating error can occur in the process of NAPLAN test scores being equated across different year levels and different years so that results can be compared on a common scale. Equating is accomplished by having a sample of students sit an equating test in addition to the main NAPLAN test. Questions on the equating test remain the same from one year to the next so comparisons can be made across time (ACARA 2015c). Errors in the equating process will mean that student test scores will not be perfectly comparable across years and across year levels even after scaling. This may affect the pooled analyses across 2013 and 2014 data in this study, but is controlled for to an extent through the inclusion of a year dummy.

Selection bias

Not all students and schools in the data were included in the analysis and this has implications on the results. In the analysis, the samples used in various regressions were restricted to students who:

- participated in the relevant NAPLAN test (see sampling error above)
- had a complete set of school data relevant to the regression[13]
- attended a school with at least one other student who participated in the relevant NAPLAN test for the same year level in either 2013 or 2014.

On the last point, the decision of the minimum number of students required in a school for it to be included in the analysis involves a tradeoff between the accuracy of estimated school effects (and thus school-level variation) and the representativeness of the estimated relationships between observed factors and achievement. If the last constraint was not put in place, then the estimated school effect would be based only on a single student for some schools, which is unlikely to be a good representation of the school effect for all students at the school. The higher the minimum threshold of students, the fewer the unreliable school effects included in the analysis. However, a higher minimum threshold also means that the estimated relationships between observed factors and achievement would be based on a smaller number of schools and students that may not reflect the characteristics of all schools, making the results less representative. Because many Indigenous students attend schools with few other Indigenous students (BP 1), the number of schools and students falls dramatically as the minimum threshold increases in the analysis of Indigenous students in particular.

The minimum threshold was set to two in order to maximise the total number of schools and students included in the analysis, while removing the most unreliable estimates of school effects. The sensitivity of the results to this choice was tested. Increasing the minimum threshold to 20 students led to only small changes in the analysis of variation attributed to school and student levels for both Indigenous and non-Indigenous students.

[13] Students missing data only for student characteristics were still included because separate 'missing data' categories were created for those variables.

In total, about 19 per cent of Indigenous students and 7 to 8 per cent of non-Indigenous students were excluded, mainly because they did not participate in NAPLAN (table A.1). About a third of the schools in the analyses of Indigenous students were excluded, primarily due to there being fewer than two Indigenous students in the school in the relevant sample. In contrast, only 7 per cent of schools in the analyses of non-Indigenous students were excluded.

Table A.1 **Size of NAPLAN reading samples, by Indigeneity and year level**
Number of students (number of schools)

	Indigenous				*Non-Indigenous*			
	Year 3		Year 5		Year 3		Year 5	
Initial total	30 811	(5 005)	29 210	(4 926)	547 054	(7 600)	529 196	(7 618)
Exclusions	5 764	(1 692)	5 455	(1 652)	42 493	(531)	39 113	(514)
Did not participate	4 112		3 811		35 760		31 174	
Missing data	298		314		6 641		7 858	
Single student	1 354		1 330		92		81	
Final sample size	25 047	(3 313)	23 755	(3 274)	504 561	(7 069)	490 083	(7 104)

Source: Commission estimates based on ACARA data (unpublished).

The main drawback of these necessary exclusions is that the results are conditional upon meeting the above requirements. The results cannot be generalised to the population of all students if students who were not included in the sample differ systematically to those who were. Descriptive analysis indicates that students who do not participate in NAPLAN differ from those who do participate in a non-random manner. For example, non-participation tends to be higher for Indigenous students in more remote areas. If the students who do not participate are those who would have been more likely to perform poorly, then applying the results of this analysis to them would overstate their performance.[14] It is important to keep this in mind when interpreting the results of this analysis.

In a similar vein, the restrictions mean that the analysis of the distribution of school effects for schools with Indigenous students is only based on about two thirds of all schools that Indigenous students attend. As mentioned above, schools with very few Indigenous students who participate in NAPLAN are not included in the analysis because the school effects would not be estimated reliably. Therefore, the analysis of schools effects may not be representative of all schools with Indigenous students, and evaluations of schools with very few Indigenous students are not considered.

[14] One possible method of seeing how much non-response matters to the results could be to perform the analysis using a binary outcome variable that indicates whether a student scored above the national minimum standard. Students who did not participate in NAPLAN could be excluded from one such analysis and treated as performing below the national minimum standard in another analysis. A comparison of the results would provide an indication of how much the non-participants might matter. This was not performed for this project due to the timeframe but could be considered in future research.

Reliability of observed factors

In addition to the issues with NAPLAN scores explained above, measurement error was considered in the variables relating to parental education, parental occupation and school finances. In undertaking the regression analysis, issues with including ICSEA as an explanatory factor were also considered.

Parental education and occupation

ACARA obtains data on parental education and occupation through schools. Schools do not usually update this data when a parent changes their education level or occupation. Although the results in this analysis could be affected if these changes are not recorded, few changes are expected to occur — there are unlikely to be many parents of school-aged children who gain additional education (at a level above that which they have reported), and even though parents may change occupations while their child is at school, changes between the broad occupation categories are less likely to occur.

School finances

Data on school finances is generally comparable across states and sectors but there may be some inconsistencies. The methodology that ACARA uses for reporting financial data takes account of differences across school sectors and across jurisdictions by excluding a range of sources of income and expenditure such as depreciation and loan interest (ACARA 2015a). Deloitte's (2015, p. 3) review of this methodology found that it provided 'a reasonable basis for the collection of materially comparable financial data by school on a national basis'. The main source of inconsistency identified by Deloitte was that states do not always consistently report capital expenditure (some jurisdictions only report capital expenditure on completed projects).

Index of Community Socio-Educational Advantage

School ICSEA was not included as an explanatory factor in the final regression results. It was used as a proxy for school SES in preliminary analyses, however, it is designed to predict school NAPLAN performance, and reflects socio-educational advantage rather than socioeconomic advantage. In addition to information about students' socioeconomic backgrounds, ICSEA incorporates factors such as school remoteness and the percentage of Indigenous students, weighted by how they relate to NAPLAN scores (ACARA 2014a). To avoid strong multicollinearity between ICSEA and the percentage of Indigenous students,[15] a number of constructed variables (described above) were used to reflect school SES in this study.

[15] The correlation between school ICSEA and the percentage of Indigenous students was about minus 0.8.

Descriptive statistics

Table A.2 presents descriptive statistics of NAPLAN scores in reading and numeracy for the Indigenous and non-Indigenous Year 3 and Year 5 samples used in the multilevel models. Table A.3 then presents descriptive statistics for each variable used in the modelling of NAPLAN reading scores for Indigenous and non-Indigenous Year 5 students. Descriptive statistics for other regressions (for other test domains and for Year 3 students) are not presented but are broadly similar. These statistics may differ slightly to those presented in BP 1 (which contains a profile of all Indigenous and non-Indigenous Year 5 students in 2014) because of the sample restrictions mentioned above. For example, Indigenous students who did not sit NAPLAN tests or who attended schools with no other Indigenous students in their year level will not be represented in this analysis. This could affect the descriptive statistics to the extent that those students differ to Indigenous students who do sit NAPLAN tests or who do attend schools with other Indigenous students in their year level.

Table A.2 **NAPLAN reading and numeracy scores, by Indigeneity and year level**

Mean score (standard deviation)

				Indigenous				*Non-Indigenous*
	Reading		Numeracy		Reading		Numeracy	
Year 3	339.66	(93.21)	332.70	(74.71)	423.75	(86.81)	403.87	(74.16)
Year 5	432.15	(80.17)	419.79	(69.03)	505.75	(75.23)	491.45	(74.65)

Source: Commission estimates based on ACARA data (unpublished).

Table A.3 **Descriptive statistics for Year 5 NAPLAN reading samples, by Indigeneity**

Mean (standard deviation)

		Indigenous		*Non-Indigenous*	
Number of students		23 755		490 083	
Number of schools		3 274		7 104	
Student factors					
Age		10.53	(0.39)	10.58	(0.38)
Gender	Female	0.50		0.49	
	Male	0.50		0.51	
Language background	English	0.79		0.74	
	LBOTE	0.16		0.23	
	Not stated	0.05		0.03	

(continued next page)

Table A.3 (continued)

		Indigenous		Non-Indigenous	
Mother's highest education level	Year 9 or below	0.10		0.03	
	Year 10 or 11	0.25		0.13	
	Year 12	0.09		0.13	
	Certificate I to IV	0.22		0.22	
	Advanced diploma / Diploma	0.06		0.14	
	Bachelor degree or above	0.05		0.27	
	Not stated	0.24		0.09	
Father's highest education level	Year 9 or below	0.07		0.03	
	Year 10 or 11	0.14		0.10	
	Year 12	0.05		0.09	
	Certificate I to IV	0.18		0.26	
	Advanced diploma / Diploma	0.03		0.10	
	Bachelor degree or above	0.03		0.23	
	Not stated	0.51		0.20	
Mother's occupation	Senior management	0.04		0.15	
	Other business manager	0.06		0.16	
	Tradesman, clerk, sales, services	0.11		0.18	
	Machine operator	0.12		0.12	
	Not in paid work	0.30		0.24	
	Not stated	0.37		0.16	
Father's occupation	Senior management	0.03		0.18	
	Other business manager	0.05		0.20	
	Tradesman, clerk, sales, services	0.10		0.20	
	Machine operator	0.17		0.15	
	Not in paid work	0.07		0.05	
	Not stated	0.57		0.23	
Same school in 2014 as in 2013	No	0.21		0.15	
	Yes	0.64		0.78	
	Unknown	0.15		0.07	
School factors					
School sector	Government	0.88		0.67	
	Independent	0.03		0.13	
	Catholic	0.09		0.20	
Combined (primary and secondary) school		0.19		0.17	
Average class size (students per teaching staff)[a]		14.75	(3.24)	16.28	(2.77)
Non-teaching staff per 100 students		3.32	(2.72)	2.20	(1.60)
Number of full-time equivalent enrolments		422.95	(315.23)	573.23	(403.16)
Percentage Indigenous students	0–5%	0.15		0.73	
	5–10%	0.19		0.15	
	10–15%	0.15		0.06	
	15–20%	0.09		0.03	
	20–30%	0.12		0.02	
	30–50%	0.09		0.01	
	50–95%	0.10		0.00	
	95–100%	0.11		0.00	
Percentage LBOTE students		21.35	(27.85)	22.21	(24.80)

(continued next page)

Table A.3 (continued)

		Indigenous		*Non-Indigenous*	
Attendance rate[b]		88.88	(8.88)	93.94	(2.05)
Recurrent income less fees per student ($100s)		142.51	(64.13)	99.65	(28.16)
Capital income deductions per student ($100s)[c]		0.96	(4.01)	4.06	(10.14)
Capital expenditure per student ($100s)		9.68	(31.67)	8.16	(27.33)
Fees per student, standardised by school sector		-0.45	(0.64)	0.05	(1.01)
Percentage of mothers by highest education level	Year 9 or below	7.18		3.43	
	Year 10 or 11	19.70		12.79	
	Year 12	11.20		12.38	
	Certificate I to IV	22.49		21.59	
	Advanced diploma / Diploma	7.99		13.31	
	Bachelor degree or above	11.51		27.53	
Percentage of fathers by highest education level	Year 9 or below	5.30		3.09	
	Year 10 or 11	12.46		9.58	
	Year 12	6.78		8.76	
	Certificate I to IV	24.13		25.89	
	Advanced diploma / Diploma	5.17		9.83	
	Bachelor degree or above	7.94		23.10	
Percentage of mothers by occupation	Senior management	6.83		14.77	
	Other business manager	8.52		15.30	
	Tradesman, clerk, sales, services	13.76		17.31	
	Machine operator	12.88		11.29	
	Not in paid work	27.19		25.24	
Percentage of fathers by occupation	Senior management	6.92		18.09	
	Other business manager	10.16		19.40	
	Tradesman, clerk, sales, services	15.50		19.14	
	Machine operator	18.04		14.94	
	Not in paid work	5.77		4.85	
Test participation rate		91.39	(8.95)	94.76	(5.35)
State	NSW	0.33		0.33	
	VIC	0.05		0.25	
	QLD	0.31		0.20	
	SA	0.05		0.07	
	WA	0.13		0.11	
	TAS	0.03		0.02	
	NT	0.09		0.01	
	ACT	0.01		0.02	
Remoteness	Metro	0.41		0.75	
	Provincial	0.40		0.24	
	Remote	0.07		0.01	
	Very remote	0.12		0.00	
Year	2013	0.50		0.49	
	2014	0.50		0.51	

[a] Teaching staff includes staff such as classroom teachers, physical education teachers, art teachers, teacher librarians and special education staff. [b] Percentage of school days attended. [c] Deductions include recurrent income allocated to current capital projects, future capital projects and diocesan capital funds, and income allocated to debt servicing.

Source: Commission estimates based on ACARA data (unpublished).

A.3 Modelling nested data

The ACARA data contain information on students, many of whom attend the same schools. It is important to consider this nesting of students within schools when analysing data on student achievement. This is because students who are from the same school are likely to have test scores that are correlated with each other in a way that cannot be fully explained by observed factors in the model. For example, a school's culture could have a positive or negative influence on the achievement of all students at the school, but this information is unavailable in the ACARA data. These correlations between students violate the assumption of independent errors in standard ordinary least squares models.

When the nested data structure is not taken into account in some way, then standard errors of the estimated coefficients tend to be underestimated (and, in rare cases, overestimated), which could result in misleading inferences (box A.2). An underestimation of the standard error increases the probability of concluding that a significant relationship between an explanatory factor and the outcome variable exists when it does not actually exist. This is especially problematic when the aim is to use data to evaluate policy and programs.

Two techniques that take into account the nested structure of data are multilevel modelling (which relies on the random effects assumption) and fixed effects modelling. Because the terms 'random effects' and 'fixed effects' have different meanings across different disciplines, a list of terminology is introduced in box A.3 to clarify how they are used in this study. These techniques rely on different modelling assumptions but both produce standard errors that are more appropriate than those produced by ordinary least squares models that do not take into account nested data structures. Random effects and fixed effects models can provide:

- estimates of the relationships between observed factors and achievement
- estimates of the proportion of variation in achievement attributable to school-level and student-level factors
- estimates of school effects, which are school-specific influences on achievement attributed to school-level factors that are not included in the model.

Multilevel models and fixed effects models have various advantages and disadvantages in producing each of these estimations. These are elaborated on below.

Box A.2 Type I and type II errors in the analysis of nested data

In making inferences from statistical analyses, there is always a probability that an incorrect inference will be made. Two types of errors exist. A **type I error** involves concluding that a relationship exists when no relationship is actually present (a false positive). A **type II error** involves failing to conclude that a relationship exists when it actually does (a false negative). Typically, measures to reduce the probability of making one type of error increase the probability of making the other type of error.

Failure to model a nested data structure usually results in underestimated standard errors, and hence an increased probability of making a type I error (but decreased probability of making a type II error).

In rare cases, standard errors could be overestimated if there is little variation in means between groups and substantial unexplained variation within groups. An overestimated standard error increases the probability of making a type II error (but decreases the probability of making a type I error). Statistical adjustments to account for the nesting will result in smaller standard errors. In these cases, Arceneaux and Nickerson (2009) recommend reporting the larger standard errors in order to avoid type I errors. However, the context of the analysis should also be considered — the consequences of committing a type I error could be larger or smaller than the consequences of a type II error in different situations.

Source: Arceneaux and Nickerson (2009).

Box A.3 Random effects and fixed effects terminology

Multilevel models originated within a number of disciplines (Raudenbush and Bryk 2002) and multilevel and fixed effects models are traditionally used in different research areas (Clarke et al. 2010). This has led to the potentially confusing situation of the same terminology having different meanings across disciplines (Chaplin 2003). The meanings of terms as they are used in this study, and as they are sometimes used in other contexts, are clarified below. The term:

- **random effects model** is used in this study to refer specifically to the multilevel model in which only the intercept varies randomly across schools, while the slopes do not. In other contexts, random effects models can describe multilevel models more generally (Raudenbush and Bryk 2002).
- **random effects** is used in this study to refer to the random school intercept and slope terms u_{0s} and u_{1s} in multilevel models. In a random effects model, and generally in economics, the term random effects is only used to refer to the random school intercept (Chaplin 2003). In other contexts, the student-level error e_{is} is also referred to as a random effect (Raudenbush and Bryk 2002).
- **fixed effects model** is used in this study to refer to a model in which school effects have been taken into account through the inclusion of school dummy variables or by differencing out school effects, rather than through modelling the error component.
- **fixed effects** is used in this study to refer specifically to the coefficients on school dummy variables in a fixed effects model. In disciplines that traditionally use multilevel models, fixed effects can refer to all the non-random components in a multilevel model (Raudenbush and Bryk 2002).

Multilevel modelling

Multilevel modelling (also known as hierarchical linear modelling and mixed effects modelling) is the method that is most commonly used by education researchers to examine how student-level and school-level factors relate to student achievement. This technique models school effects as a separate error component — part of the error is attributed to schools and part of it to students. The method takes into account the nested data problem as it recognises that there may be a correlation in the achievement of students who attend the same school due to unobserved school factors.

As an illustration, consider a simple model with only one school-level factor S and one student-level factor I. Following Raudenbush and Bryk (2002), multilevel modelling can be described using two sets of models: student-level and school-level models.

The student-level model describes the relationship between student-level characteristics and achievement.

$$A_{is} = \beta_{0s} + \beta_{1s} I_{is} + e_{is} \qquad (3)$$

where:

- A_{is} is the achievement of student i at school s
- I_{is} is the student-level factor
- β_{0s} is the student-level intercept coefficient
- β_{1s} is the student-level slope coefficient
- e_{is} is the student-level random error.

As in traditional regression models, the student-level random error e_{is} follows a normal distribution with a mean of zero and a constant variance.

The second set of models, the school-level models, describe how student-level relationships vary across schools. For example, both the student-level intercept and slope coefficients could depend on school-level factor S and could vary randomly across schools. In that case, the school-level models could be written as:

$$\beta_{0s} = \gamma_{00} + \gamma_{01} S_s + u_{0s} \qquad (4)$$

$$\beta_{1s} = \gamma_{10} + \gamma_{11} S_s + u_{1s} \qquad (5)$$

where:

- S_s is the school-level factor
- γ_{00} and γ_{10} are school-level intercept coefficients
- γ_{01} and γ_{11} are school-level slope coefficients
- u_{0s} and u_{1s} are school random effects.

Where this modelling technique differs from standard ordinary least squares is in the addition of school-specific random error terms (called random effects, u_{0s} or u_{1s}) in at least one of the school-level models. Specifically, the random effect u_{0s} allows the intercept of achievement at the student level to vary randomly according to school. The random effect u_{1s} allows the slope coefficient on the student-level factor to vary randomly across schools — that is, it enables the influence of student inputs to differ according to school (differential school effects). Both of these school random effects account for the correlations between students at the school level that cannot be explained by observed factors — any unobserved school factors relating to achievement will be captured in these random effects. The school random effects are assumed to have a mean of zero, constant variance and no correlation with the student-level error.

Putting the student-level and school-level models together forms the combined model.

$$A_{is} = \gamma_{00} + \gamma_{10}I_{is} + \gamma_{01}S_s + \gamma_{11}I_{is}S_s + u_{0s} + u_{1s}I_{is} + e_{is} \quad (6)$$

This model produces estimates of the relationships between observed school-level factors and achievement, and observed student-level factors and achievement, with appropriate standard errors that take into account the nesting of students within schools.

The random effects model

A specific type of multilevel model is one in which only the intercept in the student-level model varies randomly, while the student-level slope does not vary randomly. That is, u_{1s} is set to zero. However, the student-level slope may still vary depending on observed school-level factors. This is known as a 'random effects model' in econometric contexts and a 'random intercept model' in some other multilevel modelling contexts (box A.3).

In the simple illustration, the random effects model equation is written as above, except that the random effect u_{1s} for the student-level slope coefficient is dropped.

$$A_{is} = \gamma_{00} + \gamma_{10}I_{is} + \gamma_{01}S_s + \gamma_{11}I_{is}S_s + u_{0s} + e_{is} \quad (7)$$

In these models, the school random effect u_{0s} can be interpreted as a measure of whether a school is outperforming or underperforming relative to similar schools. Using school random effects as a measure of school performance is discussed further below.

The multilevel modelling results presented in this study are from random effects models.[16] This is not uncommon in the literature — although multilevel modelling allows student-level slopes to vary by school, a large number of papers find no evidence that slopes differ according to school and subsequently estimate random effects models

[16] Random effects models in this analysis are estimated using generalised least squares. Unlike maximum likelihood estimation, it does not require the random effects to follow a normal distribution.

(Chaplin 2003). Of the studies shown in BP 2, table 2, only a few incorporated random effects for some student-level variables in some models.[17]

In order to ease the interpretation of results for each factor, few interaction terms are included in the random effects models estimated. That is, the study does not include interaction terms between each combination of factors. This is also consistent with previous studies. As discussed below in the section on school performance, this could have implications on the estimated school random effects.

Variance partitions

One of the aims of the research is to partition the total variation in achievement into a component attributable to schools (school-level or between-school variation) and a component attributable to students (student-level or within-school variation). Early techniques for partitioning variation are based on analysis of variance. The technique used in multilevel modelling is analogous to a one-way analysis of variance with random effects (Raudenbush and Bryk 2002).

In a multilevel modelling framework, the total variation in achievement can be partitioned into school-level and student-level components by estimating a random effects model with no explanatory factors (the null model), as shown below.

$$A_{is} = \gamma_{00} + u_{0s} + e_{is} \tag{8}$$

The intraclass correlation coefficient (ICC) can be calculated from the model's outputs. This measures the amount of school-level variation as a proportion of the total variation in achievement (Raudenbush and Bryk 2002).[18]

$$ICC = \rho_N = \frac{(\hat{\sigma}_u^2)_N}{(\hat{\sigma}_u^2)_N + (\hat{\sigma}_e^2)_N} \tag{9}$$

where:

- ρ_N is the variance of the school random effects as a proportion of the total variance in the school random effects and residuals in the null model
- $(\hat{\sigma}_u^2)_N$ is the estimated variance of the school random effects in the null model

[17] For example, Gemici, Lim and Karmel (2013) included random slopes for gender and student SES in their analysis of tertiary entrance score, but not in the analysis of the probability of attending university because they were not statistically significant. Nous Group (2011) examined random slopes for SES in restricted models, in which SES was the only explanatory factor, but not in their full models. Marks, McMillan and Hillman (2001) found that most of the between-school differences in tertiary entrance performance was attributed to differences in school intercepts rather than slopes for Year 9 achievement or slopes for SES.

[18] In most cases, the ICC is non-negative. In rare cases, when there is little between-group variation in means and considerably large within-group variation, a negative ICC can occur (Arceneaux and Nickerson 2009).

- $(\hat{\sigma}_e^2)_N$ is the variance of the residuals in the null model
- $(\hat{\sigma}_u^2)_N + (\hat{\sigma}_e^2)_N$ is the total variation in achievement.

The ICC describes the extent to which students within a school resemble each other.[19] High values of the ICC indicate that school-level factors play a large role in explaining the variation in student scores, whereas small values indicate that student-level factors play a large role (Gemici, Lim and Karmel 2013). If students within a school tend to have very similar levels of achievement and the mean achievement level between different schools are very different, then the proportion of variation attributable to the school level, as represented by the ICC, will be relatively high. Alternatively, if students within a school tend to have very different levels of achievement and mean achievement levels between schools are quite similar, then the ICC will be low.

Variation can also be partitioned in terms of the amount that is explained by observed factors and that which is unexplained. The proportion of school-level variation that is explained by observed factors can be examined through the reduction in unexplained school-level variance after the observed factors are added to the null model (Raudenbush and Bryk 2002). The proportion of school-level variation explained is:[20]

$$\frac{(\hat{\sigma}_u^2)_N - (\hat{\sigma}_u^2)_F}{(\hat{\sigma}_u^2)_N} \tag{10}$$

Similarly, the proportion of student-level variation explained is:

$$\frac{(\hat{\sigma}_e^2)_N - (\hat{\sigma}_e^2)_F}{(\hat{\sigma}_e^2)_N} \tag{11}$$

where $(\hat{\sigma}_u^2)_F$ and $(\hat{\sigma}_e^2)_F$ refer to the estimated variances of the school random effects and student residuals respectively, for the full model with all relevant observed factors included.

The proportion of total variation that is explained can also be calculated. Nakagawa and Schielzeth (2013) propose two methods of calculation — the marginal R^2 and the conditional R^2.

The 'marginal R^2' describes the proportion of variation explained only by observed factors in the model, and is calculated from variance components from the full model.

[19] It is also noted that a conditional ICC can also be calculated, such that it captures the extent to which students within a school resemble each other given their observed characteristics (Raudenbush and Bryk 2002).

[20] In some cases, it is possible that the addition of a student-level factor to the null model increases the school-level variation (Rabe-Hesketh and Skrondal 2008; Raudenbush and Bryk 2002). In these cases, the proportion of school-level variation explained by the model with the student-level factor would be calculated to be negative. This is not a concern for the current analysis — school-level variation remains about the same or decreases with the addition of any student-level factor to the null model.

$$R^2_{marginal} = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_u^2 + \sigma_e^2} \qquad (12)$$

where σ_f^2 is the variance of the non-random components of the model, that is, the variance explained by observed factors. It can be estimated by calculating the variance of the predicted values based on observed factors only.

The 'conditional R^2' describes the proportion of variation explained by both observed factors and school random effects (unobserved school-level factors).

$$R^2_{conditional} = \frac{\sigma_f^2 + \sigma_u^2}{\sigma_f^2 + \sigma_u^2 + \sigma_e^2} \qquad (13)$$

Random effects as a measure of school performance

As mentioned above, when observed factors are incorporated into a random effects model, the estimated school random effects u_{0s} can be interpreted as a measure of school performance. A positive school random effect indicates that students at the school perform better than would be predicted from the observed characteristics of the school and its students on average, whereas a negative school random effect indicates that students at the school perform worse than predicted on average. School random effects have been interpreted as indicators of school effectiveness or school unobserved quality in other multilevel modelling studies (for example, Gemici, Lim and Karmel (2013) and Nous Group (2011)).

The interpretation of random effects as a measure of performance is similar to interpretations seen in the analysis of firm productivity using 'stochastic frontier models' (Greene 2005). These models split the error term into two components, with one component being interpreted as a measure of technical efficiency.

The school random effects are not explained by observed characteristics and partly capture the influence of unobserved school-level factors on achievement. These unobserved school-level factors might include important elements such as school culture and teacher quality, to the extent that they are constant within a school. The analysis cannot conclude exactly what factors result in large effects for some schools without additional information. Additional information could potentially be obtained by conducting evaluations of schools that are deemed to be performing much better than expected to see what sets them apart.

It is noted that the school random effects are shrunk according to the level of uncertainty in the estimates. More reliable estimates of school effects are given more weight, while less reliable estimates are given less weight. Estimates are less reliable when the number of students in a school is small or when the student-level variance is large relative to school-level variance. In these cases, the random effect is shrunk more so that it becomes closer to zero (Clarke et al. 2010).

School random effects are analysed in this study to see how the distributions of school effects differ for Indigenous and non-Indigenous students. The research also examines the consistency of school effects across Year 3 and Year 5 students, across reading and numeracy test domains, and across Indigenous and non-Indigenous students. Schools that perform consistently well or consistently poorly across various analyses are identified.

As mentioned, the modelled equations do not include interaction terms between all factors. A brief simulation study suggested that this is not likely to bias the estimated school random effects, but it may result in a larger variance of the estimated effects if the correlation between factors is large. That said, the results of the simulation are specific to the design of the simulation, where school effects were correlated with observed factors and the observed factors were correlated with each other. A more detailed simulation study would be necessary to identify whether the results hold under different scenarios.

Omitted variable bias

A limitation of multilevel models is that they rely on the random effects assumption — that unobserved school-level factors (captured in the school random effects) are not correlated with observed factors in the model. If there are unobserved school-level factors that are associated with achievement and that are correlated with observed factors, then the estimated coefficients of the relationships between those observed factors and achievement will be biased.[21]

The random effects assumption is unlikely to hold if students are not 'randomly selected' into schools — that is, if there are unobserved factors associated with how students (or their parents) choose which schools to attend. For example, if parents with higher education (an observed student-level factor) are better able to choose schools that have high teacher quality (an unobserved school-level factor), then part of the contribution of teacher quality to student achievement would be captured in the parent education variable and the estimated coefficient on parent education would be biased.

An alternative to multilevel modelling that can take into account the nesting of multiple students within schools, but does not rely on the random effects assumption, is fixed effects modelling.

21 Omitted variable bias would not affect the initial partitioning of the total variation into school-level and student-level components because it is estimated from the null model without any explanatory factors. It may affect subsequent partitioning of variation into explained and unexplained components, which depends on the estimated coefficients. However, as noted below, estimated coefficients from random effects and fixed effects models are similar, which suggests that there is likely little bias in the coefficients from unobserved school-level factors.

Fixed effects modelling

Economists tend to favour fixed effects models over random effects models because fixed effects models do not rely on the random effects assumption, but are still able to model the correlations between students that attend the same school. Therefore, it is able to control for the problems associated with estimating a model using ordinary least squares that does not take nesting into account (box A.2). Rather than modelling school effects as part of the error term like random effects models, fixed effects models take into account school effects by using either of two equivalent approaches. The first approach involves adding dummy variables for each school as additional predictors in the regression model. The coefficients on the school dummy variables are called school fixed effects. The other approach involves differencing out the school effects by demeaning the outcome variable and predictors at the school level, and then estimating the transformed model (Clarke et al. 2010).

A downside to the simple fixed effects model is that school-level factors cannot be included in the modelling equation because they are either collinear with the school dummy variables in the first approach, or because they are differenced out in the second approach. However, an extension to the fixed effects method enables school-level factors to be analysed in a second stage, by regressing the estimated school fixed effects from the first-stage model on observed school-level factors. This two-stage technique has been used in other areas of education research, such as in analysing teacher effectiveness (for example Leigh (2010) and Aaronson, Barrow and Sander (2007)).

Considering again a simple example with a school input S and a student input I, a two-stage fixed effects approach can be described as follows. The first-stage model regresses achievement on observed student-level factors and school fixed effects.

$$A_{is} = \alpha_s + \delta_0 + \delta_1 I_{is} + e_{is} \tag{14}$$

where:

- A_{is} is the achievement of student i at school s
- α_s is the fixed effect for school s
- δ_0 is the intercept coefficient from the first-stage student-level regression
- δ_1 is the slope coefficient on student input I_{is}
- e_{is} is the random error from the first-stage student-level regression.

The second-stage model is a school-level model that regresses estimated school fixed effects from the first-stage model on observed school-level factors.

$$\hat{\alpha}_s = \tau_0 + \tau_1 S_s + r_s \tag{15}$$

where:

- τ_0 is the intercept coefficient from the second-stage school-level regression

- τ_1 is the slope coefficient on school input S_s
- r_s is the random error from the second-stage school-level regression.

The random error from the second-stage regression r_s captures two main elements. First, it captures the influences of unobserved school-level factors on achievement, similar to that captured by u_{0s} in the random effects model. The error r_s also captures any measurement error in the dependent variable $\hat{\alpha}_s$, which itself is an estimate from the first-stage regression. Although the existence of measurement error increases the variance of r_s (meaning that estimates of the slope coefficients τ_1 will have larger standard errors), it does not bias the estimates of the slope coefficients. A consequence of the larger standard errors is that there is a larger probability of not finding a significant relationship between a school factor and achievement when a relationship exists (box A.2). That said, when compared with the random effects model, fixed effects models still have the benefit of producing estimates that are less likely to be biased, as discussed below. In the absence of any omitted variable bias, estimates from the random effects model are more efficient.

Omitted variable bias

The fixed effects model controls for some of the omitted variable bias that is present in random effects models. Because the fixed effects model incorporates school effects as dummy variables rather than through the error component, the school effects are allowed to be correlated with observed student-level factors in the first-stage model. Therefore, unlike in the random effects model, correlations between unobserved school-level factors (captured in r_s in the fixed effects model) and observed student-level factors (I_{is}) will not bias the coefficients on the observed student-level factors. Furthermore, because the influences of school-level factors are estimated in a second-stage regression that abstracts from any student-level factors, any correlations between observed school-level factors (S_s) and unobserved student-level factors (captured in e_{is}) will also not bias coefficients on observed school-level factors.

However, the two-stage fixed effects method is not completely free from omitted variable bias — there are other sources of bias that do not involve the random effects assumption. Namely, if there are unobserved student-level factors (captured in e_{is}) that are related to achievement and that are correlated with observed student-level factors (I_{is}), the effects of those unobserved factors will bias the coefficients in the first-stage regression. Similarly, unobserved school-level factors (captured in r_s) could still bias coefficients on observed school-level factors (S_s) in the second-stage regression.

As the range of observed factors in the ACARA data is limited, there are likely to be unobserved student-level (school-level) factors affecting achievement that are correlated with observed student-level (school-level) factors. For example, because cognitive ability is partially inherited, the estimated coefficient on observed parent education variables are likely to incorporate some of the effects of unobserved student cognitive ability on achievement. As a result, omitted variable bias is likely to be present in both fixed effects and random effects models, regardless of the random effects assumption.

Choosing between random effects and fixed effects

In making a choice between reporting results from random effects and fixed effects models in this project, both statistical and practical considerations were taken into account.

Clarke et al. (2010) advise that the two different approaches are appropriate in different contexts, and researchers in education should consider both. Clarke et al. focus on the implications of each method on the estimated coefficient of the relationship between a student-level policy factor and student achievement. The fixed effects model is preferred when there are only a limited range of factors that can be included in the model to take into account any bias arising from school-level factors. On the other hand, if the data are rich enough to adequately account for these potential biases, then the random effects model can also be used to make policy-relevant inferences and should be preferred because of its greater efficiency. However, Clarke et al. (2010) acknowledge that inadequate student-level data may be the greatest barrier to obtaining policy-relevant results in both fixed effects and random effects models.

One statistical method that is often used to assess which model is more appropriate is the Hausman test. In most regressions in this analysis, the test suggested that fixed effects models were more suitable. However, if omitted variable bias is present outside of the random effects assumption (that is, if there are unobserved student-level factors biasing estimated coefficients on observed student-level factors), then the Hausman test can be unreliable (Clarke et al. 2010). Given this uncertainty, both random effects and fixed effects models were analysed for this study.

The decision of which modelling results to present in the BP 2 was based on practical considerations relating to the aims of the project.

One of the aims is to estimate the relationships between observed school and student factors and student achievement. As mentioned, theoretically, the fixed effects model has the advantage of not relying on the random effects assumption (meaning it can produce estimates of the relationships for student factors that are not biased by unobserved school factors). However, research has found that estimated relationships between particular factors and achievement can be very similar in both random effects and fixed effects models when a broad set of explanatory factors is included (Clarke et al. 2010). In the current analysis, comparisons of the magnitude and significance of relationships between student factors and achievement are found to be very similar between the two models (tables A.4 and A.5). There are some differences in the estimated relationships for school factors between the random effects model and the second stage regression of the fixed effects, but these differences are not large. The difference may be a result of the fixed effects being estimated with error, as described above.

A second aim of the project is to estimate the proportion of variation in achievement attributed to school-level and student-level factors. Analysing the ICC from the null random effects model is a well-established technique for partitioning total variation in achievement into between-group and within-group components. Other studies in the education literature have also used random effects to separate total variation into school-level and student-level components (BP 2, table 2). Although the variation can be partitioned into school-level and school-level components in the fixed effects model using a formula analogous to the ICC, the validity of using fixed effects models to partition variation is less clear. Estimated variance partitions from the fixed effects models tend to attribute a larger share of total variation to schools than in the random effects models, particularly for Indigenous students in the current analysis. The random effects model was preferred to address this aim of the project because it is the more conventional method used in the literature.

Another aim of the project is to compare the relative importance of observed school-level and student-level factors with each other in the dominance analysis (section A.4). The random effects model evaluates both school-level and student-level factors in a single regression, rather than in two stages, which makes it easier to compare the relative importance of all observed factors. In contrast, using a fixed effects approach, the relative importance of school-level factors and student-level factors would have to be analysed separately, such that the importance of a school-level factor would not able to be compared with a student-level factor. Therefore, the random effects model was preferred to address this particular aim.

On balance, although the fixed effects modelling technique has its advantages, the random effects modelling technique was preferred because it was considered to better address the aims of the project.

Alternative methods of estimating school effects

A key aim of the research is to estimate school effects and identify schools that outperform other schools, given their observed characteristics. Estimated school effects from the random effects model were used for this purpose. These were chosen instead of the school effects from the two-stage fixed effects model in order to be consistent with other results presented in BP 2 that rely on the random effects model, and because the school random effects are adjusted by how reliable they are, as described above. However, it is noted that estimates of school effects from fixed effects models could also be adjusted to increase their reliability, for example, by manually shrinking the fixed effects or by attributing only part of the estimated school effect to a true school effect, as is sometimes done in the teacher effects literature (Aaronson, Barrow and Sander 2007; Leigh 2010).

Comparisons were conducted of the school effects from random effects models and unadjusted school effects from the second-stage regression of the fixed effects approach. These effects were highly correlated, with correlations of over 0.8. However, the magnitudes of the school effects from the second stage of the fixed effects modelling technique tended to be larger than those from the random effects model, especially in the tails of the distribution, that is, where the high-achieving and low-achieving schools are located. This means that the rankings of schools in the top five per cent and the bottom five per cent were not necessarily consistent between the random effects and fixed effects analyses. This observation is consistent with the school effects from the random effects models being shrunk according to the degree of reliability in the estimates, and the estimated school effects from the fixed effect approach remaining unadjusted. Reliability adjustments may lead to more similar estimates of school effects from the fixed effects approach. This was not examined within the timeframe of this project.

Other methods of measuring school effects also exist, but were not pursued for this project. They could involve extending the multilevel model or using alternative methods, such as those from the productivity literature or non-parametric methods.

Approaches involving variations or extensions to the multilevel model could consider whether the relationship between student-level factors and achievement differ depending on the school that the student attends. For example, if it is hypothesised that the relationship between parental education and achievement depends on school sector, then interaction variables between the two factors could be added to the model. Alternatively, if it is hypothesised that the relationship between parental education and achievement differs for every school, then this could be tested by estimating random slopes on the parental education factor, that is, allowing the slope coefficient on parental education to vary randomly across schools, as in equation 6 (Raudenbush and Bryk 2002). Another variation could involve estimating a value-added model, where past achievement is included in the model as a predictor for current achievement (Hanushek 1979; Todd and Wolpin 2003). Estimated school effects could then be interpreted as estimates of how a school contributes to growth in student performance, after controlling for observed factors.

Borrowing from the productivity literature, another method of measuring school efficiency could be to estimate a stochastic frontier model. Such models split the residual into a non-negative efficiency component and a random noise component with zero mean (Greene 2008; Lovell 1993).

Non-parametric approaches could also be considered. A non-parametric approach called student growth percentile analysis involves assigning students a percentile ranking, based on their current level of achievement relative to prior achievement (Betebenner 2008). School rankings can be obtained from student rankings, for example, according to the median percentile ranking of students at the school. Past research that compared school rankings from student growth percentile analysis with those from value-added fixed effects models found that they were very similar, even at the extremes (that is, for the highest-ranked and lowest-ranked schools) (Houng and Justman 2013).

While estimates of school effects using these alternative methods were not analysed in this study, it would be informative to crosscheck the rankings of schools in this study with rankings obtained from alternative methods. For schools that are deemed to perform consistently well across different methods of analyses, it may be worth examining them further to see what could be contributing to their better-than-expected performance.

Modelling results

In modelling the associations between school- and student-level factors and student achievement, pooled regressions of Indigenous and non-Indigenous students were performed first to analyse how Indigeneity is associated with NAPLAN scores. Interaction terms were added to these regressions to examine whether having a LBOTE or attending a school in a more remote area had a different relationship with NAPLAN scores for Indigenous and non-Indigenous students. Regressions were then run for Indigenous and non-Indigenous students separately to examine how observed factors were associated with NAPLAN scores for the two groups. These results are discussed in BP 2.

Some variables included in initial regressions explained relatively little of the variation in NAPLAN scores and were consequently removed from the final models. These variables include the percentage of female students at a school, an indicator for single-sex school and an indicator for whether the school was a head campus.

The results presented in this annex are for Year 5 reading and numeracy achievement for Indigenous and non-Indigenous students separately. Results for pooled regressions of Indigenous and non-Indigenous students, Year 3 students, and for other test domains are presented in annex B.

Table A.4 presents results from the null random effects models and full random effects models. The overall R^2 (the default from the modelling output) is reported for each full random effects model, as well as the marginal R^2 (analogous to the overall R^2) and conditional R^2 proposed by Nakagawa and Shielzeth (2013). The marginal R^2 is very similar to the overall R^2 for each of these models. For example, in the analysis of reading scores for year 5 Indigenous students, the overall R^2 and marginal R^2 were both estimated to be 0.31. The differences between conditional and marginal R^2s (which capture variation attributable to school random effects) are also very similar to estimates of school-level variation attributed to unobserved school-level factors, calculated from variance components of the null and full models. Modelling results reported in annex B only show the overall R^2.

Table A.5 presents results for fixed effects models, with student-level and school-level explanatory factors analysed in two stages.

Separate analyses were also conducted for Indigenous students by school remoteness. These results were sensitive to the minimum number of students required in a school for the school to be included in the analysis. However, initial analyses suggest that Indigenous students in very remote areas have a higher percentage of variation attributable to school-level factors, at about 40 per cent or more. The school-level percentage of variation in the analysis of all Indigenous students also dropped to about 17 per cent when remote schools were excluded. This suggests that students attending the same school in very remote areas are more similar to each other than are students attending the same school in less remote areas.

Consistency of school effects

The results of the analysis of school random effects are discussed in BP 2. The discussion includes an analysis of the spread of the distributions of school effects for both Indigenous and non-Indigenous students, and examines the observed characteristics of high-achieving and low-achieving schools for Indigenous students.

Comparisons of school effects were also conducted across different random effects models. In particular, the study examined the consistency of schools identified in the top and bottom five per cent of the distribution, across Year 3 and Year 5 students, across reading and numeracy scores, and across Indigenous and non-Indigenous students. The consistency of school effects is important for their interpretation as a measure of school performance. The consistency results are presented in table A.6.

This analysis found some consistency in school effects across reading and numeracy tests for both Indigenous and non-Indigenous students. That is, some of the schools that are identified in the top five per cent for reading are also in the top five per cent for numeracy. Of the schools with Indigenous students included in the analyses of both reading and numeracy achievement, about 42 per cent of schools identified in the top five per cent in numeracy are in the top five per cent in reading, and about 35 per cent of schools identified as the bottom five per cent in numeracy are also in the bottom five per cent in reading (table A.6). For schools with non-Indigenous students, these percentages are slightly higher, at about 51 per cent for the top five per cent, and 44 per cent for the bottom five per cent.

There is less consistency in school effects across Year 3 and Year 5 students when examining reading scores. For example, for Indigenous students, the degree of overlap of schools in the top five per cent across both year levels is 26 per cent. For the bottom five per cent of schools with Indigenous students, the degree of overlap is 15 per cent.

Part of the reason for the lack of consistency is because of the small spread in school effects — a few points difference between analyses could result in a school being within the top five per cent in one analysis but outside the top five per cent in another analysis. Therefore, the analysis presented in BP 2 on high-achieving and low-achieving schools identifies three categories of high-achieving and low-achieving schools based on how consistently they perform across analyses.

- Category A high-achieving (low-achieving) schools perform within the top (bottom) five per cent in *both* reading and numeracy tests for both Year 3 and Year 5 students.
- Category B high-achieving (low-achieving) schools perform within the top (bottom) five per cent in *either* reading or numeracy tests for both Year 3 and Year 5 students.

- Category C high-achieving (low-achieving) schools perform within the top (bottom) five per cent for Year 3 or Year 5 and within the top (bottom) ten per cent for the other year level, for either reading or numeracy tests.

In terms of consistency in the performance of Indigenous and non-Indigenous students at the same schools within each test domain, only a small number of schools in the top or bottom five per cent overlap across both Indigenous and non-Indigenous student groups. Of the schools that were included in the analyses of both Indigenous and non-Indigenous students, only about 17 per cent of the top five per cent for non-Indigenous students in reading were in the top five per cent for Indigenous students, and 12 per cent of the bottom five per cent for non-Indigenous students were also the bottom five per cent for Indigenous students (table A.6). These percentages are similar for numeracy. The reason for the lower consistency across Indigenous and non-Indigenous students could partly be because unobserved school factors that contribute to high achievement for non-Indigenous students may not necessarily work the same for Indigenous students, and vice versa.

Table A.4 Random effects modelling results, by Indigeneity and test domain — Year 5 students

Estimated coefficient (standard error)

	Indigenous				*Non-Indigenous*			
	Reading		Numeracy		Reading		Numeracy	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
# of students	23 755		23 424		490 083		488 454	
# of schools	3 274		3 265		7 104		7 102	
Null model								
Constant	445.75 ***	(0.88)	431.43 ***	(0.78)	500.88 ***	(0.39)	485.14 ***	(0.39)
sigma_u	39.18		35.03		29.87		30.04	
sigma_e	66.86		58.38		69.58		67.91	
rho (ICC)	0.26		0.27		0.16		0.16	
Full model								
Age	3.12 ***	(1.17)	1.74 *	(1.05)	4.80 ***	(0.28)	2.45 ***	(0.27)
Male	-14.11 ***	(0.85)	2.61 ***	(0.76)	-10.93 ***	(0.20)	12.92 ***	(0.19)
Language background (default: English)								
LBOTE	-14.05 ***	(1.81)	-13.82 ***	(1.62)	-7.60 ***	(0.29)	5.24 ***	(0.29)
Not stated	-4.79 *	(2.49)	-2.14	(2.25)	1.43 *	(0.73)	1.43 **	(0.72)
Mother's education (default: Year 9 or below)								
Year 10 or 11	8.89 ***	(1.66)	5.87 ***	(1.49)	8.49 ***	(0.63)	7.09 ***	(0.62)
Year 12	18.66 ***	(2.08)	14.59 ***	(1.86)	19.43 ***	(0.64)	17.62 ***	(0.62)
Certificate I to IV	18.12 ***	(1.77)	14.92 ***	(1.59)	16.17 ***	(0.62)	13.59 ***	(0.61)
Advanced diploma / Diploma	20.81 ***	(2.46)	17.12 ***	(2.19)	23.25 ***	(0.65)	20.21 ***	(0.63)
Bachelor degree or above	34.36 ***	(2.79)	29.27 ***	(2.50)	39.39 ***	(0.65)	36.27 ***	(0.64)
Not stated	10.19 ***	(1.90)	7.27 ***	(1.71)	21.35 ***	(0.73)	19.45 ***	(0.71)
Father's education (default: Year 9 or below)								
Year 10 or 11	7.29 ***	(2.08)	5.51 ***	(1.86)	5.87 ***	(0.67)	4.57 ***	(0.65)
Year 12	16.19 ***	(2.68)	16.53 ***	(2.40)	17.59 ***	(0.68)	15.81 ***	(0.67)
Certificate I to IV	11.13 ***	(2.11)	11.54 ***	(1.89)	12.23 ***	(0.64)	11.10 ***	(0.62)
Advanced diploma / Diploma	14.72 ***	(3.25)	7.93 ***	(2.90)	19.14 ***	(0.69)	16.86 ***	(0.68)
Bachelor degree or above	25.16 ***	(3.61)	24.11 ***	(3.23)	31.11 ***	(0.68)	30.97 ***	(0.67)
Not stated	8.32 ***	(2.12)	6.30 ***	(1.90)	10.95 ***	(0.72)	8.75 ***	(0.70)
Mother's occupation (default: Not in paid work)								
Senior management	8.14 ***	(2.64)	7.88 ***	(2.36)	2.49 ***	(0.39)	2.00 ***	(0.38)
Other business manager	6.99 ***	(2.19)	9.60 ***	(1.95)	1.36 ***	(0.34)	1.37 ***	(0.33)
Trade, clerk, sales, services	6.82 ***	(1.66)	9.20 ***	(1.48)	1.33 ***	(0.32)	2.17 ***	(0.31)
Machine operator	6.49 ***	(1.52)	7.73 ***	(1.36)	-0.88 **	(0.35)	0.10	(0.34)
Not stated	-4.38 ***	(1.39)	-2.99 **	(1.25)	-2.51 ***	(0.42)	-1.63 ***	(0.41)

(continued next page)

Table A.4 (continued)

	Indigenous				Non-Indigenous			
	Reading		Numeracy		Reading		Numeracy	
Father's occupation (default: Not in paid work)								
Senior management	19.49 ***	(3.31)	19.02 ***	(2.96)	15.59 ***	(0.58)	13.02 ***	(0.57)
Other business manager	17.91 ***	(2.73)	16.89 ***	(2.44)	11.45 ***	(0.54)	10.74 ***	(0.53)
Trade, clerk, sales, services	13.20 ***	(2.30)	9.42 ***	(2.05)	7.21 ***	(0.54)	7.08 ***	(0.53)
Machine operator	5.18 ***	(2.00)	3.96 **	(1.79)	2.39 ***	(0.54)	2.62 ***	(0.52)
Not stated	1.03	(2.14)	1.87	(1.91)	1.77 ***	(0.62)	1.31 **	(0.60)
Same school (default: No)								
Yes	8.34 ***	(1.10)	6.71 ***	(0.99)	5.82 ***	(0.29)	5.83 ***	(0.28)
Unknown	-4.15 **	(2.06)	-3.19 *	(1.85)	-6.53 ***	(0.48)	-5.36 ***	(0.47)
School sector (default: Government)								
Independent	13.75 **	(6.13)	8.32	(5.66)	4.91 ***	(1.39)	1.92	(1.44)
Catholic	1.14	(2.60)	0.46	(2.44)	-5.11 ***	(0.72)	-9.63 ***	(0.76)
Combined school	-2.76	(2.44)	-2.41	(2.28)	0.227	(0.87)	1.92 **	(0.90)
Average class size	0.43	(0.30)	0.41	(0.28)	0.41 ***	(0.12)	0.35 ***	(0.12)
Non-teaching staff per 100 students	0.29	(0.32)	0.09	(0.30)	-0.67 ***	(0.16)	-0.22	(0.16)
Enrolments	0.00	(0.00)	0.00	(0.00)	0.00 **	(0.00)	0.00	(0.00)
Percentage Indigenous students (default: 0–5%)								
5–10%	-2.40	(1.93)	-4.15 **	(1.78)	-2.22 ***	(0.58)	-2.93 ***	(0.59)
10–15%	-8.59 ***	(2.28)	-8.57 ***	(2.11)	-6.07 ***	(0.82)	-5.80 ***	(0.84)
15–20%	-11.25 ***	(2.70)	-10.98 ***	(2.51)	-9.37 ***	(1.13)	-7.30 ***	(1.16)
20–30%	-13.52 ***	(2.82)	-14.96 ***	(2.64)	-8.56 ***	(1.32)	-8.30 ***	(1.36)
30–50%	-21.46 ***	(3.35)	-20.94 ***	(3.14)	-10.34 ***	(1.88)	-6.71 ***	(1.93)
50–95%	-24.31 ***	(4.00)	-20.82 ***	(3.75)	-15.80 ***	(2.81)	-8.38 ***	(2.88)
95–100%	-28.33 ***	(5.39)	-23.72 ***	(5.07)	-77.36 ***	(12.21)	-42.41 ***	(13.35)
Percentage LBOTE students	0.07 *	(0.04)	0.05	(0.04)	-0.02 *	(0.01)	-0.02	(0.01)
Attendance rate	1.95 ***	(0.17)	1.51 ***	(0.17)	1.67 ***	(0.14)	2.40 ***	(0.14)
Recurrent income less fees per student ($100s)	-0.01	(0.02)	0.00	(0.02)	-0.014	(0.01)	-0.02	(0.01)
Capital income deductions per student ($100s)	0.06	(0.17)	-0.03	(0.16)	-0.048	(0.04)	-0.03	(0.04)
Capital expenditure per student ($100s)	-0.03	(0.02)	-0.02	(0.02)	0.000	(0.01)	0.00	(0.01)

(continued next page)

Table A.4 (continued)

	Indigenous				Non-Indigenous			
	Reading		Numeracy		Reading		Numeracy	
Percentage of mothers by highest education level								
Year 9 or below	-0.04	(0.11)	0.01	(0.10)	-0.33 ***	(0.05)	-0.31 ***	(0.05)
Year 10 or 11	-0.13	(0.10)	-0.12	(0.09)	-0.23 ***	(0.04)	-0.28 ***	(0.04)
Year 12	-0.41 ***	(0.13)	-0.28 **	(0.12)	-0.08 *	(0.04)	-0.10 **	(0.04)
Certificate I to IV	-0.11	(0.11)	-0.07	(0.10)	-0.20 ***	(0.04)	-0.25 ***	(0.04)
Advanced diploma / Diploma	0.13	(0.16)	0.07	(0.14)	-0.16 ***	(0.04)	-0.12 ***	(0.04)
Bachelor degree or above	-0.17	(0.16)	-0.05	(0.14)	-0.01	(0.04)	-0.05	(0.04)
Percentage of fathers by highest education level								
Year 9 or below	-0.08	(0.15)	-0.19	(0.14)	-0.05	(0.06)	-0.08	(0.06)
Year 10 or 11	0.13	(0.13)	0.00	(0.12)	-0.04	(0.04)	-0.01	(0.04)
Year 12	0.12	(0.17)	-0.05	(0.15)	0.03	(0.05)	0.09 **	(0.05)
Certificate I to IV	0.17	(0.12)	0.12	(0.11)	-0.01	(0.04)	0.06	(0.04)
Advanced diploma / Diploma	0.30	(0.19)	0.03	(0.17)	0.04	(0.05)	0.08 *	(0.05)
Bachelor degree or above	0.07	(0.18)	-0.03	(0.16)	0.13 ***	(0.04)	0.22 ***	(0.04)
Percentage of mothers by occupation								
Senior management	0.11	(0.17)	0.00	(0.15)	-0.07 *	(0.04)	-0.04	(0.04)
Other business manager	-0.21	(0.14)	-0.17	(0.13)	0.04	(0.04)	0.07 *	(0.04)
Tradesman, clerk, sales, services	0.10	(0.12)	0.00	(0.11)	-0.09 **	(0.04)	0.00	(0.04)
Machine operator	0.02	(0.11)	-0.12	(0.10)	0.03	(0.04)	0.02	(0.04)
Not in paid work	-0.01	(0.08)	-0.05	(0.07)	-0.08 ***	(0.03)	-0.03	(0.03)
Percentage of fathers by occupation								
Senior management	0.15	(0.17)	0.26	(0.16)	0.07 *	(0.04)	-0.07 *	(0.04)
Other business manager	0.00	(0.14)	0.17	(0.13)	0.04	(0.04)	-0.01	(0.04)
Tradesman, clerk, sales, services	-0.20	(0.13)	-0.07	(0.12)	0.03	(0.04)	-0.09 **	(0.04)
Machine operator	-0.13	(0.11)	0.01	(0.11)	-0.10 ***	(0.04)	-0.05	(0.04)
Not in paid work	-0.07	(0.14)	0.00	(0.13)	-0.08 *	(0.04)	-0.06	(0.04)
Fees per student, standardised by school sector	2.79 **	(1.24)	-0.55	(1.16)	0.70 ***	(0.23)	0.76 ***	(0.24)
Fees x Independent	8.55	(5.81)	8.76	(5.36)	7.58 ***	(0.96)	6.69 ***	(1.02)
Fees x Catholic	-3.05	(3.83)	0.20	(3.60)	0.44	(0.67)	0.33	(0.71)
Test participation rate	0.06	(0.08)	-0.03	(0.07)	-0.07 **	(0.03)	0.00	(0.03)

(continued next page)

Table A.4 (continued)

			Indigenous				*Non-Indigenous*	
	Reading		Numeracy		Reading		Numeracy	
State (default: NSW)								
VIC	11.05 ***	(2.85)	13.20 ***	(2.64)	5.02 ***	(0.75)	3.56 ***	(0.79)
QLD	6.78 ***	(2.13)	3.47 *	(1.99)	0.351	(0.83)	0.24	(0.88)
SA	-5.13 *	(3.05)	-10.19 ***	(2.84)	-6.83 ***	(1.00)	-12.78 ***	(1.06)
WA	-5.53 **	(2.67)	-7.35 ***	(2.49)	-2.34 **	(0.92)	-2.39 **	(0.98)
TAS	3.94	(3.47)	4.01	(3.25)	5.13 ***	(1.52)	-0.81	(1.61)
NT	-9.68 **	(3.77)	-5.72	(3.56)	9.41 ***	(2.55)	4.33	(2.70)
ACT	7.07	(5.96)	-2.66	(5.54)	2.875	(2.00)	-4.39 **	(2.13)
Remoteness (default: Metro)								
Provincial	0.72	(1.54)	0.66	(1.45)	3.34 ***	(0.63)	3.69 ***	(0.67)
Remote	1.48	(3.21)	-2.79	(3.05)	8.93 ***	(1.69)	8.32 ***	(1.76)
Very remote	-2.97	(4.02)	-0.16	(3.84)	12.21 ***	(2.81)	11.17 ***	(2.91)
Year 2014	-20.28 ***	(0.87)	-2.68 ***	(0.78)	-1.80 ***	(0.20)	-0.37 *	(0.20)
Constant	222.66 ***	(21.61)	258.45 ***	(20.02)	271.56 ***	(13.48)	193.96 ***	(14.05)
sigma_u	18.45		18.95		15.72		17.18	
sigma_e	63.95		56.74		66.85		65.20	
Overall R^2	0.32		0.27		0.18		0.19	
Marginal R^2	0.31		0.26		0.17		0.17	
Conditional R^2	0.36		0.33		0.21		0.23	

*** statistically significant at the 1 per cent level, ** 5 per cent level, * 10 per cent level

Source: Commission estimates based on ACARA data (unpublished).

Table A.5 **Two-stage fixed effects modelling results, by Indigeneity and test domain — Year 5 students**

Estimated coefficient (standard error)

	Indigenous				*Non-Indigenous*			
	Reading		Numeracy		Reading		Numeracy	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
# of students	23 755		23 424		490 083		488 454	
# of schools	3 274		3 265		7 104		7 102	
1st stage: Student factors								
Age	3.88 ***	(1.25)	2.60 **	(1.12)	4.78 ***	(0.28)	2.43 ***	(0.27)
Male	-14.06 ***	(0.90)	2.12 ***	(0.81)	-10.88 ***	(0.20)	12.96 ***	(0.19)
Language background (default: English)								
LBOTE	-14.31 ***	(1.95)	-14.29 ***	(1.74)	-7.49 ***	(0.29)	5.19 ***	(0.29)
Not stated	-4.15	(2.95)	0.19	(2.64)	2.17 ***	(0.77)	2.07 ***	(0.75)
Mother's education (default: Year 9 or below)								
Year 10 or 11	9.49 ***	(1.73)	6.29 ***	(1.55)	8.57 ***	(0.63)	7.01 ***	(0.62)
Year 12	18.96 ***	(2.18)	14.38 ***	(1.95)	19.53 ***	(0.64)	17.64 ***	(0.62)
Certificate I to IV	19.12 ***	(1.86)	15.17 ***	(1.66)	16.18 ***	(0.62)	13.54 ***	(0.60)
Advanced diploma / Diploma	21.77 ***	(2.61)	16.74 ***	(2.32)	23.23 ***	(0.65)	20.15 ***	(0.63)
Bachelor degree or above	34.45 ***	(3.00)	29.64 ***	(2.68)	39.46 ***	(0.65)	36.23 ***	(0.63)
Not stated	10.88 ***	(1.98)	7.68 ***	(1.77)	21.61 ***	(0.73)	19.61 ***	(0.71)
Father's education (default: Year 9 or below)								
Year 10 or 11	7.65 ***	(2.18)	5.64 ***	(1.94)	5.93 ***	(0.66)	4.60 ***	(0.65)
Year 12	15.38 ***	(2.83)	16.21 ***	(2.52)	17.58 ***	(0.68)	15.88 ***	(0.67)
Certificate I to IV	10.51 ***	(2.22)	11.75 ***	(1.99)	12.28 ***	(0.64)	11.23 ***	(0.62)
Advanced diploma / Diploma	13.84 ***	(3.48)	5.65 *	(3.10)	19.13 ***	(0.69)	16.96 ***	(0.68)
Bachelor degree or above	24.30 ***	(3.90)	22.34 ***	(3.47)	31.07 ***	(0.68)	31.06 ***	(0.67)
Not stated	8.71 ***	(2.21)	6.69 ***	(1.97)	11.04 ***	(0.72)	8.85 ***	(0.70)
Mother's occupation (default: Not in paid work)								
Senior management	8.70 ***	(2.82)	9.13 ***	(2.51)	2.43 ***	(0.39)	1.97 ***	(0.38)
Other business manager	7.25 ***	(2.35)	9.97 ***	(2.09)	1.37 ***	(0.34)	1.41 ***	(0.33)
Trade, clerk, sales, services	7.33 ***	(1.77)	10.39 ***	(1.57)	1.29 ***	(0.32)	2.16 ***	(0.31)
Machine operator	6.51 ***	(1.61)	7.85 ***	(1.43)	-0.77 **	(0.35)	0.12	(0.34)
Not stated	-3.43 **	(1.45)	-2.45 *	(1.30)	-2.39 ***	(0.42)	-1.52 ***	(0.41)

(continued next page)

Table A.5 (continued)

	Indigenous				Non-Indigenous			
	Reading		Numeracy		Reading		Numeracy	
Father's occupation (default: Not in paid work)								
Senior management	18.87 ***	(3.56)	18.63 ***	(3.17)	15.89 ***	(0.58)	12.98 ***	(0.56)
Other business manager	16.22 ***	(2.93)	15.61 ***	(2.62)	11.66 ***	(0.54)	10.70 ***	(0.53)
Trade, clerk, sales, services	13.23 ***	(2.44)	9.25 ***	(2.18)	7.40 ***	(0.54)	7.06 ***	(0.53)
Machine operator	4.43 **	(2.12)	4.26 **	(1.89)	2.49 ***	(0.53)	2.62 ***	(0.52)
Not stated	0.57	(2.24)	1.81	(2.00)	2.01 ***	(0.62)	1.42 **	(0.60)
Same school (default: No)								
Yes	8.25 ***	(1.19)	6.29 ***	(1.06)	5.81 ***	(0.29)	5.74 ***	(0.28)
Unknown	-4.34 **	(2.20)	-3.32 *	(1.97)	-6.42 ***	(0.49)	-5.33 ***	(0.48)
Year 2014	-21.62 ***	(0.90)	-3.37 ***	(0.81)	-1.97 ***	(0.19)	-0.37 *	(0.19)
Constant	378.98 ***	(13.48)	367.12 ***	(12.05)	412.29 ***	(3.10)	411.23 ***	(3.03)
Within R^2	0.08		0.06		0.08		0.08	
2nd stage: School factors								
School sector (default: Government)								
Independent	14.54 *	(7.43)	6.59	(8.15)	4.41 *	(2.65)	3.22	(2.33)
Catholic	1.56	(2.83)	-0.63	(2.81)	-6.38 ***	(0.90)	-10.85 ***	(0.91)
Combined school	-2.29	(3.18)	-1.58	(2.71)	1.82	(1.41)	0.15	(1.34)
Average class size	0.47	(0.52)	0.87 ***	(0.32)	0.50 ***	(0.19)	0.37 **	(0.17)
Non-teaching staff per 100 students	-0.75	(0.51)	-0.31	(0.36)	-1.13 ***	(0.38)	-0.76 **	(0.32)
Enrolments	0.00	(0.00)	0.00	(0.00)	0.00 ***	(0.00)	0.00	(0.00)
Percentage Indigenous students (default: 0–5%)								
5–10%	-4.36 **	(2.15)	-4.50 **	(1.98)	-2.79 ***	(0.80)	-4.00 ***	(0.81)
10–15%	-10.42 ***	(2.60)	-9.81 ***	(2.46)	-5.91 ***	(1.16)	-5.85 ***	(1.17)
15–20%	-12.51 ***	(3.14)	-8.60 ***	(2.94)	-8.20 ***	(1.69)	-7.05 ***	(1.66)
20–30%	-14.10 ***	(3.21)	-15.77 ***	(3.01)	-6.91 ***	(1.91)	-8.34 ***	(1.90)
30–50%	-22.04 ***	(3.97)	-17.19 ***	(3.84)	-12.11 ***	(2.74)	-10.42 ***	(2.68)
50–95%	-25.44 ***	(5.02)	-21.04 ***	(4.89)	-12.00 ***	(4.18)	-10.31 **	(4.37)
95–100%	-23.47 ***	(8.27)	-20.40 ***	(6.88)	-71.80 ***	(19.80)	-51.07 ***	(16.78)
Percentage LBOTE students	-0.04	(0.05)	0.00	(0.04)	-0.04 *	(0.02)	0.00	(0.02)
Attendance rate	1.92 ***	(0.29)	1.44 ***	(0.25)	1.43 ***	(0.22)	1.70 ***	(0.22)
Recurrent income less fees per student ($100s)	0.02	(0.03)	0.01	(0.02)	0.01	(0.02)	0.01	(0.02)
Capital income deductions per student ($100s)	0.07	(0.16)	0.02	(0.19)	-0.02	(0.06)	-0.02	(0.07)
Capital expenditure per student ($100s)	-0.01	(0.02)	0.01	(0.02)	-0.01	(0.01)	0.00	(0.01)

(continued next page)

Table A.5 (continued)

	Indigenous				Non-Indigenous			
	Reading		Numeracy		Reading		Numeracy	
Percentage of mothers by highest education level								
Year 9 or below	-0.02	(0.16)	-0.07	(0.15)	-0.33 ***	(0.09)	-0.32 ***	(0.10)
Year 10 or 11	-0.11	(0.12)	-0.17	(0.13)	-0.21 ***	(0.07)	-0.23 ***	(0.07)
Year 12	-0.21	(0.17)	-0.22	(0.18)	-0.09	(0.08)	-0.15 **	(0.07)
Certificate I to IV	-0.09	(0.14)	-0.06	(0.14)	-0.16 **	(0.06)	-0.24 ***	(0.07)
Advanced diploma / Diploma	0.02	(0.19)	0.11	(0.20)	-0.06	(0.07)	-0.08	(0.08)
Bachelor degree or above	-0.22	(0.20)	-0.06	(0.21)	-0.05	(0.08)	-0.03	(0.08)
Percentage of fathers by highest education level								
Year 9 or below	-0.06	(0.22)	-0.17	(0.20)	-0.12	(0.10)	-0.03	(0.10)
Year 10 or 11	-0.18	(0.18)	-0.25	(0.16)	-0.02	(0.07)	-0.01	(0.08)
Year 12	0.08	(0.20)	0.05	(0.20)	0.06	(0.08)	0.14 *	(0.08)
Certificate I to IV	0.00	(0.16)	0.01	(0.16)	0.05	(0.07)	0.07	(0.07)
Advanced diploma / Diploma	0.34	(0.24)	0.13	(0.23)	0.07	(0.09)	0.09	(0.09)
Bachelor degree or above	0.09	(0.22)	-0.06	(0.24)	0.28 ***	(0.08)	0.24 ***	(0.08)
Percentage of mothers by occupation								
Senior management	0.23	(0.21)	-0.05	(0.19)	0.01	(0.09)	0.00	(0.08)
Other business manager	0.05	(0.17)	-0.14	(0.17)	0.06	(0.07)	0.13 *	(0.07)
Tradesman, clerk, sales, services	0.23	(0.15)	0.11	(0.13)	-0.07	(0.06)	0.01	(0.06)
Machine operator	0.17	(0.14)	-0.03	(0.13)	-0.09	(0.06)	0.02	(0.06)
Not in paid work	0.09	(0.10)	0.04	(0.10)	-0.07	(0.05)	0.00	(0.05)
Percentage of fathers by occupation								
Senior management	0.22	(0.21)	0.26	(0.21)	-0.05	(0.07)	-0.03	(0.08)
Other business manager	-0.14	(0.17)	0.14	(0.16)	0.06	(0.06)	0.09	(0.07)
Tradesman, clerk, sales, services	-0.11	(0.16)	-0.17	(0.15)	0.02	(0.06)	-0.06	(0.07)
Machine operator	0.03	(0.15)	0.07	(0.14)	-0.01	(0.06)	0.01	(0.06)
Not in paid work	0.01	(0.18)	0.03	(0.19)	0.05	(0.10)	-0.01	(0.08)
Fees per student, standardised by school sector	3.67 **	(1.51)	0.35	(1.32)	0.17	(0.33)	0.12	(0.40)
Fees x Independent	8.07	(7.80)	8.71	(8.36)	9.47 ***	(1.59)	8.23 ***	0.00
Fees x Catholic	-2.96	(5.00)	0.61	(4.78)	1.09	(0.75)	1.51 **	(1.60)
Test participation rate	-0.08	(0.13)	0.07	(0.13)	0.07	(0.07)	0.19 ***	(0.70)

(continued next page)

Table A.5 (continued)

	Indigenous				Non-Indigenous			
	Reading		Numeracy		Reading		Numeracy	
State (default: NSW)								
VIC	7.84 **	(3.20)	11.89 ***	(2.99)	5.37 ***	(1.10)	4.08 ***	(1.07)
QLD	7.90 ***	(2.47)	5.97 **	(2.39)	2.16 *	(1.27)	1.75	(1.26)
SA	-2.64	(3.33)	-5.48 *	(3.09)	-4.79 ***	(1.26)	-11.19 ***	(1.27)
WA	-2.14	(3.01)	-5.86 **	(2.81)	-0.59	(1.27)	-1.03	(1.30)
TAS	5.08	(4.31)	7.50 *	(3.93)	5.75 ***	(1.75)	0.24	(1.78)
NT	2.48	(6.39)	-3.37	(5.86)	2.59	(2.07)	-4.00 *	(2.14)
ACT	-10.51 **	(4.29)	-4.71	(3.67)	8.99 ***	(3.47)	2.64	(3.19)
Remoteness (default: Metro)								
Provincial	0.15	(1.73)	1.41	(1.61)	4.04 ***	(0.77)	5.03 ***	(0.76)
Remote	3.42	(3.80)	0.18	(3.73)	10.26 ***	(2.14)	10.06 ***	(2.03)
Very remote	-2.91	(5.42)	1.13	(4.49)	10.23 **	(4.25)	9.07 **	(3.89)
Constant	-156.02 ***	(25.77)	-128.42 ***	(22.92)	-141.56 ***	(20.96)	-180.84 ***	(21.37)
R^2	0.31				0.31			

*** statistically significant at the 1 per cent level, ** 5 per cent level, * 10 per cent level

Source: Commission estimates based on ACARA data (unpublished).

Table A.6 **Consistency of school effects**

Percentage of overlap in schools in top or bottom five per cent across analyses

		Top 5%	*Bottom 5%*
Across reading and numeracy tests (Year 5)	Indigenous	41.72	34.97
	Non-Indigenous	50.56	44.38
Across Year 3 and Year 5 students (reading)	Indigenous	26.12	14.93
	Non-Indigenous	28.57	23.43
Across Indigenous and non-Indigenous students (Year 5)	Reading	17.31	12.18
	Numeracy	18.06	14.84

Source: Commission estimates based on ACARA data (unpublished).

A.4 Dominance analysis

Dominance analysis was used to examine the relative importance of observed factors in explaining NAPLAN scores. This method is based on a model's coefficient of determination (R^2), which is the proportion of variation in the outcome variable that is explained by variation in the independent variables — a measure of goodness of fit (Azen and Budescu 2003).

A rough picture of the relative importance of each factor can be investigated by comparing the change in a model's R^2 before and after the factor is included in the model. However, because of correlations between explanatory factors, the change in R^2 depends on which factors have been included in the model beforehand. An upper bound value of the proportion of variation explained by a particular factor is the change in R^2 when that factor is added to the null model. A lower bound value is the change in R^2 when the factor is the last to be added to the model.

Dominance analysis provides a more concise measure of a factor's relative importance. It involves examining the change in a model's R^2 for all possible combinations of explanatory variables. The general dominance statistic is then calculated as the average marginal contribution of a factor to the model's R^2, across all models in which that factor is included. A larger general dominance statistic indicates that the factor is relatively more important. General dominance statistics have the useful property that they sum to the full model's R^2 (Azen and Budescu 2003).

Dominance analysis was implemented in Stata with the user-written command 'domin' (Luchman 2015). The R^2 used was the overall R^2 reported from the random effects models.

Issues and limitations

The total number of regressions required to calculate dominance statistics increases exponentially with the number of explanatory factors — a model with five predictors requires 31 separate regressions while 15 predictors requires over 32 000 separate regressions (Tonidandel and LeBreton 2011). Therefore, constrained dominance analysis was used in this analysis to limit the total number of regressions required. A calendar year indicator was included in all models and dominance was examined among the remaining factors. (That is, the relative importance of calendar year was not analysed.) The remaining factors were grouped into ten sets in order to further restrict the total number of regressions required. These factor sets were created according to natural groupings (for example, student SES-related factors were grouped together), and included a set of factors that appeared to individually explain little of the total variation in achievement based on initial examinations of upper and lower bound changes in the overall R^2 of the random effects model. Each of the variables in a set were entered into the model together such that the change in R^2 reflects the contribution of all variables in the set.

Azen and Budescu (2003) and Tonidandel and LeBreton (2011) discuss some other limitations of dominance analysis. Broadly, problems that are inherent in the underlying model would also affect the dominance analysis. For example, measurement error can make dominance statistics unreliable and the results are susceptible to model misspecification.

A noteworthy issue is how dominance analysis performs in the presence of multicollinearity. For example, school attendance rate and the proportion of Indigenous students at a school are highly correlated, which could affect the dominance results. Tonidandel and LeBreton (2011) claim that concerns about excessive multicollinearity should take into account whether it is due to two constructs being similar to each other, yet still distinct, or due to one variable being redundant because it taps into the same underlying construct as another variable. Construct redundancy reduces the overall importance of a variable because the importance will be divided among the redundant predictors, thus potentially leading to a misleading result. In contrast, when there is collinearity between similar but distinct constructs, dominance analysis will still correctly partition the variance. In the context of this analysis, school attendance rate and the proportion of Indigenous students are deemed to reflect separate constructs and hence multicollinearity between these variables should not be a concern.

Another issue is the statistical significance of the dominance statistics. Bootstrapping techniques are required to derive the sampling distributions needed to determine how reliable the dominance results are (Azen and Budescu 2003), and whether dominance statistics are significantly different from each other (Tonidandel and LeBreton 2011). Due to the computationally intensive nature of bootstrapping procedures, analysis of this type was not within the timeframe of this project.

Results

Table A.7 presents the dominance analysis results for Year 5 Indigenous and non-Indigenous students in reading and numeracy. These results are discussed in BP 2. Dominance analysis results for Year 3 students in reading and numeracy are presented in annex B.

Table A.7 **Dominance analysis results, by Indigeneity and test domain — Year 5 students**

General dominance statistics

Set	*Variables included*	*Indigenous*		*Non-Indigenous*	
		Reading	Numeracy	Reading	Numeracy
1	Age, gender	0.008	0.001	0.006	0.007
2	Language background	0.016	0.016	0.002	0.001
3	Student SES: mother's/father's highest education level, mother's/father's occupation	0.051	0.057	0.088	0.084
4	Remoteness	0.023	0.019	0.002	0.003
5	State	0.021	0.017	0.002	0.004
6	Percentage Indigenous students	0.040	0.039	0.008	0.009
7	Attendance rate	0.045	0.034	0.012	0.016
8	School SES: percentage of mothers/fathers by highest education level, percentage of mothers/fathers by occupation, fees per student (standardised by school sector) and interacted with school sector	0.039	0.036	0.046	0.047
9	School finances per student: recurrent funding (less school fees), capital income deductions, capital expenditure	0.024	0.021	0.008	0.008
10	Other factors: school sector, combined school indicator, average class size, non-teaching staff per student, number of enrolments, percentage LBOTE students, test participation rate, student mobility indicator	0.037	0.031	0.009	0.011
	R^2 — year 2014 indicator only	0.012	0.000	0.000	0.000
	R^2 — full model	0.315	0.270	0.182	0.190

Source: Commission estimates based on ACARA data (unpublished).

A.5 Blinder-Oaxaca decomposition

The Blinder-Oaxaca decomposition method was used to decompose the gap in average Indigenous and non-Indigenous NAPLAN scores into a portion explained by differences in characteristics and a portion explained by differences in the returns to those characteristics (Blinder 1973; Jann 2008; Oaxaca 1973). This provides insight into whether test scores between Indigenous and non-Indigenous students differ mainly because they have different observed characteristics or because they have different relationships between those characteristics and achievement.

The decomposition for a simple example can be illustrated as follows. Suppose that achievement (A) for Indigenous students (I) and non-Indigenous students (NI) is a function of explanatory factors (X). The difference in mean score (D) between Indigenous and non-Indigenous students can be decomposed into three parts — an endowment effect, a coefficient effect and an interaction term.

$$D = E(A_{NI}) - E(A_I) = E(X_{NI})'\beta_{NI} - E(X_I)'\beta_I \tag{16}$$

$$D = \underbrace{\{[E(X_{NI}) - E(X_I)]'\beta_I\}}_{endowment\ effect} + \underbrace{\{E(X_I)'(\beta_{NI} - \beta_I)\}}_{coefficient\ effect} + \underbrace{\{[E(X_{NI}) - E(X_I)]'(\beta_{NI} - \beta_I)\}}_{interaction\ term} \tag{17}$$

The endowment effect measures the contribution of differences in characteristics to the gap in mean scores. That is, it reflects the expected change in the mean score of Indigenous students if they had the same observed characteristics as non-Indigenous students. The coefficient effect measures the contribution of differences in the coefficients (including the intercept), reflecting the expected change in the mean score of Indigenous students if they had the same coefficients as non-Indigenous students. The interaction term accounts for the fact that differences in characteristics and coefficients between Indigenous and non-Indigenous students can exist simultaneously. This term is an artefact of the decomposition method and cannot be interpreted intuitively.

Alternatively, a twofold decomposition that is often analysed in the discrimination literature apportions the interaction term into the endowment and coefficient effects according to specified weights, thus decomposing the difference in means into an explained component and an unexplained component.

$$D = \underbrace{\{[E(X_{NI}) - E(X_I)]'\beta^*\}}_{explained} + \underbrace{\underbrace{\{E(X_{NI})'(\beta_{NI} - \beta^*)\}}_{unexplained\ NI} + \underbrace{\{E(X_I)'(\beta^* - \beta_I)\}}_{unexplained\ I}}_{unexplained} \tag{18}$$

β^* represents a vector of reference coefficients, and is a weighted average of the coefficients for the two groups (Jann 2008). The explained component is the part that is explained by group differences in characteristics. The unexplained component can be split into a part that measures discrimination in favour of non-Indigenous students and a part that measures discrimination against Indigenous students (assuming that there are no relevant unobserved factors).

In addition to the threefold decomposition, this study considers two extreme scenarios using the twofold decomposition: $\beta^* = \beta_{NI}$ which places the whole interaction term into the endowment effect to form the explained component, and $\beta^* = \beta_I$ which places the interaction term into the coefficient effect to form the unexplained component.

The Blinder-Oaxaca decomposition was implemented in Stata with the user-written command 'oaxaca' (Jann 2008).

Issues and limitations

An issue with analysing Blinder-Oaxaca decompositions using random effects models is that it assumes that the Indigenous and non-Indigenous student groups are independent. However, this is unlikely to be the case when Indigenous and non-Indigenous students come from the same schools. The implication is that standard errors will be understated in the decomposition analysis. Therefore, the Blinder-Oaxaca decompositions in this study use ordinary least squares models with clustering taken into account through the standard errors, rather than estimating random effects or fixed effects. The process of managing clustering involves estimating the joint variance-covariance matrix of all used statistics and then applying a method to approximate the variance and calculate standard errors (Jann 2008).

Another issue is that the decomposition results, particularly the coefficient effect for a specific categorical variable can be sensitive to the choice of omitted default category. In effect, the default category is reflected in the intercept, and the coefficients on each categorical variable are interpreted relative to the default. Changing the default category has implications for the intercept and the estimated coefficient effect specific to that variable. To account for this, the coefficients on categorical variables are transformed to be expressed as deviations from the grand mean and the coefficient for the default category is added (Jann 2008). These results are then independent of the choice of omitted category.

Coefficient effects for particular variables are also sensitive to the scaling of the observed factor because the coefficient effect depends on the location of the mean. Even if estimated coefficients for two different factors with the same units are the same, the variable that has the larger mean will have the larger coefficient effect. To facilitate comparison between variables, continuous variables that were not based on percentages were standardised to have a mean of 0 and a standard deviation of 1.

However, coefficient effects for each factor remain difficult to interpret because the standardisation can result in the mean of a standardised variable for Indigenous students becoming negative if Indigenous students have a lower mean than non-Indigenous students, even if the original variable is strictly positive. This can obscure the true direction of the coefficient effect — a negative coefficient effect could be because the coefficient for Indigenous students is larger than that for non-Indigenous students, or because the mean of the standardised variable for Indigenous students happens to be negative. As a result of the difficulties in interpreting the coefficient effects for specific factors and the intercept, only the total coefficient effect (which is not affected by variable scaling) is discussed in this study. The endowment effects and interaction terms for each factor are also unaffected by variable scaling.

Although the unexplained component in the twofold decomposition is sometimes attributed to discrimination in the literature, it is important to note that it also captures any differences in unobserved factors that could not be included in the analysis (Jann 2008). If Indigenous and non-Indigenous students differ in unobserved factors that are related to observed factors, their effects will be captured as differences in coefficients. Therefore, the coefficients effect and unexplained component are not interpreted in terms of discrimination in this study.

Table A.8 presents results of the Blinder-Oaxaca threefold decomposition of the gap in mean reading and numeracy scores between Indigenous and non-Indigenous students in Year 5 in 2014. These results are discussed in BP 2. Annex B presents results of the twofold decompositions, and results for students in different year levels, years and for different test domains.

Of the 98-point difference in mean reading scores between Indigenous and non-Indigenous students, the threefold decomposition results suggest that about 59 points is because of differences in endowments, 51 points is because of differences in coefficients, and minus 12 points account for simultaneous differences in endowments and coefficients. This suggests that if Indigenous students had the same observed characteristics as non-Indigenous students, the gap in mean reading scores would be expected to be about 59 points smaller, and if Indigenous students had the same coefficients as non-Indigenous students, the gap in mean scores would be expected to be 51 points smaller.

The endowment effects for each particular variable can be interpreted in a similar way. For example, consider the endowment and coefficient effects on gender, which is the first factor in table A.8 that has significant endowment and coefficient effects. If the proportion of Indigenous students who were boys was the same as for non-Indigenous students, the gap in mean reading scores would be expected to be 0.23 points larger, holding all other factors constant. (The small size of this effect is due to the fact that there is a very small difference in the proportion of students who are boys between Indigenous and non-Indigenous students.) As mentioned above, the coefficient effects for each variable are not interpreted in this study because changes in scaling and omitted default categories can result in changes in the coefficient effects.

Because the interaction term is relatively small, the twofold decomposition into explained and unexplained components (presented in annex B) is fairly robust to the choice of β^*. A larger share of the gap tends to be explained — the explained component represents between 48 and 60 per cent of the gap in reading scores, and the unexplained component represents between 40 and 52 per cent of the gap, depending on β^*. The twofold decomposition of the gap in mean numeracy scores also suggests that a slightly larger share of the gap is attributed to differences in characteristics — between 52 and 58 per cent, compared with 42 to 48 per cent unexplained.

Table A.8 **Blinder-Oaxaca threefold decomposition results, by test domain — Year 5 students, 2014**

NAPLAN test score

		Reading		*Numeracy*	
		Coeff.	S.E.	Coeff.	S.E.
Difference in means		97.71***	(10.48)	84.16 ***	(8.33)
Endowment effect		58.53***	(2.62)	48.51 ***	(2.29)
Age		0.23**	(0.10)	0.15 *	(0.08)
Male		-0.23***	(0.09)	0.02	(0.02)
Language background	English	-0.20*	(0.10)	-0.21 **	(0.10)
	LBOTE	-0.45**	(0.18)	-0.41 ***	(0.15)
	Not stated	-0.05	(0.07)	-0.03	(0.05)
Mother's highest education level	Year 9 or below	1.22***	(0.17)	0.88 ***	(0.14)
	Year 10 or 11	1.12***	(0.20)	0.86 ***	(0.17)
	Year 12	0.11	(0.08)	0.06	(0.06)
	Certificate I to IV	-0.04	(0.03)	-0.04	(0.03)
	Advanced diploma / Diploma	0.59**	(0.23)	0.41 **	(0.19)
	Bachelor degree or above	4.60***	(0.76)	3.28 ***	(0.65)
	Not stated	1.21***	(0.31)	0.82 ***	(0.24)
Father's highest education level	Year 9 or below	0.61***	(0.12)	0.52 ***	(0.10)
	Year 10 or 11	0.11	(0.08)	0.19 **	(0.07)
	Year 12	0.26**	(0.12)	0.26 **	(0.10)
	Certificate I to IV	-0.08	(0.13)	0.18	(0.11)
	Advanced diploma / Diploma	0.39	(0.30)	-0.09	(0.25)
	Bachelor degree or above	2.23**	(0.99)	2.69 ***	(0.78)
	Not stated	1.20*	(0.70)	1.16 **	(0.57)
Mother's occupation	Senior management	0.55	(0.36)	0.58 *	(0.31)
	Other business manager	0.42	(0.29)	0.42 *	(0.25)
	Tradesman, clerk, sales, services	0.09	(0.15)	0.36 ***	(0.13)
	Machine operator	-0.03	(0.02)	-0.02	(0.01)
	Not in paid work	0.37***	(0.12)	0.47 ***	(0.11)
	Not stated	1.90***	(0.40)	2.05 ***	(0.34)
Father's occupation	Senior management	1.62***	(0.62)	1.01 **	(0.48)
	Other business manager	1.40***	(0.44)	1.17 ***	(0.37)
	Tradesman, clerk, sales, services	0.65***	(0.19)	0.30 *	(0.16)
	Machine operator	0.09**	(0.04)	0.08 **	(0.03)
	Not in paid work	0.39***	(0.09)	0.25 ***	(0.07)
	Not stated	2.88***	(0.73)	1.60 ***	(0.61)
Same school	No	0.07	(0.08)	0.06	(0.07)
	Yes	0.98***	(0.20)	0.79 ***	(0.16)
	Unknown	0.49**	(0.19)	0.38 **	(0.15)

(continued next page)

Table A.8 (continued)

			Reading		Numeracy
School sector	Government	0.71	(0.87)	0.69	(0.95)
	Independent	0.94	(0.69)	0.63	(0.78)
	Catholic	-0.65	(0.41)	-0.34	(0.44)
Combined school		-0.10	(0.13)	-0.03	(0.08)
Average class size		2.16**	(0.87)	0.56	(0.75)
Non-teaching staff per 100 students		-0.27	(0.59)	0.23	(0.46)
Number of full-time equivalent enrolments		-1.42**	(0.56)	-1.03 **	(0.46)
Percentage Indigenous students	0–5%	10.40***	(2.22)	9.25 ***	(1.79)
	5–10%	-0.51***	(0.14)	-0.48 ***	(0.12)
	10–15%	-0.51**	(0.22)	-0.52 ***	(0.17)
	15–20%	-0.07	(0.18)	-0.21	(0.15)
	20–30%	0.03	(0.25)	0.12	(0.22)
	30–50%	0.75***	(0.24)	0.87 ***	(0.23)
	50–95%	0.98**	(0.41)	0.78 **	(0.31)
	95–100%	2.07**	(0.83)	1.84 ***	(0.56)
Percentage LBOTE students		0.01	(0.06)	0.00	(0.05)
Attendance rate		11.06***	(1.92)	8.11 ***	(1.40)
Recurrent income less fees per student ($100s)		-0.21	(1.52)	-1.10	(0.93)
Capital income deductions per student ($100s)		0.50	(0.62)	-0.04	(0.81)
Capital expenditure per student ($100s)		0.06	(0.06)	0.07	(0.06)
Percentage of mothers by highest education level	Year 9 or below	-0.56	(0.64)	-0.42	(0.54)
	Year 10 or 11	-1.21	(0.92)	-0.14	(0.81)
	Year 12	-0.35*	(0.19)	-0.22	(0.15)
	Certificate I to IV	0.23	(0.18)	0.11	(0.15)
	Advanced diploma / Diploma	0.71	(1.08)	0.36	(0.88)
	Bachelor degree or above	0.04	(3.20)	1.18	(2.83)
	Not stated	-0.71	(1.52)	-0.41	(1.23)
Percentage of fathers by highest education level	Year 9 or below	1.04*	(0.59)	0.34	(0.46)
	Year 10 or 11	0.26	(0.52)	0.16	(0.44)
	Year 12	0.14	(0.45)	-0.32	(0.36)
	Certificate I to IV	0.25	(0.21)	0.22	(0.16)
	Advanced diploma / Diploma	1.00	(1.19)	0.73	(1.00)
	Bachelor degree or above	1.41	(3.58)	0.53	(2.94)
	Not stated	0.08	(3.21)	-0.23	(2.68)
Percentage of mothers by occupation	Senior management	-0.09	(1.78)	-0.56	(1.50)
	Other business manager	-1.36	(1.85)	-0.61	(1.04)
	Tradesman, clerk, sales, services	0.50	(0.55)	0.18	(0.44)
	Machine operator	-0.26	(0.28)	0.07	(0.22)
	Not in paid work	0.21	(0.27)	-0.08	(0.21)
	Not stated	0.00	(1.97)	-1.56	(1.51)

(continued next page)

Table A.8 (continued)

		Reading		Numeracy	
Percentage of fathers by occupation	Senior management	-0.13	(2.63)	1.44	(2.16)
	Other business manager	2.02	(1.72)	2.42 *	(1.26)
	Tradesman, clerk, sales, services	-0.71	(0.55)	-0.26	(0.42)
	Machine operator	0.78*	(0.43)	0.54	(0.35)
	Not in paid work	-0.30	(0.22)	0.08	(0.18)
	Not stated	0.45	(3.31)	1.90	(2.73)
Fees per student, standardised by school sector		2.29	(2.24)	1.63	(2.36)
Fees x Government sector		-0.48	(1.85)	-1.61	(2.03)
Fees x Independent		0.39	(0.26)	0.38	(0.29)
Fees x Catholic		-0.37*	(0.21)	-0.25	(0.18)
Test participation rate		0.51	(0.62)	-0.29	(0.50)
State	NSW	0.00	(0.01)	0.00	(0.01)
	VIC	1.69*	(0.88)	3.28 ***	(0.75)
	QLD	-1.32***	(0.34)	-0.48 **	(0.24)
	SA	-0.17	(0.11)	-0.28 ***	(0.10)
	WA	0.24*	(0.13)	0.17 *	(0.10)
	TAS	-0.04	(0.05)	-0.06	(0.04)
	NT	0.63	(0.50)	0.23	(0.35)
	ACT	0.00	(0.07)	-0.05	(0.06)
Remoteness	Metro	1.00	(0.89)	0.81	(0.73)
	Provincial	-0.49	(0.44)	-0.42	(0.32)
	Remote	0.09	(0.20)	0.14	(0.14)
	Very remote	0.49	(0.58)	0.29	(0.38)
Coefficient effect		50.95***	(10.75)	40.40 ***	(8.93)
Age		-0.02	(0.10)	-0.03	(0.08)
Male		2.48***	(0.73)	5.37 ***	(0.61)
Language background	English	-1.82	(1.53)	-4.83 ***	(1.30)
	LBOTE	0.21	(0.41)	1.59 ***	(0.35)
	Not stated	0.06	(0.16)	-0.16	(0.13)
Mother's highest education level	Year 9 or below	-0.33	(0.24)	-0.42 **	(0.20)
	Year 10 or 11	-0.34	(0.40)	-0.48	(0.34)
	Year 12	-0.19	(0.20)	-0.06	(0.17)
	Certificate I to IV	-1.47***	(0.38)	-1.54 ***	(0.32)
	Advanced diploma / Diploma	-0.09	(0.17)	-0.07	(0.14)
	Bachelor degree or above	0.15	(0.18)	0.25 *	(0.15)
	Not stated	2.64***	(0.53)	2.18 ***	(0.42)
Father's highest education level	Year 9 or below	-0.12	(0.20)	0.00	(0.17)
	Year 10 or 11	-0.81***	(0.30)	-0.35	(0.25)
	Year 12	-0.11	(0.14)	-0.16	(0.12)
	Certificate I to IV	-0.09	(0.37)	-0.71 **	(0.31)
	Advanced diploma / Diploma	0.01	(0.12)	0.15	(0.10)
	Bachelor degree or above	0.24*	(0.13)	0.13	(0.10)
	Not stated	0.50	(1.16)	-0.18	(0.95)

(continued next page)

Table A.8 (continued)

			Reading		Numeracy
Mother's occupation	Senior management	-0.11	(0.16)	-0.18	(0.14)
	Other business manager	-0.19	(0.16)	-0.20	(0.13)
	Tradesman, clerk, sales, services	-0.02	(0.23)	-0.34 *	(0.19)
	Machine operator	-0.72***	(0.25)	-0.38 *	(0.21)
	Not in paid work	1.63***	(0.53)	1.92 ***	(0.45)
	Not stated	2.35***	(0.70)	2.69 ***	(0.60)
Father's occupation	Senior management	-0.01	(0.14)	0.02	(0.11)
	Other business manager	-0.17	(0.15)	-0.15	(0.13)
	Tradesman, clerk, sales, services	-0.73***	(0.25)	-0.24	(0.21)
	Machine operator	0.32	(0.32)	0.29	(0.27)
	Not in paid work	0.41*	(0.22)	0.18	(0.17)
	Not stated	1.71	(1.24)	0.28	(1.05)
Same school	No	0.38	(0.33)	0.13	(0.28)
	Yes	-0.40	(0.91)	-0.16	(0.75)
	Unknown	-0.18	(0.33)	-0.05	(0.27)
School sector	Government	3.25	(3.69)	3.97	(3.98)
	Independent	-0.10	(0.21)	0.01	(0.23)
	Catholic	-0.04	(0.32)	-0.42	(0.33)
Combined school		-0.32	(0.84)	0.05	(0.65)
Average class size		0.77	(0.82)	-0.85	(0.72)
Non-teaching staff per 100 students		-0.92	(0.65)	-0.48	(0.53)
Number of full-time equivalent enrolments		-1.03*	(0.56)	-1.21 **	(0.49)
Percentage Indigenous students	0–5%	-0.57	(0.59)	-0.77	(0.66)
	5–10%	-0.67	(0.79)	-1.06	(0.99)
	10–15%	0.03	(0.61)	-0.36	(0.77)
	15–20%	0.21	(0.40)	-0.12	(0.51)
	20–30%	0.49	(0.51)	0.16	(0.65)
	30–50%	0.91**	(0.44)	1.16 **	(0.53)
	50–95%	0.95	(0.64)	1.11 *	(0.61)
	95–100%	-1.76	(2.43)	-0.89	(3.42)
Percentage LBOTE students		-1.31	(1.62)	0.95	(1.21)
Attendance rate		-0.19	(36.68)	85.35 **	(32.94)
Recurrent income less fees per student ($100s)		0.45	(1.61)	0.23	(1.13)
Capital income deductions per student ($100s)		0.82	(0.63)	0.15	(0.81)
Capital expenditure per student ($100s)		0.07	(0.07)	0.06	(0.06)
Percentage of mothers by highest education level	Year 9 or below	-2.65*	(1.41)	-2.20 *	(1.27)
	Year 10 or 11	-8.51***	(2.62)	-4.46 *	(2.33)
	Year 12	5.20***	(1.90)	3.63 **	(1.65)
	Certificate I to IV	2.65	(3.10)	-0.72	(2.57)
	Advanced diploma / Diploma	-0.39	(1.59)	0.52	(1.31)
	Bachelor degree or above	2.25	(2.22)	0.94	(1.98)
	Not stated	1.29	(2.84)	1.64	(2.40)

(continued next page)

Table A.8 (continued)

		Reading		Numeracy	
Percentage of fathers by highest education level	Year 9 or below	2.49*	(1.50)	0.09	(1.23)
	Year 10 or 11	0.64	(2.20)	0.93	(1.87)
	Year 12	0.00	(1.59)	1.67	(1.31)
	Certificate I to IV	-6.85**	(3.28)	-5.79 **	(2.65)
	Advanced diploma / Diploma	-1.63	(1.29)	-1.25	(1.09)
	Bachelor degree or above	0.20	(1.77)	0.94	(1.47)
	Not stated	2.25	(6.66)	0.77	(5.63)
Percentage of mothers by occupation	Senior management	-0.02	(1.48)	0.19	(1.24)
	Other business manager	2.13	(2.21)	0.40	(1.26)
	Tradesman, clerk, sales, services	-3.54	(2.27)	-1.09	(1.84)
	Machine operator	-0.12	(2.02)	2.25	(1.65)
	Not in paid work	-0.52	(3.07)	-3.08	(2.41)
	Not stated	0.94	(4.16)	-1.74	(3.26)
Percentage of fathers by occupation	Senior management	0.11	(1.54)	-0.95	(1.27)
	Other business manager	-1.67	(1.78)	-1.38	(1.33)
	Tradesman, clerk, sales, services	4.23*	(2.26)	0.87	(1.79)
	Machine operator	1.65	(2.32)	0.86	(1.87)
	Not in paid work	-1.12	(1.05)	0.80	(0.90)
	Not stated	-1.09	(7.19)	2.03	(5.98)
Fees per student, standardised by school sector		0.30	(2.05)	-0.40	(2.16)
Fees x Government sector		0.61	(1.78)	-0.57	(1.94)
Fees x Independent		0.13	(0.17)	0.14	(0.19)
Fees x Catholic		-0.17	(0.12)	-0.10	(0.11)
Test participation rate		-29.59*	(17.50)	-0.78	(12.92)
State	NSW	-0.39	(0.93)	0.11	(0.77)
	VIC	-0.26	(0.19)	-0.50 ***	(0.17)
	QLD	-3.94***	(0.88)	-0.79	(0.71)
	SA	-0.15	(0.19)	-0.03	(0.15)
	WA	0.57	(0.49)	0.79 *	(0.41)
	TAS	0.06	(0.14)	-0.08	(0.11)
	NT	1.73***	(0.61)	0.94 **	(0.46)
	ACT	-0.01	(0.07)	0.01	(0.05)
Remoteness	Metro	-3.69***	(1.14)	-3.40 ***	(0.97)
	Provincial	-2.30**	(1.16)	-1.49 *	(0.88)
	Remote	0.27	(0.27)	0.29	(0.20)
	Very remote	1.34*	(0.71)	0.96 *	(0.55)
Constant		82.27**	(38.13)	-42.08	(33.71)
Interaction effect		-11.78***	(3.52)	-4.75	(4.41)
Age		0.02	(0.10)	0.03	(0.08)
Male		0.07**	(0.03)	0.15 **	(0.06)
Language background	English	0.09	(0.08)	0.25 **	(0.11)
	LBOTE	0.08	(0.15)	0.60 ***	(0.18)
	Not stated	-0.03	(0.07)	0.07	(0.06)

(continued next page)

Table A.8 (continued)

			Reading		Numeracy
Mother's highest education level	Year 9 or below	0.23	(0.17)	0.29 **	(0.14)
	Year 10 or 11	0.17	(0.20)	0.24	(0.17)
	Year 12	-0.08	(0.08)	-0.02	(0.07)
	Certificate I to IV	0.06	(0.04)	0.07 *	(0.04)
	Advanced diploma / Diploma	-0.13	(0.23)	-0.10	(0.20)
	Bachelor degree or above	0.65	(0.76)	1.07 *	(0.65)
	Not stated	-1.66***	(0.34)	-1.36 ***	(0.27)
Father's highest education level	Year 9 or below	0.07	(0.12)	0.00	(0.10)
	Year 10 or 11	0.24***	(0.09)	0.11	(0.07)
	Year 12	-0.09	(0.12)	-0.14	(0.10)
	Certificate I to IV	-0.03	(0.14)	-0.26 **	(0.12)
	Advanced diploma / Diploma	0.02	(0.30)	0.38	(0.25)
	Bachelor degree or above	1.87*	(0.99)	1.00	(0.79)
	Not stated	-0.31	(0.72)	0.11	(0.59)
Mother's occupation	Senior management	-0.27	(0.36)	-0.42	(0.31)
	Other business manager	-0.35	(0.29)	-0.38	(0.25)
	Tradesman, clerk, sales, services	-0.01	(0.16)	-0.23 *	(0.13)
	Machine operator	0.04	(0.03)	0.02	(0.02)
	Not in paid work	-0.36***	(0.12)	-0.43 ***	(0.11)
	Not stated	-1.37***	(0.41)	-1.56 ***	(0.35)
Father's occupation	Senior management	-0.06	(0.63)	0.08	(0.48)
	Other business manager	-0.50	(0.44)	-0.44	(0.37)
	Tradesman, clerk, sales, services	-0.57***	(0.19)	-0.19	(0.16)
	Machine operator	-0.03	(0.03)	-0.03	(0.03)
	Not in paid work	-0.16*	(0.09)	-0.07	(0.07)
	Not stated	-1.03	(0.75)	-0.17	(0.63)
Same school	No	-0.10	(0.09)	-0.03	(0.08)
	Yes	-0.09	(0.20)	-0.03	(0.16)
	Unknown	0.10	(0.19)	0.03	(0.15)
School sector	Government	-0.78	(0.89)	-0.96	(0.97)
	Independent	-0.32	(0.69)	0.05	(0.78)
	Catholic	-0.06	(0.40)	-0.56	(0.44)
Combined school		0.05	(0.12)	-0.01	(0.08)
Average class size		-0.84	(0.89)	0.92	(0.78)
Non-teaching staff per 100 students		0.99	(0.70)	0.52	(0.57)
Number of full-time equivalent enrolments		1.06*	(0.58)	1.24 **	(0.50)
Percentage Indigenous students	0–5%	-2.90	(3.00)	-3.92	(3.36)
	5–10%	0.13	(0.16)	0.22	(0.21)
	10–15%	-0.02	(0.35)	0.21	(0.45)
	15–20%	-0.15	(0.29)	0.09	(0.37)
	20–30%	-0.41	(0.43)	-0.13	(0.54)
	30–50%	-0.84**	(0.41)	-1.06 **	(0.49)
	50–95%	-0.93	(0.62)	-1.08 *	(0.59)
	95–100%	1.76	(2.43)	0.89	(3.42)

(continued next page)

Table A.8 (continued)

		Reading		Numeracy	
Percentage LBOTE students		-0.05	(0.09)	0.04	(0.07)
Attendance rate		-0.01	(2.14)	4.87 **	(1.90)
Recurrent income less fees per student ($100s)		-0.48	(1.70)	-0.24	(1.20)
Capital income deductions per student ($100s)		-0.82	(0.64)	-0.15	(0.82)
Capital expenditure per student ($100s)		-0.08	(0.07)	-0.07	(0.06)
Percentage of mothers by highest education level	Year 9 or below	1.41*	(0.75)	1.17 *	(0.68)
	Year 10 or 11	3.22***	(1.00)	1.70 *	(0.89)
	Year 12	0.49**	(0.20)	0.32 **	(0.16)
	Certificate I to IV	-0.15	(0.18)	0.04	(0.16)
	Advanced diploma / Diploma	-0.27	(1.10)	0.36	(0.91)
	Bachelor degree or above	3.30	(3.25)	1.37	(2.89)
	Not stated	-0.73	(1.60)	-0.91	(1.34)
Percentage of fathers by highest education level	Year 9 or below	-1.09*	(0.66)	-0.04	(0.54)
	Year 10 or 11	-0.16	(0.54)	-0.23	(0.47)
	Year 12	0.00	(0.47)	0.49	(0.38)
	Certificate I to IV	-0.43*	(0.23)	-0.34 *	(0.18)
	Advanced diploma / Diploma	-1.53	(1.21)	-1.18	(1.02)
	Bachelor degree or above	0.42	(3.64)	1.93	(3.01)
	Not stated	-1.12	(3.31)	-0.38	(2.78)
Percentage of mothers by occupation	Senior management	-0.03	(1.81)	0.23	(1.52)
	Other business manager	1.79	(1.86)	0.34	(1.06)
	Tradesman, clerk, sales, services	-0.88	(0.56)	-0.27	(0.45)
	Machine operator	0.02	(0.29)	-0.33	(0.24)
	Not in paid work	0.05	(0.28)	0.28	(0.23)
	Not stated	-0.46	(2.03)	0.84	(1.58)
Percentage of fathers by occupation	Senior management	0.18	(2.66)	-1.65	(2.19)
	Other business manager	-1.63	(1.74)	-1.34	(1.29)
	Tradesman, clerk, sales, services	1.04*	(0.56)	0.21	(0.43)
	Machine operator	-0.31	(0.44)	-0.16	(0.36)
	Not in paid work	0.24	(0.22)	-0.17	(0.19)
	Not stated	0.51	(3.39)	-0.95	(2.81)
Fees per student, standardised by school sector		-0.33	(2.26)	0.44	(2.38)
Fees x Government sector		-0.64	(1.87)	0.61	(2.05)
Fees x Independent		-0.19	(0.26)	-0.21	(0.28)
Fees x Catholic		0.26	(0.19)	0.16	(0.17)
Test participation rate		-1.10*	(0.65)	-0.03	(0.53)
State	NSW	-0.39	(0.93)	0.00	(0.01)
	VIC	-0.26	(0.19)	-2.33 ***	(0.77)
	QLD	-3.94***	(0.88)	0.27	(0.25)
	SA	-0.15	(0.19)	-0.02	(0.09)
	WA	0.57	(0.49)	-0.15	(0.10)
	TAS	0.06	(0.14)	0.03	(0.04)
	NT	1.73***	(0.61)	-0.87 **	(0.43)
	ACT	-0.01	(0.07)	0.01	(0.06)
Remoteness	Metro	-3.69***	(1.14)	-2.93 ***	(0.84)
	Provincial	-2.30**	(1.16)	0.62 *	(0.37)
	Remote	0.27	(0.27)	-0.24	(0.17)
	Very remote	1.34*	(0.71)	-0.93 *	(0.53)

*** statistically significant at the 1 per cent level, ** 5 per cent level, * 10 per cent level

Source: Commission estimates based on ACARA data (unpublished).

Annex B — Modelling results table

An Excel spreadsheet is available online at:
www.pc.gov.au/research/completed/indigenous-primary-school-achievement

References

Aaronson, D., Barrow, L. and Sander, W. 2007, 'Teachers and student achievement in the Chicago public high schools', *Journal of Labour Economics*, vol. 25, no. 1, pp. 95–135.

Abedi, J. 2003, *Impact of Student Language Background on Content-Based Performance: Analyses of Extant Data*, Center for the Study of Evaluation Report, 603, National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles.

ABS (Australian Bureau of Statistics) 2014, *2033.0.55.001 - Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia, 2011*, Australian Bureau of Statistics, www.abs.gov.au/ausstats/abs@.nsf/Lookup/2033.0.55.001 main+features100042011 (accessed 16 March 2016).

—— 2015, *6227.0 - Education and Work, Australia, May 2015*, www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/6227.0Explanatory%20Notes1May%202015?OpenDocument (accessed 28 January 2016).

ACARA (Australian Curriculum, Assessment and Reporting Authority) 2014a, *ICSEA 2013: Technical Report*, www.acara.edu.au/verve/_resources/ICSEA_2013_Generation_Report.pdf (accessed 16 November 2015).

—— 2014b, *National Assessment Program - Literacy and Numeracy 2013: Technical Report*, Sydney, www.nap.edu.au/results-and-reports/national-reports.html (accessed 16 November 2015).

—— 2014c, *National Assessment Program — Literacy and Numeracy: National Report for 2014*, ACARA, Sydney.

—— 2015a, *Financial Data Report Methodology*, www.acara.edu.au/verve/_resources/Financial_Data_Reporting_Methodology_2014.pdf (accessed 8 December 2015).

—— 2015b, *NAPLAN Online*, www.nap.edu.au/online-assessment/naplan-online/naplan-online.html (accessed 16 November 2015).

—— 2015c, *National Assessment Program - Literacy and Numeracy 2014: Technical Report*, Sydney, www.nap.edu.au/results-and-reports/national-reports.html (accessed 16 November 2015).

—— nd, *Glossary*, Australian Curriculum, Assessment and Reporting Authority, www.acara.edu.au/myschool/myschool_glossary.html (accessed 9 March 2016).

AEU (Australian Education Union) 2010, *Submission to the Senate Education, Employment and Workplace Relations Committee Inquiry into the administration and reporting of NAPLAN testing*, Submission no. 231, www.aph.gov.au/Parliamentary_Business/Committees/Senate/Education_Employment_and_Workplace_

Relations/Completed_inquiries/2010-13/naplan/submissions (accessed 16 November 2015).

Arceneaux, K. and Nickerson, D.W. 2009, ‘Modeling Certainty with Clustered Data: A Comparison of Methods’, *Political Analysis*, vol. 17, no. 2, pp. 177–190.

Australian Government 2010, *Administration and Reporting of NAPLAN Testing*, Final Report, Senate Standing Committee on Education and Employment, Canberra.

—— 2013, *Answers to questions on notice from the Australian Curriculum, Assessment and Reporting Authority No. 6*, 1 August, The Effectiveness of the National Assessment Program - Literacy and Numeracy, Senate Standing Committee on Education and Employment, Canberra.

—— 2014, *The Effectiveness of the National Assessment Program - Literacy and Numeracy*, Final Report, Senate Standing Committee on Education and Employment, Canberra.

—— 2015, *Australian Indigenous Cultural Heritage*, australia.gov.au, www.australia.gov.au/about-australia/australian-story/austn-indigenous-cultural-heritage (accessed 10 March 2016).

Azen, R. and Budescu, D.V. 2003, ‘The dominance analysis approach for comparing predictors in multiple regression’, *Psychological Methods*, vol. 8, no. 2, pp. 129–148.

Bagshaw, E. 2015, ‘NAPLAN: students want harded exams’, *The Sydney Morning Herald*, 12 May, www.smh.com.au/interactive/2015/Reenie/NAPLAN/index.html (accessed 12 November 2015).

Betebenner, D.W. 2008, *Norm- and Criterion-Referenced Student Growth*, National Center for the Improvement of Educational Assessment, New Hampshire.

Bita, N. 2015, ‘Tutors paid to boost NAPLAN’, *The Australian*, 17 August, www.theaustralian.com.au/national-affairs/education/parents-desperate-for-results-turn-to-tutors-for-naplan-tests/story-fn59nlz9-1227486079198?sv=82a0e0abfc9253e5723155530db6d6d8 (accessed 12 November 2015).

Blinder, A.S. 1973, ‘Wage discrimination: reduced form and structural estimates’, *The Journal of Human Resources*, vol. 8, no. 4, pp. 436–455.

Biddle, N. and Cameron, T. 2012, *Potential Factors Influencing Indigenous Education Participation and Achievement*, National Centre for Vocational Education Research, Adelaide.

Boon, H.J. 2011, ‘School moves, coping, and achievement: models of possible interactions’, *Journal of Educational Research*, vol. 104, no. 1, pp. 54–70.

Chaplin, D. 2003, *Hierarchical Linear Models: Strengths and Weaknesses*, Paper prepared for the November 2003 meetings of the Association for Public Policy and Management.

Clarke, P., Crawford, C., Steele, F. and Vignoles, A. 2010, *The Choice Between Fixed and Random Effects Models: Some Considerations for Educational Research*, Discussion Paper, 5287, Institute for the Study of Labor, Bonn, Germany.

Connor, M.J. 2001, ‘Pupil stress and standard assessment tasks’, *Emotional and Behavioural Difficulties*, vol. 6, no. 2, pp. 103–111.

Bourke, P.C. 2008, *The Experiences of Teacher Aides who Support Students with Disabilities and Learning Difficulties: A Phenomenological Study*, Queensland University of Technology.

Chevalier, A. 2004, *Parental Education and Child’s Education: A Natural Experiment*, Discussion Paper, 1153, Institute for the Study of Labour.

COAG (Governments) 2011, *National Indigenous Reform Agreement*.

Cobbold, T. 2015, *A Review of Academic Studies of Public and Private School Outcomes in Australia*, Save Our Schools.

Daraganova, G., Mullan, K. and Edwards, B. 2014, *Attendance in Primary School: Factors and Consequences*, Occasional Paper, 51, Department of Social Services.

Davis-Kean, P.E. 2005, ‘The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment’, *Journal of Family Psychology*, vol. 19, no. 2, pp. 294–304.

De Bortoli, L. and Cresswell, J. 2004, *Australia’s Indigenous Students in PISA 2000: Results from an International Study*, ACER Research Monograph, No. 59, Australian Council for Educational Research, Camberwell.

—— and Thompson, S. 2010, *Contextual Factors that Influence the Achievement of Australia’s Indigenous Students: Results from PISA 2000–2006*, Australian Council for Educational Research, Camberwell.

Deloitte 2015, *Letter of Accounting Advice to ACARA*, www.acara.edu.au/verve/_resources/Deloitte_Accounting_Letter_2015.pdf (accessed 8 December 2015).

Devine, A., Fawcett, K., Szűcs, D. and Dowker, A. 2012, ‘Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety’, *Behavioural and Brain Functions*, vol. 8, no. 33, pp. 1–9.

Ehrenberg, R.G., Brewer, D.J., Gamoran, A. and Willms, J.D. 2001, ‘Class size and student achievement’, *Psychological Science in the Public Interest*, vol. 2, no. 1, pp. 1–30.

Elliott, M. 1998, ‘School finance and opportunities to learn: does money well spent enhance students’ achievement?’, *Sociology of Education*, vol. 71, no. 3, pp. 223–245.

Fergusson, D.M. and Horwood,, L.J. 1997, ‘Gender differences in educational achievement in a New Zealand birth cohort’, *New Zealand Journal of Educational Studies*, vol. 32, no. 1, pp. 83–96.

Fertig, M. and Kluve, J. 2005, *The Effect of Age at School Entry on Educational Attainment in Germany*, Discussion Paper, 1507, Institute for the Study of Labor, Bonn, Germany.

Gemici, S., Lim, P. and Karmel, T. 2013, *The Impact of Schools on Young People’s Transition to University*, Longitudinal surveys of Australian youth research report, 61, National Centre for Vocational Education Research, Adelaide, www.lsay.edu.au/publications/2541.html (accessed 11 November 2015).

Greene, W. 2005, *Efficiency of Public Spending in Developing Countries: A Stochastic Frontier Approach*, World Bank.

—— 2008, *Econometric Analysis*, 6th edn., Pearson Prentice Hall, New Jersey.

Greenwald, R., Hedges, L.V. and Laine, R.D. 1996, 'The effect of school resources on student achievement', *Review of Educational Research*, vol. 66, no. 3, pp. 361–396.

Grissom, J.B. 2004, 'Age and achievement', *Education Policy Analysis Archives*, vol. 12, no. 49, pp. 1–24.

Hancock, K.J., Carrington, C.C.J., Lawrence, D. and Zubrick, S. 2013, *Student Attendance and Educational Outcomes: Every Day Counts*, University of Western Australia, http://telethonkids.org.au/mwg-internal/de5fs23hu73ds/progress?id=SJdHXGJ4Tm (accessed 17 August 2015).

Hanushek, E. 1979, 'Conceptual and empirical issues in the estimation of educational production functions', *The Journal of Human Resources*, vol. 14, no. 3, pp. 351–388.

—— 1986, 'The economics of schooling: production and efficiency in public schools', *Journal of Economic Literature*, vol. 24, no. 3, pp. 1141–1177.

—— 1997, 'Assessing the effects of school resources on student performance: an update', *Educational Evaluation and Policy Analysis*, vol. 19, no. 2, pp. 141–164.

Harris, P., Chinnappan, M., Carter, J., de Courcy, J. and Barnett, J. 2013, *Impact and consequence of Australia's NAPLAN – Using research evidence to inform improvement*, Submission to Australia's 2013 Senate Committee Inquiry into the Effectiveness of the National Assessment Program – Literacy and Numeracy (NAPLAN), School of Education, University of South Australia, www.aph.gov.au/Parliamentary_Business/Committees/Senate/Education_and_Employment/Naplan13/Submissions (accessed 16 November 2015).

Hattie, J. 2003, *Teachers Make a Difference: What is the Research Evidence?*, Paper presented at the ACER Annual Conference on Building Teacher Quality, Melbourne.

Holmes-Smith, P. 2012, *Using NAPLAN Summative Data to Identify Instructional Successes and Challenges*, School Research Evaluation and Measurement Services, www.sreams.com.au/naplan.html (accessed 16 November 2015).

Houng, B. and Justman, M. 2013, *Comparing Least-squares Value-added Analysis and Student Growth Percentile Analysis for Evaluating Student Progress and Estimating School Effects*, Melbourne Institute Working Paper, 7/13, Melbourne Institute of Applied Economic and Social Research, www.melbourneinstitute.com/downloads/working_paper_series/wp2013n07.pdf (accessed 13 November 2015).

Howley, C. 1996, 'Compounding disadvantage: the effects of school and district size on student achievement in West Virginia', *Journal of Research in Rural Education*, vol. 12, no. 1, pp. 25–32.

Hoxby, C.M. 2000a, *Peer Effects in the Classroom: Learning from Gender and Race Variation*, NBER Working Paper, 7867, National Bureau of Economic Research.

—— 2000b, 'The effects of class size on student achievement: new evidence from population variation', *The Quarterly Journal of Economics*, vol. 115, no. 4, pp. 1239–85.

—— and Weingarth, G. 2005, *Taking Race out of the Equation: School Reassignment and the Structure of Peer Effects*, Working Paper.

Jann, B. 2008, 'The Blinder-Oaxaca Decomposition for linear regression models', *The Stata Journal*, vol. 8, no. 4, pp. 453–479.

Kuziemko, I. 2006, 'Using shocks to school enrolment to estimate the effect of school size on student achievement', *Economics of Education Review*, vol. 25, no. 1, pp. 63–75.

Lamb, S. 2015, 'Schools, funding and performance: lessons from the NSW National Partnerships', www.cese.nsw.gov.au/publications-filter/schools-funding-and-performance-lessons-from-the-nsw-national-partnerships (accessed 29 January 2016).

—— and Fullarton, S. 2001, *Classroom and School Factors Affecting Mathematics Achievement: a Comparative Study of the US and Australia using TIMSS*, TIMSS Australia Monograph Series, Australian Council for Educational Research.

——, Glover, S. and Walstab, A. 2014, *Session K - Educational Disadvantage in Regional and Rural Schools*, 2009 - 2015 ACER Research Conferences, 14.

——, Rumberger, R., Jesson, D. and Teese, R. 2004, *School Performance in Australia: Results from Analyses of School Effectiveness*, Centre for Post-compulsory Education and Lifelong Learning, www.eduweb.vic.gov.au/edulibrary/…/schoolperformance-rpt.pdf (accessed 18 August 2015).

Lee, V.E. and Loeb, S. 2000, 'School size in Chicago elementary schools: effects on teachers' attitudes and students' achievement', *American Educational Research Journal*, vol. 37, no. 1, pp. 3–31.

Leigh, A. 2010, 'Estimating teacher effectiveness from two-year changes in students' test scores', *Economics of Education Review*, vol. 29, pp. 480–488.

Leithwood, K. and Jantzi, D. 2007, *Review of Empirical Evidence about School Size Effects*, Board of Education of the Regina School Division No. 4 of Saskatchewan.

Lokan, J., Greenwood, L. and Cresswell, J. 2008, *15-Up and Counting, Reading, Writing, Reasoning: How Literate are Australia's Students?: The PISA 2000 Survey of Students' Reading, Mathematical and Scientific Literacy Skills*, Australian Council for Educational Research, http://research.acer.edu.au/indigenous_education/7/ (accessed 11 November 2015).

Lovell, C.A.K. 1993, 'Production frontiers and productive efficiency', in Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (eds), *The Measurement of Productive Efficiency*, Oxford University Press, New York.

Lu, L. and Rickard, K. 2016, *Mobility of Students in NSW Government Schools*, Centre for Education Statistics and Evaluation.

Luchman, J.N. 2015, *DOMIN: Stata Module to Conduct Dominance Analysis*, https://ideas.repec.org/c/boc/bocode/s457629.html (accessed 11 February 2016).

Mahuteau, S., Karmel, T., Mavromaras, K. and Zhu, R. 2015, *Educational Outcomes of Young Indigenous Australians*, National Institute of Labour Studies, Flinders University, Adelaide.

Marks, G., McMillan, J. and Hillman, K. 2001, *Tertiary Entrance Performance: The Role of Student Background and School Factors*, Longitudinal surveys of Australian youth research report, 22, LSAY Research Reports, Australian Council for Educational Research, www.lsay.edu.au/publications/1869.html (accessed 11 November 2015).

Marks, G.N. 2010, 'What aspects of schooling are important? School effects on tertiary entrance performance', *School Effectiveness and School Improvement*, vol. 21, no. 3, pp. 267–287.

Morgan, S.L. and Sørensen, A.B. 1999, 'Parental networks, social closure, and mathematics learning: a test of Coleman's social capital explanation of school effects', *American Sociological Review*, vol. 64, no. 5, pp. 661–681.

Morris, L. 2013, 'Parents turn NAPLAN into bestseller', *The Sydney Morning Herald*, 9 May, www.smh.com.au/national/education/parents-turn-naplan-into-bestseller-20130508-2j83n.html (accessed 12 November 2015).

Nakagawa, S. and Schielzeth, H. 2013, 'A general and simple method for obtaining R2 from generalized linear mixed-effects models', *Methods in Ecology and Evolution*, vol. 4, no. 2, pp. 133–142.

Nous Group 2011, *Schooling Challenges and Opportunities: A Report for the Review of Funding for Schooling Panel*, Nous Group, Melbourne.

Oaxaca, R. 1973, 'Male-female wage differentials in urban labor markets', *International Economic Review*, vol. 14, no. 3, pp. 693–709.

PC (Productivity Commission) 2015, *National Indigenous Reform Agreement, Performance Assessment 2013-14*, Canberra.

Perry, L. and McConney, A. 2010, 'Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003', *Teachers College Record*, vol. 112, no. 4, pp. 1137–1162.

Plug, E. and Vijverberg, W. 2003, 'Schooling, family background, and adoption: Is it nature or is it nurture?', *Journal of Political Economy*, vol. 111, no. 3, pp. 611–641.

Pomerantz, E.M., Altermatt, E.R. and Saxon, J.L. 2002, 'Making the grade but feeling distressed: gender differences in academic performance and internal distress', *Journal of Educational Psychology*, vol. 94, no. 2, pp. 396–404.

Rabe-Hesketh, S. and Skrondal, A. 2008, *Multilevel and Longitudinal Modelling Using Stata*, 2nd edn., Stata Press.

Raudenbush, S.W. and Bryk, A.S. 2002, *Hierarchical Linear Models: Applications and Data Analysis Methods*, Sage Publications, California.

Rogers, M.A., Theule, J., Ryan, B.A., Adams, G.R. and Keating, L. 2009, ‘Parental involvement and children’s school achievement’, *Canadian Journal of School Psychology*, vol. 24, no. 1, pp. 34–57.

Rothman, S. 2001, ‘School absence and student background factors: a multilevel analysis’, *International Education Journal*, vol. 2, no. 1, pp. 59–68.

—— and McMillan, J. 2003, *Influences on Achievement in Literacy and Numeracy*, Longitudinal surveys of Australian youth research report, 36, LSAY Research Reports, http://research.acer.edu.au/lsay_research/40/ (accessed 11 November 2015).

Schulz, W. 2005, *Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003*, Paper prepared for the Annual Meetings of the American Educational Research Association in San Francisco, Australian Council for Educational Research.

Song, S., Perry, L.B. and McConney, A. 2014, ‘Explaining the achievement gap between Indigenous and non-Indigenous students: an analysis of PISA 2009 results for Australia and New Zealand’, *Educational Research and Evaluation: An International Journal on Theory and Practice*, vol. 20, no. 3, pp. 178–198.

Todd, P.E. and Wolpin, K.I. 2003, ‘On the specification and estimation of the production function for cognitive achievement’, *The Economic Journal*, vol. 113, pp. F3–F33.

Tonidandel, S. and LeBreton, J.M. 2011, ‘Relative importance analysis: a useful supplement to regression analysis’, *Journal of Business and Psychology*, vol. 26, no. 1, pp. 1–9.

Topsfield, J. 2013, ‘Toys enter NAPLAN mix as the tests loom’, *The Sydney Morning Herald*, 11 May, www.smh.com.au/national/education/toys-enter-naplan-mix-as-the-tests-loom-20130510-2jde1.html (accessed 12 November 2015).

Van Ewijk, R. and Sleegers, P. 2010, ‘The effect of peer socioeconomic status on student achievement: a meta-analysis’, *Educational Research Review*, vol. 5, no. 2, pp. 134–150.

VCAA (Victorian Curriculum and Assessment Authority) 2013, *NAPLAN Frequently Asked Questions*, www.vcaa.vic.edu.au/Pages/prep10/naplan/schools/faqs.aspx (accessed 7 December 2015).

Victorian Department of Education and Training 2016, *Primary School Types*, Education Website, www.education.vic.gov.au/school/parents/primary/Pages/types.aspx (accessed 11 March 2016).

Voyer, D. and Voyer, S.D. 2014, ‘Gender differences in scholastic achievement: a meta-analysis’, *Psychological Bulletin*, vol. 140, no. 4, pp. 1174–1204.

Wiese, B.S. and Freund, A.M. 2011, ‘Parents as role models: parental behaviour affects adolescents’ plans for work involvement’, *International Journal of Behavioral Development*, vol. 35, no. 3, pp. 218–224.

Wigglesworth, G., Simpson, J. and Loakes, D. 2011, ‘NAPLAN language assessments for Indigenous children in remote communities: issues and problems’, *Australian Review* *of Applied Linguistics*, vol. 34, no. 3, www.nla.gov.au/openpublish/index.php/aral/article/view/2280 (accessed 12 November 2015).

Zubrick, S.R., Silburn, S.R., De Maio, J.A., Shepherd, C., Griffin, J.A., Dalby, R.B., Mitrou, F.G., Lawrence, D.M., Hayward, C., Pearson, G., Milroy, H., Milroy, G. and Cox, A. 2006, *The Western Australian Aboriginal Child Health Survey: Improving the Educational Experiences of Aboriginal Children and Young People*, Curtin University of Technology and Telethon Institute for Child Health Research, Perth.

——, Williams, A.A., Silburn, S.R. and Vimpani, G. 2000, *Indicators of Social and Family Functioning*, Department of Family and Community Services, Canberra.