MELBOURNE INSTITUTE
Applied Economic & Social Research

Working Paper Series

Achievement Gains from Attendance at Selective High Schools

Brendan Houng
Chris Ryan

Working Paper No. 8/18
June 2018

THE UNIVERSITY OF MELBOURNE

MELBOURNE INSTITUTE

Achievement Gains from Attendance at Selective High Schools*#

Brendan Houng and Chris Ryan

Melbourne Institute: Applied Economic & Social Research, The University of Melbourne

Melbourne Institute Working Paper No. 8/18
June 2018

* We thank the state education department for providing us with matched the data. We thank the Principals of the selective schools for their valuable input, and the individuals at the state department who helped facilitate the research. We thank seminar and conference participants at the Melbourne Institute, and at the state education department, for their helpful feedback on earlier versions of this research. This research forms part of Brendan's doctoral thesis. He is grateful to Chris Ryan (co-author), Guyonne Kalb, and Moshe Justman for their guidance as his thesis supervisors, and thanks Gigi Foster for reviewing similar research presented at the 2015 University of Queensland PhD conference. Brendan thanks Cain Polidano for earlier discussions on selective schools and Jongsay Yong for his comments. The findings and views reported in this paper are those of the authors and should not be attributed to any branches of state or Australian Commonwealth governments. For correspondence, email <bhoung@unimelb.edu.au>.

Note that due to privacy concerns the paper is written to anonymize the selective schools. Some details are intentionally less precise then might be desirable to protect the identity of the schools.

Melbourne Institute: Applied Economic & Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* +61 3 8344 2100**
***Fax* +61 3 8344 2111**
***Email* melb-inst@unimelb.edu.au**
***Website* melbourneinstitute.com**

Abstract

Selective high schools are a polarizing topic in education policy, despite only having a small presence in some Australian states. These schools perform exceptionally well when their students' educational and career outcomes are considered, but this is perhaps unsurprising because admission is based on academic performance. This paper asks whether academically selective schools improve their students' university entrance results beyond what they would have achieved otherwise.

Following a cohort of students through high school from an anonymized Australian state, we estimate the selective school effect via two methods: propensity score matching, which compares students of similar background and prior achievement, and regression discontinuity (RD), which compares marginal selective and non-selective students on the basis of the entrance exam.

Our results point to small effects in terms of university entrance ranks, which is consistent with findings from similar studies in the UK, the USA, and other Australian research. Overall, the small selective school effect appears to reflect the high levels of educational aspiration of both selective students as well as applicants who attended other schools. Both groups of students appear to be among the most driven and motivated, being disproportionately from immigrant and socio-economically advantaged backgrounds, and having implicitly signaled an aspirational intent by applying to the schools.

JEL Classification: I2, I21, J24

Keywords: Education, Selective Schools, Academic Selection, Academic Achievement

1. Introduction

Academically selective high schools[1], which admit students on the basis of an entrance exam, make up a small part of choice-based comprehensive education systems in countries like Australia, the United Kingdom and the United States. They appear successful. Selective high schools regularly top school rankings for university entrance results reported in local newspapers. Nevertheless, there is evidence from New York, Boston and the United Kingdom that selective high schools produce only scattered additional gains for those who attend them in high school test outcomes and have little impact on university attendance (Dobbie and Fryer, 2014 and Adulkadiroglu, et al., 2014 and Clark, 2010).

In Australia selective high schools make up only a small percentage of schools and students within the majority of states where they are present, including the state which is the focus of this study.[2] Overall, there were 25 schools distributed through New South Wales, Queensland, Victoria and Western Australia in 2015. The majority of these schools admit students solely on the basis of entrance exam or tests, while a few schools that focus on certain areas of study, such as science or mathematics, also include interviews as part of their admissions process. There were also 70 other high schools in these same states offering select-entry programs, where one or more classes of students are admitted in a similar process involving tests and interviews.

A key challenge for this literature is to estimate what high-achievement students who attend selective schools would have achieved had they not done so. We estimate the effect of selective high school attendance on the end of high school assessment used to determine university entrance, making comparisons between successful and unsuccessful applicants to the selective schools. We use two approaches to do this: matching, which compares students of similar background and prior achievement, and regression discontinuity (RD), which compares students who just got in to the schools, with those who missed out. The RD exploits thresholds in the entrance exam, like the studies mentioned above.

We follow a state cohort of Year 7 students through to the end of high school, including students at a set of selective high schools. The administrative data consists of nationally standardized tests in Year 7 from 2009 and in Year 9 from 2011, Year 12 university entrance results from 2014, and also test results from the selective school admissions process.

Previewing the results, the selective school effect on university entrance results appears to be small. We find estimates of approximately 2 percentile points from the matching approach, but the effect may be overstated; by making comparisons based on tests taken before the admissions process, they

---

[1] Also known as select-entry high schools or exam schools in the USA.
[2] Education policy is administrated at the state and territories level of government in Australia.

do not account for the additional positive selection between selective and non-selective students, beyond that which is implied in both from applying to the schools. There is also evidence, with the same caveat of positive bias, that the selective schools provide a floor for Year 12 achievement; estimates are between 6 percentage points (ppt) and 14 ppt increases in probability of attaining ranks in the top 10% and 15% of students.

In contrast, the regression discontinuity estimates are generally small in magnitude and not statistically significant, suggesting that marginal selective students may not experience the same gains in academic performance that the majority of selective students do. The limitation from the RD approach is that the small sample sizes arising from the reliance on one cohort of students could also be the explanation for the lack of statistically significant estimates. Our results are consistent with the insignificant effects on university entrance results found in Zen's (2016) roughly contemporaneous study of 18 selective schools in NSW, along with the findings from earlier studies from the UK and the USA.

The paper is structured as follows. Section 2 is the literature review which discusses academic outcomes, peer effects and the aspirations of immigrants. Section 3 presents the background information of the selective high schools and establishes the extent of academic selection at the selective schools. A description of the data in presented in Section 4. Section 5 is the matching approach, with subsections for: sample definition, descriptive statistics, methodology, the selection model, and then the results. Section 6 is the regression discontinuity approach, where we describe the methodology, establish its suitability and present the results. In Section 7, we discuss the results from each of the two approaches, and the influence of unobserved educational aspirations in explaining differences between the estimates from each approach. Section 8 discusses the limitations of the case study, and our findings are summarized in Section 9, the conclusion.

2. Literature

A clearer picture of the achievement gains at selective schools has emerged in recent years, with several papers having been published based in different educational institution settings. While research interest prior to this was beset with problems of how to estimate plausible counterfactuals for the academic outcomes of high-achieving selective school students, these recent papers have accounted for the problem of counterfactuals by exploiting the random variation in exams that is created by the thresholds from which offers are made.

The recent studies have produced mixed results, and institutional differences between the studies, including the nuances relating to the high school completion outcomes, mean that interpreting the combination of results can be difficult. Where there have been positive results on achievement (e.g. case studies from Pop-Eleches and Urquiola (2013) for Romania, from Jackson (2010) for Trinidad and Tobago, and from Hoekstra, Mouganie and Wang (2016) for an anonymised Chinese city), it appears that benefits from selective schools are obtained in the context of system-wide academic allocations of students, as opposed to the selective schools existing as a small part of a comprehensive education system in high-income countries like Australia, USA and the UK. [3]

The positive effects found in institutional settings with system-wide academic selection appear to correspond partially to behavioural responses, which were first identified by Pop-Eleches and Urquoila. Pop-Eleches and Urquoila find through surveys that teachers sort in a manner consistent with preferences for high-achieving students. This is demonstrated by Hoekstra et al. (2016), who were able to fully account for their positive effects of selective school attendance with differences in teacher quality between schools. Other behavioural responses identified by Pop-Eleches and Urquoila were that of parents reducing effort when their children attended better schools[4], and also that of students who were successful in attending more selective schools realising that they were weaker than their peers and subsequently feeling marginalized.

In comprehensive education systems the existence of selective schools appears to be a puzzle. Despite high levels of demand for attendance of selective schools, little positive effect has been established for academic achievement at high school completion or on college outcomes. This combination has been described aptly as ‘strong preferences and weak impacts’ by Clark and Del Bono (2016), who offer three possible explanations for the demand from students and their parents, including: a misunderstanding of achievement growth and peer effects, and the importance of non-academic or longer-term outcomes.

There are indications that all three explanations contribute to the high levels of demand; from an earlier Australian survey (Brathwaite and Kensell, 1992), students reported that an academic emphasis, including benefits for future careers, and also the social, cultural and sporting reputations of the schools were important factors influencing their decisions to attend the schools. In the UK, Clark (2010) documented students taking more advanced subjects in selective high schools, which are pre-requisites for certain university courses, but attributed this potential benefit to the academic curriculums at the schools. Later, Clark and Del Bono (2016) extended the focus of the research from short-term education outcomes to longer-term education, labour market and family outcomes.

Due to the stratification of students introduced by selective schools, whose main consequence for students is a sharp increase in their peer quality, peer effects are inevitably identified as a potential mechanism for explaining the impact of these schools on their students. In comprehensive systems, the high representation of students of immigrant background in selective schools suggests that their

[3] Estrada and Gignoux (2017) observed that the benefits were limited to low- to middle- income countries but do not explicitly make the connection to the differences between system-wide and limited academic selection within choice-based comprehensive systems.
[4] Marginally attending selective students received less homework related help from their parents.

higher levels of educational attainment are also relevant to understanding the strong preferences for selective schools. This explanation is also supported by subsample analyses producing positive local effects for minority students in the American case studies from New York and Boston (Adulkadiroglu et al., 2014). There is similar evidence that disadvantaged students receive greater achievement benefits from a North Carolina case study (Shi, 2017)[5].

There is growing research that shows there are benefits to academic achievement to students from attending higher achievement schools in institutional settings with system-wide academic selection[6]. The positive effects appear to reflect a form of resource allocation in countries where there are limited resources or an uneven distribution of resources[7]; the results have been attributed to differences in teacher quality, and the sorting of teachers consistent with the better teachers being matched with higher achievement students.

With possible exceptions of disadvantaged and minority students (discussed in peer effects, section 2.2), differences in peer quality from attending selective schools have consistently been shown to not affect achievement. In comprehensive education systems the demand for selective schools poses a different set of questions, with the absence of evidence for positive effects on achievement ruling out the explanation of benefits to students in the form of improved achievement outcomes.

2.1. Academic Outcomes

Of most relevance is Zen's (2016) concurrent study of 18 selective schools in NSW, who found little general effect on university entrance results, consistent with similar UK and USA studies relying on variation in entrance exams for estimation. With access to a much larger number of selective schools and 3 years of applicants from government schools (2007-2009), Zen analyzes the effect of selective school attendance on academic achievement over a broad range of academic selectivity.

Earlier, Coe et al (2008) summarised over twenty studies of the impact of selective schools in the United Kingdom going back to the 1980s. These selective schools, known as grammar schools, operate as part of the public school systems in the United Kingdom. While critical of the methodology of the studies they summarised, Coe et al (2008) conclude that the literature points to relatively small positive impacts of selective schools on student achievement. Their own estimates, based on regression and multi-level analyses using sex, prior achievement, ethnic background and school meal

[5] Shi (2017) follows cohorts of applicants from 4 years (2009-2012) to a selective school serving Years 11 and 12 that is also a residential high school, and observes outcome variables of SAT scores, major intentions at the end of high school, and college applications.
[6] Lucas and Mbiti with analyses from Kenya appear to be the exception that proves the rule. For an earlier stage of schooling, Zhang (2013) similarly finds little benefit from attending an oversubscribed selective middle school in China for high school entrance exams or admission outcomes, but relies on a lottery mechanism for estimation rather than variation from exam cut-offs.
[7] Positive results were obtained in low to middle income countries. Park et al. (2015) provide also the example of 'magnet' selective high schools in rural China.

status to determine attendance, similarly pointed to positive effects on later school achievement for secondary school students.

Abdulkadiroglu et al. (2014) assessed the academic contribution of selective high schools to the educational outcomes of later test scores and the college quality attended by their students using regression discontinuity methods. For selective schools in New York and Boston, they found no general effect and scattered gains to test scores to students from attendance, which included a modest localized effect on English scores for minority students. Using entrance exam scores as an instrument for peer composition, Adulkadiroglu et al. also found no effect from changes in peer composition on test scores.

In contemporaneous work, Dobbie and Fryer (2014) found little impact of attendance of selective schools in New York on college enrolment, graduation or quality for marginal students. Like Adulkadiroglu et al., Dobbie and Fryer also applied a regression discontinuity (RD) approach, where students just above and below the cut-off determining offers are compared. Similarly, Clark (2010) adopted an instrumental variable approach with entrance exam results and assessed the effect of attendance at grammar (selective high) schools in the United Kingdom, and found only small effects from attendance on test scores but some evidence for course-taking and university enrolment.

In some countries, almost all students are allocated to high schools on the basis of admissions exams. Pop-Eleches and Urquiola (2013) and Jackson (2010) investigated the effects of attending higher achievement schools for Romania and Trinidad and Tobago, respectively. Pop-Eleches and Urquiola (2013) find that students with access to higher achievement schools perform better at a high stakes graduation test[8]. With the RD approach, there are 2,000 cut-offs or 6,000 cut-offs if tracks within secondary schools are included. Instrumental variable results from Trinidad and Tobago (Jackson, 2010), where all students are assigned to schools after grade 5 on the basis of achievement tests, also showed large positive benefits on examination performance at the end of secondary school from admission to higher performing schools[9].

Hoesktra et al. (2016) conduct regression discontinuity analyses, drawing a sizeable sample from an anonymised Chinese city of 10 million, and find large and significant effects for university entrance results of 0.16 SD for the most selective tier of selective schools. They attribute all of the selective school effect to improved teacher quality, as reflected by meaningful discontinuities in access to better

---

[8] The positive effects from attending higher ranked schools and tracks within schools are as high as 0.02 to 0.10 of a standard deviation (SD) on the Baccalaureate, which is a pre-requisite for applying to university.

[9] Students attending secondary schools with peer scores half a standard deviation higher were associated with passing between 0.28 and 0.34 more graduation exams, completing 5 of which is a requirement for attending university.

quality teachers (i.e. having obtained the highest ranked and paid class of teacher in the Chinese education system).

Clark and Del Bono (2016) extended the focus on academic achievement towards longer-term outcomes, when individuals were in their mid-to-late forties from cohorts educated in the 1960s. They find large impacts on completed education from selective school attendance. For men, they find little impact on labour market outcomes in income and employment. For women, Clark and Del Bono find increases in income but also document a negative impact on female fertility. In a similar vein, Estrada and Gignoux (2017) obtain approximate effects on earnings by analysing the effects from attendance of 16 selective schools from Mexico City on college outcomes, and hence implied increases in earnings. They find that the selective students are four times as likely to graduate from an elite college, and three times as likely to obtain an engineering degree (off low bases of less than 1% probability).

2.2. Peer Effects

The positive selection of students into schools based on academic performance, and in other circumstances socio-economic advantage, might create a positive externality via peer effects, such that it leads to learning-conducive environments and better educational outcomes; peer effects may include a variety of indirect mechanisms relating to teachers, staff and school resources, which are difficult to distinguish between. Conversely, the loss or absence of high achieving students from other schools may weaken the factors which are associated with, or contribute, to positive externalities.

The vast literature on peer effects in education finds only modest effects on test scores (see the review in Sacerdote, 2011). Angrist (2014) asserts that the results from many studies are overstated due to methodological deficiencies (bias from weak instruments) that cannot overcome the non-random selection of students into schools. With respect to selective schools, the large change in the composition of peers from attending a selective high school provides an ideal setup for assessing the influence of (positive) peer effects, but little evidence of a general peer effect has been found (e.g. Adulkadiroglu et al.).

There is, however, some support from the North Carolina case study for a positive peer effect at selective schools for more disadvantaged students (Shi, 2007). With a unique set-up, the selective school from North Carolina is required to equitably admit students from 13 congressional districts, which allows for comparisons of students of different academic ability. Zen (2016) finds similar tentative evidence that the selective schools benefit students from a low socioeconomic background, with his study of 18 selective schools in NSW also allowing for the analysis of students across a broader range of academic ability.[10]

Although the relationship found between academic outcomes and peer effects has not been strong, there is accumulating evidence that there might be heterogeneous effects (although this has not been shown for test results at the end of high school). That is, higher-achieving students are disproportionately influenced by peer effects in a positive way; examples cited by Sacerdote (2011) include: Lavy, Paserman and Schlosser (2012), Duflo, Dupas and Kremer (2011), and Lavy, Silva and Weinhardt (2012) for the case of high achieving girls. Lavy, Silva and Weinhardt (2012) also observe that all students are adversely affected by low-achieving students.

Outside of cognitive outcomes, studies have found substantial effects for social outcomes such as drinking, drug use and criminal behaviour (summarised in Sacerdote, 2011). Regardless of which specific outcome is influenced by peers, students and parents act as if there are positive peer effects. A higher demand for schools with higher test scores is reflected in house prices (Davidoff and Leigh, 2008), and similarly demand for schools with a student population from an advantaged socioeconomic background is evident in commuting behaviour. [11]

2.3. Rank Effects & Self-Concept

Closely related to peer effects is the 'Big-Fish-Little-Pond' effect, posited initially by Marsh and Parker (1984) as a potential explanation for some of the heterogeneous peer effects found in the literature, especially as they relate to selective schools. Later, Marsh and Hau (2003) found a negative association between individual and school average achievement from the Program of Student Assessment (PISA, 2000) that administers tests to around 4,000 fifteen year-old students across 26 countries. They suggest that an application of social comparison theory, whereby students compare their own achievement with their peers, leads to students having lower academic self-concept if they attend schools with high-achieving peers.

Similarly, using administrative UK data from 2001 to 2005, Murphy and Weinhardt (2014) expanded on this concept by estimating a positive effect of classroom rank position of students test results in primary school on their subsequent secondary school test results. In the opposite direction, Jonsson

---

[10] Zen (2016) finds effects of 0.19 SD and 0.15 SD in university entrance results (both statistically insignificant) for selective schools with low cut-offs and with high cut-offs, respectively.

[11] A significant percentage of students travel long distances to attend primary and secondary schools with higher SES students, in the northern suburbs of Melbourne (Lamb, 2007). On a related note, one of the critiques of selective schools policy is that it results in residualisation, and contributes to the concentration of socioeconomic disadvantage due to the removal of high achievement students from other schools (Lamb, 2008). Residualisation refers to the process arising from the preferences of students for socioeconomically advantaged schools, whereby the socioeconomically disadvantaged schools gradually decrease in size and have high concentrations of disadvantage, which both potentially lower achievement.

and Mood (2008) found somewhat disconcertingly that having high-achieving peers reduces students' aspirations for attending university, by a small amount.

In addition to finding positive benefits to achievement to disadvantaged students from selective school attendance, Shi (2017) makes the similar observation to this study that attendance of selective schools confers a decrease in relative rank for some students. Shi finds that the decreases in relative rank diminishes the likelihood for selective students in applying to STEM[12] focused colleges.

2.4. Educational Aspirations & Immigrant Background

The high representation of students of foreign language background in Australian selective high schools is also a characteristic of selective schools in New York and Boston, although studies reported only the ethnicity of students. In the three older New York selective high schools, between 57% and 67% of students were Asian, and a further 3% to 8% were Hispanic (Dobbie and Fryer, 2014). American selective students were however from a lower socio-economic background than in the Australian context, which is typical of American inner city residents, between 39% and 62% of students at the New York selective schools were eligible for free or reduced lunches.

It is probable that the high representation of students of immigrant background at selective schools is closely related to the ambitions of these students; the link between the language background of student and their educational attainment has been established by several studies (see also section 7.4). Gemici et al. (2014) found language background and academic performance, along with parental expectations, to be the most influential factors for completing Year 12, from nationally representative survey data. Similarly, from a recent cohort of Australian students, overseas born students with a foreign language background were also found to be 6% more likely to complete Year 12, and this effect became statistically insignificant after controlling for surveyed educational aspirations (Homel et al., 2012).

3. Background

3.1. Selective High Schools and the Admissions Process

This study follows a state cohort of students in Year 7 in 2009, through to the end of high school; and specifically the subset of students who participate in the admissions process to enter a set of selective high schools. The selective schools were attended by less than 2% of the student population. The entrance exam for admission into the selective high schools is held annually in the year preceding entry. Results from the entrance exam determined the vast majority of places, with the balance of students comprising of: applicants who were exceptions to a quota rule designed to limit the impact of

[12] Science, Technology, Engineering and Mathematics.

selective schools on other schools; and also applicants from a disadvantaged background who also performed well on the exam.

3.2. Academic Selection

The level of academic competitiveness for attending selective schools is illustrated in Figure 1, which presents histograms of Year 7 standardised test scores for the general population, in the left panels, and the students who attend the selective schools, in the right panels; one selective school is excluded for being in its first year of operation (discussed later in section 5.1, the Data Sample). Histograms for boys and girls are presented in the top and bottom panels, respectively. The test scores are the average of Reading and Numeracy, which are each standardized normal, with means of zero and standard deviations of one.

Figure 1: Prior Achievement for General Population and Selective Students*

*based on Year 7 standardised test scores. Red and green vertical lines indicate minimum and mean values at selective schools respectively: for girls is -0.19 and 1.71; and -0.05 and 1.71 for boys.

For boys and girls separately, the minimum and mean Year 7 test scores for students at selective schools are indicated by the red and green vertical lines. The minimum selective student for girls is at the 42nd percentile of all students, and corresponding value for boys is the 48th percentile. The means are far higher, placing in the top 5% of all students, at almost the 96th percentile, with the same value for boys and girls. At the most academically competitive selective school, the minimum score (0.58) was at the 72nd percentile of all students.

Note that more boys were admitted to selective schools, which likely reflects the greater proportion of boys among students with higher levels of achievement. [13]

4. Data Description

4.1. Demographic Information

The demographic information of the student population is shown in Table 1. There were 67,687 students aged 14 to 16 in Year 9, of which 42,193 (63.4%) appear in the Year 12 data. Details of the fuzzy matching involved in combining the administrative and standardized data with the data from the selective schools admissions process are included in the Appendix, A2.

Table 1: Demographic Information of Student Population

	YR7, 2009	YR9, 2011
Number of students	66,509	67,687
Male (%)	51.5	52.4
Aged 14 / 15 / 16 in 2011 (%)	17.6/76.8/5.6	17.7/76.3/6.0
Language Background Other Than English (%)	25.0	23.4
Aboriginal and Torres Straits Islanders (%)	1.3	1.3

4.2. Standardised Tests and Year 12 Assessment

The data consists of nationally standardized tests from 2009 and 2011 and Year 12 results from 2014. Introduced in 2008, the National Assessment Program – Literacy and Numeracy (NAPLAN) test is conducted annually in Years 3, 5, 7 and 9 and covers five learning domains. The NAPLAN data contains information relating to students' parental occupation and education and also language background, indicated by English Speaking Background (ESB) or Language Background other than English (LBOTE). From the parental education and occupation information, we derive an index of socio-economic background (SES) using principal components analysis with the students in each of Year 7 from 2009 and Year 9 from 2011.

The Year 12 results are derived from the state's high school qualification to produce a nationally comparable rank, which we use for the outcome variables. Specifically, Australian Tertiary Admission Ranks (ATAR) values are percentile ranks derived from a combination of the results from subjects completed by students in Year 12. Though the aggregated subject results and the ATARs are equivalent, the ATARs are widely reported and used for admission to university courses.

We derive three binary Year 12 achievement outcomes from the ATAR ranks: ATAR $\geq$ 95, ATAR $\geq$ 90, and ATAR $\geq$ 85; effectively scores within the top 5%, 10%, and 15% of results, respectively.

---

[13] From the sample of applicants, the percentage male, as a local weighted average, increased from 50% to around 60% over the full range of exam results (see Figure 20 in Appendix A1, which shows a locally weighted regression of the percentage male over percentile ranks of entrance exam results).

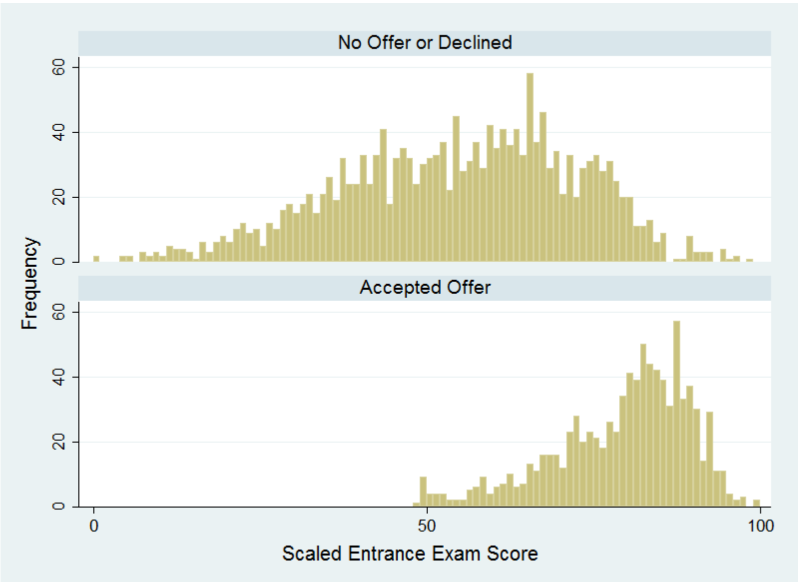
These derived Year 12 outcomes provide a meaningful measure of success for the group of students of interest, those with very high prior achievement, in addition to assessing ATAR directly. For instance, a result of ATAR $\geq$ 95 would allow a student to enter many university courses, while scores of ATAR $\geq$ 85 also provide access to a wide range of university courses; requirements for entry to courses and universities varies with demand from applicants, which reflects to some extent their reputation.

4.3. The Entrance Exam

Applicants to the selective schools sit the same entrance exam and indicate their specific school preferences. The selective schools entrance exam data included information relating to the school preferences of applicants, their equity status, and their exam scores and outcomes. The outcomes information consisted of whether applicants received offers and what types of offers were accepted.

Figure 2 presents the histograms of the entrance exam scores by the outcomes of the admissions process. The presentation is of a composite measure of the tests undertaken and are rescaled to from zero to hundred. It is clear that the distribution of entrance scores for those who attend the selective schools lies substantially to the right of those who do not attend the schools. There are two main groups who do not attend – those whose test results gave them a rank in the distribution too low to be offered a place, and those whose test results were such that they received offers, but they declined them. [14]

Figure 2: Histograms of Entrance Exam Scores by Selective Attendance



[14] A small number of students also missed out on receiving offers due to quotas intended to limit the loss of students from other schools.

5. Propensity Score Matching

5.1. The Data Sample

With access to student-level background characteristics and standardized test scores for the general population (a state cohort), our challenge is to find a group of students comparable to those who attended selective high schools. We identify applicants to the selective schools who go on to attend other (i.e. non-selective) schools as the students who most resemble the selective students in their motivations; sitting the test reveals a high level of academic ambition and motivation among these students, which we think deals to some extent with otherwise unobserved factors that are problematic in using matching estimators. Given that admission is based on academic performance, we additionally limit the comparison group to students with Year 7 achievement above the minimum level observed among those who subsequently attended the selective high schools[15]. The decision to use NAPLAN as the basis of comparison over the entrance exam is discussed in detail in a later section (5.3.2. Prior Achievement).

We define the matching data sample, the "exam sample", as students who sit the entrance exam who have prior achievement of at least the minimum level observed among those who subsequently attended selective schools. We choose to exclude the disadvantaged intake and exclude the discretionary-intake, whose inclusion is likely to positively bias the results, from the pool of students with high prior achievement for our base case analysis. We also exclude students who attend one of the selective schools for the reason that they entered the school in its first year of operation as a new selective school (discussed next).

The small number of students admitted to selective high schools on a discretionary basis are excluded due to concerns they have been selected for other characteristics that are not reflected in the data. These were students affected by the quota rule to protect the loss of students from other schools. Including these positively selected students is problematic for estimating the overall effect from selective school attendance due to the anticipated positive selection bias from the unobserved factors that led to their selection. We also excluded students from a disadvantaged socioeconomic background admitted through another intake due to limited sample size. Further details of both intakes are included in Appendix A3.

5.1.1. Selective Schools

In the main analyses, we estimate the selective school effect at the aggregate level, pooling the selective schools. In order to present reliable estimates, we exclude one of the selective schools from the analyses as it is a new school for which the intake of students we observe is the first of its cohorts

[15] With one selective school excluded, discussed in the next section, 5.1.1. Selective Schools.

to reach Year 12; there are several aspects about this cohort of students that differ from those observed for the other selective schools.

First, few students who sat the select entry test identified this school as their first preference school to attend, meaning that relatively few of those who finished up attending it had placed it first. Second, the prior achievement and social background profile of its first intake differed substantially from other selective school applicants. In particular, the minimum entrance exam result required to receive an offer was much lower than the other schools.

In addition, when the minimum NAPLAN score at the excluded selective school was used to define the selective potential sample, the covariate balancing tests[16] of the intake and the comparison group who sat the test were not passed for that sample, which we ran as a robustness check. Failing the balancing tests is further confirmation that the students at the excluded selective school are unlike the other selective students. For all these reasons, we decided to exclude this school from the analyses that follows.

5.2. Descriptive Statistics

In this section we review the prior achievement, student characteristics and achievement outcomes of selective and non-selective students in the exam sample that we defined in Section 5.1, the Data Sample.

5.2.1. Prior Achievement

For the purposes of defining the data sample we derive a NAPLAN variable that is the average of Numeracy and Reading, which are themselves standardised with mean zero and standard deviation of one from the whole student population. Table 2 shows the average and minimum NAPLAN scores of selective students who were admitted to selective school strictly on the basis of their entrance exam results. That is, it excludes the discretionary intake and disadvantaged intake students. The table also shows the average NAPLAN score for non-selective students from the exam sample.

Table 2: Standardised Year 7 NAPLAN Scores by Selective Attendance

	Mean	SD	Missing (%)	Count
Selective	1.76	0.63	0.7	592
Non-Selective	0.94	0.64	0.5	1,550

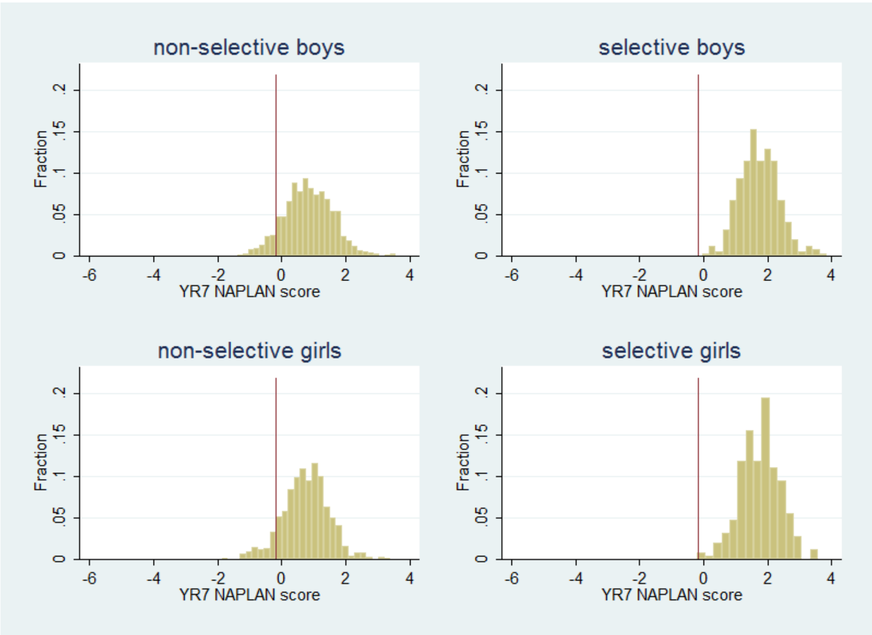
It is apparent that the non-selective students, who sat the exam but did not attend selective schools, had lower NAPLAN scores than selective students. The mean NAPLAN score for the non-selective exam sample is 0.8 standard deviations below that of the selective sample (0.94 compared with 1.76),

[16] From the matching approach, refer to Methodology, section 5.3.

whereas the difference among applicants whose first preference was the most competitive selective school between non-selective and selective samples is smaller at 0.55 (not shown).

Non-selective students that sit the exam are nonetheless closer to the selective school students in terms of Year 7 NAPLAN than the broader school population. This is perhaps unsurprising as only around 4% of the student population sat the admissions test, and in addition, it is probable that the students that do are informed by their perceived likelihood of getting into the selective schools.

Figure 3: Histograms of YR7 NAPLAN Scores for the Exam Sample



Here, the minimum Year 7 NAPLAN scores at the selective schools is -0.19.

The suitability of the exam sample as a comparison group is demonstrated visually in the histograms in Figure 3. The histograms of Year 7 NAPLAN scores are shown by selective school attendance and by sex. The red vertical lines indicate the lowest minimum NAPLAN score among the remaining selective schools, which are used to limit the sample of both selective and non-selective students. Reassuringly, the numbers of students to the left of the red vertical line for the non-selective students are relatively few (left panels) – meaning that the non-selective students in the exam example sample are reasonable in approximation to the group of selective students.

5.2.2. Student Characteristics

Table 3 presents average characteristics for the selective and non-selective students in the exam sample (defined in the Data Sample, section 5.1)[17]. The rows contain information of the language

[17] In the analyses we assess the girls and boys separately. The student characteristics for those samples are included in Appendix A4, in Table 10.

background and socio-economic background in quartiles of these groups, as well as sector of the school attended by students in each group.

Table 3: Characteristics of Selective and Non-Selective Student in the Exam Sample

	Selective	Non-Selective
N	592	1,520
Male (%)	57.4	51.4
Foreign Language Background (%)	80.1	75.8
SES Category (%)		
Quartile 1	11.3	14.9
Quartile 2	13.0	17.0
Quartile 3	27.0	26.8
Quartile 4	48.6	41.2
YR7 Sector (%)		
Government	54.9	54.5
YR9 Sector (%)		
Government	100	51.8

^ above minimum score among students at Selective Schools (excludes one school).

Selective and non-selective students are quite similar in terms of socio-economic backgrounds, with similar distributions in the SES quartiles, although a slightly larger percentage of selective students in the top quartile, at 48.6%, as compared with the non-selective students, at 41.2%. In terms of language background, selective school students are overwhelmingly from a foreign language background; over 75% of both the selective school students and non-selective students from foreign language backgrounds, which is much greater than the Year 9 student population percentage of 23.4% (from Table 1 earlier). There was also a slightly greater proportion of males in the selective (57.4%) group, compared with non-selective group (51.4%).

Since an important part of the research design for this project is to find the right comparison group for the selective school students, it is clear that an important first step can be made by focusing on the group who sat the test as the relevant comparison group. From Table 3, their observed socio-demographic characteristics were already like those of the group who attended the school in terms of language background and social background. We posit that their motivation levels are similar, since they were prepared to sit the test, it is likely that any important unobserved characteristics are also likely to be more like those of the group who attend selective schools.

5.2.3. Achievement Outcomes

Table 4 presents mean ATARs and aggregated year 12 scores and the percentage of missing values for the selective and non-selective groups[18]. Despite similarities in prior achievement and background characteristics, the selective students had higher values in ATAR and ATAR derived outcomes, consistent with the difference between groups in Year 7 NAPLAN from Table 2.

Table 4: Achievement Outcomes of Selective and Non-Selective Students

	ATAR	% Missing	ATAR ≥ 95 (%)	ATAR ≥ 90 (%)	ATAR ≥ 85 (%)
Selective	92.6	5.7	51.2	69.8	81.1
Non-Selective	83.9	8.6	23.2	39.4	52.4

For students without missing ATARS (first column), selective students had 8.7 points higher ATARs on average than the non-selective students in the exam sample. The differences in the derived achievement outcomes were more marked, with for example only 23.2% of non-selective students attaining an ATAR above 95, whereas 51.2% of the selective school students did so.

Note the percentage of missing ATARs (second column) is fairly similar between the selective and non-selective students who sat the entrance exam, at 5.7% and 8.6%, respectively. Students have missing ATAR values for various reasons, including: completing the qualification the following year, leaving high school early, moving interstate, or completing alternative qualifications, such as the International Baccalaureate. There is potential bias to the extent that there is differential selection into missing ATAR values between the selective and non-selective students, but the magnitude of the bias is relatively small with a difference of 2.8 ppt, when compared with the size of the statistically significant matching estimates for the binary outcomes, which are over 6.0 ppt.

5.3. Methodology

Propensity score matching makes like with like comparisons, where the probability of treatment is used to capture similarities between the groups. Formally, the propensity score is defined by Rosenbaum and Rubin (1983) as the probability of treatment (attending a selective school) given a set of characteristics **x**, the function p(**x**):

$$p(\mathbf{x}) = \Pr(D = 1|\mathbf{x})$$

[18] In the analyses we assess the girls and boys separately. The achievement outcomes for those samples are included in Appendix A.4, in Table 11.

where D = {0, 1} where 1 means the individual attended a selective school, and **x** are pre-treatment characteristics (Becker and Ichino, 2002). Propensity score matching requires the *ignorability of treatment* assumption, such that treatment and outcomes are independent, conditional on **x.**

In the first stage, a probit or logit model is fitted to estimate the propensity score (or predicted probability) from the selection equation:

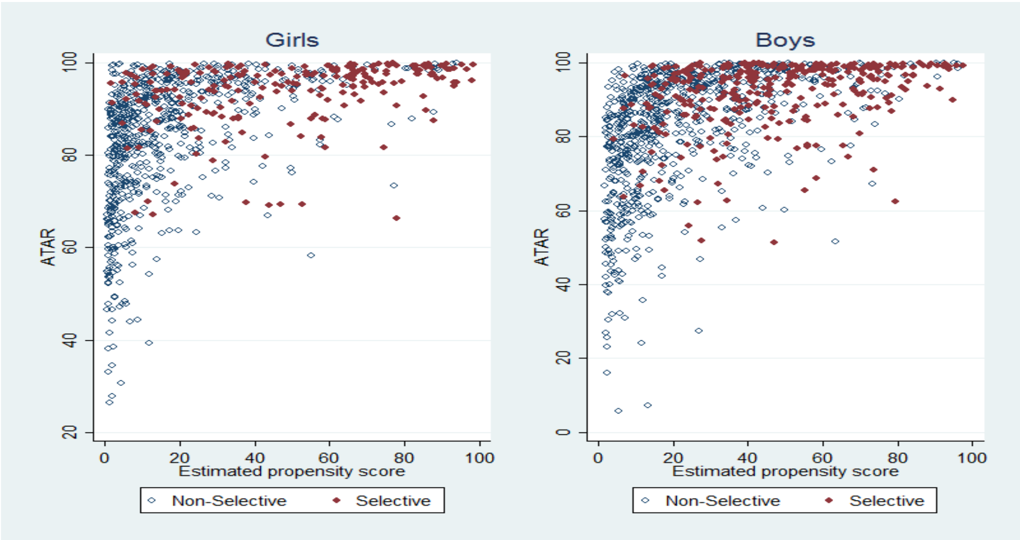
$$\Pr(D = 1|\mathbf{x}) = cdf\{h(\mathbf{x})\}$$

where *cdf* is the normal or logistic cumulative distribution function, and the specification of $h(\mathbf{x})$ is such that selection is independent of **x** given p(**x**), a requirement known as the balancing hypothesis. The estimation details relating to satisfying the balancing hypothesis as well as the first stage specifications are provided in the next sub-section (The Balancing Hypothesis).

In the second stage, we adopt several approaches to match control to treatment observations: a) nearest neighbour, b) kernel based matching with a bin width = 0.06 and c) stratification matching (see Becker and Imbens 2002). These various approaches weight the control observation outcomes according to the predicted probability of being in a selective school to match to selective school student outcomes, such that the difference in outcomes produces the estimated treatment effect.

Scatterplots of joint ATAR and propensity scores for girls and boys separately appear in Figure 4 below. Selective students are indicated by red dots, while non-selective students are indicated by hollow blue circles. The selective dots are clustered above ATARs of 80, with less variance than the non-selective dots, which have a large representation below 20 in propensity score, with ATARs across the whole range of 20 to a 100.

Figure 4: Scatterplot of ATARs and Propensity Scores by Sex and Selective School Attendance*



*From selection model specification in Table 5.

5.3.1. The Balancing Hypothesis

Under Propensity Score Matching, the specification of the selection model requires the balancing hypothesis to be satisfied. That is, when observations with the same probability of treatment are compared, the group who obtained treatment cannot be distinguished from the group who did not. Stated formally, it is required that selection is independent of covariates given the predicted probability of treatment.

To assess whether the data appear to satisfy this requirement, we follow Becker and Ichino (2002) and test for various specifications that the characteristics of treatment and control groups are not different from each other within a given number of subgroups; tests are conducted at the 1% level of statistical significance, and subgroups may be split into smaller units until the requirement is satisfied. Note: these subgroups are also used within stratification matching to produce the estimated effect of attending a selective school by averaging the within-subgroup estimates.

5.3.2. Prior Achievement

To measure prior achievement in the matching approach, we adopt Year 7 NAPLAN over the alternatives of Year 9 NAPLAN, or the entrance exam, due to both the timing and interpretation of the achievement tests. The Year 7 NAPLAN tests precede the commencement of the selective schools by two years, and precede the admissions process by one year, meaning that the tests are likely not to reflect any influence from attending the selective schools, thus making it suitable as the basis for comparison of later achievement outcomes between selective and non-selective students.

By the same reasoning, we rule out Year 9 NAPLAN as students sit the tests after having attended the selective schools for several months, which could bias the estimates from the immediate influence of selective schools on the achievement of their students. There would be a downwards bias from an expected positive selective school effect.

Although like Year 7 NAPLAN the entrance exam is sat by applicants before the selective schools start, when applicants are in Year 8, the entrance exam cannot be used as prior achievement for the matching comparisons because the stronger applicants generally accept their offers to the selective schools. This results in an uneven proportion of non-selective and selective students at given levels of performance in the entrance exam, which also means that the balancing hypothesis cannot be satisfied due to differences in characteristics between groups for the same probability of selective school attendance.

The different achievement tests, however, pose the question of how to interpret what they measure. Our approach is informed by Sternberg (2007), who suggested that achievement tests can differentiated qualitatively according to their purpose. The NAPLAN tests, as nationally standardized tests, appear to emphasize accumulated curriculum knowledge due to their accountability function of benchmarking student and school performances. In contrast, the entrance exam assessment is designed to differentiate between high intellectual ability students, appearing closer in form to intelligence tests which have been described as measuring achievement that individuals should have accomplished several years earlier (Sternberg 2007, citing Anastasi and Urbina, 1979).

We secondly distinguish between the assessments by the incentives that they offer to the students, consistent with Duckworth (2016), who defined achievement as the output of skill and effort, where skill itself is a function of effort and talent. This is a departure from the conventional economic interpretations of education, as a cumulative process in which prior achievement is an input (Todd and Wolpin, 2003), or as a product of cognitive and non-cognitive skills (Heckman and Rubinstein, 2001, and Cunha and Heckman, 2008), both of which do not account for incentives in test assessments.

When considered by the incentives offered by the tests, NAPLAN is a low stakes outcome for most students such that it is likely that students exert much less effort on these tests, which allows for a clearer interpretation of the matching estimates. The achievement gain produced by the matching approach is a fair comparison in that the university entrance results are a high stakes outcome and the Year 7 NAPLAN are a low stakes outcome for both selective and non-selective students. The qualifier to the interpretation is that some (potentially large) part of the expected gain likely reflects positive unobservables in the selective students.

In comparison, the potentially large incentive of attending a selective school suggests that there is variable effort in taking the entrance exam by applicants to selective schools, conditional on their motivations[19]. Using the entrance exam for prior achievement, if it had satisfied the balancing requirement of the matching approach, consequently suffers from the likely comparison of students who were trying against some of those who were not, leading to a negatively biased selective school effect.

The interpretation of the selective school effect (as the value-added) changes from using the entrance exam in place of Year 7 NAPLAN as prior achievement. Due to variable effort from applicants, a large part of the selective school effect would be more accurately described as the difference in remaining improvement in achievement between selective and non-selective students, with a small part of the effect being attributable to the difference in schools. In other words, by making comparisons based on the entrance exam, the difference in achievement gains between selective and non-selective students would already exclude the gain in academic performance that occurred before the applicants sat the entrance exam.

[19] We explore reports of additional effort and resources on the part of selective school applicants in Educational Aspiration (section 7.4).

5.4. The Selection Model

The earlier descriptive statistics showed that students at selective schools were more likely to be from a foreign language background, have higher prior achievement scores, and were more likely to be from a higher socio-economic background. Modelling selection into treatment, we estimate a logistic regression on selective school attendance with standardized Numeracy and Reading included as linear terms, language background as an indicator variable, and socio-economic background as a continuous variable, derived from parental education and occupation. The average marginal effects for these variables are presented in Table 5, separately for boys and girls.

Table 5: Logistic Regressions of Selection into Treatment by Sex

	Boys	Girls
Std. NUM	0.1531***	0.1555***
	(0.0141)	(0.0135)
Std. RDG	0.1253***	0.1269***
	(0.0183)	(0.0163)
LBOTE	0.0226	0.0754*
	(0.0320)	(0.0308)
Std. SES	0.0049	0.0206
	(0.0131)	(0.0117)
Government	0.0164	
	(0.0264)	
Number of blocks	8	9
Tjur	0.221	0.324
N		
Selective	328	224
Non-Selective	702	675
Total	1030	899

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Recall the differences in prior achievement and background characteristics, presented in Tables 1 and 2: selective school students were different from non-selective students who sat the exam in terms of their prior achievement (on Year 7 NAPLAN), but they were not different in terms of their language background or SES. As a result, in regressions distinguishing selective school students against non-selective students, we should expect only achievement to distinguish selective school students from non-selective students who sat the exam.

This is largely what we see in Table 5. Year 7 Numeracy and Reading were large and statistically significant, at over 12 ppts for a 1 SD increase. For the boys in the exam sample, Numeracy had a stronger effect than Reading on selection into treatment (selective attendance); a one standard deviation increase in standardized Numeracy translates to a 15.3% percentage point (ppt) increase in the probability of attendance, while the equivalent increase for standardized Reading is 12.5 ppt.

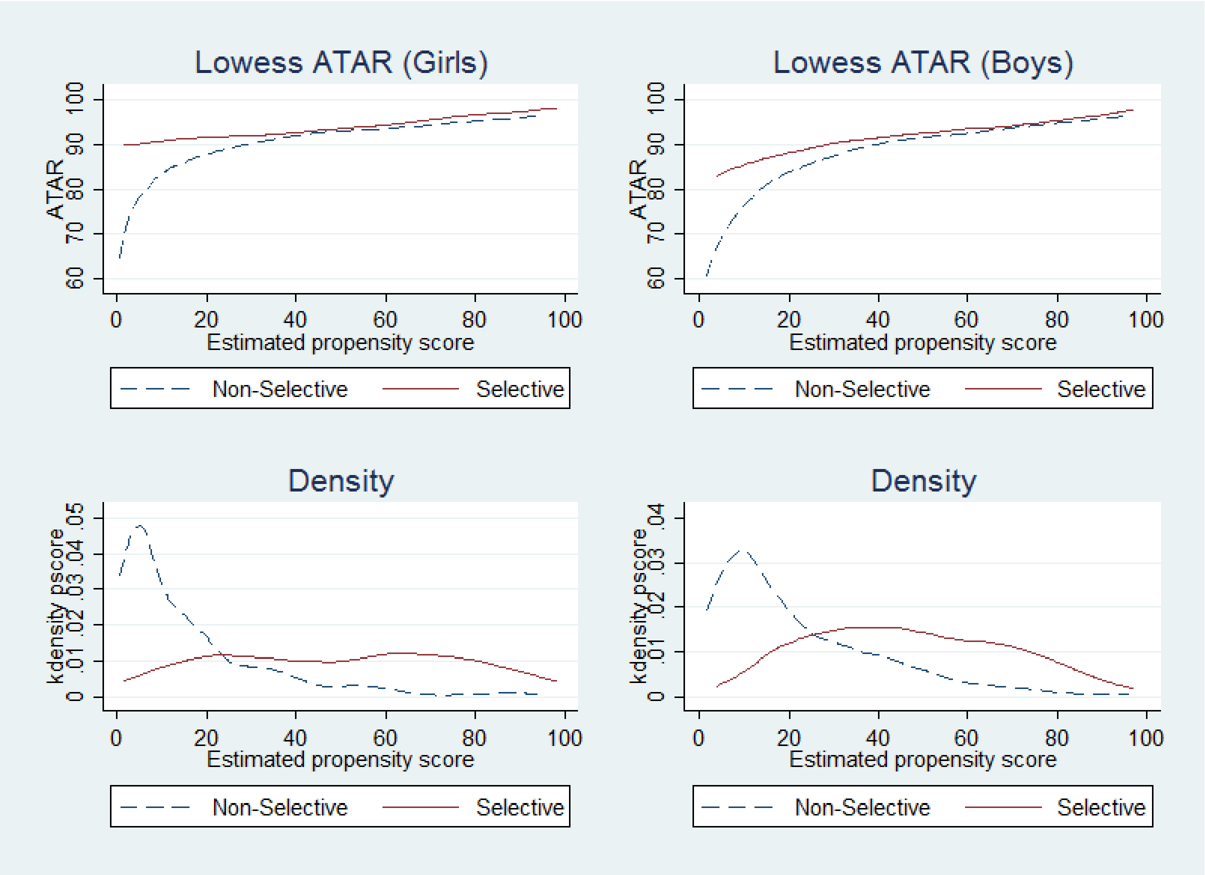
Language background is not significant for boys, but remains significant for girls at the p-value of 0.05. Socio-economic background was statistically unimportant in size and statistical significance, for both boys and girls, with magnitudes less than around 2% for girls and less than 0.5% for boys for a one standard deviation increase in SES.

Lastly, a Year 7 government sector indicator was also included for the samples for the boys in the exam sample; the inclusion of sector did not satisfy the balancing tests for the other samples. The Tjur goodness-of-fit values, which are differences in predicted probability between treatment and control, were reasonable, at 0.32 for the girls sample and at 0.22 for the boys sample.

5.5. Results

This section presents the main results from the propensity score matching for the girls and boys, from the pooled sample of included selective schools. These results relate to the exam sample, as we found that the non-selective students who sat the exam most resembled students attending selective schools (in the Descriptive Statistics, section 5.2).

Figure 5: Lowess ATAR and Density of Propensity Scores by Selective Status and Sex^



^ Scatterplots of the student observations were much noisier, shown in Figure 4, earlier.

Figure 5 presents the densities of the propensity scores from the selection model in section 5.4, the Selection Model, for girls and boys depending on their selective school attendance status in the lower panels, and lowess plots of ATAR ranks against the propensity scores in the upper panels[20]. The densities indicate that the distribution for selective school students is relatively uniform, while that for the group who do not attend is quite bunched at low values of the propensity score. That is because the typical prior achievement levels of the group who did not attend selective schools are lower than those who do. The panels on the left correspond to the girls and the panels on the right are for the boys.

The graphs in the top panels provide guidance in identifying which selective students are likely to outperform the non-selective students; it shows that the non-selective students with lower propensity scores, those below 40 percent, have lower average ATAR ranks than selective students with the same propensity scores. Above that point, there is little difference in localized average ATAR ranks of the two groups for either sex. In other words, it appears that the students who are least likely to attend selective schools appear to gain the most. These are the students who get into the schools but whose prior achievement is lower than their peers, as is implied by the selection model.

Table 6: Propensity Score Analysis, Boys

	Outcomes				Frequencies		
	ATAR	ATAR ≥ 95	ATAR ≥ 90	ATAR ≥ 85	NS	S	Total
Nearest Neighbour Matching	1.509 (1.107)	0.047 (0.052)	0.043 (0.050)	0.076* (0.044)	184	328	512
Kernel Matching (bin width= 0.06)	1.979*** (0.582[b])	0.052 (0.042[b])	0.046 (0.033[b])	0.066* (0.03[b])	684	328	1,012
Stratification Matching	1.640* (0.721[b])	0.038 (0.039[b])	0.040 (0.035[b])	0.062* (0.03[b])	684	328	1,012
Expected Success (Non-Selective)	83.6	0.325	0.520	0.638			

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors indicated in parentheses,

[b] indicates bootstrap standard errors.

[20] Lowess, derived from ‘locally weighted scatterplot smoother’, is a smoothing technique applied to scatterplots to help visualise the relationship between two variables, and is also a non-parametric regression model (Fox and Weisberg, 2010). Locally weighted regressions are applied within bandwidths over the range of values to produce polynomial fits of the data, where the local weighting refers to horizontal distance of sample observations from the observation for which the fit is being estimated (Cleveland, 1979).

The average treatment effects for the boys are detailed in Table 6 and in Table 7 for the girls. Estimates from each of the nearest neighbour, kernel and stratification matching methods are presented along the rows. Along the columns are the Year 12 outcomes and also frequencies of non-selective ("NS"), selective ("S") and total students used in each method.

Table 6 shows increases in the success rates from attending selective schools were significant for boys, with the most consistent results for ATAR and ATAR >= 85; selective students benefited by 1.6 to 2.0 ATAR percentile points depending on matching type, and were more likely to attain ATAR >= 85 by between 6.2 ppt and 7.6 ppt. The estimates were consistent in magnitude and statistical significance across the different types of matching, except for nearest neighbor matching not being statistically significant for the ATAR outcome.

Table 7: Propensity Score Analysis, Girls

	Outcomes				Frequencies		
	ATAR	ATAR ≥ 95	ATAR ≥ 90	ATAR ≥ 85	NS	S	Total
Nearest Neighbour Matching	1.171 (1.335)	-0.022 (0.071)	0.098 (0.067)	0.056 (0.057)	117	224	341
Kernel Matching (bin width= 0.06)	2.201** (0.852[b])	0.056 (0.05[b])	0.135** (0.055[b])	0.091** (0.03[b])	642	224	866
Stratification Matching	1.755* (0.762[b])	0.030 (0.047[b])	0.110** (0.043[b])	0.079** (0.032[b])	642	224	866
Expected Success (Non-Selective)	84.3	0.319	0.487	0.640			

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors indicated in parentheses,

[b] indicates bootstrap standard errors. Without sector in selection equation.

On the whole, the results for the girls (presented in Table 7) were stronger than those for the boys. Unlike with the boys, none of the estimates from the nearest neighbour matching were statistically significant but this could be because there are fewer treatment and control observations. From kernel and stratification matching, the propensity score matching produced effects of 9.1 ppt and 7.9 ppt respectively in attaining ATAR >= 85, and effects of 13.5 ppt from kernel matching and 11.0 ppt from stratification matching in attaining ATAR >= 90; girls attending selective schools had higher ATARs by 1.8 to 2.2 percentile points.

In summary, the results from propensity score matching showed that the girls increased their likelihood of attaining each of the Year 12 outcomes from attending selective schools, with the exception of ATAR >= 95. The effect sizes were largest for ATAR ≥ 90, with estimates suggesting an increase in likelihood of up to 13.5 ppt. For context, the expected success rates for ATAR ≥ 90 from probit regressions was 48.7% for the base, that of non-selective students.

These results are positive for the girls in that there is evidence of gains across three of four Year 12 outcomes, and gains are largest for the highest outcome of ATAR ≥ 90, meaning selective students were more likely to be able to attend courses and universities with greater demand. The results for boys were not as strong as that for girls, but there was evidence for boys improving their ATAR ≥ 85 and ATARs in general from attending selective schools, which would still allow students to enter a wide range of courses; the matching produced statistically significant estimates of between 6.2 ppt and 7.6 ppt for ATAR >= 85, off a base expected success rate of 63.8% for non-selective students.[21]

The differences are statistically significant and larger for the more attainable Year 12 outcomes of ATAR ≥ 90 and ATAR ≥ 85 for girls and ATAR ≥ 85 for boys. These differences suggest that the gains are larger for those with lower propensity scores, which roughly corresponds to those with lower achievement levels who were admitted to the selective schools. [22]

6. Regression Discontinuity

We discuss the methodology and review the suitability of the data for the regression discontinuity approach, describing the samples, the mean student characteristics at the cut-offs and the relationship between bandwidth and sample sizes, before presenting the results.

6.1. Methodology

The regression discontinuity approach relies on the assignment of treatment from a rating variable that is continuous and measured before treatment. At the same time, treatment is randomly determined by a cut-point, and because treatment is random the intuition of the approach is that the characteristics of participants below and above the cut-point are the same aside from the treatment. The regression discontinuity approach consequently produces a local treatment effect as comparisons are made on marginal students; students that just get in compared with students that just miss out.

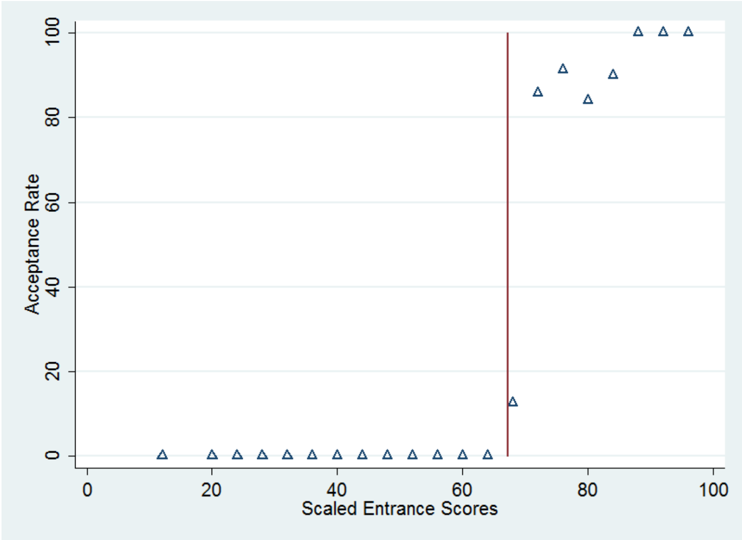
In this institutional context, the entrance exam results determine the treatment of attendance at selective schools, and the cut-points are determined arbitrarily by the size of the schools, but are nevertheless high. Figure 6 shows the acceptance rates against bins of entrance exam scores for applicants that listed one of the selective schools as their first preference. The minimum entrance

---

[21] As would be expected, much larger effects are estimated if the comparisons are made against the selective potential sample via matching. Simple unweighted regression estimation on the same sample also points to positive selective school effects of about the magnitude estimated via matching (which does not account for performance on the entrance exam).
[22] Regression estimates also pointed to positive effects from attendance of around 3 ATAR points for those students who gained entry via the disadvantaged background intake, consistent with the estimated effects being larger for those least likely to get into the selective schools. Students admitted through the discretionary intake were found to receive similar effects to other students in the selective schools.

exam score is indicated by the red vertical line which we assume to be the cut-points; students admitted via the discretionary or disadvantaged background intakes are excluded. The acceptance rates are similar to that of the other selective schools, which have different cut-offs.

Figure 6: Acceptance Rates and Entrance Exam Scores



The minimum scores at the selective school was 67.3. Bin-size of 4 points.

The variability in acceptance rates to the right of the threshold suggests the existence of a fuzzy discontinuity rather than a strict discontinuity, where all the offers are accepted by students. Up to 20% of students declined offers despite placing the school as their first preference.

Note that the appearance of the acceptance rates in Figure 6 is sensitive to the size of the intervals, which can be varied. The observation below an acceptance rate of 20, but not zero, is explained by students falling within the interval that includes the cut-off. In the empirical estimation, the bandwidth is adjusted to include students above and below the cut-offs (i.e. within a certain number of points of the cut-off). Another variation to the approach is to apply different weightings associated with distance from the cut-offs.

Non-compliers and Instrumental Variables

To account for the students who decline offers, defined as *non-compliers* in the literature (from Jacob et al. 2012), we apply two-stage least squares (2SLS) using the rating variable, the entrance exam scores (r_i). This divides the difference in outcome from the discontinuity by the difference in treatment from the discontinuity. The specification of the 2SLS is:

$$S_i = \alpha_1 + \gamma D_i + f_1(r_i) + \varepsilon_i$$

$$Y_i = \alpha_2 + \beta S_i + f_2(r_i) + \mu_i$$

S is selective school attendance, D is an indicator variable above the cut-point, while f is the different functional form for stages 1 and 2. The remaining variables are constants (α) and error terms (ε and μ) that are assumed to be identically and independently distributed.

For continuous outcome variables we apply linear regressions to both stages, and for binary outcome variables (ATAR >= 95, ATAR>= 90 and ATAR>= 85) we apply linear and probit regressions for stages 1 and 2, respectively. For the latter, replacing the first stage with probit regressions did not change the results, while the estimates were generally more negative with larger errors for the continuous outcomes. For all outcomes, we estimate cubic and quadratic functional forms of the rating variable (r) in both stages, and report the cubic estimates as it has no bearing on the results.

6.2. Regression Discontinuity Samples

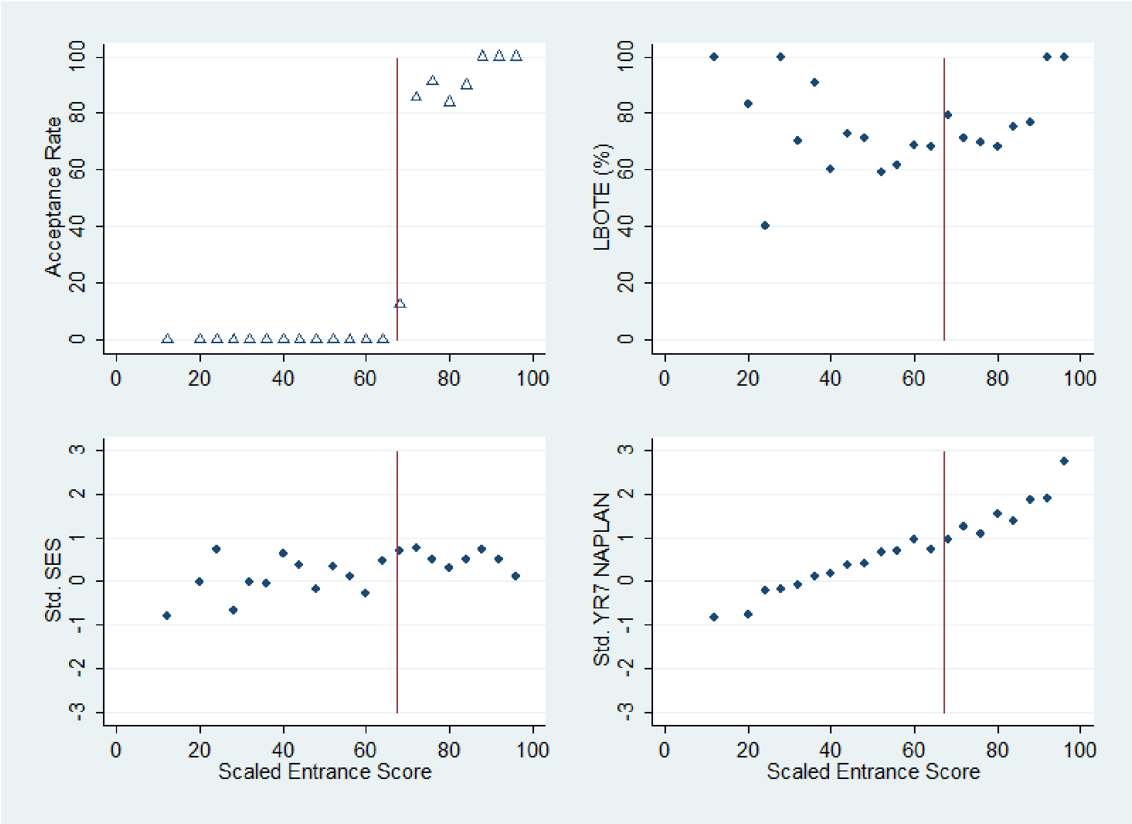
We estimate the selective school effect over the pooled sample of selective schools, recentering the entrance exam scores around the cut-offs. As offers are allocated from entrance exam results according to students' first preferences, the cut-offs for each of the schools differ according to the demand for the respective schools. For the analyses, we exclude students who accepted offers for selective schools that were their second preferences because the acceptance rates for students that received offers for second preferences were much lower. The lower acceptance rates imply that there are unobserved influencing factors or characteristics that differ between applicants who accepted or rejected their second preference offers, which is likely to lead to biased estimates from comparing the two groups. Because preferences were much weaker for certain schools, this reduced the sample size for the pooled analyses.

As with the matching approach, we excluded the new selective school from the analyses, due to the small number of applicants expressing the school as a first preference, and also because of the lower rates of acceptances from students receiving offers. Specifically, from the few applicants choosing the school as their first preference, the significant proportion who decline offers to attend the omitted school introduces doubt that the marginal students attending and not attending are essentially the same.

6.3. Student Characteristics

The regression discontinuity approach relies on random variation of assignment such that the characteristics of the students are the same in the vicinity of the cut-point. For applicants expressing as their first preference one of the selective schools, Figure 7 presents the acceptance rates (from Figure 6), and the mean student characteristics within bins of entrance scores for the selective schools: the percentage with foreign language background, the socio-economic index of parental education and occupation, and Year 7 NAPLAN scores.

Figure 7: Student Characteristics by Entrance Scores



The minimum entrance exam score at the selective school was 67.3. Bin-size of 4 points.

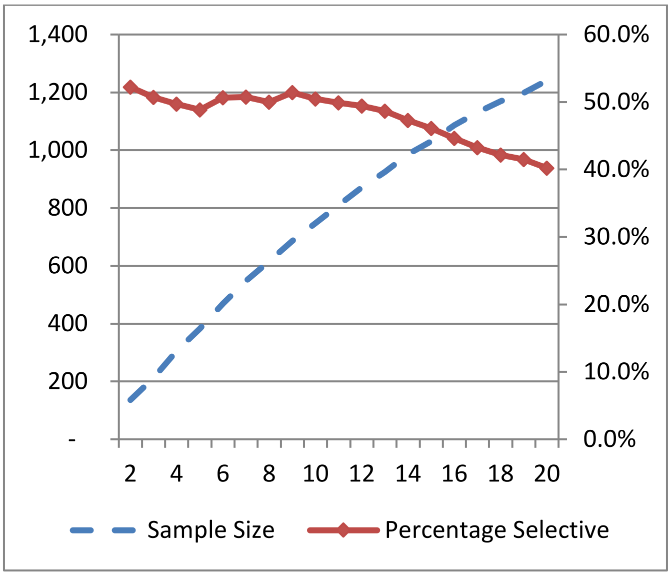
Although slightly noisier in the behaviour of student characteristics against the entrance scores, the example selective school is reflective of all the selective schools in the pooled analyses; the language and socio-economic background and prior achievement of students behave in a continuous manner in the vicinity of the cut-offs, and the visualisations support the assertion that the cut-offs were arbitrarily determined and that students are essentially the same on either side of the cut-offs. From the mean student characteristics, it appears that the data is suitable for the regression discontinuity analyses.

6.4. Bandwidths and Frequencies

In the comparisons between treatment and control groups, the bandwidth or range chosen for the pooled selective schools analysis may influence the results as this determines the number of observations surrounding the cut-offs on either side of the threshold. A trade-off exists between the increases in the sample size from a larger bandwidth, which is balanced against the weakening of the assumption that the students on either side have the same mean characteristics. This is most apparent, for example, in the trend of increasing Year 7 NAPLAN score with increasing entrance exam (Figure 7).

Figure 8 shows how the sample size increases with increasing bandwidth, as measured in entrance scores, for the pooled selective schools sample. For example, the sample size represented by the dashed blue line, increases from 136 to 384 observations moving from a bandwidth of 2 to 5 (on the left vertical axis). The red diamonds, measured on the right vertical axis, indicate the percentage of students attending selective schools within the bandwidth dependent samples; the percentage of selective students hovers around 50% before steadily decreasing with increasing bandwidth.

Figure 8: Frequencies by Bandwidth and Percentage Selective



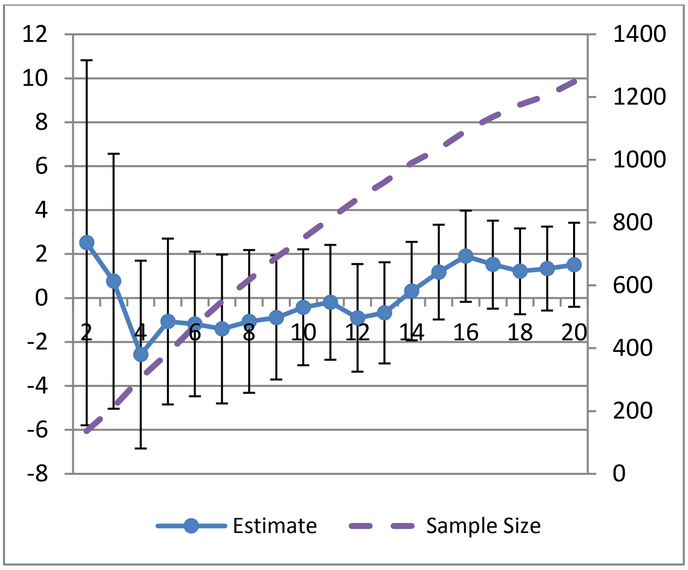
As a point of reference, when the bandwidth approaches 20 (entrance exam) points, the sample includes almost all the students who attend two of the selective schools. The sharp increases in acceptance rates (from Figure 6) fall within a bandwidth of around 5 points, while a bandwidth over 10 points appears not to be very meaningful from the perspective of identifying the marginal students; those marginally attending and those just missing the cut-off. The precise bandwidth chosen for the analyses can be an important decision but, previewing the results, in our study the results are not sensitive to the bandwidth chosen.

6.5. Results

The estimates from the fuzzy regression discontinuity analyses of the pooled selective schools sample were not statistically significant. The estimates are presented in a visual format in Figures 9 to 15 to show how they vary with bandwidth choice, on the horizontal axis; the blue dots are the estimates (left vertical axis) and the dashed purple lines represent the sample size (right vertical axis). Note that the standard errors represented by the bars narrow with increasing bandwidth as sample sizes increase. For statistically significant results, the tops or bottoms of 95% confidence intervals, which are almost twice (1.96) the size of the error bars shown, would need to be completely below or above zero respectively. These estimates are also included in Appendix A.5 in Table 12 for reference.

The lack of statistical significance in estimates across the different outcomes is indicated by the magnitude of the error bars and could be attributed to the small sample sizes. For the ATAR outcome (Figure 9), the statistically insignificant estimates are positive at 2.5 and 0.8 percentile points at bandwidths of 2 and 3 entrance exam points, respectively. With increasing bandwidth, between 4 and 14 points, estimates of the selective school effect on ATAR fall below zero, before moving upwards with increasing bandwidth to 2 ATAR points at a bandwidth of 16. This pattern of increasing estimates is expected, consistent with an increasing differential between levels of prior achievement in the marginal selective and marginal non-selective students (see positive relationship between Year 7 NAPLAN and entrance scores in Figure 7).

Figure 9: ATAR Estimates by Bandwidth



6. Regression Discontinuity

Figure 10: ATAR ≥ 95

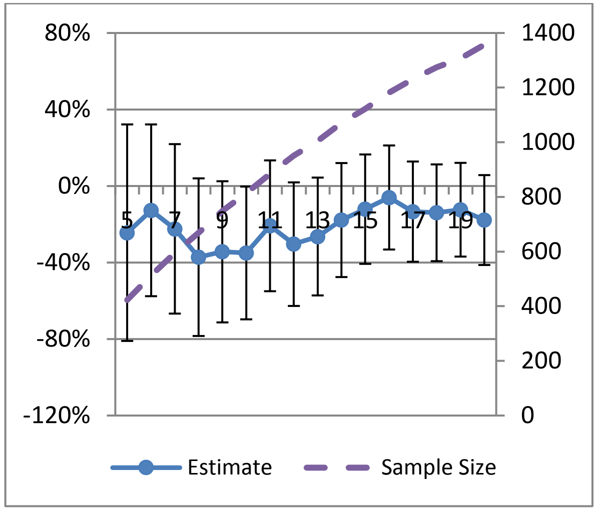


Figure 11: ATAR ≥ 90

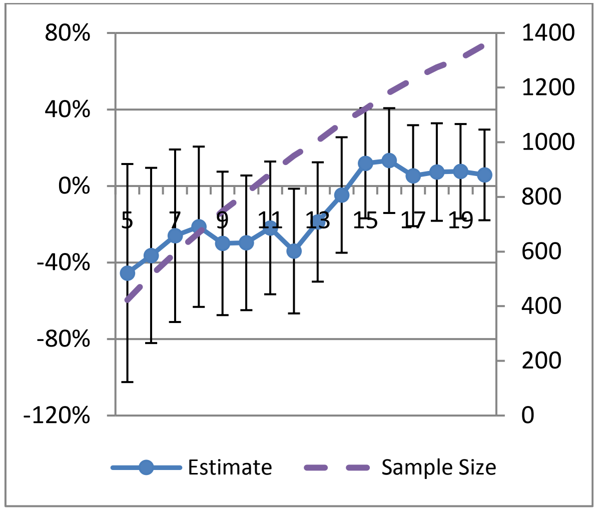


Figure 12: ATAR ≥ 85*

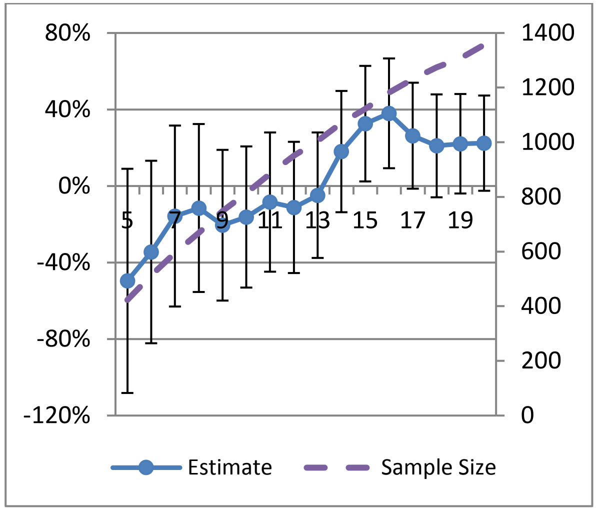


Figure 13: Mathematics

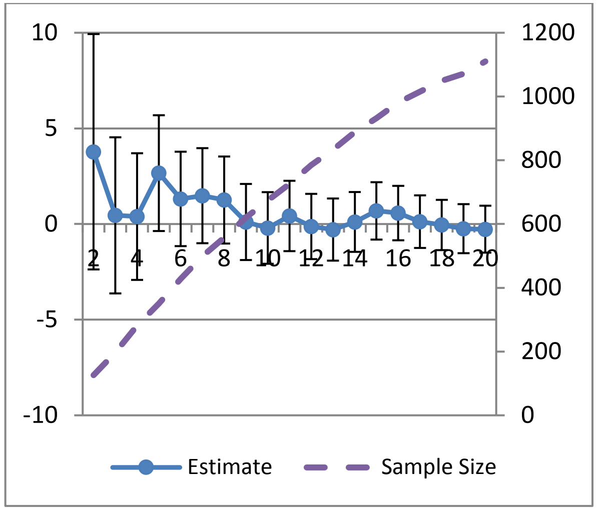


Figure 14: English

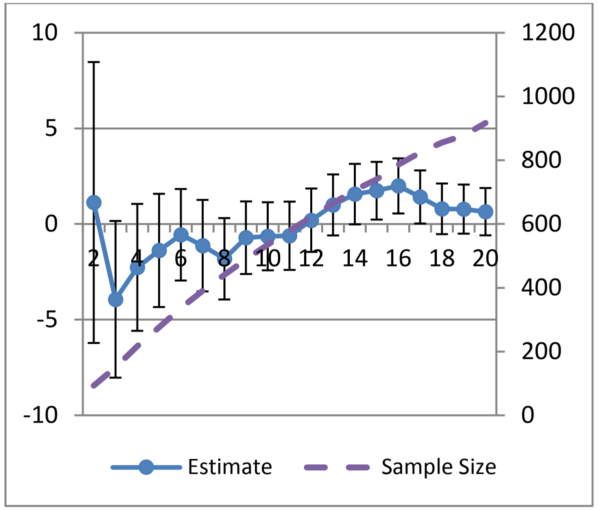
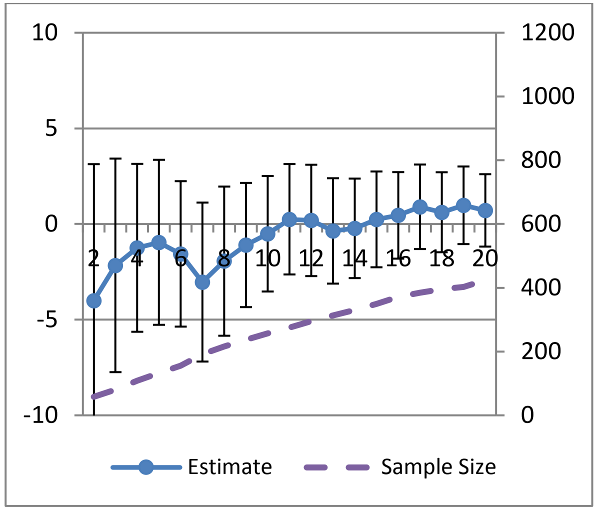


Figure 15: Physics^



^ Truncated error bar for first estimate (bandwidth = 2) ranges from -11.1 to 3.1.

The estimates for the binary outcomes of ATAR ≥ 95, ATAR ≥ 90 and ATAR ≥ 85 are mostly negative and large in magnitude but again not statistically significant with very large error bars (Figures 10 to 12). The estimates for ATAR >= 95 (Figure 10) for the bandwidth range between 5 and 7 are negative, ranging between -12.7 ppt and -24.4 ppt, but less negative compared with estimates for ATAR >= 90 and for ATAR >= 85, which ranged between -26.0 ppt and -45.5 ppt and -15.7 ppt and -49.6 ppt, respectively. Note the same vertical and horizonal scales across the binary outcomes (Figures 10 to 12), and across continuous subject outcomes (Figures 13 to 15), discussed next.

The standard errors for the binary outcomes (Figures 10 to 12) are much larger than for continuous outcomes, such as ATAR (Figure 9), due to probit regressions in the second stage of the instrumental variables which are less efficient than the linear regressions. This is also the reason, in combination with the small samples, for the lack of estimates for the binary outcomes for bandwidths less than 5 points.

The results for Mathematics, English, and Physics, as continuous outcomes, are shown in Figure 13, Figure 14, and Figure 15 for reference. These Year 12 subjects contribute to students' final university entrance results[23], and were taken by the greatest number of selective students. Note subject scores have a maximum of 50. Estimates across the three subjects were also not statistically significant, and the pattern in estimates for English and Physics were similar to those relating to ATAR and the ATAR derived outcomes, being mostly negative at smaller bandwidths before trending upwards at larger bandwidths. The pattern in estimates for Mathematics (Figure 13) was different with mostly positive estimates at smaller bandwidths, including a maximum estimate of 3.8 at bandwidth 2, before fluctuating around zero from bandwidths above 9.

7. Discussion

7.1. Matching

The results from the matching approach suggested there were small effects of selective school attendance on achievement at the end of high school. The analyses produced estimates of approximately 2 percentile points in ATAR, and estimates of between 6 ppt and 14 ppt increases in probability of attaining ATARs in the top 10% and 15% of students. The estimates are of significant magnitude for high-achieving students, and suggest that the schools provide a floor for Year 12 achievement, but they are likely overstated for not accounting for the differences in unobservable motivation between selective and non-selective students implied by the results from the entrance exam (discussed further below).

[23] This is intentionally vague in order to retain anonymity of the state and identities of the selective schools.

Two points in ATAR at the top of the distribution may be the difference in whether students get into their preferred course at their desired higher education institution, so even effects of this size may have substantial implications for individuals. Estimates of around 2 ATAR points are also similar to value-added estimates of New South Wales selective schools for Year 12 outcomes (Lu and Rickard, 2014). Between Year 9 NAPLAN and measured Year 12 outcomes, the selective school effect was just 0.04 of a standard deviation. This translates into around 1.2 of one ATAR point (0.04 x 28.9 – the SD of the uniform ATAR distribution).

An advantage of the matching approach is that it helps with conceptualizing the selection issues, of which there are two forms; first, students express a preference for attending the schools, and second, they are sorted into the schools from their performances on an academic entrance exam. We identified the comparison group from their behaviour of sitting the entrance exam, and excluded students whose achievement on Year 7 NAPLAN was below that of the minimum level at the selective schools. For the second form of selection, we identified selective students from this restricted sample as more likely having higher levels of earlier achievement, and a foreign language background for girls (Selection Model, section 5.4).

A limitation from the matching approach was that it was not able to account for both prior achievement in Year 7 NAPLAN and achievement on the entrance exam due to the balancing requirement (see Methodology, section 5.3); there was not equal representation between selective and non-selective groups, against the predicted probability of attendance, because the former outperformed the latter on the entrance exam for similar levels of prior achievement. Including Year 7 NAPLAN in the selection model does, however, increase the ease of interpretation in that Year 7 NAPLAN provides the same incentive to both selective and non-selective students as a low stakes outcome.[24]

We find matching estimates of 2 percentile points in ATAR from attending the selective schools, but they are likely overstated as they partially reflect pre-existing differences between the selective and non-selective students, such as additional motivation or latent academic ability. That is, the selective students might view admission to the selective schools as a high stakes outcome, from their outperformance over non-selective students with similar levels of prior achievement on the entrance exam. We investigate this assertion further in section 7.4 (Educational Aspiration), asking whether applicants have different levels of motivation.

[24] Replacing Year 7 NAPLAN with entrance exam in the selection model, similarly potentially leads to unfair comparisons between applicants with differing levels of motivation. Students who attain sufficiently high scores and decline offers appear to have less incentive to try on the exam because they are less invested in the outcome (admission to the selective schools).

7.2. Regression Discontinuity

The results from the regression discontinuity (RD) analyses were not statistically significant, which is consistent with the 2014 US studies (Adulkadiroglu et al. and Dobbie and Fryer), though in our case this could be attributed to the small sample of marginal students taken from one cohort. In isolation the regression discontinuity results are inconclusive, but there is information from the negative direction of the estimates for the binary outcomes of ATAR >= 85, ATAR >= 90, and ATAR >= 95.

Although regression discontinuity is a convincing method due to its resolution of the unobservables problem, there are differences in the characteristics of students for which the local effect is relevant as compared with the characteristics of selective students in general. Most notably, the marginal selective students are among the academically weakest students at the selective schools, giving rise to the possibility of a decrease in their academic performance arising from diminished self-concept. Further, these marginal selective students are compared with students who just missed out on admission that are likely to be academically stronger students in the schools they attend, which could also help explain the absence of positive effects that we found in the matching analyses.

Another possible explanation for the lack of effect is that the high levels of educational aspiration of marginal selective and marginal non-selective students means that they are both less influenced by their school environment[25]. An interpretation of applicants to selective schools as being driven and conscientious is supported by anecdotal evidence, including media reports of the coaching of applicants, and is also consistent with a large survey of applicants and their parents (Braithwaite and Kensell, 1992), which indicated that the most important reasons for applying were related to educational aspiration.

Overall, the local effect that applies to marginal students is non-existent or at least weaker than the positive effect from the matching analyses**,** and can be attributable partially to the impact of relative status in the comparisons of marginal students, but mostly from the similarity in the characteristics of both successful and unsuccessful applicants to selective schools. This refers to their similar foreign language background, their academic performance, and the reportedly high levels of both educational aspiration and also coaching.

In terms of institutional context, there was contrasting evidence of positive effects on achievement from other studies of selective schools which cater to a broader range of academically ability. As noted in the literature review (section 2.2), Shi (2017) found evidence that the selective schools benefited students of disadvantaged background and students who previously attended low-

[25] Note that we think that their educational aspirational is also relevant to the matching results, such that they positively biased for not having accounted for the additional motivation of selective students over non-selective students on average, rather than locally just for marginal students, as is the case in here.

achievement schools; SAT scores[26] were improved by 3 to 5 percentile points from selective attendance, and the share of preferences to more selective colleges increased by 4 to 6 percentage points for disadvantaged students. Zen (2016) similarly found tentative evidence (statistically insignificant estimates) of benefits to university entrance results from attendance of selective schools with lower admissions cut-offs.

7.3. Rank Effects & Self-Concept

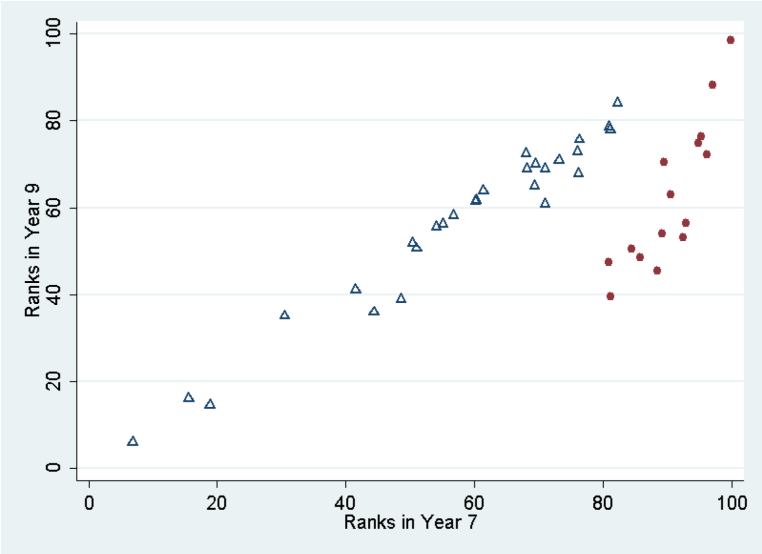
The relevance of rank effects to the selective schools stems from the potentially large changes in academic status that students experience when they enter the selective schools, by virtue of sorting into an academically competitive environment. It is possible that a negative effect on achievement from diminished academic self-concept, i.e. students' views of their own ability as informed by others, could offset potential benefits to later achievement anticipated by students from attending the selective schools. The importance of relative status in academic achievement was first documented by Marsh and Parker (1984), who found an inverse relationship between the achievement levels of schools and students' academic self-concept[27].

Figure 16 shows the extent of the change in relative status implied by changes in ranks between Year 7 and Year 9 from attending one of the selective schools, as an example[28]. It is a scatterplot of achievement ranks in Year 7 and in Year 9 for applicants of one of the selective schools; ranks are measured as within-school ranks of averaged standardized Numeracy and standardized Reading from Year 7. Non-selective and selective students are represented by blue triangles and red dots respectively.

---

[26] The Scholastic Assessment Test, a test of academic ability used in college admissions.
[27] More recently, Murphy and Weinhardt (2014) documented the influence of "rank effects" on later achievement using within-school ranks to reflect academic self-concept.
[28] This is the same selective school as was presented in comparing the student characteristics at the entrance exam cut-off for the regression discontinuity analyses (section 6.3). See Figure 7 which corresponds to Figure 17, discussed next.

Figure 16: Local Ranks in Year 7 against Local Ranks in Year 9 by Selective Attendance^



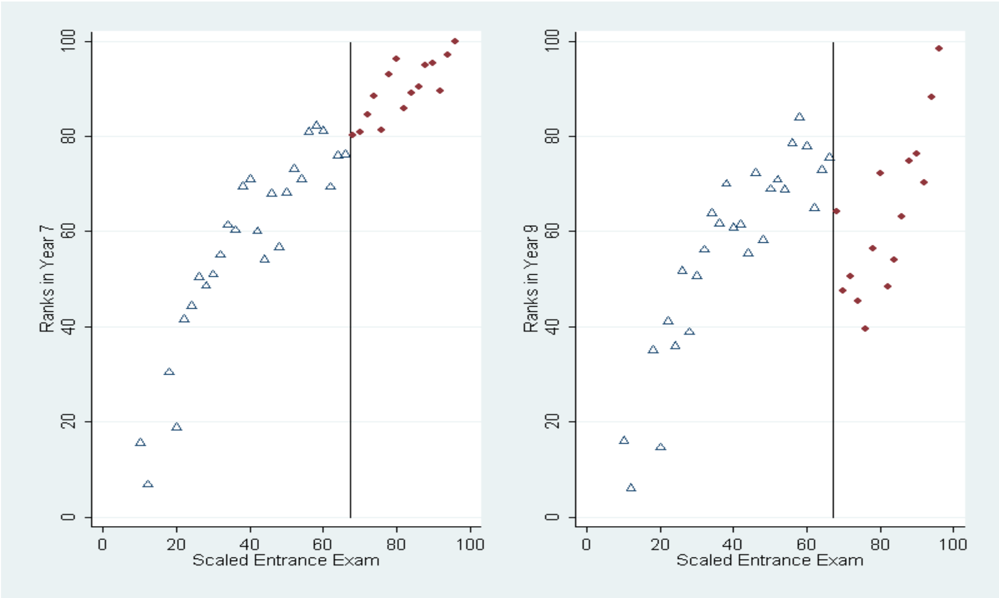
^ Relates to applicants for one selective school only. Ranks are based on results from Year 7 NAPLAN and discretionary selective students are excluded. Red dots and blue triangles represent selective and non-selective students, respectively, and each point is an aggregation of students over small intervals of Year 7 local ranks. Applicants who decline offers not shown.

There is a clear linear relationship between ranks in Year 7, along the horizontal axis, and ranks in Year 9, along the vertical axis, for the non-selective students. This hypothetical line for non-selective students ends at around 80 percentile points in Year 7 ranks, where observations of the selective students begin, but with a different trajectory. Note the difference in Year 9 ranks between non-selective and selective students near this point, where the Year 7 ranks overlap, which is as large as 40 percentile points. This difference in ranks would be even larger but for the exclusion from the plot of selective students with lower achievement, who were admitted through the secondary discretionary and disadvantage background intakes (for consistency with the analyses); these secondary intake students explain the absence of observations between 0 and 40 in ranks in Year 9.

The potential influence of rank effects at selective schools is also evident in the regression discontinuity approach, which compares marginal selective students, who are likely to be among the academically weakest at the selective schools, with marginal non-selective students, who are likely to be academically stronger than others in the school they attend.

Figure 17 shows again the extent to which there is a change in relative status from attending a selective school, but against performances on the entrance exam. The left panel is a plot of Year 7 ranks and the right panel is a plot of Year 9 ranks, which together show the change that is experienced by marginal selective, compared with marginal non-selective students. Based on their Year 7 ranks, the selective students follow a similar ascending pattern of within-school ranks with increasing entrance exam results. Like in Figure 16 before, there is a decrease in ranks in Year 9 for the selective students, with a difference between marginal selective students and marginal non-selective students, which is between 20 and 40 percentile ranks, depending on weighting.

Figure 17: Local Ranks in Year 7 and Local Ranks in Year 9 against the Entrance Exam^



^ Ranks in Year 7 are shown in the left panel, and ranks in Year 9 are shown in the right panel. Ranks are based on Year 7 NAPLAN for both year levels. Blue triangles and red dots represent non-selective and selective students respectively, and each observation represents students aggregated over small intervals of the entrance exam, as per Figure 16.

An indirect consequence of sorting into academically competitive selective schools is that many students experience a sharp increase in peer quality, which results in downgrades in their own relative academic status, and could negatively affect their later achievement. That is, marginal selective students are likely to be among the academically weakest at the selective schools.

While the offsetting influence of negative rank effects can help explain how the selective school effect might be lower than expected, the overall influence from rank effects for selective students may be less problematic as they are high-achievers and because they enter the schools at a relatively later stage of schooling. Research on academic self-concept and academic selection across schools indicates that the negative effects on self-concept from having high-achievement peers were weaker for high-achieving students (Trautwein et al. 2009). Similarly, Salchegger (2016) found that countries with earlier explicit school-level tracking produced larger negative effects on academic self-concept, which is not the case here.

For non-academic outcomes, the relative achievement rank of students has been shown by Elsner and Isphording (2015) to increase the likelihood of students engaging in risky behaviours, such as smoking and drinking. In the selective schools context, Shi (2017) finds that relative rank appears to influence the college enrolment decisions of students. Specifically, students with higher relative ranks expressed greater preferences for attending STEM-intensive colleges, with average increases of 1.8 to 2.4 percentage points from selective school attendance.

7.4. Educational Aspiration

We sought to ascertain the extent of unobserved educational aspiration in selective school applicants relative to the general population, and also between selective and non-selective students, via exploratory analyses and also a synthesis of the relevant research and empirical facts.

An earlier study of the surveyed reasons for applying to selective high schools by students and their parents strongly supports the interpretation that selective students have greater levels of educational aspiration. Braithwaite and Kensell (1992) found that academic emphasis, including the perceived advantage for future careers, and the reputation of the schools were the most important factors for accepting offers to selective high schools; the reputation of the schools referred to perceived social, cultural and sporting reputations as well as academic standing, while reputation effects were much weaker for newly created selective high schools.

The high representation of students from a foreign language background at selective schools (at over 75% in the exam sample, Table 3) suggests that for these students their ambitions are closely related to their immigrant background, especially given the positive association documented between educational attainment and foreign language background[29]. It could be argued that the selective schools symbolise the successful positive selection that is designed in immigration policy. E.g. Miller (1999) noted that Australian immigration policy allowed for the positive selection of individuals for adapting to the labour market.

Linking the selective schools to historical immigration patterns, a significant proportion of migrants to Australia were of Asian background for the time period relevant to this case study[30]. It is possible then that the positive associations between educational attainment and foreign language background are inter-related with the well-documented outperformance in academic achievement by students of Asian background (see for example, Jerrim 2015 [31], Hsin and Xie, 2014 [32], and Dandy and Nettelbeck, 2002). The research attributes the academic outperformance mostly to educational and occupational aspiration and cultural factors, providing guidance for how we interpret the exploratory analyses

---

[29] Gemici, et al. (2014) and Homel et al (2012), refer to section 2.4, Literature.

[30] Individuals from Asia (from South-East, North-East, Southern and Central Asia) comprised 55.3% of permanent settlers in 2016, up from 21.1% in 1996, 3401.0 Overseas Arrivals and Departures, Dec 2016, Australia Bureau of Statistics and Asian Immigration, Current Issues Brief 16 1996-97, Parliamentary Library.

[31] Jerrim (2015) documents the outperformance on the Programme for International Student Assessment (PISA) of students from an East-Asian background in Australia, explaining most of the variance in performance with multiple factors relating to subjective norms, instrumental motivation, attitudes to school and future aspirations.

[32] Relying on teacher-evaluations of student behaviours and attitudes, Hsin and Xie (2014) explained the outperformance of Asian-Americans on academic achievement by their exertion of greater academic effort and not to cognitive abilities or socio-demographics.

presented next. There are indications that the relationship between achievement and geography is more pronounced for subgroup of students from East and South-East Asia.

Research from the UK supports the interpretation that for many selective students their high levels of education aspiration are closely related to their immigrant backgrounds, and that they can be attributed to cultural attitudes. Abbas (2007) interviewed students and their parents from South Asian backgrounds with reference to the selective schools, noting that South Asian parents are especially encouraging of their children in education[33]. Abbas found that social class was the strongest factor in gaining entry; middle class families were highly motivated and possessed the requisite economic, cultural and social capital, while some working-class parents possessed strong attitudes towards selective education, regardless of resources.

Research on the link between socioeconomic background and educational aspiration also supports the interpretation that both selective and non-selective applicants to selective schools have high levels of educational aspiration, which can help explain the lack of effect on achievement from selective attendance. Many selective school applicants were from an advantaged socio-economic background, with almost half of selective students and over 40% of applicants in the exam sample were in the top quartile of SES (from Table 3). For example, Bowden and Doughney (2010) find from a survey of Year 9 to Year 12 high school students from the Western suburbs of Melbourne that those from higher socio-economic backgrounds were more likely to aspire to attend university than expected if SES were an independent variable, while parental educational attainment has also been shown to have a strong positive association with high school completion in the Australian context (Todhunter, 2009)[34]. There was similar research from Chevalier et al. (2009) of expectations of 15 year-old students for attending university in the UK using PISA data[35]; estimates were around 10 ppt increases in probability for having a socio-economic background in the top category among 5, compared to lowest baseline category of SES.

Finally, there is also anecdotal evidence of the additional effort exerted by applicants to selective school. Private tutoring for preparation for selective entry tests in Australia is endemic, if the popular press is to be believed[36]. Another phenomenon that has been observed within the schools is that many

---

[33] Taylor & Hegerty (1985) and Ghuman (1997, 2002) cited in Abbas (2007).

[34] The association was as large as 29 ppt increases in likelihood of high school completion for male students at the half median level of real household equivalized income, if both parents were well-educated, being: university-educated for father and completed Year 12 for mother.

[35] Programme of International Student Assessment (PISA) - triennial international standardized tests of mathematics, reading and science and context questionnaires on family backgrounds, students' lives, aspects of schooling and context of instruction.

[36] See for example "Testing times: selective schools and tiger parents", Anna Broinowski, Sydney Morning Herald, 01/24/2015 http://www.smh.com.au/good-weekend/testing-times-selective-schools-and-tiger-parents-20150108-12kecw.html (accessed 28 June 2017). We note that tutoring services for the specific purpose of attaining entry to the selective schools were easily found.

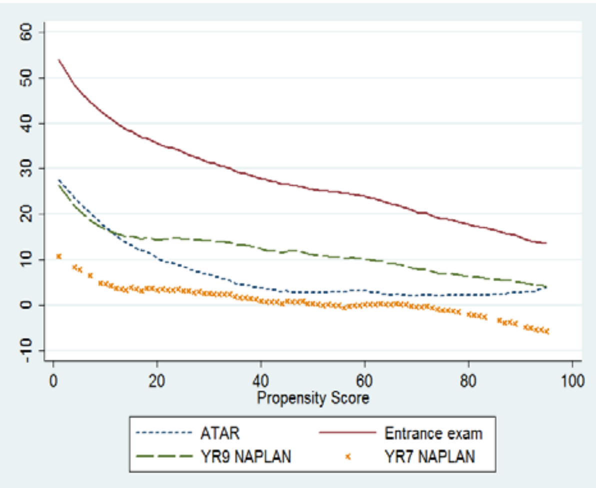
students who make it into the schools find their academic performances in the school diminished. Both phenomena point to applicants of selective schools of being conscientious, having high educational aspirations, as well as having access to educational and family resources (in terms of time and financially).

7.4.1. Exploratory Analyses

We undertake additional exploratory analyses which support the interpretation of selective students having greater levels of educational aspiration than non-selective students who sat the entrance exam. First, percentile rank analyses of the different measures of achievement based on propensity scores showed that selective students outperformed non-selective students the most on the entrance exam, followed by Year 9 NAPLAN, ATAR and Year 7 NAPLAN (discussed next). Second, value-added regressions of ATAR showed that the inclusion of the entrance exam improved the predictive power for selective students but not for non-selective students (refer to Appendix A6). We interpret this to suggest that selective students treated both ATAR and the entrance exam as high stakes outcomes.

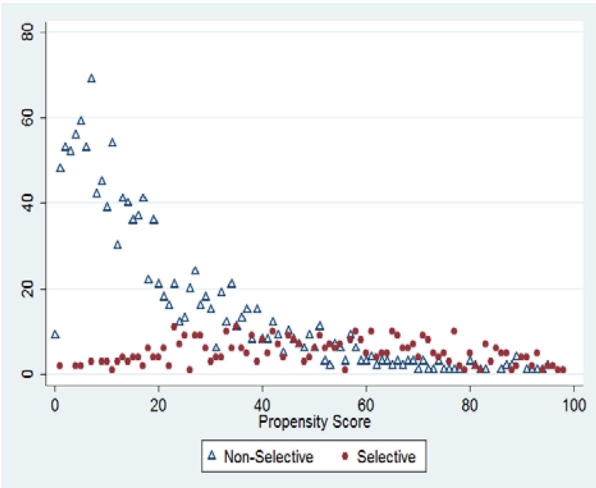
To compare the four achievement measures between selective and non-selective students, we calculate percentile ranks over the exam sample for each of the achievement measures. By creating relative measures from the same sample of students, this allows us to make comparisons between achievement measures that reflect different points in time and have different purposes.

Figure 18: Differential in Percentile Ranks (S – NS) by Achievement Measure*



* Exam sample with prior achievement above min.

Figure 19: Frequencies by Propensity Score Intervals*



* over 1 pt intervals in propensity score

Figure 18 plots the differential in percentile ranks between selective and non-selective groups - on the verticial axis - along propensity scores - on the horizontal axis. Each line represents a different achievement measure: the yellow crosses reflect Year 7 NAPLAN, the red line is the entrance exam, the dashed line is Year 9 NAPLAN, and the small dotted line is ATAR. Figure 19 presents the frequency counts of selective (red dots) and non-selective students (blue triangles) over 1 pt intervals in propensity scores between 0 and 100.

The smaller differences in percentile ranks against the propensity scores for Year 7 NAPLAN, between -10 and 10 percentile points, were as expected given that the matching approach requires that treatment (selective) and control (non-selective) groups not differ in pre-treatment characteristics, which include Year 7 NAPLAN as prior achievement. In practice, the requirement was satisfied was by splitting the propensity scores into 8 and 9 blocks for boys and girls respectively, and conducting t-tests at the 1% level of probability (see sections 5.3 and 5.4).

For the other measures, the difference in relative performance between selective and non-selective students was largest for the entrance exam, and smaller for Year 9 NAPLAN and generally even smaller for ATAR. The outperformance by selective students across all measures was higher at lower propensity scores, which is to be expected due to the admission mechanism where offers are conditional on entrance exam performance, and because the propensity scores are based only on Year 7 NAPLAN, in addition to language and socioeconomic background. This pattern is consistent with the lowess plots of ATAR against the propensity scores (Figure 5, section 5.5), where the largest differences in ATAR between selective and non-selective students were for those with the lowest predicted probabilities of attendance.

The differences in ranks for the entrance exam are of primary interest for the reason that they can help with interpreting the results from the matching approach. Differences were as large as between 25 and 35 percentile ranks for the entrance exam between 20 and 60 pts in propensity score, which is the range that provides a more reliable indication of the differential in achievement, with greater numbers of selective and non-selective students. Note that there were few selective students in the lower and upper ranges (between 0 and 20 pts and 80 and 100 pts), at below 10 students, and that there similarly few non-selective students above 60 pts.

As we are inclined to interpret Year 7 NAPLAN as reflecting academic ability as the stakes are low, it is somewhat surprising that the gap between selective and non-selective students on the entrance exam remains as large as over 10 percentile points in propensity scores above 90 pts. This can plausibly be interpreted as additional effort reflecting education aspiration as a response to the incentive of attending the selective schools. The small differentials for Year 9 NAPLAN, when successful applicants have only attended the selective schools for a few months, also supports the effort interpretation of the differential in entrance exam between selective and non-selective students with the same propensity scores.

In additional analyses in Appendix A6, we estimate regressions of ATAR separately for selective and non-selective students, while varying the specification to include the different measures of prior achievement. When Year 7 NAPLAN was replaced by the entrance exam as prior achievement in regressions of ATAR, the R-squares improved by 0.09, from 0.18 to 0.28, for selective students, but were improved only marginally, by 0.02, for non-selective students; only basic student characteristics were included in the regressions.

Due to the incentive structure of the admissions process where offers to the selective schools are tied to performance on the entrance exam, the entrance exam provides an example of when effort appears to be variable among applicants[37]. From this we conclude that the entrance exam can be considered a high stakes outcome for selective students in the same manner as the university entrance results.

From the analyses of both achievement differentials based on propensity scores and the value-added regressions, we interpret much of the difference in results between the entrance exam and Year 7 NAPLAN of selective students, compared with their non-selective counterparts, to reflect greater short-term effort and motivation that is consistent with having greater levels of educational aspiration.

8. Limitations

We note that the results from our study are subject to certain limitations. First, the study relied on only one cohort of students, and replicating the analyses across several cohorts would assuage concerns relating to the generalisability of the results. Our results are, however, consistent with those of other studies on selective high schools with better quality data, including Zen's (2016) NSW study, and studies on related education topics (see Literature Review, section 2).

A related concern is that our selective school effect is an aggregate of a relatively small set of schools, meaning that our results are influenced by the circumstances of specific schools. As such, we interpret the estimates in its appropriate context of comparing the characteristics of the selective students in relation to their differences from the student population, discussed in Academic Selection (section 3.2) and Descriptive Statistics (section 5.2). We excluded one selective school for which applicants and students were noticeably different from the other selective schools (detailed in Selective Schools, section 5.1.1).

More importantly, there is a question of interpretation of what the achievement tests measure. We interpreted the achievement tests as reflecting a combination of ability and effort associated with the incentives provided by the test (see Prior Achievement, section 5.3.2). For instance, the university entrance results can be thought of as reflecting fully expressed potential in achievement performance in individuals, as a multiplicative of effort and ability, while for applicants who tried their best on the entrance exam, the performance differential between the entrance exam and Year 7 NAPLAN could arguably be described as previously unexpressed latent ability or achievement potential.

---

[37] There has been greater discussion conceptually about the role that effort (grit or persistence) plays in achievement; Duckworth (2016) defines achievement as the output of skill and effort, where skill itself is a function of effort and talent, whereas the well-cited conceptual framework of an education production function (Todd and Wolpin 2003) describes achievement as a cumulative process in which prior achievement is an input.

The methodological limitation of the research is the existence of unobservable characteristics of applicants to selective schools, and those of students who are admitted to selective schools, which was also incidentally the key challenge of the research. This concern was highlighted by the initial background information that the demand for selective school attendance was high (e.g. from the percentage of successful applicants, roughly 33% in our study), and media reports that applicants have been coached and exert greater effort in preparation for the entrance exam.

We sought to address the concern of positive selection in applicants by identifying the sample of students who sat the entrance exam. Additional exploratory analyses, however, suggested that the educational aspirations of applicants who were accepted into the selective schools were much higher than those of the unsuccessful applicants. The positive effects from the matching were consistent with that expected from the use of low stakes standardized test scores, which reflected academic ability but did not account for educational aspiration, so we anticipate a large part of the effect on ATAR to be positive selection in the form of increased effort.

While applying the fuzzy regression discontinuity method with the entrance exam appeared to account for positive unobservables (i.e. access to resources and increased effort on the part of selective school applicants), as both marginal selective and marginal non-selective were identical in characteristics by design, the small sample sizes meant that the standard errors were very large, leading to estimates which were not statistically significant; the estimates were generally not positive for smaller bandwidth or range, with the exception of those for Mathematics (Figure 13) and for only two point estimates with large standard errors, for ATAR (Figure 9).

Finally, estimating value-added effects is an exercise in abstraction, one which involves the measurement of academic achievement, and relies on an interpretation of the results in its institutional context. As noted by Clark and Del Bono (2016), the effect size is dependent on the counterfactual school environment, which may tilt the advantage towards selective schools in different areas and for certain subgroups of students. For example, larger effects were found for disadvantaged students with lower quality outside options attending a selective high school from the North Carolina case study (Shi 2017). While we emphasize the importance of university entrance results on subsequent university and course enrolments, we acknowledge the narrow focus of the study in ignoring other outcomes from schooling. [38]

[38] We note again the path of research inquiry pursued by Clark and Del Bono (2016) on longer-run outcomes. In addition, the poor performance of selective students for their first year grades in university remains an open question (from Dobson and Skuja, 2005).

9. Conclusion

This study asks whether selective schools improve the Year 12 achievement of their students, beyond what they would achieve in non-selective schools. We follow a cohort of students through high school from an anonymized Australian state, comparing the end-of-high school results of students who attend selective schools with other students who sat the entrance exam. We estimate the selective school effect via two methods: propensity score matching, which compares students of similar background and prior achievement, and regression discontinuity (RD), which compares marginal selective and non-selective students on the basis of the entrance exam.

Our results point to a small effect in terms of university entrance ranks, consistent with similar studies from the UK and the USA and other Australian studies. In a roughly contemporaneous study for 18 selective schools in NSW, Zen (2016) finds limited and insignificant effects on university entrance results by conducting in-depth regression discontinuity analyses. Both this and Zen's study improve upon Rickard and Lu (2014), who previously estimated the selective school effect on university entrance results with value-added regression analyses.

In this study we find mostly small or negative statistically insignificant estimates from the RD, which can likely be attributed to the small sample sizes from the reliance of only one cohort of students. Although we find estimates of 2 percentile points from the matching, they are likely overstated for not having accounted for additional positive selection between selective and non-selective students with similar levels of prior achievement, beyond that which exists from applying to the schools. There is also evidence, with the same caveat, that the selective schools provide a floor for Year 12 achievement; estimates are between 6 ppt and 14 ppt increases in probability of attaining ranks in the top 10% and 15% of students.

A small selective school effect is consistent with higher levels of educational aspiration of both selective students and non-selective applicants who performed well on the entrance exam. The exploratory analyses indicated that selective students performed better on the entrance exam than applicants who went on to attend non-selective schools, and who were also identified as similar in their probability of selective attendance from Year 7 information. Meaning, the matching estimates likely reflect pre-existing differences between selective and non-selective students, such as additional motivation or latent academic ability. In support, value-added regressions suggested that both ATAR and the entrance exam were high stakes outcomes for selective students, but not for non-selective students.

A large survey of applicants also showed that the most important reasons for accepting offers to selective schools were for their academic emphasis, including perceived advantage for their future careers, and reputational benefits (Kensell and Braithwaite, 1992). In addition, applicants to selective schools were disproportionately from immigrant and socio-economically advantaged backgrounds, both of which has been found to be correlated with higher levels of high school completion and educational aspiration. Reports that applicants have been coached and exert great effort in preparation for the entrance exam provide further indication of the positive selection in the applicants.

This study contributed to the growing selective schools' literature that finds little general effect from selective school attendance on academic achievement. There are some indications, however, that selective schools confer advantages to certain subgroups of students. Shi (2017) finds from a North Carolina case study that attendance benefits disadvantaged students in performances on SATs and in their preferences for attending more selective colleges[39]. Zen (2016) also finds tentative evidence of positive effects on university entrance results from attendance of lower-achievement intake selective schools.

Overall, it is perhaps not so surprising that the effects of selective schools on ATARs might be small for high-achieving students who are among the most driven and motivated, having been positively selected from a group of students who have already expressed these positive character traits, by applying to the selective schools. The findings are consistent with the interpretation that the selective schools' policies encourage the educational and career aspirations of both their students as well as applicants who attended other schools.

[39] It appears that the North Carolina selective school's district-based quota admission system, which gives rise to a broader range of academic ability in the student intake, may play an important role for the result.

References

Abbas, T. (2007), British South Asians and Pathways Into Selective Schooling: Social Class, Culture and Ethnicity, British Educational Research Journal, 33(1), 75-90.

Abdulkadiroglu, A., Angrist, J., and Parag, P. (2014), The Elite Illusion: Achievement Effects at Boston and New York Exam Schools, Econometrica, 82(1), p137-196.

Angrist, J. (2014), The Perils of Peer Effects. Labour Economics (30), p98-108.

Becker, S. and Ichino, A. (2002), Estimation of Average Treatment Effects Based on Propensity Scores, The Stata Journal, 2(4), p358-377.

Bowden, M. and Doughney, J. (2010). Socio-Economic Status, Cultural Diversity and the Aspirations of Secondary Students in the Western Suburbs of Melbourne, Australia. Higher Education, 59(1), 115-129.

Chevalier, A., Gibbons, S., Thorpe, A., Snell, M., and Hoskins, S. (2009). Students' academic self-perception. Economics of Education Review, 28, 716-727.

Clark, D. (2010), Selective Schools and Academic Achievement, The B.E. Journal of Economic Analysis & Policy, 10 (1).

Clark, D. and Del Bono (2016), The Long-Run Effects of Attending an Elite School: Evidence from the United Kingdom, American Economic Journal: Applied Economics, 8(1): 150-176.

Cleveland, W. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. Journal of American Statistical Association, 74 (368), 829-836.

Coe, R., Jones, K., Searle, J., Kokotsaki, D., Mohd Kosnin, A., Skinner, P. (2008), Evidence on the effects of selective educational systems: A report for the Sutton Trust, Centre for Evaluation and Monitoring, Durham University.

Cunha, F. and Heckman, J. (2008). Formulating, Identifying, and Estimating the Technology of Cognitive and Noncognitive Skill Formation. The Journal of Human Resources, 43 (4), 738-782.

Dandy, J. and Nettelbeck, T. (2002). The Relationship between IQ, Homework, Aspirations and Academic Achievement for Chinese, Vietnamese and Anglo-Celtic Australian School Children, Educational Psychology, 22(3), p267-275.

Davidoff, I. and Leigh, A. (2008). How Much do Public Schools Really Cost? Estimating the Relationship between House Prices and School Quality, The Economic Record 84(264) 193-206.

Dobbie. W. and Fryer, Jr R. (2014), The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools, American Economic Journal: Applied Economics 6(3), p58-75.

Dobson, I. and Skuja, E. (2005), Secondary Schooling, Tertiary Entry Ranks and University Performance, People and Place, 13(1), p53-62.

Duckworth, A (2016). Grit, The Power of Passion and Perseverance. Simon and Schuster.

Education and Training Committee, Inquiry into the Education of Gifted and Talented Students (2012), Parliamentary Paper of Victorian Government, No. 108 Session 2010-2012.

Duflo E., Dupas P., and Kremer, M., 2011, American Economic Review, 101 (August 2011): 1739-1774.

Estrada and Gignoux (2017), Benefits to Elite Schools and the Expected Returns to Education: Evidence from Mexico City, Development Bank of Latin America, Working Paper 2017/02.

Elsner, B., and Isphording, I. (2015). Rank, Sex, Drugs and Crime. IZA Discussion Papers, No. 9478.

Fox, J. and Weisberg, S. (2010). Nonparametric Regression in R: An Appendix to an R Companion to Applied Regression, 2nd edition. Revised December 2010.
Gemici, S., Bednarz, A., Karmel, T., and Lim, P. (2014), The factors affecting the educational and occupational aspirations of young Australians. Longitudinal Surveys of Australian Youth, Research Report 66.
Heckman, J. and Rubinstein, Y. (2001). The importance of Noncognitive Skills: Lessons from the GED Testing Program. AEA Papers and Proceedings, 91(2), 145-149.
Hoekstra, Mouganie and Wang (2016), Peer Quality and Academic Benefits to Attending Better Schools, NBER Working Paper No. w22337.
Homel, J., Mavisakalyan, A., Nguyen, HT., Ryan, C. (2012), School completion: what we learn from different measures of family background. Longitudinal Surveys of Australian Youth, Research Report 59.
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, Working Paper 07/13.
Hsin, A. and Xie, Y. (2014). Explaining Asian Americans' academic advantage over whites. Proceedings of the National Academy of Sciences of the United States of America, 111(23), p8416-8421.
Imbens, G. (2004), Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A review, The Review of Economics and Statistics, 86 (1), p4-29.
Jacob, R. Zhu, P, Somers, M-A, and Bloom, H.. (2012), A Practical Guide to Regression Discontinuity, MDRC (www.mdrc.org) working paper.
Jackson, CK. (2010), Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago, Economic Journal, 120 (549), p1399-1429.
Jerrim, J. (2015). Why do East Asian children perform so well in PISA? An investigation of Western-born children of East-Asian descent. Oxford Review of Education, 41(3), p310-333.
Jonsson, J. and Mood, C. (2008), Choice by Contrast in Swedish Schools: How Peers' Achievement Affects Educational Choice. Social Forces 87(2), 741-765.
Lamb, S. (2007). School Reform and Inequality in Urban Australia. A Case of residualising the Poor. International Studies in Educational Inequality, Theory and Policy Volume 3: Inequality: Educational Theory and Public Policy, 1 – 37.
Lamb, S. (2008). Selective-Entry Schools: the Need for a Re-think. Professional Voice (magazine for the AEU Victoria Branch), Volume 5 Issue 3, 17-22.
Lavy, Paserman and Schlosser (2012), Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom, The Economic Journal, 122 (March), 208-237.
Lavy, V., Silva, O., and Weinhardt, F. (2012). The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools. Journal of Labour Economics, 30(2), 367-414.
Lu, L. and Rickard, K. (2014). Value added models for NSW government schools. Centre for Education Statistics and Evaluation. Technical Paper / July 2014. Education and Communities, Office of Education, NSW government.
Marsh, H. and Hau, K. (2003), Big-Fish-Little-Pond Effect of Academic Self-Concept, American Psychologist, 58(5), p364-376.

Marsh, H. and Parker, J. W. (1984), Determinants of Student Self-Concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well?, Journal of Personality and Social Psychology, 74(1), p213-231.

Miller, P. (1999), Immigration Policy and Immigrant Quality: The Australian Points System. The American Economic Review, 89 (2), 192-197.

Murphy, R. and Weinhardt, F. (2014), The Importance of Rank Position, CEP Discussion Paper No 1241.

NSW Government, (April 2015). Selective High Schools Entry in Year 7 2015 – minimum scores, Education and Communities.

Park, A., Shi, X., Hsieh, C. and An, X. (2015), Magnet high schools and academic performance in China: A regression discontinuity design. Journal of Comparative Economics, 43, p824-843.

Pop-Eleches, C. and Urquiola, M. (2013), Going to a Better School: Effects and Behavioral Responses, American Economic Review, 103 (4).

Sacerdote, B. (2011). Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? In: Hanushek, E., Machin, S., Woessman, L. (Eds.), Handbook of the Economics of Education, Volume 3, Elsevier.

Salchegger, S. (2016). Selective School Systems and Academic Self-Concept: How Explicit and Implicit School-Level Tracking Relate to the Big-Fish-Little-Pond Effect Across Cultures. Journal of Educational Psychology, 108(3), 405-423.

Shi, Y. (2017). Who Benefits from Selective Schools? Evidence from High-Achieving Math and Science Students in North Carolina. Working Paper, Society of Labor Economists Conference, May5-6, Raleigh, North Carolina.

Todhunter, J. (2009, July). The Effect of Permanent Income on Completed Schooling in Australia. HILDA Survey Research Conference, Melbourne (pp 16-17).

Todd, P. and Wolpin, K. (2003), On the Specification and Estimation of the Production Function for Cognitive Achievement, The Economic Journal, 113 (February), F3-F33.

Trautwein, U., Ludtke, O., Marsh, H., and Nagy, G. (2009). Within-School Social Comparison: How Students Perceive the Standing of Their Class Predicts Academic Self-Concept. Journal of Educational Psychology, 101(4), 853-866.

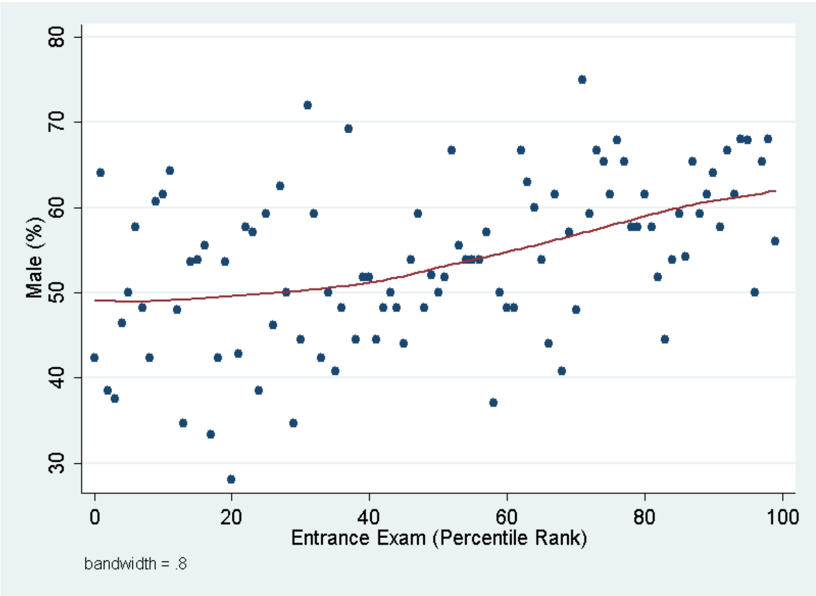
Zen (2016). The Impact of Selective High Schools on Student Achievement. Honours Research Essay in Economics (newly accessible in March 2018), University of New South Wales. https://www.business.unsw.edu.au/about/schools/economics/degrees/honours/current-past-honoursstudents.

Zhang, H. (2013), The Mirage of Elite Schools: Evidence from Lottery-based School Admission in China, unpublished manuscript, dated April 2013.

Appendix

A1. Academic Selection

Figure 20: Locally Weighted Regression of Percentage Male & Entrance Exam



Each dot represents at least 26 students along percentile ranks of the entrance exam.

We do not distinguish between selective and non-selective students.

A2. The Matched Data

Adding the Entrance Exam Data

To allow for comparisons of education outcomes and control for prior achievement, we match the entrance exam data to the NAPLAN and Year 12 data. Due to the absence of a unique identifier between the datasets there is some loss of observations, as described below.

Table 8: Fuzzy Matching between the Entrance Exam and NAPLAN/Year 12 datasets

		N	(%)	Cum.
(1)	Exact match	2,432	85.2	85.2
(2)	Exact match with cleansed data*	37	1.3	86.5
(3)	Fuzzy match with cleansed first name, surname and birthdate	15	0.5	87.0
(4)	Fuzzy match with cleansed surname and birthdate	202	7.1	94.1
(5)	Fuzzy match with cleansed first name and birthdate	10	0.4	94.4
(6)	Fuzzy match with cleansed first name and surname	45	1.6	96.0
(7)	Unmatched	115	4.0	100.0
		2,856	100	

* removing spaces, non-alphabetic characters and brackets

Fuzzy matching, which matches observations where the spelling in names are very close, was applied to students' first names, surnames and birthdates, producing match rates of over 85.2% (Table 8). We opt for using the data from match types (1) to (4), where the most lenient form of matching is a fuzzy match between cleansed surname and birth date. Dropping 170 observations (including 15 selective school students), there are 2,686 observations in the resulting sample.

Note that 4 students (1 selective school student) in the entrance exam data were not in the Year 9 NAPLAN dataset with a restricted age range of 14 to 16, leaving 2,682 observations.

A3. Data Sample - Further Details

Secondary Intake Students

From the sample of students who sat the entrance exam, we first exclude students who were not admitted strictly on the basis of their entrance exam results (Table 9). These are students admitted through a discretionary intake or a disadvantaged background intake. For the discretionary intake we exclude these students to reduce the possibility of positive bias in the event that these students were selected for favourable characteristics that are not reflected in the data.

We distinguish between students in the disadvantaged background intake who attained sufficiently high entrance exam scores and those who did not. Excluding the latter group of 56 and also the 46 students admitted in the discretionary stream, the remaining pool of students in the matched sample is 2,580.

Table 9: Selective Students by Intake

		N
Matched Sample		2,682
less Discretionary Intake	46	2,636
less Disadvantaged Background Intake (below min)	56	2,580

A4. Matching Approach

Table 10: Characteristics of Non-Selective ("NS") and Selective ("S") Students

	Boys		Girls	
	NS	S	NS	S
N	781	340	739	252
LBOTE (%)	76.3	77.1	75.2	84.1
SES Category (%)				
Quartile 1	14.6	10.3	15.3	12.7
Quartile 2	15.9	13.8	18.3	11.9
Quartile 3	27.3	26.5	26.4	27.8
Quartile 4	42.3	49.4	40.1	47.6
YR7 Sector (%)				
Government	58.5	58.2	50.3	50.4
YR9 Sector (%)				
Government	54.7		48.8	
Year 7 NAPLAN Results				
Std. NUM				
Mean	1.4	2.3	1.0	2.0
SD	0.9	0.9	0.8	0.9
%Missing	0.8	0.6	0.3	0.8
Std. RDG				
Mean	0.6	1.2	0.8	1.5
SD	0.7	0.7	0.7	0.7
%Missing	0.4	0.9	0.1	0.0
Score				
Mean	1.0	1.8	0.9	1.8
SD	0.6	0.6	0.6	0.6

^ above minimum score among students at Selective Schools. Pooled without School D

Table 11: Achievement Outcomes of Selective and Non-Selective Students

	ATAR	% Missing	ATAR ≥ 95	ATAR ≥ 90	ATAR ≥ 85
Selective Boys	91.8	2.1	49.1	68.2	81.5
Non-Selective	83.6	9.0	23.9	40.8	52.0
Selective Girls	93.8	10.7	54.0	71.8	80.6
Non-Selective	84.2	8.3	22.3	37.9	52.9

A5. Fuzzy Regression Discontinuity

Table 12: Fuzzy RD Estimates of Selective School Effect by Bandwidth^

	Bandwidth								
	2	3	4	5	6	7	8	9	10
ATAR	2.5	0.8	-2.6	-1.1	-1.2	-1.4	-1.1	-0.9	-0.4
	(8.3)	(5.8)	(4.3)	(3.8)	(3.3)	(3.4)	(3.2)	(2.8)	(2.6)
	136	213	306	384	471	549	618	689	750
ATAR ≥ 95				-0.24	-0.13	-0.22	-0.37	-0.34	-0.35
				(0.57)	(0.45)	(0.44)	(0.41)	(0.37)	(0.35)
				423	513	595	669	748	814
ATAR ≥ 90				-0.46	-0.36	-0.26	-0.21	-0.30	-0.30
				(0.57)	(0.46)	(0.45)	(0.42)	(0.38)	(0.35)
				423	513	595	669	748	814
ATAR ≥ 85				-0.50	-0.35	-0.16	-0.12	-0.21	-0.16
				(0.59)	(0.48)	(0.47)	(0.44)	(0.39)	(0.37)
				423	513	595	669	748	814
English	1.1	-3.9	-2.3	-1.4	-0.6	-1.1	-1.8	-0.7	-0.6
	(7.3)	(4.1)	(3.3)	(3.0)	(2.4)	(2.4)	(2.1)	(1.9)	(1.8)
	93	151	217	276	336	391	441	489	535
Maths	3.8	0.5	0.4	2.7	1.3	1.5	1.3	0.1	-0.2
	(6.2)	(4.1)	(3.3)	(3.0)	(2.5)	(2.5)	(2.3)	(2.0)	(1.9)
	126	197	281	350	428	499	557	621	674
Physics	-4.0	-2.2	-1.2	-1.0	-1.6	-3.0	-1.9	-1.1	-0.5
	(7.1)	(5.6)	(4.4)	(4.3)	(3.8)	(4.2)	(3.9)	(3.2)	(3.0)
	58	80	108	133	156	192	216	237	257

^ estimates, standard errors, and number of observations for each achievement outcome.

A6. Value-Added Regressions

This appendix investigates with value-added regressions the idea that selective students have high levels of educational aspiration from their outperformance on the entrance exam compared with the non-selective students. We estimate regressions of ATAR separately for selective and non-selective students, while varying the specification to include the different measures of prior achievement: using Year 7 NAPLAN first; the entrance second; and lastly, together. For the other explanatory variables we include the basic student characteristics of sex, language and socio-economic background.

Table 13 presents results from regressions of ATAR for non-selective students in columns 1 to 3 and selective students in columns 4 to 6. For the non-selective students, we reduced the sample to those who sat the entrance exam and received scores above the minimum at selective schools; the goodness of fit is higher using the results from the standardized tests than for the entrance exam score; R-square of 0.180 for Year 7 NAPLAN compared with 0.158. There is a slight increase to 0.206 with the inclusion of both achievement measures.

Table 13: ATAR Regressions by Selective Attendance while varying Prior Achievement

	Non-Selective			Selective		
	(1)	(2)	(3)	(4)	(5)	(6)
Entrance Exam	0.41***		0.24***	0.53***		0.46***
	(0.04)		(0.04)	(0.03)		(0.05)
YR7 NAPLAN		7.45***	5.28***		6.03***	1.55*
		(0.58)	(0.69)		(0.53)	(0.68)
Male	-2.29***	-2.41***	-2.60***	-2.99***	-2.12**	-2.86***
	(0.69)	(0.68)	(0.67)	(0.70)	(0.75)	(0.71)
Std. SES	1.45***	1.50***	1.31***	0.53	1.21***	0.51
	(0.34)	(0.34)	(0.33)	(0.35)	(0.37)	(0.35)
LBOTE	3.00***	4.13***	3.65***	1.69	2.40*	1.77*
	(0.82)	(0.82)	(0.81)	(0.87)	(0.93)	(0.88)
Constant	58.62***	76.18***	63.64***	48.07***	79.46***	51.19***
	(2.38)	(1.07)	(2.42)	(2.84)	(1.31)	(3.18)
N	992	985	985	721	713	713
R-squared	0.158	0.180	0.206	0.275	0.184	0.279

Entrance exam scores above min at selective schools, includes excl. selective school.

For the selective students the relationship between prior achievement variable and goodness of fit is reversed for selective students (columns 4 and 5): the entrance exam variable pushes the goodness of fit up to 0.275, which is much higher than for the non-selective students, and when Year 7 NAPLAN is included in the regression instead, the goodness of fit is similar to that for the non-selective sample. When both measures of prior achievement are included in the regression for selective students, the goodness of fit is near unimproved, at 0.279, from just the inclusion of the entrance exam scores by itself.

For both non-selective and selective groups, the same general relationship in goodness-of-fit and prior achievement holds when the selective school in its first year of operation is excluded from the sample[40]; using the entrance exam produced a lower r-square than the when NAPLAN was used for non-selective students and the opposite was true for selective students but the increase in predictive power from the entrance exam was smaller, at around 0.03 instead of 0.09. When the non-selective and selective students were pooled, the goodness-of-fit from ATAR regressions was similar to that of for the selective students, being higher for the use of entrance exam than for NAPLAN[41].

The value-added analyses, which consisted of regressions of ATAR while varying the measures of prior achievement, showed that the entrance exam had greater predictive power for selective students than for non-selective students, which is consistent with an interpretation that selective students view the entrance exam as a high stakes outcome. We found that the entrance exam variable has greater predictive power for selective students in terms of their subsequent performance in Year 12, as indicated by higher values in R-square (larger by 0.09, 0.28 compared with 0.18), whereas the inclusion of the entrance exam makes only a small contribution to the goodness of fit for non-selective students (difference of 0.02).

[40] There was both weaker demand from applicants, and lower academic competitiveness for attendance, at the new selective school (see section 5.1. The Data Sample, Selective Schools).

[41] We also run regressions for a combined sample of all applicants restricted to students with Year 7 NAPLAN greater than the minimum at selective schools, without the excluded school (not shown); the R-square value is higher with the inclusion of the entrance exam variable, at 0.304, compared with 0.279 with Year 7 NAPLAN, and is highest at 0.329 when both are included. The negative coefficient of -2.4 for selective attendance when we use the entrance exam to reflect prior achievement is consistent with what we found from the exploratory analyses of achievement by selective attendance; selective students outperform non-selective students in the entrance exam in general, and there is a subsequent regression-to the mean pattern for Year 9 NAPLAN and for ATAR.