Late Start: The Effect of Increase in School Entry-Age on Educational Achievement

Nikhil Jha*

Melbourne Institute of Applied Economic and Social Research
University of Melbourne, Australia

Abstract

This paper analyses the effect of school entry-age on educational achievement. Achievement is assessed using the score across a range of subjects on national standardized tests. The analysis uses change in state policy across two states in Australia to estimate the effect. Policies were enacted in different years and affect different cohorts in the data. Using a difference-in-difference specification, I find positive effects consistent with within-grade human capital accumulation.

JEL Classification: *I28, J24*
Keywords: *Educational Achievement, Human Capital*

The effect, if any, of optimal school starting age has important practical implications on educational policies. If the effects sustain over time, it has important consequences for policies that impact or seek to affect human capital as well. The most common institutional set-up where schools accept students once a year generates variation in the actual school starting age of students. Nonetheless, if students generally perform better when they are older at school-entry, a simple policy change that increases the minimum school starting age would raise academic achievement. Yet, the benefits must be weighed against any costs, such as potentially fewer years of education if compulsory school-leaving-age remains unchanged.[1]

Within the same cohort, older students seem to do better academically (Bedard and Dhuey, 2006). This "entry-age achievement gap" may emanate from distinct sources or under different scenarios. Three interpretations provide a more nuanced view of the maturity effects (see Cascio, 2008). First, the gap arises due to older kindergarten being bigger and smarter "relative" to their classmates. Thus, they may be sorted into, for instance, more advanced reading group. This is the notion of relative maturity. As the gain for the older kids come at the expense of younger students, it implies that, on average, there is no gain. The second interpretation, "age at entry" argues that the advantage older students have is because they are better equipped to succeed at school (due to proportionally greater experience/exposure) regardless of being sorted into advance groups. Finally, the third interpretation, "age at test", argues that entry-age gap is essentially an artifact of positive correlation between age and skill accumulation outside of school (Elder and Lubotsky, 2009).

*Level 5, Faculty of Business and Economics Building, 111 Barry Street, The University of Melbourne, Australia. Phone: +61 3 9035 6389. Fax: +61 3 8344 2111. Email: nikhil.jha@unimelb.edu.au

[1] Increase in the school entry-age resulted in fewer years of education for the delayed entrant in the US (Deming and Dynarski, 2008).

While the "relative age" interpretation implies no average gain for the cohort, the "age at entry" interpretation implies that increase in minimum age is likely to increase academic outcomes of a cohort on average. The"age at test" interpretation suggests that the advantage of the additional stock of knowledge of the older kids diminishes over time and therefore, increasing the school entry age is unlikely to have any long-term impacts.

Two problems complicate the analysis of the maturity effects. The first is the practice of "red-shirting": parents delay their children's enrollment in the first year of schooling if they are too young within their cohort. Second, as age at test has a one-to-one correspondence with age at school entry and years of schooling, it is difficult to disentangle the combined effect of school entry age and age at test. Following Bedard and Dhuey (2006), Kawaguchi (2011) uses the Trends in International Mathematics and Science Study (TIMSS) data but with specific focus on Japan where length of compulsory education does not vary by birth month, and children follow fixed grade progression. He extends the analysis to look at long-term outcomes by using wage data from Employment Status Survey (ESS). He finds evidence for the "relative age" effect wherein older children in a school cohort obtain higher test score and more education which in turn, leads to higher wages. Black et al. (2011), exploit data from Norway, where time in school is fixed by age at school entry and age at test vary. They find that the entry-age gap is due to the "age at test" and not due to the "age at entry." Furthermore, for earnings as well, they find modest effects. Bedard and Dhuey (2012), however, find positive wage effects for U.S. although they do not find any effect on educational attainment. Fredriksson and Öckert (2013), using Swedish data, find that there is a positive effect of increase in school starting age on education but none for the discounted life-time earnings. In short, the evidence is mixed.

In this paper, I exploit policy changes that exogenously affected the age at entry in two different states across different periods in Australia to examine the maturity effect on educational achievement while still in school. This is naturally constrained to identifying the combined effect of age at test and age of school entry.[2] However, the value-added here is that although the spread of the distribution of relative age for a cohort did not change, the location of the age distribution at test did change. As such it provides estimates of the net effect of the policy change. Furthermore, this increase in age is unlikely to be endogenously determined by parents in the present context but is policy driven.[3] Using the average achievement of a particular grade within a school, I find that academic achievement was boosted by the policy change. This suggests the "age at entry" effect interpretation. The effects are examined for two grades—grade 3 and grade 7. Increasing the minimum age at entry affects academic outcomes positively across a range of subjects for both grades. As positive effects are observed even at grade 7, it raises some concerns about the "age at test" interpretation that implies the effects should

[2] Indeed, given the one-to-one relationship between entry age and age at test, this will always be the case for any policy—past, present or future—that changes school entry age.

[3] The issue of "red-shirting" for the states examined is discussed below (section 1.2). Note, however, that not all students are not affected uniformly.

dissipate over time. Data limitations constraint the analysis at the grade level. Yet, besides ameliorating measurement error concerns, using the grade as the unit of analysis is informative. In particular, the analysis speaks to policy relevant issue about net potential gains. As the benefits are observed for the entire grade, "relative age" effect, as currently understood, is an insufficient interpretation since it posits that academic gains to individual students are at the expense of other students within the grade.[4] Rather, the results seem to be consistent with findings of Bedard and Dhuey (2012). They find that policy change which increases the school entry-age across board leads to higher earnings via increases in within grade human capital accumulation. Indeed, I can not rule out positive spillovers from the combined effects of higher age at entry with relative age effect.[5] Cascio and Schanzenbach (2007) find evidence for relative age gap *adversely* affecting older-for-cohorts. They argue that positive peer effect (such as, better behaved peers) on young-for-cohort drive these results.[6] The policy change in Australia conceivably would not vary the overall relative age but increased the age at a particular grade for the affected cohort. Therefore, if the older students, for instance, receive positive reinforcement for academic prowess, it may foster greater individual effort by other students as well. For early primary grades, a higher school entry-age perhaps also implies a less disruptive class. At the very least, results here indicate that even if relative age difference affects academic achievement of older students positively and those of younger students adversely, the effect of relative age difference may be quite different if age distribution shifts to the right.

1 Institutional Background

This paper uses data from Australia. As Kalil et al. (2012) highlight, Australia is quite similar to the United States[7] with policy in both countries addressing similar socio-economic concerns. Yet, institutional settings are quite distinct. This provides a unique opportunity to examine effects of school age-entry on education.

In regard to the educational context, akin to the US, there is virtually no tracking in schools and primary years of schooling is compulsory across Australia. Grade-retention is not very common[8] and declining; proportion of student who have normal grade progression increased by 10 percentage points between 2003 and 2012 (see p.88 OECD, 2013). Furthermore, streaming of children, albeit not uncom-

[4]It could, of course, be that some students experience gains and other students losses with a net positive gain. However, the magnitude of these gains (or losses) will have to be exceptionally large and unrealistic to be consistent with the evidence presented here. Furthermore, the interpretation of relative age effect is unsatisfactory even if, as commonly observed, a particular grade within a school is composed of several classes. This is because generally tracking does not begin before secondary school in Australia.

[5]Notwithstanding the older entry-age and as noted earlier, the single admission intake of school ensures that some children will be older relative to other students.

[6]This would suggest that, on average, the gains to young-for-grade far outweigh the losses to old-for-grade.

[7]The authors point out, for instance, that both, Australia and US are: rich developed countries with similar per-capita income; both were former colonies of UK and have an Anglo-Saxon culture but also substantially large immigrant population.

[8]Less than 10 percent of students repeat a grade in Australia (OECD, 2013).

mon, usually occurs in secondary school. Although similar to other western countries, both private and public schools co-exist, in Australia, non-government sector is substantially larger. Private schools[9] enroll 34 per cent of students (DEEWR, 2011).[10] Education in Australia is the responsibility of the States and of the Territories.[11] Regulations governing public and private schooling is therefore managed by the states and territories. Policy change in two states—Queensland and Western Australia—is used in the analysis here. Queensland and Western Australia have, respectively, the third and the forth largest student population in Australia. Compared to other states, both Queensland and Western Australia experienced increases in the number of students in 2011-2012 that was among the highest as well as proportionally similar (ABS, 2012).[12] An important similarity is a common structure of schooling where primary school, unlike other Australian states, includes grade 7.[13]

1.1 Policy Changes

1.1.1 School Entry-Age

Although states are responsible for their respective school policies, states nonetheless react to, learn from the experiences of, and co-operate with other states. In fact, the role of the Council of Australian Governments (COAG) is to promote policy reform that requires coordinated action. The school entry age policy changes, however, were not concurrent. It was increased in the state of Western Australia in 2002 and in Queensland in 2008 (ABS, 2012). In this analysis, the cohort in grade 7 in the year 2010 is affected by policy change in Western Australia. For Queensland, it is the cohort observed in grade 3 in 2011 that is affected by increase in school entry-age. Since 2008, children in Queensland need to be six by 30 June of the year they enroll in grade 1.[14] Prior to 2008, the cut-off was December. The change essentially meant that the compulsory school starting age increased by six months. Earlier in 2002, policy change in Western Australia had similarly increased school entry-age by six months such that the student had to be six by June 30 in the year of enrollment (see p.36 DET, 2003).

Official data from the Australian Bureau of Statistics shows that the policy increased average age for affected students (grade 3) in Queensland increased from 7.7 to 8 year. The corresponding increase in age for treatment group in Western Australia (grade 7 in 2010) was from 11.7 to 12 year.[15]

Policy change may be a response to underlying characteristics. Thus, it is important that the policy

[9]These are either Catholic schools or Independent schools. The later may be denominational or non-denominational and generally charge fees that are substantially higher than that of Catholic schools.

[10]The corresponding proportion is 10 per cent for US (Ewert, 2013) and 6.5 per cent for the UK (Ryan and Sibieta, 2011).

[11]Australia has six mainland states and two mainland territories.

[12]An increase of 2.8 percent for Queensland and an increase of 2.1 percent for Western Australia.

[13]In most of the other states, students transition to secondary school in grade 7.

[14]Note that although the first cohort affected in Queensland and Western Australia were the corresponding grades in 2010 and 2009 respectively, the subsequent year is used to isolate the effect of increase in school starting age from other confounding factors, for example, smaller class size. Furthermore, smaller cohort may not necessarily imply a smaller class size for all schools especially for popular schools.

[15]See NSSC Table 45a: Average Age of Students, 2012 at:
http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4221.02012?OpenDocument.

change be plausibly exogenous to educational achievement to assess its impact. Queensland, for instance, generally performed below most of the other states across several grades and subjects on the standardized national tests (ACARA, 2011). If policy change that increased school entry-age was enacted to increase educational achievement, the estimates below will be biased upward. These policy changes, however, were enacted to align the education experience in these states with the rest of Australia (Kronemann, 2001; Dowling and O'Malley, 2009). In discussing the policy change, the Queensland Department of Education website states "This means that Queensland children will be starting school at about the same age as children in other states."[16] Thus, the policy changes were not a response to different academic achievement across the states. Other initiatives to further align educational structure across states is being operationalized. For instance, from 2015, grade 7 will be moved to secondary school in both, Western Australia as well as in Queensland. Pre-primary school policies have been undergoing changes to provide uniform early childhood education across all states.

1.1.2 Pre-Primary School

Pre-primary year education is quite prevalent in Australia[17] and although composite grades[18] are not uncommon, children are generally sorted by age. Insofar as pre-school attendance affects educational outcomes, policy changes affecting pre-school provision, may be a potential source of bias. Particularly for Queensland, this may be a concern. Indeed, the school entry age change in Queensland was preceded by the introduction of preparatory/pre-primary year. The pre-primary year care in Australia is often, correctly, considered the first year of schooling. Consequently, students in Queensland affected by school entry-age changes were older than the previous cohort but also had an additional year of schooling. Dowling and O'Malley (2009), however, state that the policy change was essentially a re-configuration of grades and that it "caused preschools in that state [Queensland] to be re-badged as the first year of school." This is evident, they argue, from the observed Australian Bureau of Statistics data documenting the dramatic fall in number of preschools in 2008. As such then, the estimate of the effect of increase in school entry-age in Queensland is not confounded by the introduction of "additional year" of schooling.

Although Western Australia also saw concurrent change in pre-school policy, it is also, arguably, not plagued by simultaneous policy change affecting educational achievement. The primary reason is because even before the school entry-age changed, universal access to government-funded educational programs was already available in Western Australia before the start of grade 1 schooling. The change was in the amount of pre-primary hours available. Furthermore, the increase in the hours available was marginal. Pre-primary provision availability changed from four days to five days starting in 2002 (Kronemann, 2001). Given the high participation rate, the increase is unlikely to have affected the

[16]See http://education.qld.gov.au/strategic/advice/earlychildhood/faq-prep1.html

[17]A large proportion of eligible student population is enrolled in pre-primary (ABS, 2012). See Dowling and O'Malley (2009) for an overview of pre-school provision in Australia.

[18]In composite grades, students across two grades, for instance, grade 2 and grade 3, may be in a combined class.

decision to participate in child-care. Indeed, it may increase the take-up of hours of child-care. However, it is unlikely that a family that chose one-day-a-week of child-care suddenly switched to five-day-a-week childcare as the family already had an option of using up to four days even before the policy change occurred. In short, the affects of the concomitant increase in child-care hours availability in Western Australia should be minimal.

1.1.3 School-Quality Information

Another recent policy intervention relating to school quality information has implications for this study. In January of 2010, "myschool" website went public. The website had information on the average NAPLAN test scores for every school—public and private—in Australia. It also had information on other characteristics of school such as percent of indigenous students and other contextual information. In fact, this analysis uses the "myschool" data.

It is possible that the dissemination of school quality information affected the academic outcome of schools. Three reasons, however, suggest it is unlikely that the effect will be manifest before 2011 for the grades examined. First, the school quality information was made publicly available only two days before school opening date. This would, conceivably, limit any "voting by feet" type behavior by parents in response to quality information. Second, as primary school ends in grade 7 in these states, grade 8 become a natural point for exercising school choice. Finally, the standardized tests were administered in May—just three months after school quality information was released. Thus, it seems reasonable to assume that the effect of dissemination of public schools would have limited effect, if any at all, on educational outcomes in 2010. This has two implications for this analysis. First, if school-quality dissemination affects grades[19] differentially, this analysis can only inform about the combined effect of the school entry-age change and the dissemination of school quality information in Queensland. This is because, in the data, only the 2011 cohort was affected by policy change. Second, for Western Australia, the effect of change in school entry-age can be isolated by using only 2010 data as post-policy period.[20]

1.2 "Red-shirting"

The other important concern mentioned above relates to the practice of "red-shirting"—parents delay child's enrollment if the child is too young for their cohort. Aliprantis (2012) presents evidence of red-shirting in the US and also shows that the practice is potentially driven by the heterogeneity in the treatment effect of educational attainment. Barua and Lang (2009) show that when legal entry-age is not strictly enforced, i.e. red-shirting occurs, the estimates examining effect of entry-age may not identify

[19]The specifications that employs school-by-grade fixed-effects will likely go someway toward addressing this concern.

[20]Recall that, in Western Australia, grade 7 cohort for both years, 2010 and 2011, were affected by policy change. The main specification below is not sensitive to inclusion to 2011. Results available on request.

relevant policy parameters and/or may not be consistent.[21] Fiorini et al. (2013) show that the practice of red-shirting warrants caution in interpreting effects of school entry-age even when using discontinuity design related to school-entry cut-off.

For the policy change states in Australia, evidence suggests that parents holding back young-for-cohort kids from school enrollment is less than 2 percent (see Table 2. Taylor and Fiorini, 2011).[22] Furthermore, the delayed entry is, in general, to ensure that children are older than they would have been if enrolled when eligible. Therefore, Taylor and Fiorini (2011) conclude that parents are more likely to respond if change in entry age potentially reduces their child's age. In fact, for some children, the new cut-off will make them younger for the cohort then otherwise. The data does not allow me to assess this. However, this is a concern only insofar as these children are held back *and* if they have significantly different outcome than the children held back before the policy change. Assuming that a particular child being younger for cohort under new entry-age cut off is random and the continued trend of near full compliance to mandated entry-age cut-off, the estimates presented may be treated as population average treatment effect.

2 Data

2.1 Test Scores

The analysis uses grade-level school data for two states in Australia—Queensland and Western Australia. The data spans from 2008 through 2011 where one (or more) year observation is prior to or is after the policy change. The outcome of interest is the mean grade test score for each individual school. The scores are results from National Assessment Program - Literacy and Numeracy (NAPLAN) tests. Beginning in 2008, these standardized tests are administered to all[23] students in grades 3, 5, 7 and 9 in each of the following 5 domains: reading, writing, spelling, grammar and punctuation, and numeracy. The analysis here uses all subjects except writing for Queensland.[24] All five subjects are analyzed for Western Australia.

NAPLAN test measures absolute competence, therefore scores are higher in higher grades. Scores range from 0 to 1000[25]. On average, the learning for a single grade year is roughly 50 points on the test for lower grades and about 25 points for higher grades. Scores are vertically scaled and subject-specific to allow for comparison across grades. Within subject scores are comparable across time. Thus, the

[21]The concern is the violation of monotonicity that requires that those affected by the instrument should be affected unidirectionally.

[22]For Queensland and for Western Australia, the percentage of children who enter school on time (i.e. neither early nor delayed entry) is 99.8 and 98.3, respectively.

[23]Some parents may withdraw their children from participation and some students may be exempt or absent.

[24]Since 2011, the writing test also assesses persuasive writing in addition to narrative assessment. Therefore, writing results for 2011 are not comparable with prior year writing scores (see `http://www.nap.edu.au/information/faqs/naplan--writing-test.html`).

[25]Analysing NAPLAN Data: `www.vcaa.vic.edu.au/Documents/naplan/analysingnaplandata.pdf` (retrieved on December 13, 2013).

educational achievement represented by a particular score does not change over time. These features make the NAPLAN test scores an appropriate measure of academic achievement for the empirical strategy employed (see Section 3).

Figures 1 and 2 show the distribution of NAPLAN scores across Queensland and Western Australia for pre- and post-treatment period. It suggests that the distribution of scores for the treated grade has shifted to the right for both subjects and for both states. However, the distribution of the control group for numeracy in Queensland and for reading in Western Australia also suggest overall improvement in scores. Table 1 and table 2 shows the average subject specific NAPLAN scores for the two policy change states in greater detail. Two patterns emerge. First, generally large positive growth (except in Numeracy) is observed for the treated group across pre- and post-treatment period. Moreover, the positive growth is observed across both states for schools across all sectors. The negative growth in writing for the Catholic sector in Western Australia is an exception. Second, the growth is generally the smallest for numeracy. In contrast to the control group, the change in mean score for the treatment group is substantially larger and positive for all subjects.

For state of Queensland (Table 1), grade 3 cohort is the treatment grade which is older than prior cohorts in the year 2011.[26] The corresponding control group is grade 5 cohort. The growth for the treatment group ranges from 6.6 points (numeracy) to 26.7 points (Grammar). The highest growth for the control group across any subject within any school sector is 15.9 points (numeracy). In fact, numeracy showed positive gains for the control group across all sectors. The contrast between the treated and non-treated is far greater for subjects other than numeracy.

In Western Australia (Table 2), grade 7 students of cohort 2010 were, on average, older than the previous cohort students.[27] The control group for Western Australia are students from grade 3. Similar to Queensland and with the exception in writing for Catholic schools noted above, the treatment group shows large positive gains across all subjects. In Western Australia, the growth ranges from 5.2 points (numeracy) to 23.5 points (grammar). And for the treatment group in Western Australia also, numeracy generally experienced the lowest growth among all subjects. Within the control group, the highest growth is in grammar (15.8 points).

Tables 1 and 2 about here

2.2 School Characteristics

Table 3 shows that for the state of Queensland, the data is comprised of 1133 primary schools and 224 combined schools—schools that serve primary as well as secondary grades. According to official statistics, the corresponding number of unique primary schools and combined schools in Queensland are 1152 and

[26] The pre-treatment period is the cohort from the year 2008. As noted earlier, although 2010 is the first treatment cohort, it is excluded as it was a smaller cohort.
[27] The pre-treatment period is the the cohort from the year 2008.

251 respectively (ABS, 2012). Similarly, in Western Australia, there were 663 primary schools and 221 combined schools (ABS, 2012) while the data here has 529 unique primary schools and 181 unique combined schools observed across pre- and post-treatment periods. The requirement that school report scores for the pre- and post-treatment period results in the smaller sample. Naturally, this means that the analysis excludes new schools or schools that closed in either period. Over 75 percent of independent schools in Queensland and Western Australia are combined schools. Majority of Catholic schools (about 92 percent) and public schools (about 90 percent) are primary schools in Queensland. A similar pattern is also observed for Western Australia. For both states, there are large variation in the location of school depending on whether the schools are public or private. For instance, in Queensland, approximately 85 percent of government schools are located in the metropolitan area or suburbs (provincial). The corresponding number of Catholic schools and Independent schools are 90 and 99 percent respectively. In Western Australia, 75 percent of public schools and 90 percent of independent schools are located in metropolitan or suburban areas. The location of Catholic schools observed in the sample is anomalous with substantially greater proportion in suburbs than in metropolitan areas. Main results are, therefore, presented using public schools only that in any case constitutes two-thirds of the sample. Finally, for either states, public schools and Catholic schools, as may be expected, are more likely to be located in remote or very remote location. However, in Western Australia, over 9 percent of Independent schools are located in remote or very remote areas compared to less than 1 percent in Queensland.

Table 3 about here

Table 4 presents the average school characteristics in the pre- and post-treatment period for private and public school in Queensland and Western Australia. The ICSEA[28] is an index of socio-economic measure of student population of the schools where a lower value is indicative of a lower socio-economic status. The independent school sector has, on average, the highest ICSEA value while public school sector has the lowest.[29] On average, the ICSEA has been increasing for private schools while declining (or increasing very marginally) for public school. For Queensland, the difference between the ICSEA before- and after-treatment period is statistically significant for all private school sectors.[30] An important factor in assessing socio-economic index is the percent of indigenous students in the school. On average, approximately 10-14 percent of student population in public schools, are indigenous in either states in the sample used. Across Australia, public schools accounted for 85 percent of all Indigenous student enrollment in 2010 (DEEWR, 2011). Although the number of indigenous students in non-government schools

[28]Index of Community Socio-Educational Advantage (ICSEA) is a means of making a comparison of the levels of educational advantage or disadvantage that students bring to their academic studies. ICSEA values range from around 500 (extremely educationally disadvantaged backgrounds) to about 1300 (very educationally advantaged backgrounds).

[29]In 2010, 36 percent of public school students were from the lowest quarter of socio-educational advantage. The corresponding number for Catholic schools and Independent schools were 21 and 13 percent, respectively (DEEWR, 2011). The trend is reversed for the proportion of students from the highest quarter of socio-educational advantage.

[30]Watson and Ryan (2010) found that almost 60 per cent of the decline in government school enrollments between 1975 and 2006 occurred in the top half of the socioeconomic status distribution.

has been increasing in recent years (ABS, 2012), the percent of indigenous students observed for private sector in Western Australia in the sample is substantially high (24 percent and 14 percent for Catholic and independent sector, respectively) possibly due to higher proportion observed in primary schools. As anticipated, proportion of indigenous students has been increasing in public schools. Notwithstanding these general trends, there is variation in the student composition with each sector. For instance, even some Independent schools may draw their total enrollment from students of low socio-economic status (p.11 DEEWR, 2011).

In terms of the number of students enrolled, Independent schools, on average, are larger than public schools or Catholic schools. This is due to, in part, majority of them being combined schools rather than primary schools (DEEWR, 2011). In the sample as well, the total enrollment has been increasing. The increase is substantially greater in independent schools in the post-treatment period across both states. The increase was approximately 36 students in Queensland, and about 34 students in Western Australia. In contrast, enrollment in public schools increased by about 13 students in Western Australia and by 7 students in Queensland. As may be anticipated, the proportion of female students is quite similar across the two states and the three sectors. Finally, table 4 also documents the student-teacher ratio across the two state in the relevant periods. In general, changes in student=teacher ratio have been minor across states and sectors.

Table 4 about here

3 Estimation Strategy

I use the policy induced variation in the age at school entry to identify the effect of increase in school entry age on the NAPLAN test scores. The policy change lends itself to the difference-in-difference framework. A standard specification, for the present context, is:

$$A_{gst} = \alpha + \beta X_{st} + \gamma Treatment_{gs} + \tau Post_t + \theta Treatment_{gs} * Post_t + \epsilon_{gst} \tag{1}$$

A_{gst} is the average academic achievement of grade g students in school s in year t. Academic achievement is measured using score on standardized test (discussed in Section 2) across a range of subjects including numeracy and reading. *Treatment* is an indicator variable for student in the treated grade (3 and 7 for Queensland and Western Australia, respectively). *Post* is an indicator variable for the period after treatment. The cohort in the same grade as the the one affected by the policy change but for period other than *Post* is the pre-treatment measure of academic achievement (A_{gst}). X are school-level, time-varying as well as time-invariant, controls that account for observable differences. These are: ICSEA (socio-economic index), percent of indigenous students, percent of female students, total enrollment, indicators for school sector (for e.g. private or public) and for school type (for e.g. combined or primary).[31] The coefficient of interest, θ, identifies the effect of increase in school entry-age.[32]

The difference-in-difference estimator assumes that the potential outcome is a linear function of the time dummy (indicator for post) and the group dummy (treatment). For instance, in the case of Queensland policy change, the difference-in-difference estimator eliminates the systematic effect across the treatment group across time as well as eliminates the general time effects common to both the treatment group and the control group. The identification is based on the well known, albeit inherently untestable, "common trend" assumption: the change in average outcome across the treatment and the control group would be same in the absence of treatment.

Therefore, selection of the appropriate control group is important. One potential, and commonly used, control group is the corresponding treatment group in another state. Notwithstanding the similarity for students in same grade, states may operate in varying educational contexts and framework. Thus, a more appropriate control group is an alternate grade in the same state across the pre- and post-treatment period. In particular, the alternate grade is preferably common to school with the treatment grade. It seems reasonable to assume that each state is likely to follow similar strategies to improve achievement especially for grades within school types such as primary or secondary. Finally, the composition is less likely to be changing across cohorts within a state than across states.

An additional requirement for unbiased estimates is that the composition of the treatment group is stable. This is especially a concern when using cross-sectional data. Here, the pre- and post-treated group belong to different cohorts. However, it is reasonable to assume that, on average, the performance of cohorts would be similar. That is, academic performance is unlikely to be systematically different across these cohorts for both treatment as well as the control grades. Note that the general year-on-year increasing trend in test scores would be captured by the variable *Post*. Systematic manipulation of grade assignment is also unlikely to be a concern as the unit of analysis is grade. A closely related concern is potentially endogenous treatment. However, as indicated, the context for policy change was alignment of policies across states rather than response to any performance concerns.

3.1 Threats to Identification

Although the increase in school entry-age is driven by an arguably exogenous policy change, there are two concerns. First, it is possible that schools even within a state and with a common policy environment, pursue different strategies or draw enrollments from disparate population that differentially affects educational achievement. For instance, strategies for primary schools may differ from combined

[31] Estimates are qualitatively similar when X is not included. Results available on request.

[32] In particular, under full compliance to the treatment, this is the average treatment effect on the treated (ATT). In the present context, the assumption of full compliance requires that all units in the cohorts affected by treatment should be older for the grade. In the absence of perfect compliance, the difference-in-difference identifies the Intention-to-Treatment (ITT) effect. ITT estimates the average effect of the treatment on the outcome of all eligible units, regardless of their participation.

schools. Indeed, since 2001, growth of combined schools have exhibited a strong upward trend unlike primary or secondary schools (ABS, 2012). Furthermore, the relatively high proportion of private schools in Australia—approximately 34 per cent[33] of students in Australia are enrolled in the private sector—suggests other differences across the education sectors. In fact, even among public schools there is heterogeneity. In Western Australia, under the Independent Public School initiative, increasing number of public schools are being given additional management autonomy.[34]

In order to mitigate these concerns, school fixed effects is included. Consequently, the control group is a different grade *within* the same school. It is reasonable to assume that schools will have similar strategy for grades that are close i.e. treatment grades and control grades differ by only 2 years. Thus, it is conceivable that unobserved difference between the treatment and control group would be same over time. Similarly, the other important source of potential bias—sorting of students across public-private schools due to time-invariant school characteristics would also be ameliorated by the inclusion of school fixed-effects. The updated equation, then, is:

$$A_{gst} = \alpha + \beta X_{st} + \gamma Treatment_{gs} + \tau Post_t + \theta Treatment_{gs} * Post_t + \kappa_s + \epsilon_{gst} \quad (2)$$

where κ_s is the school fixed effect and other variable are as defined above.

The implication of including school fixed effects is that the estimates are identified using within school variation across the two grades under consideration. In particular, the identifying assumption is that variation in the score for the treatment grade in a specific subject over and above the variation in average score for treatment and control grade within a particular school is orthogonal to unobservables across the treatment and control group.

A second potential threat to identification relates to time-varying unobservables. For the estimates above to be biased, the time-varying factors would need to differentially affect different school types and schools in different sectors so as to be systematically related to the specific treatment and control grades. For instance, quality of students that are enrolled by private schools or combined schools may vary over time. If so, it may be that the unobservables differ by grade over time. For instance, the 2011 cohort affected by the entry-age increase in Queensland would be in the first grade at the inception of the introduction of standardized testing. Therefore, schools may have a different strategies preparing them for their tests two years hence. Similarly, the 2011 treated cohort of Western Australia coincides with the last grade for primary schools and so may be differentially affected by public availability of school-quality information. Another possibility is that the individual schools, when making teaching and allocation decision for specific grades, weigh the performance of competing schools in those grades.

[33] See DEEWR (2011).

[34] This initiative provides the principals with greater autonomy to develop staffing profiles, and select and appoint staff. As of 2013, there are 255 Independent Public Schools operating; 34 of these commenced in 2010 (see http://www.det.wa.edu.au/independentpublicschools/detcms/portal/).

Consequently, the treatment cohort of 2011 for Western Australia is excluded[35] to deal with the potential effect of the nation-wide school quality information dissemination in 2010. Nonetheless, some schools, especially private schools, may be more likely to experience and respond to parental pressures to improve outcomes for specific grade. Furthermore, the data does not permit such exclusion for Queensland as the 2011 cohort is the only treated group. Thus, school-grade fixed effect is incorporated in the analysis. It will capture any systematic difference across grades in a school that is common across all time periods. Thus, it accounts for school specific teaching practice or curriculum across grades. Naturally, it also absorbs the indicator for treatment which is grade specific. Note, however, that to deal with bias due to time-varying unobservables, the preferred specification would need to account for time-varying factors across grades within school. This is not feasible since the grade-cohort variation is used for identification. Consequently, the following specification with school specific grade effects is estimated :

$$A_{gst} = \alpha + \beta X_{st} + \tau Post_t + \theta Treatment_{gs} * Post_t + \mu_{gs} + \epsilon_{gst} \quad (3)$$

In the equation above, μ_{gs} is school-by-grade fixed effect. The effect of policy change is identified by the variation in treated grade over and above the average score for each specific grade within the school.[36] Insofar as schools are comprised of grades and to the extent that time-varying factors are differentially affecting specific grades, these factors would also be affecting schools and therefore this estimation equation is similar to the school fixed-effect specification.

4 Results

4.1 Main Results

Table 5 presents result for the coefficient of interest for the basic difference-in-difference specifications for Queensland. Full results are relegated to the Appendix (section A). The large negative coefficient on the variable *Treatment* is expected to reflect the approximately 90 points lower score that is expected in grade 3 (treated grade) relative to the score in grade 5 (control group) due to the vertical scaling of test scores. Average time-trend across both grades is negative as shown by coefficient for the variable *Post*. The variable *Treatment*Post* captures the effect of increase in school entry-age. It shows a positive effect of increase in school entry-age (Panel A of Table 5). The effect ranges from 3.3 points in numeracy to 25.5 points in grammar. Reading and spelling also show statistically significant effect of 24 points and 18 points, respectively. The subsequent estimate incorporates school fixed-effects (Panel B). This controls for all, observed and unobserved, time-invariant school characteristics. Finally, Panel C present

[35]Results are qualitatively similar with the inclusion of the year 2011. Available on request.

[36]An alternate approach would use school-by-year fixed effects. This would isolate variation across year for all common grades. Results are similar to grade-by-year fixed effects and is available on request.

results that include school specific grade fixed effects to account for potentially grade-specific bias. All specifications produce virtually same estimates for the coefficient of interest.

The pattern for Western Australia is quite similar although in absolute terms, the magnitude of the effect is smaller. For the basic specification, gain of over 12 points is observed in numeracy and of 5 points in reading. The estimated effect on spelling is 10 points while that on writing is 8 points. Only the estimate for grammar is not statistically significant (Panel A of Table 6). Again, augmenting the basic specification with school fixed-effects and, subsequently, with school-by-grade fixed effects does not change the results (Panel B and C of Table 6).

Tables 5 and 6 about here

4.2 Robustness

A natural falsification test would attribute the treatment to a different grade and use the same control group. Therefore, the treatment for Queensland is attributed to grade 7 while the grade 5 remains the control group. As grade 7 cohort was not older, no positive effect is anticipated. Panel A of table 7 shows the result using the basic specification reported in Panel A of table 5 using the fake treatment grade. The coefficient on the difference-in-difference term for all but grammar is statistically significant. However, it is positive only for reading with an estimate of 2.3 points compared to the estimate of 24 points for the true treatment grade (compare Panel A, column 2 of table 5). Panel B of the same table (7) reports results from an alternate robustness check. Here, the treatment group is the true treated group (grade 3 of cohort 2011) but the control group is grade 7 rather than grade 5. Results show large and statistically significant positive gains across board. This suggests that the increase in school entry-age indeed lead to higher achievement and is not an artifact of a specific control group. Table 8 report results of analogous exercise for Western Australia. In Panel A, the treatment is attributed to grade 9 while the grade 3 continues to be the control group. Again, results of the placebo test are reassuring. Although the difference-in-difference estimate is significant for most of the subjects, the estimate is negative. The coefficient of interest is positive for numeracy but is not statistically significant. Panel B of table 8 presents results using grade 9 as the alternate control group for the true treated group (grade 7) using the main specification. All coefficients are positive, although the coefficient of interest is not statistically significant for numeracy and spelling.

Tables 7 and 8 about here

5 Conclusion

This paper examines the effect of policy mandated increase in school entry-age on standardized test scores in Australia. On average, the policy change manifest large gains across two different states. For policy, understanding the economic significance of these effects is important. Two reasons suggest that policy decisions affecting school entry-age matter. First, recall that on the NAPLAN tests, approximately 25 points on the test scores is equivalent to one year of learning. For lower grades such as grade 3, one year of learning equates with roughly 50 points on NAPLAN test score. Therefore, as the estimates of positive effect for Queensland are around 20 points (except for 3 point effect on numeracy), these are large effects. The gains are smaller in magnitude at grade 7 for Western Australia. They range from 5.8 points for reading to 12.5 points for numeracy. No effect—statistical or economic—is detected for grammar. The economic significance ranges from one-quarter to one-half of a year worth of learning. This is roughly similar to that observed for Queensland. Second, recall that the unit of analysis here is a grade. Thus, when minimum school entry-age is increased for all students, benefits—even if restricted to a subgroup within the grade—are sufficiently large to manifest an overall positive effects. In fact, it is most likely the case that the majority of the students within the grade experience positive effects. For students who are older due to policy change, there are obvious gains. Students, who are younger among the cohort due to policy gain are likely to benefit from older peers. However, this can not be ascertained with the current data. It is important to understand if these gains are evenly distributed across students and across gender[37] and if not, the underlying mechanisms that generate differential returns. Although, the effect observed is consistent with the "age at grade" effect, positive peer effects may further reinforce results when the all students are, on average, older. Results here find support, in a different context, for the within-grade human capital formation reported by Bedard and Dhuey (2012). Does the higher human capital formation lead to greater educational attainment? Or to higher returns in the labor market? And how large is the impact, if any at all? The current study can not explore these questions but it underscores the importance of continued research on the issue of school entry-age.

References

ABS (2012): "Schools, Australia," Tech. Rep. 4221.0, Australian Bureau of Statistics, Canberra.

ACARA (2011): "National Assessment Program Literacy and Numeracy Achievement in Reading, Persuasive Writing, Language Conventions and Numeracy: National Report for 2011," Australian Curriculum, Assessment and Reporting Authority, Sydney.

Aliprantis, D. (2012): "Redshirting, compulsory schooling laws, and educational attainment," *Journal of Educational and Behavioral Statistics*, 37, 316–338.

Barua, R. and K. Lang (2009): "School Entry, Educational Attainment and Quarter of Birth: A Cautionary Tale of LATE," Working Paper 15236, National Bureau of Economic Research.

Bedard, K. and E. Dhuey (2006): "The persistence of early childhood maturity: International evidence of long-run age effects," *The Quarterly Journal of Economics*, 121, 1437–1472.

——— (2012): "School-Entry Policies and Skill Accumulation Across Directly and Indirectly Affected Individuals," *Journal of Human Resources*, 47, 643–683.

[37] Fiorini et al. (2013) find that to the extent red-shirting occurs, boys are more likely to be held back. And, Bedard and Dhuey (2012), do not find evidence of positive effect on females.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2011): "Too young to leave the nest? The effects of school starting age," *The Review of Economics and Statistics*, 93, 455–467.

Cascio, E. and D. W. Schanzenbach (2007): "First in the class? Age and the education production function," Tech. rep., National Bureau of Economic Research.

Cascio, E. U. (2008): "How and why does age at kindergarten entry matter?" *FRBSF Economic Letter*.

DEEWR (2011): "Review of Funding for Schooling," Report, Department of Education and Workplace Relations (DEEWR).

Deming, D. and S. Dynarski (2008): "The Lengthening of Childhood," *The Journal of Economic Perspectives*, 22, 71–92.

DET (2003): "Annual-Report 2002-2003," Tech. rep., Department of Education and Training, Government of Western Australia.

Dowling, A. and K. O'Malley (2009): "Preschool education in Australia," *Policy Briefs*, 1–16.

Elder, T. E. and D. H. Lubotsky (2009): "Kindergarten entrance age and children's achievement impacts of state policies, family background, and peers," *Journal of Human Resources*, 44, 641–683.

Ewert, S. (2013): "The Decline in Private School Enrollment," SEHSD Working Paper FY12-117, U.S. Census Bureau.

Fiorini, M., K. Stevens, M. Taylor, and B. Edwards (2013): "Monotonically Hopeless? Monotonicity in IV and fuzzy RD designs," Tech. rep.

Fredriksson, P. and B. Öckert (2013): "Life-cycle effects of age at school start," *The Economic Journal*.

Kalil, A., R. Haskins, and J. Chesters (2012): *Investing in Children: Work, Education, and Social Policy in Two Rich Countries*, Brookings Institution Press.

Kawaguchi, D. (2011): "Actual age at school entry, educational outcomes, and earnings," *Journal of the Japanese and International Economies*, 25, 64 – 80.

Kronemann, M. (2001): "The Western Australian model of preschool education," Tech. rep., Australian Education Union.

OECD (2013): "PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices (Volume IV)," OECD Publishing.

Ryan, C. and L. Sibieta (2011): "A comparison of private schooling in the United Kingdom and Australia," *Australian Economic Review*, 44, 295–307.

Taylor, M. and M. Fiorini (2011): "Who gets the 'gift of time' in Australia? Exploring delayed primary school entry," *Australian Review of Public Affairs*, 10, 41–60.

Watson, L. and C. Ryan (2010): "Choosers and Losers: The Impact of Government Subsidies on Australian Secondary Schools." *Australian Journal of Education*, 54, 86–107.

6 Figures and Tables

A Appendix

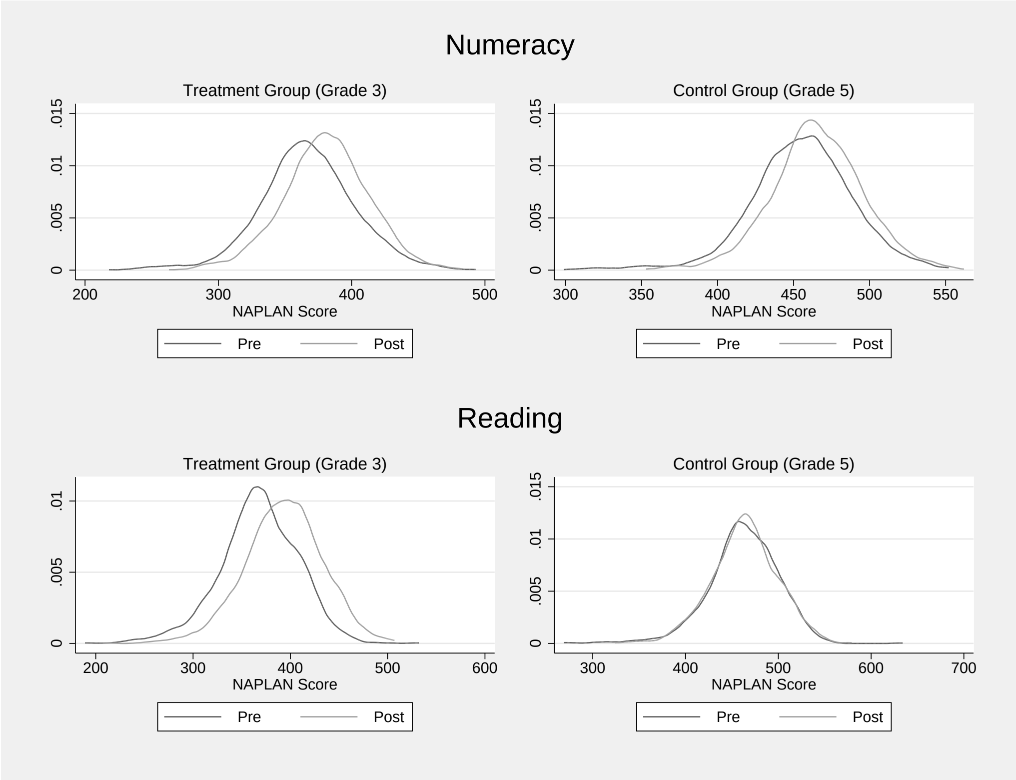

Figure 1: Queensland

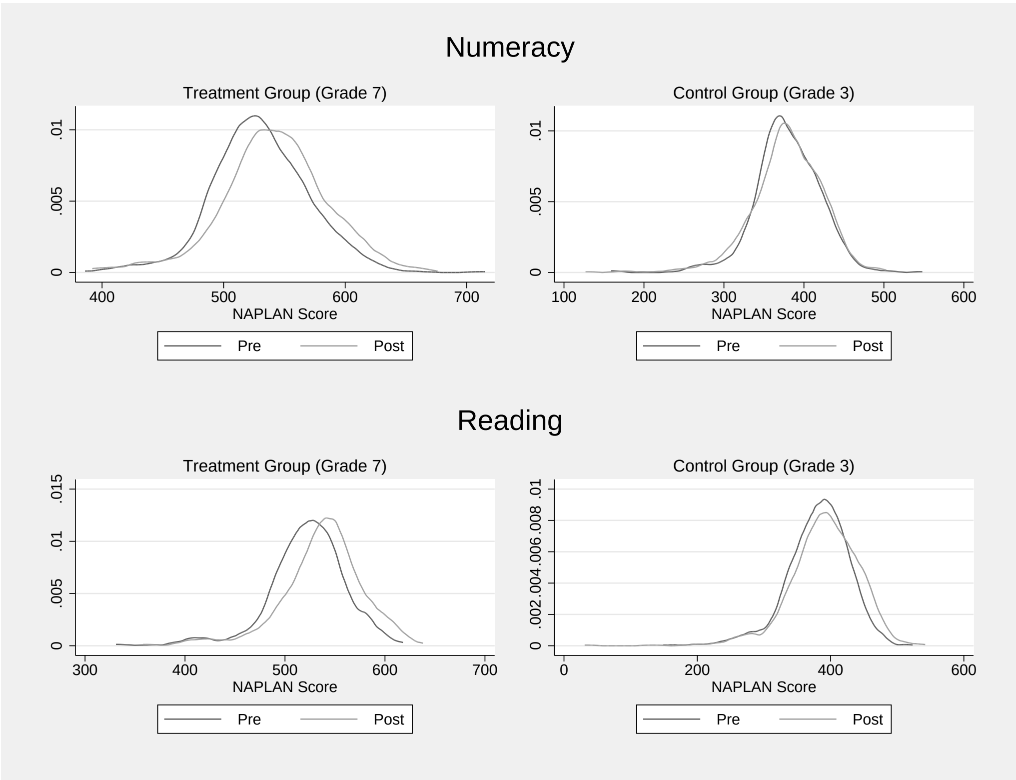

Figure 2: Western Australia

Table 1: Average Test Scores by School Sector for Queensland

	Catholic Schools					
	Treatment Group (Grade 3)			Control Group (Grade 5)		
	Pre	*Post*	***Diff*** *(se)*	*Pre*	*Post*	***Diff*** *(se)*
NUMERACY	376.830	383.470	6.640 (2.670)	465.949	470.359	4.410 (2.692)
READING	384.530	404.550	20.020 (3.210)	478.515	476.171	-2.344 (2.991)
GRAMMAR	383.220	406.610	23.390 (3.830)	489.031	492.010	2.979 (3.429)
SPELLING	376.590	389.040	12.450 (2.780)	469.663	469.217	-0.446 (2.565)
	Government Schools					
	Treatment Group (Grade 3)			Control Group (Grade 5)		
	Pre	*Post*	***Diff*** *(se)*	*Pre*	*Post*	***Diff*** *(se)*
NUMERACY	359.870	375.860	15.980 (1.690)	448.864	460.314	11.450 (1.571)
READING	358.470	384.830	26.350 (1.930)	453.926	455.566	1.640 (1.773)
GRAMMAR	357.570	389.130	31.560 (2.270)	464.403	469.457	5.053 (2.034)
SPELLING	354.810	375.480	20.670 (1.740)	451.907	452.269	0.362 (1.476)
	Independent Schools					
	Treatment Group (Grade 3)			Control Group (Grade 5)		
	Pre	*Post*	***Diff*** *(se)*	*Pre*	*Post*	***Diff*** *(se)*
NUMERACY	393.420	402.780	9.360 (3.930)	477.177	487.816	10.639 (3.847)
READING	396.270	422.990	26.710 (4.900)	490.824	491.036	0.212 (4.184)
GRAMMAR	391.770	427.280	35.510 (5.800)	499.691	505.184	5.493 (4.861)
SPELLING	384.060	407.380	23.310 (4.100)	474.574	479.922	5.348 (3.413)

Table 2: Average Test Scores by School Sector for Western Australia

	Catholic Schools					
	Treatment Group (Grade 7)			Control Group (Grade 3)		
	Pre	*Post*	***Diff*** *(se)*	*Pre*	*Post*	***Diff*** *(se)*
NUMERACY	525.192	530.575	5.383 (8.570)	367.212	366.100	-1.112 (10.090)
READING	518.271	532.604	14.333 (9.837)	375.212	390.080	14.868 (10.861)
GRAMMAR	492.388	515.979	23.591 (15.693)	369.519	380.360	10.841 (13.615)
SPELLING	516.163	526.042	9.878 (10.489)	368.289	368.800	0.512 (10.576)
WRITING	519.894	514.938	-4.956 (14.064)	384.404	385.080	0.676 (12.277)
	Government Schools					
	Treatment Group (Grade 7)			Control Group (Grade 5)		
	Pre	*Post*	***Diff*** *(se)*	*Pre*	*Post*	***Diff*** *(se)*
NUMERACY	527.607	540.025	12.418 (2.643)	376.576	376.633	0.057 (2.589)
READING	518.463	534.683	16.219 (2.483)	376.938	386.497	9.559 (3.071)
GRAMMAR	504.148	518.946	14.799 (3.043)	370.643	386.538	15.895 (3.446)
SPELLING	517.533	531.897	14.364 (2.341)	370.877	373.698	2.821 (2.782)
WRITING	512.291	522.017	9.726 (2.807)	388.905	390.127	1.222 (2.720)
	Independent Schools					
	Treatment Group (Grade 7)			Control Group (Grade 5)		
	Pre	*Post*	***Diff*** *(se)*	*Pre*	*Post*	***Diff*** *(se)*
NUMERACY	551.301	560.639	9.337 (6.459)	404.131	399.689	-4.442 (6.403)
READING	542.952	555.929	12.978 (5.949)	409.333	416.396	7.062 (7.163)
GRAMMAR	527.723	546.631	18.908 (7.027)	403.405	408.901	5.496 (8.036)
SPELLING	536.060	544.083	8.023 (5.985)	394.119	391.341	-2.778 (6.418)
WRITING	523.274	536.659	13.385 (8.554)	407.869	402.868	-5.001 (5.650)

Table 3: School Type and Location for Treatment States

		Queensland		
	Catholic	**Government**	**Independent**	**Total**
School Type: Count†				
Combined	15	91	118	224
Primary	194	904	35	1,133
Total	209	995	153	1,357
Location (% in)				
Metropolitan	54.07	37.29	64.05	42.89
Provincial	36.84	48.54	35.29	45.25
Remote	4.78	8.34	0.00	6.85
Very Remote	4.31	5.83	0.65	5.01
		Western Australia		
	Catholic	**Government**	**Independent**	**Total**
School Type: Count				
Combined	19	82	80	181
Primary	38	463	28	529
Total	57	545	108	710
Location (% in)				
Metropolitan	14.04	52.48	68.52	51.83
Provincial	54.39	24.04	22.22	26.20
Remote	12.28	14.13	1.85	12.11
Very Remote	19.30	9.36	7.41	9.86

Table 4: **School Characteristics by School Type for Queensland and Western Australia**

	Queensland								
	Catholic School			**Public School**			**Independent School**		
	Mean		**Diff**	**Mean**		**Diff**	**Mean**		**Diff**
	Pre	Post		Pre	Post		Pre	Post	
ICSEA	992.470	1038.505	46.036***	965.272	958.669	-6.603	999.829	1053.451	53.622***
Total Enrollment	362.515	395.677	33.162	349.110	356.894	7.783	626.738	662.933	36.195
Percent Indigenous	4.791	5.337	0.546	10.303	11.147	0.844	5.469	5.463	-0.006
Percent Female	0.502	0.503	0.000	0.482	0.480	-0.003	0.507	0.528	0.021
Student-Teacher Ratio	16.764	16.899	0.135	14.422	14.720	0.298	14.357	14.409	0.052
	Western Australia								
	Catholic School			**Public School**			**Independent School**		
	Mean		**Diff**	**Mean**		**Diff**	**Mean**		**Diff**
	Pre	Post		Pre	Post		Pre	Post	
ICSEA	913.477	942.236	28.760	967.002	970.207	3.204	993.594	1017.664	24.069
Total Enrollment	322.782	334.946	12.164	289.540	302.937	13.398	466.514	500.710	34.196
Percent Indigenous	25.867	24.745	-1.122	13.893	14.083	0.190	15.974	14.963	-1.011
Percent Female	0.507	0.506	0.000	0.482	0.484	0.001	0.492	0.512	0.020
Student-Teacher Ratio	14.680	14.331	-0.349	14.368	14.516	0.147	13.248	13.132	-0.116

*** $p<0.01$, ** $p<0.05$, * $p<0.10$

Table 5: Queensland Results

	NUMERACY	READING	GRAMMAR	SPELLING
	Panel A			
Treatment	-88.163***	-94.672***	-106.285***	-95.405***
	(0.911)	(0.941)	(1.068)	(0.806)
Post	-5.590*	-15.227***	-10.616***	-13.005***
	(2.999)	(3.427)	(3.795)	(2.848)
*Treatment*Post*	3.327***	24.113***	25.524***	18.473***
	(1.113)	(1.227)	(1.483)	(1.083)
Observations	4,434	4,434	4,434	4,434
R-squared	0.822	0.810	0.798	0.824
School FE	No	No	No	No
	Panel B: With FE			
Treatment	-88.608***	-95.015***	-106.925***	-95.527***
	(0.867)	(0.879)	(1.013)	(0.771)
Post	10.744***	0.292	4.472***	0.777
	(1.017)	(1.075)	(1.300)	(1.003)
*Treatment*Post*	3.813***	24.588***	26.377***	18.701***
	(1.074)	(1.170)	(1.426)	(1.055)
Observations	4,434	4,434	4,434	4,434
R-squared	0.873	0.844	0.835	0.878
Number of schools	1,216	1,215	1,216	1,216
School FE	Yes	Yes	Yes	Yes
	Panel C: With School by Grade Fixed Effects			
Post	7.506***	-0.077	5.210	7.023
	(2.376)	(2.641)	(3.201)	(9.550)
*Treatment*Post*	3.678***	23.964***	26.156***	18.333***
	(1.061)	(1.158)	(1.436)	(1.059)
Observations	4,434	4,434	4,434	4,434
R-squared	0.201	0.277	0.295	0.220
Number of school-grades	2,358	2,354	2,356	2,356
SchoolGrade FE	Yes	Yes	Yes	Yes

Post = Year 2011 dummy; Treatment is Grade 3; Control is Grade 5

Standard errors (in parentheses) clustered at school level

*** p<0.01, ** p<0.05, * p<0.1

Table 6: Western Australia Results

	NUMERACY	READING	GRAMMAR	SPELLING	WRITING
			Panel A		
Treatment	149.531***	140.350***	130.957***	145.728***	122.516***
	(1.305)	(1.255)	(1.626)	(1.379)	(1.298)
Post	-14.919***	6.810	12.500**	-4.745	-8.096*
	(4.114)	(4.973)	(5.839)	(4.515)	(4.443)
*Treatment*Post*	12.585***	5.822***	0.623	10.766***	8.382***
	(1.726)	(1.851)	(2.169)	(1.803)	(1.697)
Observations	2,368	2,374	2,371	2,371	2,371
R-squared	0.916	0.908	0.868	0.913	0.900
School FE	No	No	No	No	No
		Panel B: With School Fixed-Effects			
Treatment	150.759***	140.704***	132.337***	146.431***	123.691***
	(1.231)	(1.225)	(1.527)	(1.334)	(1.233)
Post	0.863	10.886***	16.730***	4.017***	2.146
	(1.339)	(1.676)	(1.829)	(1.392)	(1.306)
*Treatment*Post*	11.258***	5.083***	-1.070	10.119***	7.397***
	(1.612)	(1.769)	(2.072)	(1.746)	(1.619)
Observations	2,368	2,374	2,371	2,371	2,371
R-squared	0.948	0.931	0.888	0.943	0.923
Number of schools	648	650	649	649	651
School FE	Yes	Yes	Yes	Yes	Yes
		Panel C: With School by Grade Fixed Effects			
Post	-5.838	11.357**	16.733***	2.426	0.826
	(3.772)	(5.029)	(5.477)	(4.274)	(4.333)
*Treatment*Post*	11.965***	6.538***	0.378	11.044***	8.597***
	(1.634)	(1.798)	(2.080)	(1.742)	(1.602)
Observations	2,368	2,374	2,371	2,371	2,371
R-squared	0.097	0.173	0.163	0.116	0.079
Number of school-grades	1,250	1,253	1,251	1,251	1,252
SchoolGrade FE	Yes	Yes	Yes	Yes	Yes

Post = Year 2010 dummy; Treatment is grade 7; Control is grade 3
Standard errors (in parentheses) clustered at school level
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Queensland Placebo Test

	NUMERACY	READING	GRAMMAR	SPELLING
	Panel A: Alternate Treatment (Gr7); Same Control Group (Gr5)			
Treatment	80.318***	63.295***	42.208***	65.252***
	(0.839)	(0.842)	(0.942)	(0.777)
Post	3.190	-9.969***	-4.251	-0.529
	(3.827)	(3.809)	(4.422)	(3.401)
*Treatment*Post*	-12.338***	2.382**	-0.710	-1.885*
	(1.076)	(1.095)	(1.273)	(1.055)
Observations	4,436	4,439	4,439	4,439
R-squared	0.787	0.782	0.679	0.762
School FE	No	No	No	No
	Panel B: Same Treatment (Gr3); Alternate Control Group (Gr7)			
Treatment	-168.523***	-158.011***	-148.663***	-160.757***
	(0.947)	(0.942)	(1.052)	(0.883)
Post	-8.984**	-3.002	-1.790	-0.307
	(4.061)	(3.828)	(4.674)	(3.746)
*Treatment*Post*	15.804***	21.871***	26.536***	20.443***
	(1.163)	(1.216)	(1.399)	(1.142)
Observations	4,388	4,383	4,384	4,384
R-squared	0.924	0.917	0.879	0.921
School FE	No	No	No	No

Post = Year 2011 dummy; True Treatment is grade 3; True Control is grade 5
Standard errors (in parentheses) clustered at school level
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Western Australia Placebo Results

	NUMERACY	READING	GRAMMAR	SPELLING	WRITING
	Panel A: Alternate Treatment Group (Gr-9); Same Control Group (Gr-3)				
Treatment	182.638***	182.163***	173.078***	186.540***	160.292***
	(3.278)	(3.164)	(3.379)	(2.640)	(3.456)
Post	-12.496***	6.930	3.826	-0.434	1.027
	(4.561)	(5.419)	(5.823)	(4.861)	(4.606)
*Treatment*Post*	3.277	-21.301***	-10.009*	-4.946	-7.417*
	(4.235)	(4.547)	(5.138)	(4.168)	(4.379)
Observations	1,608	1,608	1,610	1,610	1,608
R-squared	0.894	0.881	0.856	0.891	0.888
School FE	No	No	No	No	No
	Panel B: Same Treatment Group (Gr-7); Alternate Control Group (Gr -9)				
Treatment	-35.593***	-43.114***	-49.552***	-36.588***	-38.505***
	(2.200)	(2.203)	(3.273)	(2.666)	(3.184)
Post	-11.797***	-16.227***	1.339	-7.591**	-15.553***
	(3.463)	(3.270)	(5.352)	(3.699)	(3.880)
*Treatment*Post*	3.977	21.975***	13.544***	5.597	9.246**
	(3.255)	(3.225)	(4.290)	(3.653)	(3.984)
Observations	1,492	1,502	1,499	1,499	1,499
R-squared	0.697	0.762	0.730	0.696	0.751
School FE	No	No	No	No	No

Post = Year 2010 dummy; True Treatment is grade 7; True Control is grade 3
Standard errors (in parentheses) clustered at school level
*** p<0.01, ** p<0.05, * p<0.1

Table A.1: Queensland

	NUMERACY	READING	GRAMMAR	SPELLING
ICSEA	0.264***	0.311***	0.304***	0.207***
	(0.010)	(0.010)	(0.012)	(0.009)
% Indigenous	0.174***	0.214***	-0.161*	-0.098
	(0.063)	(0.065)	(0.092)	(0.075)
% Female	3.150	35.066***	53.192***	35.816***
	(6.929)	(6.730)	(8.924)	(7.019)
Student-Teacher ratio	0.325	0.465*	0.584	0.315
	(0.326)	(0.281)	(0.388)	(0.238)
Total Enrolment	0.004***	0.008***	0.007***	0.011***
	(0.001)	(0.001)	(0.002)	(0.001)
Primary	-0.679	2.346	2.124	3.769
	(2.686)	(2.745)	(3.344)	(2.457)
Post	-5.590*	-15.227***	-10.616***	-13.005***
	(2.999)	(3.427)	(3.795)	(2.848)
Primary*Post	-4.272	-4.301	-4.345	0.039
	(2.658)	(3.030)	(3.335)	(2.469)
Public	-9.266***	-15.639***	-12.847***	-10.967***
	(1.900)	(1.938)	(2.282)	(1.780)
Independent	11.262***	9.662***	8.120**	5.175*
	(2.956)	(3.177)	(3.584)	(2.812)
Public*Post	20.998***	20.448***	19.669***	13.881***
	(1.900)	(2.098)	(2.296)	(1.778)
Independent*Post	-1.957	-1.965	-0.617	4.549
	(3.178)	(3.708)	(4.060)	(2.978)
Treatment	-88.163***	-94.672***	-106.285***	-95.405***
	(0.911)	(0.941)	(1.068)	(0.806)
*Treatment*Post*	3.327***	24.113***	25.524***	18.473***
	(1.113)	(1.227)	(1.483)	(1.083)
Constant	195.615***	139.421***	147.980***	234.146***
	(9.882)	(10.663)	(12.439)	(10.015)
Observations	4,434	4,434	4,434	4,434
R-squared	0.822	0.810	0.798	0.824
School FE	No	No	No	No

Post = Year 2011 dummy; Treatment = Grade 3; Control is Grade 5

Standard errors (in parentheses) clustered at school level

*** p<0.01, ** p<0.05, * p<0.1

Table A.2: Western Australia

	NUMERACY	READING	GRAMMAR	SPELLING	WRITING
ICSEA	0.304***	0.297***	0.331***	0.233***	0.227***
	(0.013)	(0.012)	(0.013)	(0.011)	(0.010)
% Indigenous	0.122	-0.099	-0.248***	-0.226***	-0.608***
	(0.085)	(0.087)	(0.096)	(0.075)	(0.089)
% Female	-5.419	22.940***	24.609***	27.198***	46.145***
	(6.803)	(6.726)	(8.045)	(7.093)	(7.156)
Student-Teacher ratio	-0.053	0.334	0.358	0.814**	0.237
	(0.354)	(0.331)	(0.347)	(0.323)	(0.267)
Total Enrolment	-0.001	-0.001	0.001	0.005**	0.004*
	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)
Primary	-1.983	-1.181	-0.279	-0.262	1.363
	(3.224)	(3.086)	(3.827)	(3.159)	(3.146)
Primary*Post	3.976	-1.260	-4.015	-0.838	-0.147
	(3.590)	(4.033)	(4.472)	(3.430)	(3.744)
Public	-2.779	-8.072**	-5.030	-5.751	-7.092
	(3.581)	(3.684)	(5.444)	(3.832)	(4.411)
Independent	4.694	5.973	4.620	-2.521	-6.032
	(4.348)	(4.512)	(5.795)	(4.463)	(5.133)
Public*Post	10.920***	3.492	5.210	7.906**	9.367**
	(3.687)	(4.311)	(5.338)	(3.900)	(3.873)
Independent*Post	2.215	-9.524	-11.195*	-4.894	-2.559
	(5.191)	(5.829)	(6.536)	(5.609)	(5.005)
Post	-14.919***	6.810	12.500**	-4.745	-8.096*
	(4.114)	(4.973)	(5.839)	(4.515)	(4.443)
Treatment	149.531***	140.350***	130.957***	145.728***	122.516***
	(1.305)	(1.255)	(1.626)	(1.379)	(1.298)
*Treatment*Post*	12.585***	5.822***	0.623	10.766***	8.382***
	(1.726)	(1.851)	(2.169)	(1.803)	(1.697)
Constant	88.078***	82.919***	40.195***	125.880***	154.306***
	(14.620)	(13.672)	(14.490)	(12.203)	(11.878)
Observations	2,368	2,374	2,371	2,371	2,371
R-squared	0.916	0.908	0.868	0.913	0.900
School FE	No	No	No	No	No

Post = Year 2010 dummy; Treatment = Grade 7; Control is Grade 3

Standard errors (in parentheses) clustered at school level

*** p<0.01, ** p<0.05, * p<0.1