How much of the variation in literacy and numeracy can be explained by school performance?

Andrew Leigh & Hector Thompson[1]

Family background is known to have a substantial impact on students' literacy and numeracy results. This raises questions about whether any of the remaining differences in results are due to school performance — or whether they are merely due to random noise. This article reviews research from the OECD's Programme for International Student Assessment (PISA) study, based on student-level analysis. It then presents new evidence based on publicly reported school-level data from Western Australia. Combining test results with data on schools' socioeconomic characteristics, this study estimates the degree to which some schools outperform those with similar characteristics. On a 'like schools' basis, school differences are shown to be persistent across subjects, grades and years.

1 The authors are from Social Policy Division, the Australian Treasury. This article has benefited from comments and suggestions provided by Robert Breunig, Bill Burmester, Nigel Ray, Catherine Thompson and Tony Zanderigo. The authors are particularly grateful to Damien Moore and Ralf Steinhauser for sharing Western Australian school performance data with us, and to Kris Erwood and Jordan Korda from the Australian Government Department of Education, Employment and Workplace Relations for sharing data on schools' socioeconomic status. The views in this article are those of the authors and not necessarily those of the Australian Treasury.

Introduction

For decades, education policymakers have recognised that socioeconomic factors such as income, race, and parental education are crucial determinants of students' outcomes. Not only do these factors matter; they also seem to matter more than school inputs. Since at least the 1966 Coleman Report in the United States, researchers looking at large samples of student test scores have found that family background characteristics explain a larger share of the variation in student performance than school characteristics.

Such findings — and the many subsequent studies that have looked at this question — have profound implications for education policy. The more that children's academic achievement is determined in the home, the less chance that policies to improve schools' performance will have a transformative impact on the life chances of disadvantaged students. At the extreme, if socioeconomic status entirely explains academic performance, it is pointless to think about reforming schools in order to raise educational outcomes.

This article considers this question by looking at the relationship between socioeconomic status and school performance in Western Australian public schools. Western Australia is the only Australian state that publishes the test score performance of all government schools. Although it would theoretically be possible to do so, this study does not identify any individual schools.

Using the relationship between socioeconomic status and school performance across Western Australian government schools, this article considers whether there are schools that consistently perform better or worse than their socioeconomic status would predict.

Naturally, test scores will vary somewhat from year to year, and from test to test. From a policy perspective, it is important to distinguish random fluctuations (for example, variation caused by a barking dog outside the classroom) from systematic overperformance/underperformance by a school which persists from subject to subject, grade to grade, or year to year.

This analysis is also helpful in thinking about how the socioeconomic status of a school's student body might be used in constructing a measure of the performance of like schools in literacy and numeracy assessments.

The article is structured as follows. The next section discusses the evidence from international PISA tests on socioeconomic status and performance. The article then describes the sources of testing and socioeconomic data that is used for Western Australia. Next, the study looks at the relationship between socioeconomic status and

test scores in Western Australia. The article then turns to consider whether the unexplained differences in school performance are lasting or random. The following section analyses the implications of the findings for performance of like schools in literacy and numeracy assessments. The final section concludes.

The OECD's Programme for International Student Assessment (PISA)

To date much of the domestic debate around the link between socioeconomic status and student outcomes has been informed by the OECD's PISA, a test administered to a sample of 15 year-old students in OECD and non-OECD countries. The use of PISA reflects the lack of publicly reported data on student outcomes and characteristics in Australia, rather than any quantitative advantages that PISA offers.

At the broadest level PISA provides two important insights. The first is that socioeconomic status matters. The second is that while it matters, it is by no means deterministic. Table 1 highlights both of these points. It shows the relationship between test scores and the composite PISA measure of socioeconomic status — the index of economic, social and cultural status (ESCS). This index is based on parental occupation and education, family wealth, home educational resources, and cultural possessions. Across all OECD countries, the index is standardised to have a mean of zero and a standard deviation of one.

Table 1: Socioeconomic status and student achievement of Australian students in PISA

	Average score	Slope	Share of variation explained
Scientific literacy			
PISA 2000	528	42	14%
PISA 2003	525	47	15%
PISA 2006	527	43	11%
Reading literacy			
PISA 2000	528	50	17%
PISA 2003	525	44	14%
PISA 2006	513	41	12%
Mathematical literacy			
PISA 2000	533	44	17%
PISA 2003	524	42	14%
PISA 2006	520	38	12%

Source: Thomson and De Bortoli 2008, p 229.

Table 1 (taken from Thomson & De Bortoli 2008) shows three characteristics of Australia's PISA results — the average score, the socioeconomic gradient (slope), and the share of variation across students that is explained by socioeconomic status.

In terms of its average score, Australia scored above the OECD mean (approximately 500) in all three tests. However, over the period from 2000 to 2006,

Australia's average PISA scores fell on each test (this drop was only statistically significant in the case of reading).

The next column shows the socioeconomic gradient, or slope. This is the effect on test scores of a one-unit increase in the ESCS index. The results show that a one-unit increase in socioeconomic status is associated with a 38-50 point increase in test scores; slightly over one school year of achievement. Between 2000 and 2006, the socioeconomic slope has flattened for reading and mathematics, with the change being statistically significant for reading.

Table 1 also shows the share of variation across students that can be explained by this single socioeconomic status measure. In the 2006 PISA tests, socioeconomic status explains 11-12 per cent of the variation in student results, leaving 88-89 per cent to be explained by other factors. The explanatory power of socioeconomic status in PISA has declined over the period 2000-2006, with this drop being statistically significant for reading and mathematics.

From Table 1, it can be seen that socioeconomic status does not explain much of the variation in student outcomes. Another way of thinking about this is that if one was to plot the relationship between ESCS indices and test scores, the dots would not cluster particularly closely to the line. There are many students whose PISA scores place them a long way above the line (performing better than expected on the basis of their socioeconomic status score) and lots of students sitting a long way below the line (performing worse than expected on the basis of their socioeconomic status).

There are many reasons why some students perform better or worse than their socioeconomic status predicts. Students of similar backgrounds might attend schools that differ in quality. There may also be aptitude differences across individuals (driven by genetics, environment, or other factors) that are not related to socioeconomic status. Other reasons could include mismeasurement of socioeconomic status, or measurement error in student testing.

The main focus of this article is on school, rather than individual, performance. It is important to distinguish how much variation is explained by socioeconomic status at the level of the individual student compared to the school level. Because variation amongst individual students is idiosyncratic, it will average out at the school level. As such, a much greater proportion of the variation will be explained by socioeconomic status at the school level than when the analysis is conducted at the individual level.

In the same way that researchers have used PISA to demonstrate how much variation can be explained by a single socioeconomic measure, this study uses Australian literacy and numeracy testing to demonstrate how much variation is caused by school performance alone.

Sources of test score and socioeconomic data

The test results analysed in this article are based upon benchmark data from the Western Australian Literacy and Numeracy Assessments, undertaken by government school students in grades 3, 5 and 7 in August of each year (grade 7 is a primary school grade in Western Australia). In each grade, students are tested in four areas — numeracy, reading, spelling and writing. The benchmark is set at a level such that students who do not meet it are deemed to be at risk of not making adequate progress in literacy and numeracy.

On its 'Schools Online' website (http://www.det.wa.edu.au/schoolsonline/), the Western Australian Government publishes the results for each test, grade, and school — provided that at least 10 students take the test. The reported result is the share of students achieving the benchmark. These results are publicly reported in bar charts, and this analysis converts them to numbers by measuring the heights of the bars. Across all tests and grades, the benchmark pass rate in Western Australia ranged from 10 per cent to 100 per cent, with a mean of 84 per cent.[2]

As a measure of family background, this analysis uses a socioeconomic status index calculated by the Department of Education, Employment and Workplace Relations (DEEWR). This is based on three variables — occupation, education and income — combined in the same manner as for the socioeconomic index used in the Commonwealth's non-government school funding formula. For more details on the methodology, see Farish (2008). The occupation, education and income variables are taken from the 2006 Census, and are measured at the collection district level. In urban areas, collection districts comprise approximately 220 dwellings. The DEEWR procedure aggregates scores by averaging the scores of the collection districts within a 2.5 kilometre radius of the school. The schools' SES scores range from 69 to 129, with a mean of 98.

While this socioeconomic indicator is a reasonable measure for the purposes of this analysis it is important to note that it suffers from two weaknesses.

The first limitation is that this measure omits students who live more than 2.5 kilometres from the school, since it is based upon the socioeconomic status of collection districts around the school's location rather than on the basis of the school's enrolments. To the extent that a school's enrolments are drawn from addresses outside

2 For the purposes of a 'like school' analysis, the use of a benchmark measure has some limitations. One is that the measure is insensitive to changes in performance at the top and bottom of the distribution. Another is that within the group of schools where all students meet the benchmark, the like schools ranking is simply an inverse function of schools' socioeconomic status.

the immediate vicinity of the school, they should ideally be included in the socioeconomic status index. In certain instances this difference would be substantial, for example in the case of selective or non-government high schools. However, because this analysis is only concerned with government primary schools, the socioeconomic status of the surrounding neighbourhoods should be a reasonably good proxy for the composition of the student body.

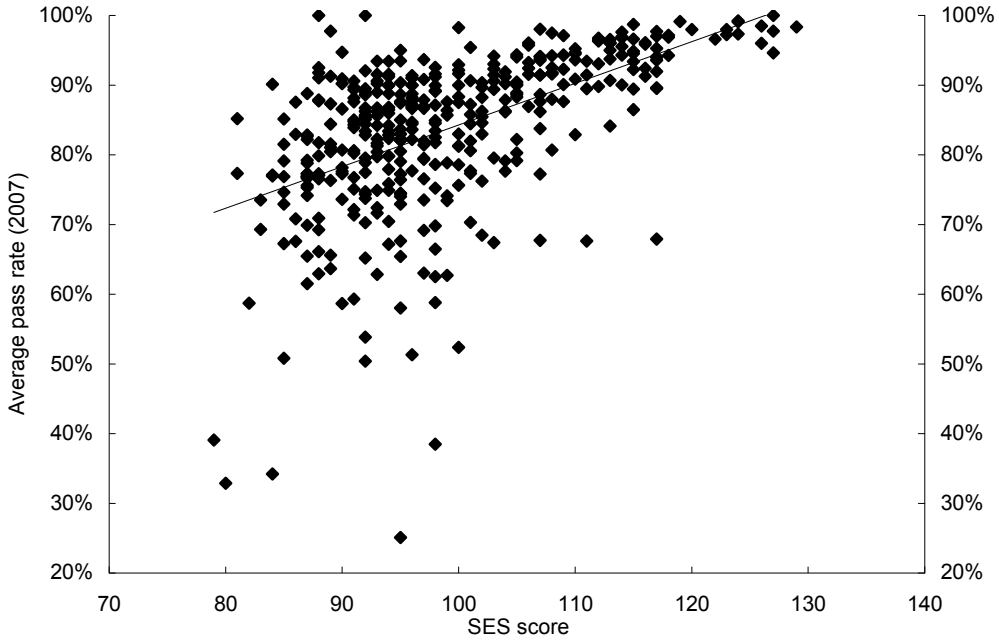
A related weakness is that the index may understate disadvantage in unequal neighbourhoods, since it measures student's socioeconomic status as the average in a geographic area. In certain instances this could cause a significant difference. For example in a neighbourhood that comprises both high and low socioeconomic status groups the average may overstate or understate a particular student's true socioeconomic status. Such an outcome might occur where public housing and expensive private housing sit within the same neighbourhood. The only way to address this issue would be to measure individual students' socioeconomic status, rather than rely on aggregates across those living in the same area.

In addition, this study also controls for two additional variables available via the Schools Online database — the number of children in a given school, and the share of pupils who are Indigenous. Note that while the SES measure captures the occupation, education and income mix in the neighbourhoods surrounding the school, the share of Indigenous students and size of the school are direct measures of the student body.

Explaining school performance with socioeconomic characteristics

Chart 1 shows a plot of the relationship between socioeconomic status and school performance, using test results from 2007, and taking a simple average across the four types of tests (numeracy, reading, spelling and writing) and three grades (grades 3, 5 and 7). There is a strong positive relationship between academic outcomes and socioeconomic status scores (on average, a 10 point increase in socioeconomic status scores is associated with a 6 percentage point increase in the pass rate). However, there is also considerable dispersion around this line, with the linear relationship between socioeconomic status scores and test results explaining only 28 per cent of the variation in student test scores.

Chart 1: Socioeconomic status and school performance[(a)]
(Each dot represents one school)



(a) Pass rate is the average for all tests and grades.

In Table 2, this study looks to see whether it is possible to explain more of the variation in socioeconomic status scores by allowing for a nonlinear relationship between socioeconomic status and student performance, or by including other variables, such as the share of students who are Indigenous, or the size of the school. Allowing for a nonlinear (quartic) relationship between SES and performance only increases the explained variation from 28 per cent to 31 per cent. Similarly, controlling for school size only increases the explained variation from 28 per cent to 29 per cent. However, when taking account of the share of Indigenous students in a school, the share of explained variation rises to 70 per cent.[3] Including all these variables in the regression together (a quartic in socioeconomic score, school size, and the share of pupils who are Indigenous), the model still cannot explain more than 70 per cent of the variation — leaving another 30 per cent that is not explained by these background characteristics.

3 For the schools in this sample, the share of Indigenous students in a school is a better predictor of performance than its socioeconomic status. In part this is driven by schools in regional areas. Restricting the regression to urban schools, the share of explained variation in the specification shown in column 5 of Table 2 falls from 70 per cent to 58 per cent.

Table 2: Explaining school performance with school-level controls[(a)]

Variables included:	(1)	(2)	(3)	(4)	(5)
Linear SES score	Yes	Yes	Yes	Yes	Yes
Quartic in SES score	No	Yes	No	No	Yes
School size	No	No	Yes	No	Yes
Share Indigenous	No	No	No	Yes	Yes
Share of variation explained	28%	31%	29%	70%	70%

(b) Sample size for all regressions is 394 schools.

As noted previously, when one moves from the level of the individual student to the level of the school, much of the individual-specific variation is averaged out. Consequently, these three socioeconomic variables are able to explain 70 per cent of the variation across schools in Western Australian data (Table 2), while a single socioeconomic index only explained 11-17 per cent of the variation across individual students in PISA data (Table 1).

It might also be the case that the present measure of socioeconomic status has not fully captured the variation across schools. Although the socioeconomic status score used here is a composite, it is only a single number, and therefore cannot capture the multidimensional variation across the individual characteristics that make up the index. Some evidence of the limits of such an approach can be seen from the fact that adding a variable denoting the share of Indigenous children at the school boosts the explanatory power of the model to about 70 per cent — leaving around 30 per cent to be explained by cross-school variation within the same socioeconomic groups. Given that school-level measures of socioeconomic status are less precise than the student-level socioeconomic information in PISA, it seems unsurprising that the present approach leaves 30 per cent of unexplained variation across schools, while a study using PISA data has found 18 per cent of unexplained variation across schools.[4]

Are the unexplained differences in school performance lasting or random?

Table 2 showed that about 70 per cent of the observed variation in school performance can be explained by schools' socioeconomic characteristics, leaving another 30 per cent that is not explained by these background characteristics. This raises the question of whether the remaining 30 per cent of 'unexplained' variance is merely random variation, or whether more systematic patterns can be discerned. The present analysis

4 Using multilevel modelling on the Australian PISA data, Thomson and De Bortoli (2008, p 229) estimated that 18 per cent of the variance in achievement between students was due to differences across schools.

tests this in three ways. Defining 'overperforming schools' as those that achieve better test scores than their socioeconomic characteristics would predict, one can ask:[5]

- do schools that overperform in one subject (for example, reading) also tend to overperform in other subjects (for example, numeracy)?
- do schools that overperform in one grade (for example, grade 3) also tend to overperform in another grade (for example, grade 5)?
- do schools that overperform in 2007 also tend to overperform in 2006?

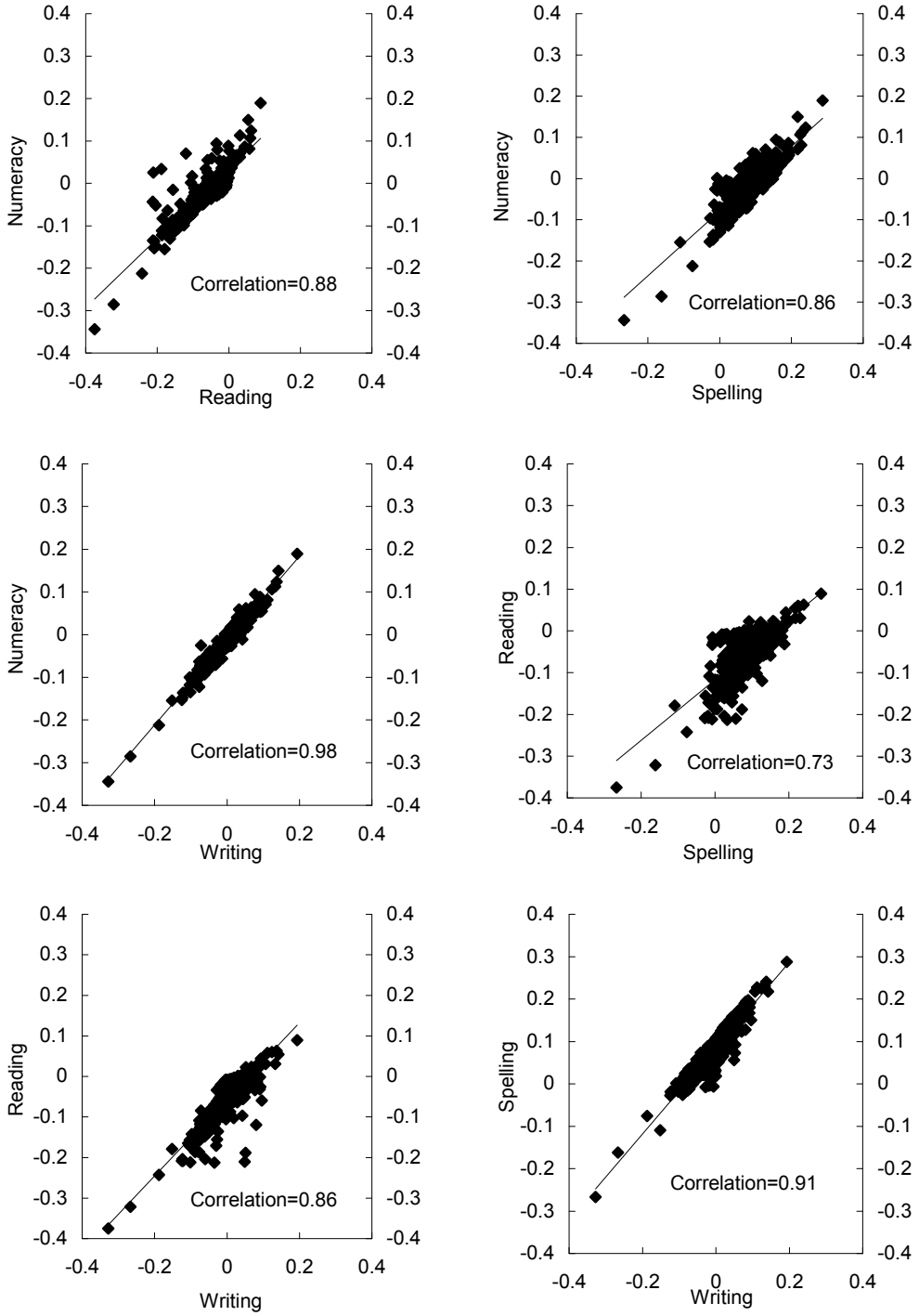
If overperformance is simply due to random variation between tests, one should expect the answers to these three questions to be 'no'. To the extent that there are systematic patterns across subjects, grades and years, it is more likely that this analysis is capturing true differences between schools.

Chart 2 shows the relationship between measures of overperformance across different subjects. With four subjects, there are six possible pairwise comparisons. Schools which overperform in one subject (relative to their socioeconomic status) are very likely to also overperform in another subject. The correlations range from 0.73 (reading and spelling) to 0.98 (numeracy and writing).

5 This article uses 'overperformance' for simplicity; but one could just as easily use 'underperformance'.

How much of the variation in literacy and numeracy can be explained by school performance?

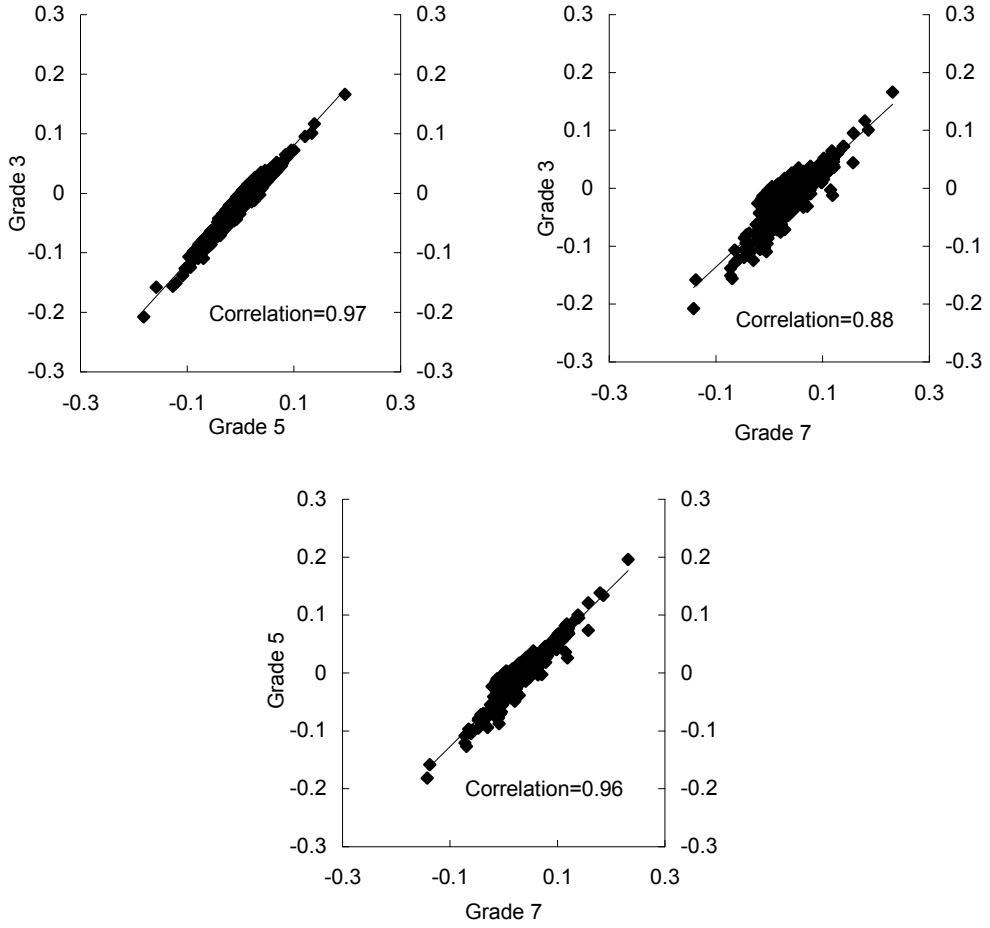
Chart 2: Comparing overperformance across subjects[(a)]
(Each dot represents one school)



(a) Graphs show the residual of a regression of test scores on a quartic in SES score, share Indigenous and school size.

The next question is whether overperformance is systematically correlated across grades 3, 5 and 7. It is important to recognise that this not only removes idiosyncratic school quality differences, but also within-school variation in teacher quality. While the same teacher typically provides instruction in reading, writing, spelling and numeracy, most schools have different teachers instructing grades 3, 5 and 7 (the exception will be composite classes in small remote schools, who are most likely not in the dataset). Chart 3 compares overperformance across grades. The three cross-grade correlations are 0.88, 0.96 and 0.97; not noticeably lower than the cross-subject correlations in Chart 2.

Chart 3: Comparing overperformance across grades[(a)]
(Each dot represents one school)

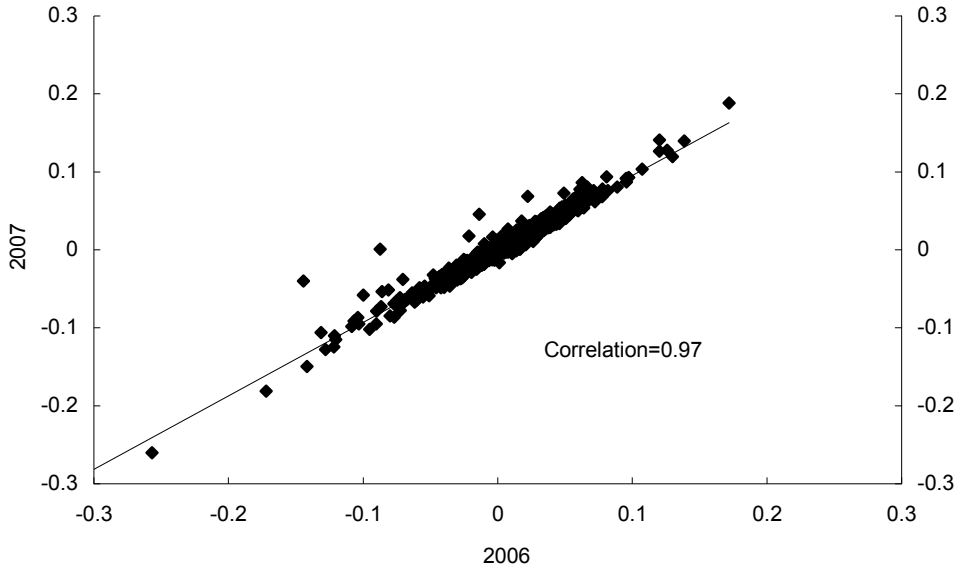


(a) Graphs show the residual of a regression of test scores on a quartic in SES score, share Indigenous and school size.

The third question is whether overperformance differs systematically from one year to the next. In other words, if in one year a school performs better than its socioeconomic status would predict, how likely is it that the school will repeat this the following year?

From Chart 4, it is possible to say that the answer is 'extremely likely'. Comparing schools' benchmark test results, there is a 0.97 correlation between tests in 2006 and 2007. Note that since the tests are administered to students every second year, there is no overlap between the cohorts who took these two tests (with the exception of students who skip or repeat a grade). This analysis therefore indicates that if students in a particular grade and school outperform those in a school with the same socioeconomic status, then it is very likely that students in an adjacent grade in the same school also overperform (relative to socioeconomic status).

Chart 4: Comparing overperformance across years[(a)]
(Each dot represents one school)



(a) Graphs show the residual of a regression of test scores on a quartic in SES score, share Indigenous and school size.

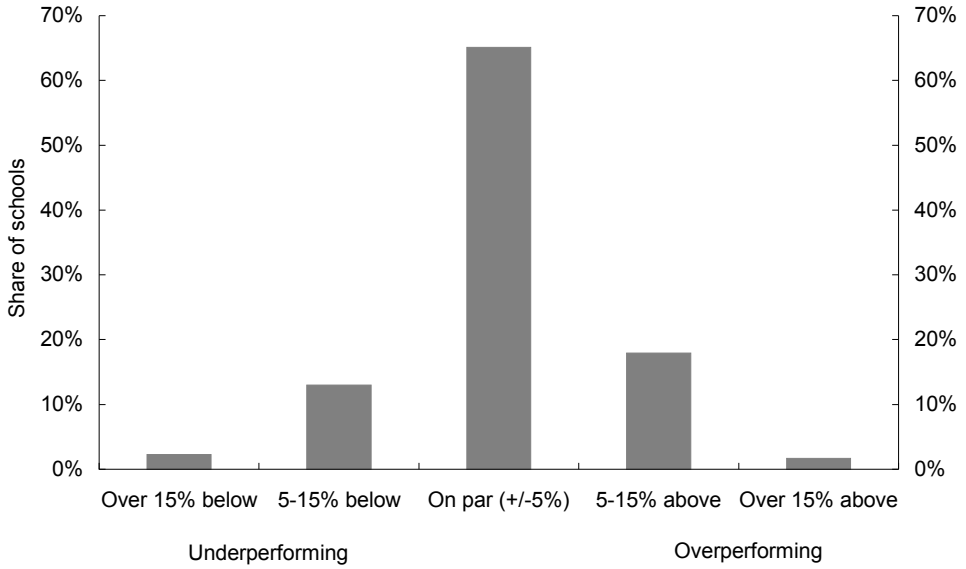
What are the implications for performance of like schools in literacy and numeracy assessments?

Chart 5 shows the distribution of underperforming and overperforming schools, relative to what one would expect, given the socioeconomic status score, school size, and share of Indigenous pupils. While 65 per cent of schools have a benchmark score that is within 5 percentage points of what would be expected given their socioeconomic status mix, 13 per cent are below the benchmark score for similar

socioeconomic status schools by 5-15 percentage points, and a further 2 per cent underperform their socioeconomic status peers by more than 15 percentage points.

Conversely, 18 per cent of schools outperform those of a similar socioeconomic status by 5-15 percentage points, and 2 per cent outperform schools of a similar socioeconomic status by more than 15 percentage points. To see the impact of such overperformance, recall that in a typical school, 84 per cent of students met the benchmark. To overperform by 15 percentage points would be equivalent to a school in a neighbourhood of average socioeconomic status in which 99 per cent of students met the benchmark.

Chart 5: Share of schools that are underperforming or overperforming compared with others of similar SES[(a)]



(a) Graphs show the residual of a regression of test scores on a quartic in SES score, share Indigenous and school size.

The schools which perform best on a 'like schools' analysis are not typically those that perform best on a comparison of test scores that is unadjusted for socioeconomic status. To illustrate this, Table 3 sets out the results for the five government primary schools in Western Australia that perform best on a 'like schools' basis and compares this to their raw ranking. There is a substantial difference between the two measures.

On a like schools basis, the school that performs best has a pass rate that is 26 per cent above what would be expected, given its size and socioeconomic characteristics. Yet its raw pass rate — 84 per cent — would place it 230th out of the 394 schools for which there are adequate data. Put another way, the socioeconomic characteristics of

School A predict that only 58 per cent of its students would meet benchmark, yet the school exceeded this by 26 per cent, with 84 per cent of students meeting benchmark.

A similar pattern can be seen across Schools B-E, which include a school that ranked equal 1st and a school that ranked 382nd. Across all schools, the Spearman rank correlation between raw rank and 'like schools' rank is 0.6, indicating that there are substantial differences between the two sets of rankings (that is, only 36 per cent of the variation in 'like schools' rank can be explained by raw rank.)

Table 3: How does a like schools comparison match up with a raw comparison?[(a)]

	Rank (like schools basis)	Overperformance	Rank (raw)	Pass rate (raw)
School A	1	26% above	230	84%
School B	2	21% above	125	90%
School C	3	20% above	382	59%
School D	4	19% above	1 (equal)	100%
School E	5	16% above	128	90%

(a) Note: Comparison is based on 394 schools. Overperformance measure is based on the specification shown in column 5 of Table 2.

Conclusion

This study has looked at the relationship between socioeconomic status and school performance in Western Australian public schools. The results suggest that a simple linear relationship between school results and socioeconomic status score explains less than one-third of the variation between schools. However, once one allows for a nonlinear relationship between socioeconomic status and performance, and takes account of the share of students who are Indigenous, these background characteristics account for over two-thirds of the variation across schools.

One possible explanation of the remaining variation is simply that it is due to random fluctuations. To test that hypothesis, this study looked at whether a school's 'overperformance' (relative to its socioeconomic status) was sustained across subjects, grades and years. In all cases, the evidence pointed to strong evidence of persistence. This suggests that it is likely that this 'residual' component of school performance captures something important about a school.

In terms of the how much impact a school can have, this analysis suggests that it can explain around one-third of the variation in literacy and numeracy skills. To the authors' knowledge, the only other publicly available analysis is that of the OECD which suggests that between-school variance can explain around one-fifth of the variation in PISA testing.

This level of variation has significant implications for students. Recall from Chart 5 that 20 per cent of Western Australian government schools outperform those of a similar socioeconomic status by more than 5 percentage points. Assuming that these schools would otherwise have been at the state average, this means that in these schools, at least one-third of students who would otherwise not have met the benchmark, do meet the benchmark. This highlights that for students who are at risk of not meeting the benchmark, being in a better performing school can make a difference.

What might affect whether a school performs better or worse than its socioeconomic status would lead one to expect? There are many possibilities, ranging from permanent features of the school itself to transitory aspects such as the current principal and cohort of teachers. However, the first step in determining this is ascertaining which schools are performing better or worse. This knowledge is important in assisting governments, principals, teachers, parents and the community to improve the performance of all schools.[6]

These results provide some evidence in favour of the proposition that socioeconomic status does not determine a school's destiny. Using the data available, there appears to be considerable variation between schools of a similar socioeconomic status. 'Like schools' (on the socioeconomic dimension) do not invariably produce 'like results'.

6 There is good international evidence that that the publication of school-level test scores tends to improve the performance of all schools. See for example Carnoy and Loeb (2002); Hanushek and Raymond (2005); Hanushek and Raymond (2006); OECD (2007).

References

Carnoy, M and Loeb, S 2002, 'Does External Accountability Affect Student Outcomes? A Cross-State Analysis', *Educational Evaluation and Policy Analysis*, Winter, vol. 24, no. 4, pp 305–331.

Coleman, J 1966, *Equality of Educational Opportunity,* U.S. National Centre for Educational Statistics, Washington DC.

Farish, S 2008, *Funding Arrangements for Non-government Schools 2009-2012. Recalculation of the Modified A Socioeconomic Status (SES) Indicator using 2006 Australian Bureau of Statistics Census Data,* Consultant's Technical Report, Department of Education, Employment and Workplace Relations, Canberra.

Hanushek, EA and Raymond, ME 2006, 'School Accountability and Student Performance', *Regional Economic Development*, vol. 2 no. 1, Federal Reserve Bank of St Louis, pp 51-61.

Hanushek, EA and Raymond, ME 2005. 'Does School Accountability Lead to Improved Student Performance' *Journal of Policy Analysis and Management,* vol. 24 no. 2, pp 297-327.

OECD, *PISA 2006: Science Competencies for Tomorrow's World*. OECD, Paris.

Thomson, S and De Bortoli, L 2008, *Exploring Scientific Literacy: How Australia Measures Up. The PISA 2006 Survey of Students' Scientific, Reading and Mathematical Skills,* Australian Council for Educational Research, Melbourne.