# School Expenditure and School Performance: Evidence from New South Wales schools using a dynamic panel analysis

G. Pugha\*, J. Mangana, V. Blackburnb and D. Radicica

a Staffordshire University, UK;

b NSW Department of Education and Communities,

Australia

\* Corresponding author. Institute for Education Policy Research, Staffordshire University, Leek

Road, Stoke-on-Trent, ST4 2DF, UK. Email: [g.t.pugh@staffs.ac.uk](mailto:g.t.pugh@staffs.ac.uk)

**This is an Accepted Manuscript of an article published online on 21 MAY 2014 before print by Wiley in the *British Educational Research Journal,* Volume 41, Issue 2, pages 244–264, April 2015, available online:** <http://onlinelibrary.wiley.com/doi/10.1002/berj.3146/abstract> ***.***

Pugh, G., Mangan, J., Blackburn, V. and Radicic, D. (2015), School expenditure and school performance: evidence from New South Wales schools using a dynamic panel analysis. *British Educational Research Journal*, 41: 244–264. doi: 10.1002/berj.3146

**Abstract**

This article estimates the effects of school expenditure on school performance in government secondary schools in New South Wales, Australia over the period 2006-10. It uses dynamic panel analysis to exploit time series data on individual schools that only recently has become available. We find a significant but small effect of expenditure on relative school performance and significant dynamics in the school improvement process. In addition, our estimates support previous studies in suggesting that the size of schools is an important factor in their performance. Regarding methods for analysing school expenditure effects, this article argues the relevance of a dynamic specification for the investigation of school performance and recommends a continuing research agenda using panel data and dynamic models.

**Key words:**

School expenditure; school performance; panel data; dynamic models; school size.

# 

# Introduction

Given the large investment by governments in schooling, the effects of education on diverse areas such as equity and the increasing desire of governments for ‘value for money’, the effect on school performance of school expenditure is of continuing concern. This includes Australia where over the last decade increasing attention has been paid to the performance and public accountability of Commonwealth and State government funds devoted to school education.

The implementation of the Karmel Committee recommendations by the Australian Commonwealth government in 1973 to fund both government and non-government schools on a needs basis was a watershed moment in Australian School Finance policy development. Constitutionally in Australia government school education is a State government responsibility. Initial studies of the budgetary effects of such Commonwealth school education funding initiatives concentrated on the Federal Fiscal dimensions of Australian Intergovernmental Relations (Blackburn, 1983; Hinz, 2010). In 2010, two-thirds of all children in Australia went to government schools and the combined Commonwealth and State funding for government schools was $26.3 billion and for independent schools was $2.1 billion (Hinz, 2010).

The need for school efficiency and performance studies in Australia as measured by the academic performance of students vis-à-vis the money spent, while considering socio-demographic variables outside the control of the schools, has been recognized since 2008. Indeed, since the introduction by the Commonwealth Government of the “My School” website in 2008, which reports student test scores and financial variables for each school, it has become more important to measure the overall performance of each school and report the cost efficiency index for public school funding policy. The recent Gonski Inquiry report (2011) into the funding of Australian Schools has increased the demand for well-constructed school efficiency and performance studies. However, there is a lack of studies that examine the effects of school and non-school inputs such as financial resources, teacher characteristics, family socio-economic status, and student composition on student outcomes in the context of Australian schools. The focus of this paper is on the analysis of the relative performance of New South Wales (NSW) secondary schools, where performance is measured by schools’ median Year 12 University entrance Australian Tertiary Admission Rank (ATAR) scores, utilizing dynamic panel econometric methods similar to a recent study on English secondary schools (Pugh *et al.* 2011). This article further contributes to establishing the case for investigating school performance by using panel data and a modelling strategy that allows for the influence of past performance on current performance.

This paper seeks to address some of the requirements for the future research directions into school performance assessment outlined in the Gonski report. The aim of our study is to identify the factors that account for performance differentials among schools, in particular the role of expenditure and dynamic effects, which have not been taken account of fully in previous Australian studies. The policy context of this study is set by reforms currently initiating a movement away from centrally determined school spending towards decentralised, school-level spending decisions. This study is an evaluation of the effectiveness of school spending over the period 2006-10, hence under the outgoing (centralised) system. In turn, this study will provide benchmark estimates for future evaluations of the effectiveness of school spending under the incoming (decentralised) system. Accordingly, the results reported below are more informative for policy makers who determine per pupil spending than for school managers who operate within budgetary constraints set by policy makers.

In the next section we examine the characteristics of the NSW secondary school system to provide the setting of our study and as a background to developing our model. Section 3 discusses the existing literature on school expenditure/resources and performance including a brief survey of previous efficiency and performance studies using Australian data. Section 4 considers the data, methodology and model; and Section 5 discusses estimates of the determinants of school performance, including expenditure effects for sub-groups of schools. Section 6 concludes and considers policy implications.

# The New South Wales secondary school system

NSW operates a centralized system of funding to government schools, which are the subject of this study. Approximately 82.5% of school recurrent resources are through NSW state allocations. Commonwealth government allocations make up 13%, this amount having grown since 2009 through increased Federal funding under the Building the Education Revolution and National Partnership programs (Keating *et al*., 2011). School derived revenue makes up about 5% of school funding. Most schools in the government sector in NSW are non-selective, but a small proportion select on performance in an examination at the end of primary school (year 6).

The expenditure that is incurred at the school level from these State and Commonwealth allocations is met at the present time through two basic methods: (1) central allocations of resources (including staff) and funds that schools can utilize; and (2) direct central payments of school-based costs. This is provided through two core mechanisms, centralized staffing allocations and via grants, which are either ‘tied’ or ‘untied’. All staff positions are allocated centrally upon the basis of formulae, with some capacity for variation based on negotiations between the school and Department of Education and Communities personnel. Schools may seek additional staff if they have a budget surplus. Staffing constitutes about 81% of the operational costs of a school, and the effective budget allocations using the same formula across schools will vary due to the different salary steps of teachers. Low socio-economic status (SES) schools also receive allocations under the Priority School Funding Scheme. Global funding allocations are calculated annually for each school at the beginning of each school year and at the commencement of Semester 2 and are intended to help schools meet operational costs.

Special factor loadings are additional entitlements to compensate schools affected by specific circumstances such as urgent minor maintenance and geographical isolation, an important factor in the Australian environment. A Global Funding enhancement element also operates to take account of rural location and socio-economic considerations. Beyond the above allocations a range of services and grants are delivered by central and regional staff including school cleaning and maintenance and professional development programs. Additional equity and needs allocations are also delivered to schools mainly through the staffing formulae.[[1]](#footnote-1)

# Recent literature

Considering briefly the international literature, much of the early research concludes that school expenditure is not closely related to student performance (see for instance Hanushek’s surveys, 2002 and 2003). More recent evidence has generally suggested a small, but positive effect (see Haegland *et al*., 2005, and Holmlund *et al*., 2008, for England; Heinesen and Graversen, 2005, for Denmark). Most recent studies have been at student level, using multilevel modelling with instrumental variables (IV) estimation, reflecting concerns over endogeneity.[[2]](#footnote-2)

Most previous studies of education production functions have not incorporated dynamic elements that capture persistence.[[3]](#footnote-3) An exception is the study of Key Stage 2 performance in England using school-level data by Gibbons (2002), which uses a short time dimension of two years and includes lagged performance as an explanatory variable to estimate the model conditional on previous performance. Since lagged performance is endogenous, the second lag is used as an instrument. A significant but small positive effect of resources on performance is estimated. A more recent study by Pugh *et al* (2011), using English data at Key stage 4, uses a dynamic panel model estimated on five years of data. It reports a generally significant but small effect of expenditure on school performance in the short run, with larger effects in the long run, but it also finds that the effect varies between specialist and non-specialist schools, with the effect on the latter being larger.

Studies of Australian school efficiency that examine the effect of financial resourcing are limited. Mante and O’Brien (2002) and Bradley *et al*. (2004) conduct Data Envelopment Analysis (DEA) and include measures of teacher inputs in their models. The former assesses the technical efficiency of 27 secondary schools in Victoria in 1996. They find that most of the schools were operating at a fairly high level of efficiency, although most were in a position to increase their outputs through a more efficient use of their available resources, and that it was important to take into account the socioeconomic characteristics of students. Bradley *et al*. (2004) discuss the role of league tables in providing signals and incentives in a school education quasi-market framework, extending their DEA analysis with an investigation into the determinants of the estimated efficiency. They compare a range of unadjusted and model-based league tables for primary school performance in Queensland government schools. Their results indicate that model-based tables which account for SES and student intake quality vary significantly from the unadjusted tables.

A report for the Victorian Department of Premier and Cabinet (Lamb *et al*. 2004) examines school performance using school and student level analysis. The school level analysis, which includes student background characteristics, school size and school location, suggests that is important to control for social intake and that many government schools are performing well against non-government schools after taking account of the populations of students they serve. A second step of the investigation considers whether the achievement levels ‘unexplained’ by the primary analysis (i.e. the residuals) are related to various measures of expenditure, and finds that locally raised funds were of considerable importance in secondary schools. However, this two stage procedure is likely to lead to biased estimates, because of endogeneity in the primary regression arising from omitted variables.

Other Australian studies exploring school performance and effectiveness focus on the relationship between the socio-economic status of students and their academic achievements, without considering the effects of school resourcing. Perry and McConney (2010) using Australian data from the Program for International Student Assessment found: (1) increases in the mean SES of a school are associated with consistent increases in student academic achievement; (2) the relationship between school SES and academic achievement is similar for all students regardless of their individual social background; and (3) the strength of the relationship between school SES and achievement becomes stronger as the SES of the school increases. Lamb & Teese(2005) review programs in New South Wales designed to reduce social disparities in student performance and outcomes and conclude that these vary in their effectiveness. They argue that an important fundamental is the need to promote continuity in teaching staff in disadvantaged schools and the recruitment of quality teachers.

Miller and Voon (2011) examine Australia’s National Assessment Programme (Literacy and Numeracy) results for 2008 and 2009 using an education production function approach. Test score data for 3rd, 5th, 7th, and 9th graders are regressed against SES characteristics, type of school, percent of female students, student attendance, school size, and state and region. They find large differences in educational outcomes by state and school type.

# Data, methodology and model

## *The Estimation Method*

As considered above in relation to the literature, it is important in any investigation of the effect of resources on school outputs to consider the possible endogeneity of resources. This is particularly the case when school resourcing systems are partly compensatory for a more disadvantaged student intake or other aspects of the school environment that may otherwise give rise to lower results. For NSW, one aspect of this is illustrated by Table 1, which gives the mean of the median ATAR score and the real per student school expenditure by type of school. The funding per head is lower while results are considerably higher in selective schools; if, however, selective schools’ intakes differ from other schools it is important to account for this in estimation.

***Table 1. Mean ATAR Median Scores and Per Student School Expenditure (in Aus$, at 2005 prices), 2006-10\****

|  |  |  |
| --- | --- | --- |
| **School type** | **Median ATAR Score** | **Expenditure** |
| Non-selective | 55.3 | 11,147 |
| Part selective | 58.7 | 10,589 |
| Fully selective | 93.0 | 9,952 |

\* Because we estimate a model including the lagged dependent variable, the data for 2005 is ‘lost’. Hence, this table and the following charts reflect the sample used for estimation.

Much of the literature has used a regression approach and in order to address potential endogeneity with this method best practice is both to include a wide range of control variables to reduce the number of possible missing variables and to instrument the resource variable(s). However, despite considerable ingenuity in some studies, a major difficulty with this approach is finding appropriate instruments and this has sometimes been shown to be problematic (see for instance the discussion in Wooldridge, 2002, pp. 104-5).

An alternative approach to using instrumental variable estimation is to use matching estimation. Matching is frequently used for the evaluation of public policies, usually where we can distinguish between those units that received a treatment (a treatment group) and those that did not receive a treatment (a control group) and ideally with data from at least two periods (before and after a treatment assignment). Matching estimators are based on the premise that participants (treatment group) should be matched with non-participants (control group), conditional on pre-treatment observed characteristics. Outcomes are then compared between matched units and the difference in outcomes is attributed to the treatment. A treatment assignment can be measured as a binary variable (e.g. a school participates in some programme or it does not) as well as a categorical or a continuous variable (as in our case, where a treatment variable would be school expenditure). Hirano and Imbens (2004) develop an extension to matching estimators based on the propensity score in a setting with a continuous treatment.

However, the main disadvantage of matching is that it can only control for observed characteristics (in this investigation, of schools and pupils). Consequently, matching is inferior to any estimator that can allow for both observed and unobserved heterogeneity (Imbens and Wooldridge, 2009; Khandker *et al*., 2009; Morgan and Harding, 2006). School performance is influenced by factors that are inherently unobservable - such as the ability and effort of pupils, parents, staff and management - and these ‘missing variables’ in a regression give rise to endogeneity if these factors are correlated with expenditure. This study specifies a dynamic panel model and estimates by the system generalised method-of-moments estimator (Arellano and Bover, 1995) that controls for such unobserved heterogeneity.[[4]](#footnote-4) In panel estimation the group-specific effects control for unobservable factors at school level, on the assumption that these are stable throughout the sample period. School level variables that are not measurable (such as effort by pupils and teachers) that are constant (or slowly moving) are factored out (or largely factored out) by the school specific effect in the estimation. In addition, the dynamic panel model provides ‘internal instruments’, not available in cross-sectional studies, from lagged and differenced values of the included variables, which allows the resource variable(s) to be treated as endogenous, given that some heterogeneity still may not be fully accounted for by the other aspects of the modelling.

Our specification as a *dynamic* panel model (i.e. the inclusion of the lagged dependent variable in the model) is, however, based on theoretical grounds. *A priori* there are strong reasons for expecting lagged responses in schools to spending changes and associated improvement measures in education because, for instance, changes take time to implement and to take effect. It is also the case that the effect of an expenditure change may be more important if it is there throughout a student’s time in the school rather than for one year towards the end.

## *The model*

Equation 1 gives the general formulation, where school performance is hypothesised to depend on – i.e. to be a function (*f*) of - previous performance, school expenditure, a set of school level controls (that reflect characteristics of the school such as size, the average level of experience of the teachers and school location) and of the aggregate characteristics of the students attending the school (such as school-level measures of deprivation) that may be related to performance. Each observation on each variable is uniquely identified as belonging to school *i*=1…N (where N is the number of schools in the sample) and year *t*=1…T (where T is the number of years for which we have data on schools).



A wide range of control variables are included in the model to take account of the *observed* heterogeneity of schools and of their pupil cohorts. Accordingly, our estimates of the performance effects of school spending take account of the influence of each of the observed control variables. Similarly, if we are interested in the particular influence of one or more of the control variables, the respective estimates factor out the influence of spending as well as of each of the other control variables. The *unobserved* heterogeneity at school level is controlled for by *vi*, the group-level effects; and the addition of sets of time and regional dummies also controls both for time invariant (or, at least slowly moving) characteristics of the regions that may affect school performance and for developments in some particular period that may affect all schools in the sample in much the same way. The observation-specific error term is *εit*.

## *Data and Variables*

The data for this study came from the Departmental Annual Financial Statements in the state of New South Wales. The dataset contains detailed information on several financial inputs/outputs and other socio-economic variables for all secondary schools in NSW. In NSW the Higher School Certificate (HSC) is the school certificate awarded to students at the end of 13 years of schooling who have completed a pattern of Preliminary and HSC units. However, students take a range of subjects with different means and standard deviations. For this study an overall student achievement measure is needed and the ATAR score is employed, as this is the single unit of comparison in the Australian context that allows comparison between students who have completed different combinations of HSC courses. The ATAR is a ranking of examination performance for HE selection, but ATARs are calculated for all ATAR-eligible students. To be eligible students need to have completed 10 units of ATAR courses, including two units of English. The scaling to derive the score takes into account differences in the abilities between students taking various courses to enable comparison across students, that is to give equivalence. The ATAR is given as a percentile, and represents how many students obtained a lower aggregate than the given student. In the period of this study close to 80 per cent of the HSC candidature received an ATAR score (NSW Vice-Chancellor’s Committee – Technical Committee on Scaling, 2012).

The recent dynamic panel data study (Pugh *et al.* 2011) covered a five year period (2003-2007) in England when real (inflation-adjusted) expenditure per student rose by more than 20 per cent and there also was a strong upward trend in the performance measure. The latter cannot be the case here, given that we have a relative performance measure. However, compared to England, the increase in expenditure in NSW was considerably smaller over the sample period (2006-2010) at a little less than 7 per cent (from around A$10,500 to around A$11,200 per student at 2005 prices). Thus analysis of the relative performance of schools, rather than the effect of overall expenditure increases, is of more importance here than in the English study.

Total expenditure per student includes Central Department determined salaries for teachers and support staff, as well as other operating expenses (maintenance, cleaning etc.). Since the centrally determined payments for staff were a large proportion of total expenditure during this period,[[5]](#footnote-5) we favoured a wider measure of expenditure that allowed more discretion at school level, particularly given the findings of Lamb (2004) for Victoria that found locally raised funds of importance. School Own Source expenditure was identified as a separate expense item and has been added to total expenditure to derive the aggregate school expenditure variable.[[6]](#footnote-6) However, the amount of discretion available to schools was limited at this time, a consideration in our policy discussion in Section 6.

Figure 1 gives the distribution of the median ATAR scores in the sample, which is approximately normal, except for a cluster of high scoring schools that are largely selective (all schools with a score of over 90 are selective and most with scores over 80).

***Figure 1. Histogram of Median ATAR Scores, 2006-10***

******

In order to compare expenditure over time it is important to remove the effect of inflation on expenditure, and the data was deflated using the CPI for Sydney, to give real expenditure with a base year of 2005. The expenditure variables are expressed per full-time equivalent student (FTE). Figure 2 gives the distribution of the aggregated expenditure variable in the sample; given this was strongly right skewed the natural logarithm is used in the estimations. As well as this statistical reason, there is also a good theoretical reason to transform the expenditure variable into a natural logarithm. The model, set out in Equation 1, is an educational production function: namely, schools transform inputs (captured by real per student expenditure) into some output (attainment measured by test scores). It is to be expected that this transformation is subject to diminishing returns: in other words, an increase in expenditure at some low level of expenditure yields a larger gain in attainment than would a similar size increase at some higher level of expenditure. The corresponding functional form for estimation is defined by regressing the level of attainment on the natural logarithm of expenditure and, given the discussion above, this variable is treated as endogenous in our estimation.

***Figure 2. Histogram of Aggregated Expenditure by Schools, per FTE, 2006-10, deflated***

**

It is important to include in the specification a range of control variables as this may also reduce potential endogeneity. Attendance is included as this may in itself affect performance, but it may also be a proxy for student effort and/or parental involvement (the association of attendance with social class in Australia is illustrated in Lamb et al, 2004). This variable is also treated as endogenous in the estimation (since there is likely to be two-way causation; attendance affecting performance, but also poor performance may reduce attendance).

The other control variables that may affect performance are considered to be exogenous.[[7]](#footnote-7) This choice is in line with most other studies that have used instrumental variable estimation (for example Levacic *et al*., 2005), and which have only instrumented the expenditure variable.In this set of control variables are measures that relate to the characteristics of the student body. The socio-economic status of the students is measured by two indices. The Index of Community Socio Educational Advantage (ICSEA) is national, computed by the Australian Curriculum Assessment and Reporting Authority (2010). The mean index is 1000: schools above this number are declared to be more advantaged, those below are less advantaged. Although the ICSEA is a long standing national measure of disadvantaged background, the Family Occupation, Employment and Income (FOEI) index is the preferred measure of the NSW authorities, being based only on NSW government school students while the former includes both government and non-government schools (NSW Education and Communities, 2013). Missing data is dealt with differently in the two indexes and FOEI is more robust to outliers. ICSEA also includes two more school characteristics than does FOEI: whether a school is in a metropolitan, regional or remote area; and the proportion of indigenous student enrolments, although the latter is already controlled for separately in our model. However, given the importance of controlling for disadvantage in school performance estimation and the large size of our sample, both indices are included in the model.[[8]](#footnote-8) These measures were calculated as the respective averages of the 2009 and 2010 values (or just the latter if the former was unavailable for a school), since data was not available for earlier years. The other variables used to control for the characteristics of the student body are: the percentage of students enrolled in English as a second language (ESL); the percentage in special education; and the percentage of indigenous students.

School size, measured by FTE student enrolment may be an important variable in the Australian context (e.g. reflecting the provision of small, rurally based schools). Mok and Flynn (1996) analysing a sample of 4,949 Year-12 students from NSW Catholic high schools suggested that students from larger Catholic High Schools, on average tended to achieve more highly than their peers from smaller schools, even after controlling for students’ background, motivation and school-culture variables. Accordingly, our model is specified with the size of the school, to reflect the rising costs per student of small schools (i.e. scale economies), along with the square of school size, to allow for the possibility of a non-linear relationship (i.e. diseconomies of scale beyond a certain size threshold).Particularly important, given the discussion on endogeneity above, are indicator variables for whether the school is fully- or part-selective (with non-selective schools as the base). In addition, a variable is included to indicate whether the school was a Priority and/or Priority Action School in 2012. As well as showing access to particular streams of funding, this variable may also reflect the socio-economic factors that were the criteria for inclusion in these policies. The average years of service are used to proxy for the quality of the teaching staff.

Finally, regional dummies are included (with Sydney as the base), along with time dummies for each year (with 2010 as the base) to reduce cross-sectional dependence. The full list of variables is given in Table 2, along with standard descriptive statistics. The table also includes the abbreviations of the variables used in the detailed results (included to facilitate replication).

***Table 2. Variable Descriptions and Descriptive Statistics***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable Description** | **Abbrev.** | **Mean** | **SD** | **Min** | **Max** |
| Median ATAR score (percentage) | Median\_ATAR | 57.5 | 14.6 | 11.5 | 98.3 |
| ***School Characteristics*** |  |  |  |  |  |
| Natural log of aggregate expenditure per FTE at 2005 prices \* | Ln\_PPExpenditure | 9.283 | 0.184 | 8.938 | 9.998 |
| * Untransformed \* | PPExpenditure | 10953 | 2203 | 7615 | 21991 |
| Natural log of total expenditure per FTE (i.e. net of school discretionary spending) at 2005 prices \* | Ln\_PPtotexp\_deflated | 9.137 | 0.196 | 8.792 | 10.090 |
| * Untransformed \* | PPtotexp\_deflated | 9480 | 2064 | 6580 | 24111 |
| Boys school | Boys | 0.05 |  | 0 | 1 |
| Girls school | Girls | 0.07 |  | 0 | 1 |
| Attendance, proportion | Attend | 0.89 | 0.03 | 0.72 | 0.97 |
| Student enrolment (FTEs) | stud | 794 | 289 | 152 | 2029 |
| Specialist school | Specialist~V | 0.08 |  | 0 | 1 |
| Fully selective school | Sel\_full | 0.05 |  | 0 | 1 |
| Partially selective school | Sel\_part | 0.06 |  | 0 | 1 |
| Priority/Priority action school 2012 | Priority | 0.23 |  | 0 | 1 |
| FOEI, Average 2009-10 | foei\_ave | 1006 | 80 | 847 | 1216 |
| ICSEA, Average 2009-10 | icsea\_ave | 986 | 86 | 509 | 1202 |
| ESL students, proportion | ESL\_prop | 0.21 | 0.29 | 0 | 1.00 |
| Special educational needs, proportion | Spec\_Ed\_prop | 0.03 | 0.03 | 0 | 0.19 |
| Indigenous students, proportion | Abor\_prop | 0.05 | 0.07 | 0 | 0.61 |
| Teacher’s years of service, average | Yrs\_of\_Service2 | 15.7 | 3.6 | 3.7 | 25 |
| ***Years*** |  |  |  |  |  |
| 2006 | Per\_2 | 0.20 | 0.40 | 0 | 1 |
| 2007 | Per\_3 | 0.20 | 0.40 | 0 | 1 |
| 2008 | Per\_4 | 0.20 | 0.40 | 0 | 1 |
| 2009 | Per\_5 | 0.20 | 0.40 | 0 | 1 |
| ***Districts*** |  |  |  |  |  |
| Hunter/Central Coast | new\_1 | 0.12 | 0.32 | 0 | 1 |
| Illawara and the South East | new\_2 | 0.11 | 0.31 | 0 | 1 |
| New England | new\_3 | 0.04 | 0.20 | 0 | 1 |
| North Coast | new\_4 | 0.10 | 0.30 | 0 | 1 |
| Northern Sydney | new\_5 | 0.10 | 0.30 | 0 | 1 |
| Riverina | new\_6 | 0.07 | 0.26 | 0 | 1 |
| South Western Sydney | new\_7 | 0.17 | 0.38 | 0 | 1 |
| Western New South Wales | new\_9 | 0.08 | 0.26 | 0 | 1 |
| Western Sydney | new\_10 | 0.11 | 0.32 | 0 | 1 |

\* Total Expenditure: Teacher salaries; SASS salaries; Depreciation; Maintenance; Insurance; Cleaning; and Other salaries. Aggregate Expenditure: Total Expenditure plus School Own Resource Expenditure. The statistics for the two expenditure variables reflect the samples used to obtain the estimates reported in Columns 1 and 3 respectively of Table 3 below. The statistics on the other variables reflect the Column 1 sample only, as there is little difference.

Information was collected on all variables for all secondary schools in NSW for the years 2005-2010 and we started with 395 schools in our database. Twenty-seven were excluded on the grounds of their atypical characteristics. Excluded were 24 schools designated as agricultural, which are seen as having their own unique characteristics, and one Community College that had multiple enrolments across several study streams and was not a typical High School. Further inspection of our data indicated two other schools as large outliers which were also excluded: a very small Performing Arts School with considerably higher per student aggregate expenditure in any year than any other school; and a school that, compared to any other in the sample, had considerably higher per student school own source expenditure in all years as well as the highest percentage of Indigenous and ESL students. Non-availability of exam results data and missing information on some other variables prohibit inclusion of a smaller number of schools. As a result we have full information on all variables for 348 or 350 schools, depending on the model estimated (see Table 3 below). The data is annual for each school year in the period 2005 to 2010.

# Findings

Our estimates are reported in Table 3, below: columns 1 and 2 report the specification with the natural logarithm of real aggregate school expenditure per FTE; and columns 3 and 4 report the specification with the natural logarithm of real total expenditure per FTE. Columns 1 and 3 report the original full models and columns 2 and 4 the slightly more parsimonious versions, which omit variables that were consistently insignificant across specifications (and thus do not appear to contribute to explaining variations in school performance). Across all four models reported in Table 3, the diagnostics are satisfactory and the following discussion applies to all four estimates. The Wald test indicates the joint significance, hence explanatory power, of the variables in the respective models. The validity of the instruments for the lagged dependent variable, the expenditure variable and attendance is not rejected by: the *m1* and *m2* tests (i.e. the AR tests for, respectively, first- and second- order serial correlation in the first-differenced residuals); the Sargan test and the heteroskedasticity-robust Hansen test; the difference-in-Hansen tests (not reported but available on request), which establish the validity of each group of instruments separately; and the combination of the *m2* test and the difference-in-Hansen tests, especially for the IV instruments, which does not suggest a problem with cross-school correlation of the error-terms. Finally, a check on the consistency – hence, the validity - of system GMM estimates is that the point estimate on the lagged dependent variable is expected to lie below the corresponding Ordinary Least Squares (OLS) estimate and above the corresponding Fixed Effects (FE) estimate (Bond, 2002). All four estimates reported in Table 3 comfortably satisfy this consistency check.

The estimates reported in Table 3 are stable across both definitions of expenditure, our variable of interest, and across our fully specified models and parsimonious models: there are very few differences in the levels of significance, or indeed in the sizes of the estimated coefficients. Thus the detailed interpretation in this section is for the parsimonious model of aggregate school expenditure (column 2), our preferred wider measure. The quantitative interpretation of each coefficient is to be understood, in each case, as subject to the condition of “other factors held constant”.

The estimated coefficient on the lagged dependent variable suggests that relative school performance, as measured by the median ATAR score, is subject to a statistically significantly positive but small persistence effect. This indicates that change occurs fairly quickly in the secondary schools of NSW. The greater the persistence the more the estimated effects of the explanatory variables accumulate over time. In this case, although the lagged dependent variable is significant, the estimated difference between the short-run and the long-run effects is relatively small (unlike the estimates for England by Pugh *et al.* 2011). The multiplier is estimated at 1.2: i.e. the long-run effects are 20 percent higher than the short-run or impact effects. Thus in what follows the estimated short-run impacts are discussed (that is the immediate effects in one year of a change in the variables).

Other than the influence of recent performance, the factors determining higher relative school performance and estimated at conventional levels of statistical significance are in line with the theoretical expectations and the empirical literature:

* high per student spending, high student attendance, a predominantly girls school, large school size (enrolment), a fully selective admissions policy, a high average socio-economic background of students; and
* not being a Priority School, and having a low proportion of students with special educational needs.

***Table 3. Estimated Effects on the Median ATAR Score (Dynamic Panel Estimation), 2006-10***

Dependent Variable: Median ATAR score

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 |
| Lagged median ATAR score | 0.163 | 0.167 | 0.173 | 0.176 |
|  | (2.20)\*\* | (2.47)\*\* | (2.09)\*\* | (2.22)\*\* |
| Log real aggregate expenditure, per FTE | 42.402 | 42.120 |  |  |
|  | (2.88)\*\*\* | (3.16)\*\*\* |  |  |
| Log real total expenditure, per FTE |  |  | 41.691 | 40.421 |
|  |  |  | (2.21)\*\* | (2.31)\*\* |
| Boys school | -1.562 | -1.364 | -0.869 | -0.771 |
|  | (0.96) | (0.85) | (0.50) | (0.46) |
| Girls school | 4.610 | 5.052 | 5.011 | 5.255 |
|  | (3.94)\*\*\* | (4.64)\*\*\* | (3.64)\*\*\* | (4.30)\*\*\* |
| Attendance | 102.164 | 93.778 | 92.676 | 85.633 |
|  | (2.35)\*\* | (2.48)\*\* | (2.06)\*\* | (2.17)\*\* |
| Student enrolment (FTEs) | 0.061 | 0.060 | 0.059 | 0.056 |
|  | (3.22)\*\*\* | (3.46)\*\*\* | (2.59)\*\*\* | (2.68)\*\*\* |
| Student enrolment (FTEs) squared | -0.00002 | -0.00002 | -0.00002 | -0.00002 |
|  | (2.75)\*\*\* | (2.84)\*\*\* | (2.30)\*\* | (2.29)\*\* |
| Specialist school | 0.224 |  | 0.478 |  |
|  | (0.16) |  | (0.33) |  |
| Fully selective school | 11.432 | 12.049 | 13.452 | 13.629 |
|  | (5.12)\*\*\* | (5.74)\*\*\* | (5.55)\*\*\* | (6.59)\*\*\* |
| Partially selective school | 0.929 | 1.228 | 0.737 | 1.109 |
|  | (0.39) | (0.52) | (0.28) | (0.44) |
| Priority/Priority action school | -2.741 | -2.589 | -2.451 | -2.279 |
|  | (2.38)\*\* | (2.31)\*\* | (1.83)\* | (1.72)\* |
| Socio-economic status (1) FOEI | 0.058 | 0.062 | 0.062 | 0.067 |
|  | (6.12)\*\*\* | (6.41)\*\*\* | (5.62)\*\*\* | (6.13)\*\*\* |
| Socio-economic status (2) ICSEA | 0.004 |  | 0.006 |  |
|  | (1.14) |  | (1.47) |  |
| ESL students, proportion | 1.611 |  | 0.884 |  |
|  | (1.23) |  | (0.62) |  |
| Special education students, proportion | -39.557 | -40.346 | -38.283 | -37.668 |
|  | (2.53)\*\* | (2.75)\*\*\* | (2.01)\*\* | (2.03)\*\* |
| Indigenous students, proportion | 5.946 |  | 5.036 |  |
|  | (0.54) |  | (0.40) |  |
| Teachers’ years of service, average | 0.101 |  | 0.057 |  |
|  | (0.74) |  | (0.36) |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 2006 | 0.244 | -0.290 | 0.159 | -0.162 |
|  | (0.28) | (0.36) | (0.18) | (0.19) |
| 2007 | -0.639 | -0.688 | -0.746 | -0.788 |
|  | (1.26) | (1.37) | (1.52) | (1.64) |
| 2008 | -0.875 | -0.910 | -1.518 | -1.514 |
|  | (2.09)\*\* | (2.22)\*\* | (3.04)\*\*\* | (3.30)\*\*\* |
| 2009 | 0.136 | 0.109 | -0.185 | -0.153 |
|  | (0.22) | (0.20) | (0.21) | (0.19) |
| Hunter/Central Coast | 3.452 | 2.955 | 2.448 | 2.103 |
|  | (2.35)\*\* | (2.07)\*\* | (1.65)\* | (1.46) |
| Illawara and the South East | 3.335 | 3.043 | 2.377 | 2.169 |
|  | (2.01)\*\* | (1.92)\* | (1.34) | (1.28) |
| New England | 8.772 | 8.345 | 8.136 | 7.812 |
|  | (3.33)\*\*\* | (3.32)\*\*\* | (3.00)\*\*\* | (3.02)\*\*\* |
| North Coast | 5.566 | 5.340 | 4.874 | 4.672 |
|  | (3.26)\*\*\* | (3.27)\*\*\* | (2.66)\*\*\* | (2.63)\*\*\* |
| Northern Sydney | 1.329 | 1.008 | 2.110 | 1.722 |
|  | (1.01) | (0.86) | (1.50) | (1.37) |
| Riverina | 5.786 | 5.079 | 5.067 | 4.487 |
|  | (3.13)\*\*\* | (2.81)\*\*\* | (2.73)\*\*\* | (2.42)\*\* |
| South Western Sydney | -0.725 | -1.053 | -1.374 | -1.624 |
|  | (0.52) | (0.70) | (0.93) | (1.05) |
| Western New South Wales | 4.602 | 4.151 | 3.749 | 3.533 |
|  | (2.25)\*\* | (2.18)\*\* | (1.78)\* | (1.77)\* |
| Western Sydney | 0.569 | 0.076 | -0.465 | -0.855 |
|  | (0.36) | (0.05) | (0.31) | (0.55) |
| Constant | -535.56 | -522.38 | -518.47 | -497.35 |
|  | (3.39)\*\*\* | (3.59)\*\*\* | (2.68)\*\*\* | (2.73)\*\*\* |
| Observations  Number of schools  Number of instruments  Wald  AR(1), No serial corr. in 1st differences  AR(2), No serial corr. in 2nd differences  Sargan test for instrument validity  Hansen (robust) test for instrument validity | 1,701  348  54  p=0.000  p=0.000  p=0.839  p=0.156  p=0.747 | 1,701  348  49  p=0.000  p=0.000  p=0.799  p=0.201  p=0.822 | 1,711  350  54  p=0.000  p=0.000  p=0.388  p=0.084  p=0.765 | 1,711  350  49  p=0.000  p=0.000  p=0.364  p=0.095  p=0.822 |

t-statistics from cluster- robust standard errors in parentheses

\* *p*<0.1; \*\* *p*<0.05; \*\*\* *p*<0.01

The effect of real school expenditure per student is estimated to be that a one per cent change in spending per student (approximately an increase in spending of $110 in this period) leads to an increase in the median ATAR score of 0.42 percentage points.

This study has argued for the importance of regarding school performance as the outcome of a dynamic process and that ignoring this may lead to biased estimators. Moreover, for dynamic panel modelling the available performance data can be analysed only at school level. (For each individual pupil, the ATAR score is recorded once, so – in the absence of annual comparable test scores for each individual pupil - lagged values do not exist and dynamic analysis is precluded.) However, it should be noted that school-level analysis itself may give rise to problems.[[9]](#footnote-9) The “ecological fallacy” is concerned with the possibility of spurious inferences when interpreting the results of analysis at some aggregate or group level (“ecological analysis”) ‘in terms of the individuals who gave rise to the data’ (Piantadosi, 1988, p.893). When interpreting ecological analyses, ‘*inferences should be confined to the level of observation*’ (Piantadosi, 1988, p.902; emphasis added), even if ecological analyses ‘may offer valuable clues about individual behaviour’ for further investigation using individual data (Freedman, 1999, p.5; also Piantadosi, 1988, p.902). Our “level of observation” is the school, not only because we use a dynamic panel model, but also because this is the level at which data on school expenditure is available. (Indeed, it is difficult to envisage even in principle how actual expenditure at the individual pupil level could be identified.) Thus inference on the subject of our investigation above, the association between school level expenditure and school average performance, is arguably valid.

However, in school performance research, the ecological fallacy may arise when schools are the units of analysis and ‘aggregate results are used to explain differences between individual students’ (May et al., 2003). Accordingly, we have acknowledged (Introduction, above) that that our study is directed towards policy makers who determine levels and changes in school funding rather than towards school managers who allocate their funding according to the heterogeneous needs and behaviours of their pupils. In our data, control variables such as attendance, socio-economic status, special educational needs and gender are derived by aggregating individual pupil data, so that inference on these variables would best be conducted in the context of pupil-level data. Accordingly, we do not interpret the estimated effects of these variables in detail, but our findings are in line with the literature considered above.

School size, being a school-level variable, where the effect on average school performance may also be of interest to policy makers, is interpreted. The effect of the school size variables is complicated by not being able to keep “other factors held constant” (as school size changes simultaneously with its squared value). However, the effects of both are estimated at high levels of significance and the pattern of signs is sensible: i.e. if enrolment increases from very low levels the squared term has little influence so that the aggregate performance effect is positive; however, at some higher level of enrolment the squared term must dominate and so further increases in enrolment tend then to reduce school performance. In this case, maximum median ATAR is achieved by a school with around 1,500 students, which is considerably larger than the mean enrolment of a little less than 800 (only four schools in the sample were larger than 1500). The estimates suggest a considerably lower performance for small schools (compared to a school with 1200 students the estimated effect of this variable for a school with 600 students is a reduction in the median ATAR score by 14 percentage points).

Schools in several of the regions have significantly higher estimated median ATAR scores than does Sydney. There may be differences in aggregate school ATAR scores between regions because of factors taken account of in this estimation, such as whether schools are selective and size. However, these differences are already factored out by the regression. These differences between regions are arising from factors not, or not fully, accounted for in the model.

**5.1 Expenditure Effects for Sub-Groups of Schools**

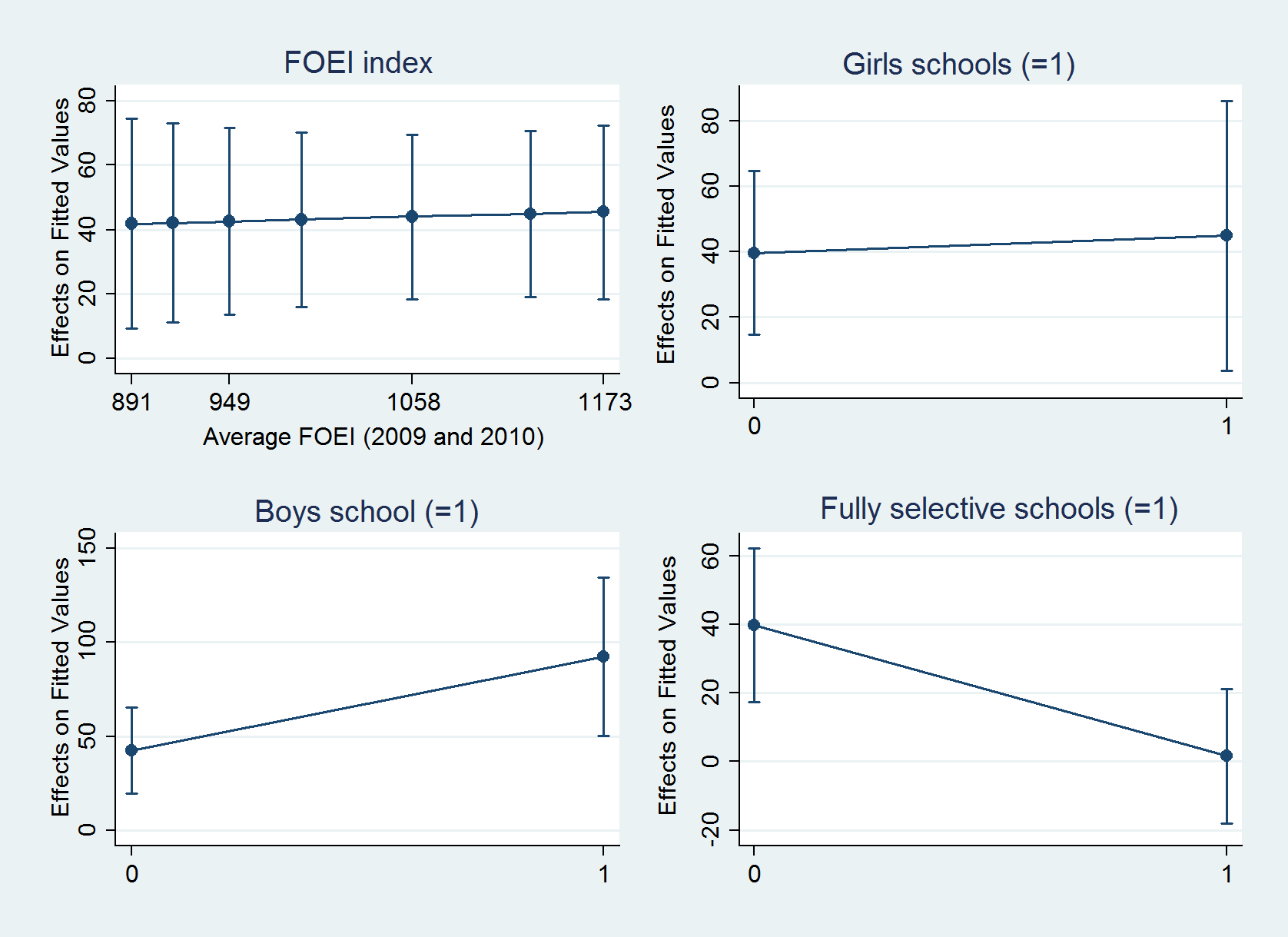
Results reported in Table 3 show the average marginal effect of changes in per pupil school expenditure for all schools. This section checks whether this average effect is sufficiently representative for use in policy decisions on school spending; or, conversely, whether this average effect disguises differences by types of schools that might have important policy implications and thus render the use of an average result problematic. To this end, we explore the possibility of significant differences in school expenditure effects by estimating differences between:

1. schools with student bodies with varying levels of socio-economic status measured by FOEI;
2. schools enrolling 90 per cent or more of girls and mixed schools;
3. schools enrolling 90 per cent or more of boys and mixed schools; and
4. fully selective and all other schools;

The method of investigation proceeds by augmenting the parsimonious regression reported in Table 3, column 2, with interaction terms.[[10]](#footnote-10) These are constructed by multiplying the aggregate expenditure variable with the respective variables for the potential heterogeneities of interest specified above. These interaction terms were included in separate regressions in order to avoid undue multicollinearity. For each of the four estimated augmented regressions, all model diagnostics were uniformly satisfactory. From each regression, we then calculated the extent to which, if any, each source of potential heterogeneity moderates the expenditure effect (how this is accomplished in Stata is comprehensively explained by Williams, 2012).

The results of this investigation are displayed graphically in Figure 3. For concreteness, we explain in detail the graph headed “FOEI index”. In this graph the horizontal axis has representative values of the FOEI index. The vertical axis is labelled “Effects on Fitted values” and refers to the estimated marginal effect on Median ATAR of small changes in aggregate per pupil spending. This graph shows that the marginal spending effect on performance is approximately the same across the different values of FOEI, with no significant differences between the estimates (the 95 per cent confidence intervals around these estimates are shown by the vertical lines symmetrically above and below each dot and these all overlap).[[11]](#footnote-11) In the other graphs the change of the indicator variable is from 1 to 0; for example, in the graph headed “Girls’ schools” 1 is for Girls’ schools and 0 for mixed. In all cases there is overlap in the confidence intervals, thus there is no significant difference in the estimates. Although not a significant difference, the estimate for fully selective schools of the expenditure effect is considerably lower and is not itself statistically significant (this may be because many of these schools are operating close to the maximum median ATAR score). Given this we also estimated our preferred regression (as given in Table 3, column 2) excluding fully selective schools from the sample.[[12]](#footnote-12) There was little difference in the estimated effect of aggregate per pupil expenditure, although the significance level increased using this restricted sample.

***Figure 3. Expenditure Effects for Sub-Groups of Schools (Estimated Marginal Effects with 95% Confidence Intervals)***



Given the lack of statistically significant systematic differences in the estimated marginal performance effects of per pupil spending changes across a variety of important school and pupil characteristics, the overall effect reported in Table 3 can be taken as a representative estimate for the population of NSW schools. This estimated effect therefore provides robust guidance for policy-makers in deciding on overall per pupil spending, with the caveat that this may be lower for selective schools.

# Conclusions, policy implications and policy context

We conclude by considering the objective set out in the introduction of identifying the effect of different factors, in particular expenditure, that affect the performance differentials between schools. Our main conclusion on the effect of expenditure on relative school performance in NSW is that increases in resources do translate into improved performance, although only to a rather limited extent. This finding of a significant effect of expenditure is despite the system in place for the estimation period, which gave limited discretion to schools. Generally recent studies in other developed countries are on absolute measures of school performance, yet there is a similar finding of a statistically significant, but educationally small effect.

The relevance of a dynamic specification for the investigation of school performance is demonstrated in this application: the validity of the estimates is supported by the standard statistical tests associated with this approach; and there are the uniformly significant coefficients estimated on the lagged values of the dependent variables across all our specifications. Taken with the similar findings in Pugh *et al*. (2011), firstly the use of a dynamic specification in this area of education research is supported; and, secondly, the use of panel data together with GMM methods may provide an approach to the thorny question of endogeneity that is endemic in this area. We hope that this study will contribute to a continuing research agenda using panel data and a dynamic model.

The finding on the effect of size on performance is relevant in the Australian context given the relatively small size of schools on average in our sample. As illustrated in Section 5, the estimated effect of size on performance is large. This supports the findings of Mok and Flynn (1996) for NSW Catholic high schools. Interestingly the size that maximises the ATAR score in this study, at around 1500 students, is the same size as estimated by Bradley and Taylor (1998) for 11-18 schools England using the examination performance of 15-16 year-old students.[[13]](#footnote-13) However, there are sound geographical and social reasons for the small size for some of the schools in NSW. Even in districts where this does not apply the results do not imply that every school should aim for this “optimum”. For instance, results may vary with different measures of performance and there may be educational benefits, such as improved development of personal and social skills, associated with smaller schools. However, the results suggest that school size is an important factor that should be considered, along with others, as part of the school improvement agenda.

Finally, there is a policy context for this study, which is transferable to other authorities undertaking major reforms. The NSW Government Secondary school system is currently a predominantly homogenous one that is mainly funded from State Government resources across a common curriculum. However, the analysis in this study suggests that schools do vary across a wide range of dimensions that have a profound potential to affect their respective levels of performance. Individual schools now have a very 39limited degree of control over decision making, over teacher hiring and firing and resource allocation processes. Nevertheless, under the “Local Schools Local Decisions” policy of the newly elected State Government to devolve decision making, each school in NSW will have control over some 70 per cent of their total school budget as well as control over hiring and firing teachers and other school personnel. This new policy will commence in the 2013 school year with 229 schools participating in the program with the balance of the remaining 2,000 schools being integrated into the program of decentralized school decision making by the start of the 2015 school year. This study provides a platform for “before” and “after” assessment aimed at measuring any significant changes in relative school performance arising from such budgetary and staffing devolutionary reforms, to evaluate the implementation and effectiveness of these school decentralization policies. Similar studies could also be undertaken for the other seven State and Territory Government School Systems across Australia. Likewise such studies could be undertaken for the Catholic and other independent school authorities across each State in Australia.

**References**

Arellano M. & Bover O. (1995) Another look at the instrumental variable estimation of error-components models, *Journal of Econometrics*, 68(1), 29-51.

Australian Curriculum, Assessment and Reporting Authority (2010) *NAPLAN Achievement in Reading, Writing, Language Conventions and Numeracy: National Report for 2010,* ACARA, Sydney*.*

Blackburn, V.C. (1983) An Econometric Model of School Finance in Australia with special emphasis on measuring the Impact of Commonwealth School Grants on Government School Budgets in the States from 1973/74 to 1982/83. Paper presented to the 1983 Annual Conference of the Australian Association for Research in Education, Canberra, November 23, 1983, Canberra.

Bond, S. (2002) Dynamic panel models: A guide to micro data methods and practice, Working Paper 09/02 (London: Institute for Fiscal Studies).

Bradley, S., Draca, M. & Green C. (2004) School Performance in Australia: Is There a Role for Quasi-Markets? *Australian Economic Review*, 37(3), 271-286.

Bradley, S. & Taylor, J. (1998) The effect of school size on exam performance in secondary schools, *Oxford Bulletin of Economics and Statistics*, 60 (3), 291-324.

Freedman, D. (1999) Ecological Inference and the Ecological Fallacy, Technical Report No.549, International Encyclopedia of the Social and Behavioural Sciences 6, 4027-30. Available online at: <http://www.stanford.edu/class/ed260/freedman549.pdf>

Gibbons, E. (2002) Geography, Resources and Primary School Performance, Centre for the Economics of Education. Available online at: <http://cee.lse.ac.uk/cee%20dps/CEEDP25.pdf> (accessed 8 March 2013).

Gonski, D. (2011) Review of Funding for Schooling - Final Report, December, Commonwealth Government, Canberra, Australia. Available online at: <http://foi.deewr.gov.au/system/files/doc/other/review-of-funding-for-schooling-final-report-dec-2011.pdf> (accessed 8 March 2013).

Haegland, T., Raaum, O. & Salvanes, K. (2005) Pupil achievement, school resources and family background, IZA (Bonn) Discussion Paper No. 1459.

Hanushek, E. (2002) Publicly provided education, in: A. Auerbach & M. Feldstein (Eds) *Handbook of Public Economics* (vol. 4) (Amsterdam, Elsevier), 2045-2141.

Hanushek, E. (2003) The failure of input-based schooling policies, *Economic Journal*, 113 (February), F64-F98.

Heinesen, E. & Graversen, B. (2005) The effects of school resources on educational attainment: evidence from Denmark, *Bulletin of Economic Research*, 57, 109-143.

Hinz, B. (2010) Australian Federalism and School Funding Arrangements: An Examination of Competing Models and Recurrent Critiques. Canadian Political Science Association Annual Conference, Montreal 1-3 June 2010.

Hirano, K., & Imbens, G. (2004) The propensity score with continuous treatments,

in: A. Gelman & X.-L. Meng (Eds*) Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives* (West Sussex, Wiley

InterScience), 73–84.

Holmlund, H., McNally, S. & Viarengo, M. (2008) Impact of School Resources on Attainment at Key Stage 2, Department for Children, Schools and Families, Research Report DCSF-RR043. Available online at: <http://www.dcsf.gov.uk/rsgateway/DB/RRP/u015147/index.shtml> (accessed 8 March 2013).

Imbens, G. & Wooldridge, J. (2009) Recent Developments in the Econometrics of Program Evaluation, *Journal of Economic Literature*, 47(1), 5-86.

Keating, J., Annett, P., Burke, G. and O’Hanlon, C. (2011) Mapping Funding and Regulatory Arrangements Across the Commonwealth and States and Territories, Melbourne Graduate School of Education, Melbourne.

Khandker, S., Koolwal G. & Samad, H. (2009) *Handbook on Impact Evaluation: Quantitative methods and practices*, World Bank Training Series, World Bank Publications.

Lamb, S. & Teese, R. (2005) Equity Program for Government Schools in New South Wales: a review. Centre for Post-compulsory Education and Lifelong Learning University of Melbourne. Available online at: <https://www.det.nsw.edu.au/media/downloads/research/completedprojects/nswequityrev.pdf> (accessed 8 March 2013).

Lamb, S., Rumberger, R., Jesson, D. & Teese, R. (2004) School Performance in Australia: Results from Analyses of School Effectiveness. Centre for Post-compulsory Education and Lifelong Learning. University of Melbourne. Available online at: <http://www.eduweb.vic.gov.au/edulibrary/public/govrel/reports/schoolperformance-rpt.pdf> (accessed 8 March 2013).

Levacic, R., Jenkins, A., Vignoles, A., Steele, F. & Allen, R. (2005) The effect of school resources on student attainment in English secondary schools, Institute of Education, University of London, DfES Research Report RR679. Available online at: http://eprints.ioe.ac.uk/1319/1/Levacic2005estimatingfullreport.pdf (accessed 8 March 2013).

Mante, B. & O’Brien, G. (2002) Efficiency measurement of Australian public sector organizations: The case of state secondary schools in Victoria, *Journal* *of Educational Administration*, 40(3), 274–296.

May, H., Boe, E. and Boruch R. (2003) The Ecological Fallacy in Comparative and International Education Research: Discovering More from TIMSS through Multilevel Modelling, Research Report No.2003-TIMSS1, Centre for Research and Evaluation in Social Policy, Graduate School of Education, University of Pennsylvania. Available online at: <http://www.gse.upenn.edu/cresp/pdfs/CRESP%20RR%202003-TIMSS1.pdf>

Miller, P. & Voon, D. (2011) Lessons from My School, *Australian Economic Review,* 44(4), 366-386.

Mok, M. & Flynn, M. (1996) School Size and Academic Achievement in the HSC Examination: Is there a relationship? *Issues in Educational Leadership,* 6(1), 57-78.

Morgan, S. & Harding, D. (2006) Matching Estimators of Causal Effects: Prospects and Pitfalls in Theory and Practice, *Sociological Methods & Research*, 35(1), 3-60.

NSW Education and Communities (2013) Getting the funding right: Using Family Occupation and Education Index (FOEI) to identify disadvantage in NSW schools, Learning Issue 5, August. http://www.cese.nsw.gov.au/publications/learning-curve/item/38-learning-curve-5

NSW Vice-Chancellor’s Committee (2012) Technical Committee on Scaling Report on the Scaling of the 2011 NSW Higher School Certificate, Universities Admission Centre http://www.uac.edu.au/documents/atar/2011-ScalingReport.pdf

Piantadosis, S., Byar, D. and Green S. (1988) The Ecological Fallacy, *American Journal of Epidemiology*, 127(5) 893-904.

Perry, L. & McConney, A. (2010) Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003, *Teachers College Record*, 112 (4), 1137-1162.

Pugh, G., Mangan, J. & Gray, J. (2008) *Resources and attainment at Key Stage 4: estimates from a dynamic methodology*. Publication Code: DCSF-RR056. Available online at: <https://www.education.gov.uk/publications/eOrderingDownload/DCSF-RR056.pdf> (accessed 8 March 2013).

Pugh, G., Mangan, J. & Gray, J. (2011) Do increased resources increase educational attainment during a period of rising expenditure? Evidence from English secondary schools using a dynamic panel analysis, *British Educational Research Journal*, 37(1), 163-189.

Roodman, D. (2009) How to do xtabond2: An introduction to difference and system GMM in Stata. *The* *Stata Journal*, 9(1), 86–136.

Williams, R. (2012) Using the margins command to estimate and interpret adjusted predictions and marginal effect, *The* *Stata Journal*, 12(2), 308-331.

Wooldridge, J. *Econometric Analysis of Cross Section and Panel Data* (Massachusetts: The MIT Press).

1. Student population factors utilised include SES, ESL and New Arrivals, Indigenous, Isolated and Disability characteristics. School circumstances recognized include location, enrolment size (diseconomies of scale) and complexity. The disabilities and SES dimensions contribute the most. [↑](#footnote-ref-1)
2. Endogeneity arises when an explanatory variable in a regression model is correlated with the error term. This may be caused by a simultaneous relationship with the dependent variable, by measurement error, or by the correlation of one or more of the explanatory variables included in the model with an explanatory variable that is omitted from the model (and which thus enters the error term). Whatever the reason for endogeneity, estimation by ordinary least squares (OLS) gives biased estimators. Conversely, instrumental variables estimation may provide valid results so long as for each potentially endogenous variable at least one valid “instrument” is available, which is a variable not included in the model, not correlated with the model errors, and highly correlated with the potentially endogenous variable. [↑](#footnote-ref-2)
3. The model estimated in this paper – see Equation 1, below – is an education production function, and the concept is explained briefly in Section 4.3. [↑](#footnote-ref-3)
4. For further details of the model specification and of the estimation method, see Pugh et al., 2008, Appendix 6. Roodman, 2009, gives a ‘pedagogic’ (p.86) introduction to the dynamic panel system generalised method-of-moments estimator and explains how to apply it with *xtabond2*, which is a user-written programme for STATA 9.2 or higher. [↑](#footnote-ref-4)
5. The share of salaries of teachers and support staff in total expenditure is 93.7 per cent (83.3 per cent is the share of teachers’ salaries and 10.4 per cent of supporting staff salaries). [↑](#footnote-ref-5)
6. The share of salaries of teachers and support staff in the aggregate expenditure is 80.5 per cent (71.6 per cent is the share of teachers’ salaries and 8.9 per cent of supporting staff salaries). [↑](#footnote-ref-6)
7. Exogenous variables are explanatory variables that are not correlated with the error term. [↑](#footnote-ref-7)
8. The correlation coefficient between the two indices is 0.76 in the sample used to estimate the model reported in Table 3, Columns 3 and 4. [↑](#footnote-ref-8)
9. We thank an anonymous referee for initial discussion of this point. [↑](#footnote-ref-9)
10. The full results for this sub-section are available from the authors on request. For brevity we summarise the main results only in this section. [↑](#footnote-ref-10)
11. In contrast, Pugh et al. (2011) found that the performance effects of expenditure in England were greater in schools with moderate disadvantage (i.e. those in Quartile 3 with respect to disadvantage, where Quartile 1 is the most advantaged). We also replicated the methodology in this paper but found no such effect – or any other effect - for New South Wales by quartile of the FOEI index. [↑](#footnote-ref-11)
12. The sample was too small to separately estimate the relationship for fully selective schools. [↑](#footnote-ref-12)
13. Percentage in the schools achieving 5+ A\*-C grades in the General Certificate of Secondary Education. [↑](#footnote-ref-13)