Author Manuscript

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/1475-4932.12470](https://doi.org/10.1111/1475-4932.12470)

This article is protected by copyright. All rights reserved

I Introduction

Recent policy reform of schooling in Australia has seen greater emphasis placed on the role of ‘needs-based’ funding models to improve student outcomes. A needs-based funding formula establishes a *base* level of funding that is provided to all schools. This level is intended to reflect the cost required to meet a benchmark level of student performance, in the absence of the effect of educational disadvantage on outcomes (Gonski 2011).

In this paper, we discuss the methodological and data issues presented to researchers in determining base funding requirements for Australian schools. In particular, we outline the commonly used ‘reference school’ approach to determining a benchmark level of school effectiveness and efficiency, and discuss some of its limitations. In contrast to this method, we propose a new approach to estimating base funding requirements that—with the benefit of more granular data—may provide a more robust understanding of school effectiveness and efficiency.

The new method developed in this paper proposes the use of school ‘value-added’ to student learning outcomes as a measure of school effectiveness, with schools’ underlying efficiency in achieving value-added estimated using Data Envelopment Analysis (DEA). This paper focuses, in particular, on how DEA can be used to estimate the frontier of school efficiency, and then subsequently to derive a base funding formula that accounts for key factors that affect school efficiency (such as school size). We demonstrate the use of this new method using student and school-level data from an anonymous Australian school jurisdiction, and (qualitatively) compare the results to alternative methods.¹

II Context and scope of this study

(i) The Gonski Review of Funding for Schooling

The Gonski *Review of Funding for Schooling* in Australia represents a landmark moment in schooling policy in Australia—with the Australian Government using its findings as the basis for sig-

¹Due to reasons of confidentiality, the specific schooling jurisdiction is not named in this paper (this also includes outlining any data or evidence that would otherwise identify the jurisdiction).

nificant new investment in schooling, and with this, a further increased role in setting the policy agenda for school education nationally (an area of policy that has historically been the focus of individual states and territories, particularly with respect to public schooling).

Among other things the Gonski Review advocated for, and subsequently developed, was a new nationally consistent Schooling Resource Standard (SRS) that was intended to apply to all Australian schools, both government and non-government. A SRS can be defined as a formula which relates the observed characteristics of a school (including its students) to an allocation of funding, often defined on a per-student basis. Typically, this formula takes the form of: (i) a base per-student allocation, which reflects the average minimum cost required to achieve current school performance for schools with minimal educational disadvantage; and (ii) additional student- or school-based ‘loadings’, which are based on factors of educational disadvantage and need.²

The findings and recommendations of the Gonski Review were met with a mixed reception among education policy experts and researchers (and the Australian public more broadly). While there has been widespread support behind the idea of a nationally consistent, needs-based approach to resourcing schools, there have been a number of criticisms of the specific approaches and evidence that underpin the design of the Australian SRS. A brief summary of some of the key issues and challenges with the findings and recommendations of the Gonski Review is provided by Justman & Ryan (2013).

(ii) The reference school approach to determining base funding requirements

This paper does not seek to overcome the full extent of the limitations posed by Justman & Ryan (2013) and others; however, in recognising that the notion of a needs-based funding formula is well established among Australian policy makers, we examine the specific approach that was used to determine the *base funding* requirement for schools that forms part of the Gonski SRS (that is, not the component of additional needs-based loadings).

For the purpose of this paper, we define the original approach to determining the base component

²In addition to this formula, the SRS developed by the Australian government also incorporates a measure of parental capacity to pay, which effectively tapers the contribution of the government to school funding based on a school’s socioeconomic status.

of the Gonski SRS formula as a ‘reference school’ approach. In simple terms, this approach involves identifying a subset of schools that meet certain criteria in terms of student achievement. From within this set of schools, a benchmark standard of efficiency is then defined in terms of a per-student funding requirement.³ The specific definition adopted by the Gonski Review was that:

Reference schools [are] those where at least 80 per cent of students are achieving above the national minimum standard, for their year level, in both reading and numeracy, across each of the three years 2008 to 2010.

Given the set of reference schools that meet this criteria, the Gonski SRS formula was estimated based on the relationship between school characteristics and a measure of school funding. The measure of funding used in the Gonski Review was Net Recurrent Income Per Student (NRIPS), which represents the income available to a school for the provision of educational services.⁴ A quantile regression approach was adopted to estimate the relationship between NRIPS and various school characteristics. There are two primary limitations to this approach. The selection of percentile is somewhat arbitrary, and the original approach adopted the 25th percentile. Further, the formula was estimated based on a set of schools facing minimal educational disadvantage. The majority of those used in the sample are efficiently sized, and were located in metropolitan areas with a relatively high socioeconomic status. The results were then used to estimate funding requirements across all schools, extrapolating costs associated with scale and location to a range of settings in which the reference school set may not be representative.

(iii) Overview of method

This study addresses the two primary limitations of the Gonski Review’s approach to determining a resourcing standard for schools. The reference school selection criterion was intended to define a set of schools that allocated resources to effectively deliver educational outcomes. However, the proportion of students that met a national minimum standard is not an adequate measure of the

³It should be noted that Gonski Review acknowledged aspects of this approach as being only *initial* in nature, pending further refinement to the available methodology and data—advancements to which (by the Australian government) are not publicly available at the time of writing this paper.

⁴For an explanation of the scope of NRIPS see: <https://www.acara.edu.au/reporting/national-report-on-schooling-in-australia-2011/funding-Australias-schools/overview-of-my-school-financial-information>.

performance of a school in delivering outcomes. It is highly correlated with student socioeconomic background and does not reflect the relative effectiveness of schools. Further, the 25th percentile of school income used to determine the efficient standard of school resourcing is somewhat arbitrary. It does not necessarily reflect the level of resources required by a school that is allocating resources efficiently.

In simple terms, these two limitations of the reference school approach can be described as relating to the approach to determining school *effectiveness* and, subsequently, school *efficiency* to inform school base funding requirements. The focus of this study is to consider both of these aspects in an approach to determining the base level for a school funding formula. We propose an alternative method which seeks to overcome these specific limitations.

Our proposed approach follows two steps:

1. Estimate the relative *effectiveness* of schools in terms of their value-added to learning outcomes—where the measure of value-added is based on estimates of the school contribution to students NAPLAN outcomes, controlling for student and school characteristics.
2. Estimate the relative *efficiency* of schools in achieving current levels of value-add—where Data Envelopment Analysis (DEA) is used to measure the relative efficiency of each school. DEA allows for an efficiency frontier to be mapped with each school's distance from the frontier reflecting the potential efficiency gains it could achieve, based on comparisons to like schools in terms of both inputs and outputs.

Following these steps, a base funding formula for schools is calibrated by estimating a parametric model relating the efficient level funding and school characteristics, such as school size.

(iv) Data

Student achievement is measured through the 'National Assessment Program - Literacy and Numeracy' (NAPLAN), which is a standardised achievement test that has been used in schools nationwide since 2008. The test covers five learning domains: Numeracy, Writing, and Literacy (which comprises Reading, Grammar and Punctuation, and Spelling), and is administered annually to all

Australian school students in grades 3, 5, 7, and 9. Results for each domain are reported on a common scale ranging from 0 to 1000 with a midpoint of 500 NAPLAN score points. The mean score varies from year to year, across grades and test domain (Australian Curriculum, Assessment and Reporting Authority 2016). For this study, data on all student scores was available from 2012-2016 for one Australian schooling system.

To ensure value-added measures only capture relative improvements to student outcomes that can be attributed to a given school, students who changed schools both within a NAPLAN grade and across NAPLAN years were dropped from the sample. Estimation focused on the learning progress of students from grades 3 to 5, and grades 7 to 9, given that a significant number of students change schools between grades 5 and 7 (that is, between Primary and Secondary school).

Student standardised test scores were linked to student level characteristics, school characteristics, and school financial data to create a panel of student achievement. School expenditure, and contextual factors are used as controls in school value-added models. Parental education and occupation is an important measure of a student socio-economic status, a factor which has been found to significantly impact student achievement internationally. Our measure of parental education and occupation was taken as the highest (non missing) level for both parents.

Financial data used in this study included school expenditure, rather than school funding or income. Data on expenditure included teacher salaries and allowances (among other items), and excluded any capital investment (due to the requirement to focus on annually recurrent resourcing and funding). Expenditure was used instead of funding as it better reflects the resourcing decisions of schools.

III School effectiveness

The method employed for estimating the effectiveness of schools involves calculating school value-added to students learning outcomes. Value-added models of school performance attempt to control for the role of confounding factors such as socio-economic background on student outcomes. The application of value-added measures is used to assess school performance in their ability to lift student outcomes. That is, to identify the conditional effect that schools have on lifting student

outcomes.

We employ a random effects model of student outcomes, with students nested within schools, to estimate the relative effectiveness in producing academic outcomes. These models stem from recent international and Australian research, including Chetty et al. (2014) and Cobb-Clark & Jha (2016).

(i) Method

For the purpose of estimating the effectiveness of schools conditional on contextual factors, calculating value-added involves estimating the following equation for student outcomes,

$$y_{ij} = f(Y'_{ij}; \alpha) + X_{ij}\beta + Z_j\Gamma + E_j\Theta + u_j + \epsilon_{ij} \quad (1)$$

where y_{ij} represents the score of student i in school j . To control for ability and other innate characteristics of the student, we allow the current score to be a flexible function of past scores on all five NAPLAN domains (denoted Y'_{ij}). In practice, we estimate a third order polynomial in each domain score, and include interaction terms between each domain. This is to account for the non-linear nature of NAPLAN improvement.⁵

The inclusion of lagged outcomes to control for ability substantially reduces residual variation in contemporaneous student outcomes. If the focus of high performing students is on improving outcomes in the early years, the estimated value-added as a measure of relative performance may be biased. However, omitting previous outcomes also introduces bias into the measure of outcomes, as schools are limited in the extent to which they can affect student outcomes resulting from innate characteristics.

An advantage of the efficiency analysis presented in the following section is that it is less limited in the range of outcomes that could be considered. Indeed, the cognitive outcomes measured by NAPLAN represent only a small proportion of the role that schools play for students. To the extent that they are measurable, the efficiency analysis could feasibly allow for a range of different outcomes. Several notable examples could include attendance and alternative test measures.

⁵We omit any time dimension from these expressions, as separate regressions are estimated for each calendar year.

Student characteristics are represented by the vector X_{ij} . This includes parental education and occupation, whether the student has a language background other than English (LBOTE), whether they are Indigenous, and their gender.

Further, Z_j is a vector of school (grade) characteristics, including averages of the student level variables, an indicator of remoteness, an indicator for whether the school has both Primary and Secondary students, and a second order polynomial in enrolments. The $E_j\Theta$ term represents the per-pupil expenditure level for school j .

Our interest is in estimating the portion of student outcomes that a school can control, excluding any contextual factors. That is, a prediction of $E_j\hat{\Theta} + \hat{u}_j$.⁶ We assume that $u_j \sim N(0, \sigma_u^2)$ and $\epsilon_{ij} = N(0, \sigma_\epsilon^2)$.

The estimated value-added reflects the average level of student outcomes that is not explained by observed characteristics (excluding the estimated effect of expenditure). In the absence of omitted variables, this is a causal estimate of the relative effect of each school. However, there are a number of additional unobserved factors, such as parental time and investment, that may be correlated with school quality. This is a general limitation of value-added analysis that our study does not seek to address. Derivation of a funding formula discussed in the following section could incorporate additional measures of academic and non-academic outcomes to ensure robustness to the measure(s) of performance.

The relative ‘importance’ of schools is expressed as the intraclass correlation (ICC). The ICC is the fraction of unexplained variation in student scores at the school level, calculated as,

$$\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\epsilon^2} \quad (2)$$

where σ_u^2 and σ_ϵ^2 are the variance of the school and student level residuals.

Value-added in NAPLAN reflects the role of the school in lifting outcomes at the average. As with many standardised testing measures, there are ceiling and floor effects. For example, schools with many students who are at the top end of the distribution may be penalised by this measure of value-added. The efficiency analysis could further be extended to include measures that incorporate

⁶This prediction is scaled by its estimated standard error.

variation in the tails of the distribution.

(ii) Results

The impact of student level and school level contextual factors on student achievement is shown in Table 1. This table highlights a subset of contextual factors tested, and excludes the reporting of NAPLAN domains other than Numeracy and prior achievement.

We find a strong negative relationship between prior achievement in Numeracy and current achievement, reflecting that students with lower prior achievement display greater levels of learning gain between years than those who are higher achieving.⁷ There is a strong negative relationship between female students and learning gain in Numeracy, however, these effects are significantly positive when estimating Literacy learning gain for females. Refugee students are more likely to achieve less growth in Numeracy achievement, and in other NAPLAN domains in Primary schools, however, this effect does not appear in Secondary school. In contrast, LBOTE students, of which Refugee students are a subset, are likely to show greater levels of learning gain in both Primary and Secondary school, showing an important distinction between the outcomes of the two groups.

Figure 1 depicts the distribution of value-added estimated across all years (2012-2016) and all domains for both Primary and Secondary schools. The value-added scores can be interpreted as the average NAPLAN scale score points that a school is estimated to raise student achievement between tests. High performing Primary schools are estimated to improve student outcomes by up to 35 NAPLAN points. For Secondary schools, high performing schools can improve student outcomes by up to 10 NAPLAN points.

The scale of variation in Primary school value-added is flatter than in Secondary schools due to the convex nature of NAPLAN scores. Students in higher grades are more likely to have greater differences in NAPLAN achievement. These widening gaps in achievement are reflected through the averaging process, creating a slightly more variable distribution in Secondary schools. The different distributions reflect the importance of estimating both Primary and Secondary school value-added scores separately, especially when concerned with more than one student outcome measure.

⁷Although not presented, the third order term dominates the marginal effect of prior achievement across the distribution.

Table 1: Regression results, Numeracy (2012-16)

	(1)	(2)	(3)	(4)	(5)	(6)
	2014		2015		2016	
	Year 5	Year 9	Year 5	Year 9	Year 5	Year 9
LBOTE	1.161 (0.958)	2.621*** (0.877)	2.619*** (0.906)	2.948*** (0.828)	2.093** (0.859)	3.288*** (0.773)
ATSI	-1.526 (2.186)	-4.940** (2.035)	-4.405** (2.181)	1.824 (2.08)	-5.881*** (2.016)	-6.121*** (1.796)
Female	-13.63*** (0.524)	-5.526*** (0.503)	-13.02*** (0.502)	-3.818*** (0.482)	-11.26*** (0.488)	-4.902*** (0.453)
Refugee	-8.316*** (2.667)	-0.754 (2.528)	-12.86*** (2.73)	3.952* (2.389)	-6.856*** (2.501)	-0.480 (2.353)
Parental occupation level 1 (highest)	9.874** (4.976)	0.696 (6.937)	-4.477 (4.838)	7.041 (5.682)	8.450* (5.1)	2.677 (6.433)
Parental occupation level 5 (lowest)	6.309 (5)	-4.303 (6.931)	-8.440* (4.859)	3.653 (5.671)	4.887 (5.135)	-1.973 (6.43)
Parental education level 1 (lowest)	-10.49** (4.628)	-5.056 (5.088)	3.266 (5.376)	2.337 (5.518)	6.754 (8.278)	-4.509 (6.622)
Parental education level 7 (highest)	-1.281 (4.385)	2.191 (5)	11.53** (5.144)	8.896 (5.446)	18.11** (8.138)	0.481 (6.574)
Lagged Numeracy	-0.629*** (0.132)	-4.863*** (0.252)	-1.729*** (0.138)	-3.192*** (0.241)	-1.022*** (0.102)	-3.686*** (0.224)
Regional	1.539 (1.584)	-0.533 (1.67)	0.531 (1.534)	0.432 (1.442)	3.697** (1.448)	-1.912 (1.465)
Rural	5.480** (2.422)	0.935 (2.493)	1.723 (2.401)	5.399** (2.165)	8.124*** (2.267)	2.670 (2.265)
Remote	20.52* (11.09)	-2.419 (9.618)	17.090 (11.01)	4.387 (7.797)	22.24** (11.28)	0.637 (9.774)
Combined school	-8.607*** (2.538)	2.193 (1.614)	-3.745 (2.351)	3.801*** (1.454)	-4.330* (2.254)	2.367 (1.481)
Concentration of disability	0.145 (11.87)	-19.530 (24.27)	14.990 (11.2)	14.030 (18.34)	-15.430 (11.29)	-10.480 (18.04)
Enrolments	-0.005 (0.0077)	-0.0149*** (0.00492)	-0.0225*** (0.0075)	-0.0134*** (0.00431)	-0.005 (0.00666)	-0.0100*** (0.00374)
Enrolments ²	0.000 (0.000)	6.45e-06*** (0.000)	1.57e-05** (0.000)	5.45e-06*** (0.000)	0.000 (0.000)	3.76e-06** (0.000)
School LBOTE	11.470 (7.888)	4.309 (9.84)	8.465 (7.455)	-5.660 (8.321)	18.49*** (6.967)	-8.114 (8.356)
School ATSI	-2.500 (13.01)	26.720 (25.41)	-15.330 (13.16)	4.223 (22.12)	14.130 (11.15)	-26.750 (22.81)
School Female	1.085 (4.493)	-3.193 (5.479)	0.674 (4.354)	3.962 (4.732)	-9.449** (4.178)	6.317 (4.529)
School Refugee	25.160 (18.05)	1.113 (26.73)	14.930 (17.48)	-23.340 (23.35)	8.091 (14.6)	11.100 (22.95)
School Numeracy	-0.384 (1.078)	3.615 (4.91)	0.995 (1.146)	2.291 (4.57)	1.078 (0.811)	-1.794 (3.457)
School Expenditure	0.000348* (0.000)	0.000579*** (0.000)	0.000 (0.000)	0.000438** (0.000)	0.000 (0.000)	0.000 (0.000)
Constant	312.100 (229.100)	303.000 (1316.000)	645.9*** (218.100)	554.100 (1007.000)	335.0** (166.000)	1,789** (844.200)

Standard errors in parentheses

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

Figure 1: Average NAPLAN domain value-added, by school type (2012-2016)

To test the practical importance of our value-added estimates, we calculate the ICC for each regression in Table 2. The ICCs range between 0.01-0.11, and vary across both NAPLAN domain and school type. In general, value-added explains a larger proportion of the total variation in Primary and Secondary Numeracy achievement, and Secondary student Writing. Estimated ICCs were consistently higher for Primary schools than Secondary schools across all domains, by an average of 0.05. In relative terms, the effects of student prior achievement and student and school contextual characteristics tend to account for a larger share of the variation in student outcomes than school level value-added measures.

Table 2: ICC estimates, by NAPLAN domain and grade (2012-2016)

	2012	2013	2014	2015	2016
<i>Year 5</i>					
Reading	0.043	0.054	0.053	0.042	0.041
Grammar	0.066	0.090	0.055	0.067	0.060
Spelling	0.089	0.082	0.053	0.050	0.053
Writing	0.083	0.064	0.060	0.053	0.036
Numeracy	0.112	0.103	0.094	0.091	0.078
<i>Year 9</i>					
Reading	0.017	0.017	0.012	0.015	0.012
Grammar	0.014	0.010	0.013	0.013	0.014
Spelling	0.018	0.020	0.026	0.018	0.013
Writing	0.033	0.028	0.028	0.014	0.022
Numeracy	0.044	0.042	0.033	0.027	0.032

The proportion of variation across schools is relatively small, and reflects the fact that there

is substantial heterogeneity *within* schools, both within and across classrooms. This limits the validity of value-added scores as a measure of individual school performance. However, the efficiency analysis and calibration of a funding formula relies more on the distribution of value-added scores overall, and the estimation of value-added scores serves as an adjustment to school performance that accounts for across school heterogeneity in contextual factors that affect outcomes. The calibrated formula is also robust to alternative measures of school performance, such as gain scores (the simple difference in outcomes between testing periods).

(iii) Discussion of value-added results

A key assumption in the calculation of value-added scores is that unobserved school performance is uncorrelated with explanatory variables used in the value-added models. Correlation between omitted school performance (value-added) and control variables would lead to bias in the estimated measure. Value-added based measures of the effect of schools on student outcomes would therefore be understated relative to the true effect of schools on students.

For example, if higher ability students (assumed to be observed through prior performance) are more likely to attend schools with better practice (which is unobserved), then the role of prior performance would be overstated. However, additional analysis suggests that the impact of omitted school-level variables do not substantially impact the results observed.⁸

The relatively low proportion of variation in student outcomes attributable to the school value-added is likely to be a valid observation. It may be the case that the unobserved school-level factors do relatively little to influence the average performance of students, relative to the within school effects of different teachers.

Other research measuring value-added in schools has found variations in teaching quality explain a much larger share of the variation in student outcomes than school level effects. For example, Hanushek et al. (2013) discusses how research in the United States has found a great deal of

⁸These results are omitted for brevity, although estimates of school value-added align closely to estimates using school level fixed effects due to the use of many school level averages of student level controls. A more practical concern that could influence results was NAPLAN participation, which is for the majority of schools around 90% but can be as low as 10% for some schools. This is of concern if selection into NAPLAN is non-random. However, tests between NAPLAN and school value-added show very little correlation, suggesting that the value-added results are not driven by underlying student selection to participate in NAPLAN.

variation in estimates of teacher effectiveness, even within schools.

IV School efficiency

Given an estimate of school effectiveness, the value-added measures developed in the previous section, the relative efficiency of schools at producing those outcomes is then estimated. The base funding of schools is designed to effectively replicate the current system using the minimum required inputs. Given this base level of resources required to replicate the observed distribution of school value-added, remaining resources can be allocated towards other goals of the system.

The method proposed in this section assumes that the level of school practice can be replicated effectively across the system according to the benchmark established, with an efficient standard of resourcing. In the presence of additional restrictions, such as a limited supply of quality teachers and principals, it may not be possible to achieve this benchmark. The method of adjustment and efficiency analysis could incorporate additional structural aspects of the sector, to the extent that these can be observed by researchers. However, the presence of unobservable differences that underlie the costs required to efficiently achieve a performance standard remains a barrier for the field.

Our estimates of technical efficiency in the schooling sector tend to lie between 70 and 80 percent. This implies that, given an appropriate distribution of resources across the sector, 20 to 30 percent of funding can be redistributed for alternative objectives, such as lifting academic achievement of students with disadvantaged backgrounds. These results are somewhat lower than other Australian studies (Mante 2001, Blackburn et al. 2014, Chakraborty & Harper 2017), who tend to find efficiency scores between 85 and 95 percent. We believe there are a number of reasons for this result, including the jurisdiction and the choice of method to control for non-discretionary inputs in the production process (see Muñoz et al. 2006, for an overview of these methods).

The following section proposes a non-parametric linear programming technique, Data Envelopment Analysis, for measuring the relative efficiency of schools at producing student outcomes.⁹ We

⁹Although the non-parametric nature of DEA is advantageous in this context, its inability to readily allow measurement error and inference is problematic. We considered estimating a cost function using stochastic frontier analysis, however, we were unable to obtain sensible estimates (the estimated cost of Literacy and Numeracy were of

then apply this technique to the Australian context, utilising expenditure patterns from government schools in one state, to calibrate a simple base funding formula.

(i) Method

To obtain estimates of the minimum resources required to replicate the existing distribution of school value-added, we employ DEA. DEA combines the concepts of defining a performance standard for schools, and evaluating performance against the current resourcing standard.

The central idea behind DEA as applied to the schooling context is to find the set of schools that are using the least resources for a given output standard (input oriented), or achieving the highest value-added for a given level of resources (output oriented). As the aim of our analysis is finding the base funding requirements of schools, we focus on input oriented measures of inefficiency. Here, the (in)efficiency of a given school corresponds to the proportion by which inputs could be scaled back whilst maintaining current levels of value-added. This distance is based on a linear combination of (potentially many) other schools who are observed to achieve a comparable level of output using fewer inputs.

Figure 2 illustrates the DEA concept. For a given level of output, points 1, 2, and 3 are calculated to be operating using an efficient level of the input. The horizontal distance between the input that point 4 is using, x_4 , and its efficient level of input x_4^* , is its inefficiency. That is, technology is such that point 4 could feasibly scale its inputs back to x_4^* whilst maintaining the same level output by increasing its productive efficiency.

the opposite sign in a number of different specifications). This is likely due to the optimisation nature of the problem, DEA explicitly accounts for trade-offs between outputs without imposing a uniform technology across schools. The correlation of estimated value-added scores within schools ($\rho_{Pri} = 0.5945$ and $\rho_{Sec} = 0.6957$) may be a reason for this observation.

Figure 2: Input oriented Data Envelopment Analysis illustration

Now, consider a general programming problem applied to the schooling context. Schools, indexed by $j = 1, \dots, K$, use expenditure E_j to produce Literacy value-added, denoted y_{jL} , and Numeracy value-added, denoted y_{jN} . Schools are subject to a given set of environmental factors that are non-discretionary. As the problem is to produce value-added, not raw academic outcomes, the effect of non-discretionary factors such as student socio-economic background is already accounted for in output. However, in some contexts it may be more costly to produce a given level of value-added, even after controlling for observable heterogeneity in the estimates. In the context of our application, we found that school size affected the cost distribution independently of its effect on student performance.

We modify a standard DEA problem to account for the presence of an environmental input by looking for the largest proportional reduction in only expenditure, taking enrolments (n_j) as given

(Banker & Morey 1986).¹⁰ This leads to the following DEA problem,

$$\begin{aligned}
& \min_{\{\theta_j, \lambda\}} \theta_j \\
& \text{s.t. } \theta_j E_j \geq \sum_{k=1}^K \lambda_k E_k \\
& \quad n_j \geq \sum_{k=1}^K \lambda_k n_k \\
& \quad y_{jL} \leq \sum_{k=1}^K \lambda_k y_{kL} \\
& \quad y_{jN} \leq \sum_{k=1}^K \lambda_k y_{kN} \\
& \quad \sum_{k=1}^K \lambda_k = 1, \lambda_k \geq 0
\end{aligned} \tag{3}$$

where λ_k is the weight placed in the convex combination of all schools $k = 1, \dots, K$. This specification permits variable returns to scale in production technology, as depicted in Figure 2.

It is worth noting that there are a number of alternative methods proposed to account for the role of non-discretionary inputs in the production process using DEA methods (see Muñiz et al. 2006, for a survey and simulation evidence). The method applied here is a hybrid of several proposed in the literature. In a first stage, estimating school effectiveness, we adjust inputs to account for the differing ability of schools to produce student outcomes given the observable characteristics of each cohort. In the second stage, we explicitly incorporate enrolments as a non-discretionary input in the DEA.

Further, this method differs from a recent study applying DEA to Australian government schools in the state of NSW. Blackburn et al. (2014) apply a conditional DEA estimator that does not allow schools facing a ‘better’ environment to be benchmarks for schools with a worse outcome. In their case, an index of socio-economic status of the school serves to stratify the problem. Schools with a higher value of the index, implying a better environment and therefore student outcomes, do not serve as a benchmark for schools with lower values in the programming problem. We are unable to

¹⁰Our results are robust to including additional non-discretionary inputs, such as socio-economic status and location.

employ this technique in our application as, a priori, we do not know whether increasing enrolments is more favourable for schools.

The solution to the programming problem given in Equation (3) provides a set of efficiency scores for each school j , θ_j . These scores represent the proportion by which expenditure could be scaled back whilst maintaining the current distribution of value-added. As discussed above, the goal of this analysis is to calibrate a formula for the distribution of funds that reflects the base expenditure requirements of schools. To this end, denote this base funding requirement as $\underline{E}$, calculated as,

$$\underline{E}_j = \theta_j E_j \quad (4)$$

Given this distribution of base funding requirements, the final step in our analysis is to fit a relatively simple formula that can accurately describe this distribution.

Whilst we have included a relatively small set of inputs and outputs in our analysis, it is relatively straightforward to include additional inputs and/or outputs. Schools focus on more outcomes than achievement in standardised tests, and the analysis presented here could readily be extended to include these additional factors. For example, student engagement and wellbeing, measured through variables such as attendance and retention, or responses to student surveys of engagement, which may be a precursor to academic achievement. Excluding these measures could bias efficiency estimates downward, as we have restricted the dimensions along which schools can achieve productive efficiency. A judicious choice of inputs and outputs in the analysis should therefore strike a balance between the actions of schools and the goals of the policy maker in calibrating a funding formula.

Finally, as DEA is a non-parametric deterministic method of calculating relative efficiency, it is not straightforward to make statistical inference about the efficiency score distribution. Further, measurement error in either inputs or outputs could bias this distribution. Expenditure and enrolments are likely not measured with error, as they are taken from administrative data. However, the first stage of our analysis involves estimating the value-added of schools, which most likely involves measurement error. Ruggerio (2006) presents simulation evidence suggesting that this bias can be reduced by averaging inputs and outputs across a higher hierarchical dimension. We adopt

a comparable approach in calibrating the funding formula by pooling the estimated base funding requirements over the sample period and allowing for heterogeneity in the average requirements over time.

(ii) Efficiency results

This section briefly describes results from the application of DEA to Australian schools from one state. As discussed above, schools are assumed to produce value-added outcomes in Literacy and Numeracy. We measure Literacy as an average of value-added estimate in the NAPLAN domains of Grammar and Punctuation, Reading, and Spelling. Value-added scores are standardised prior to the efficiency analysis. In producing Literacy and Numeracy value-added, we use per-student expenditure of schools as the discretionary input. Enrolments are considered as the only non-discretionary input.

In order to ensure a reasonable set of comparison schools in the analysis, DEA is run separately across Primary and Secondary Schools. Further, a separate DEA program is run each calendar year in our sample data, from 2012 to 2016. Table 3 presents the results of this analysis. The results suggest that, on average, Primary schools are more efficient than Secondary schools. Pooling scores across the sample period (2012-16), the average efficiency of Primary schools is 0.77, indicating that approximately 23% of expenditure could be allocated to more efficient activities. Overall, the efficiency estimates for Secondary schools suggest that 29% of resources are not used effectively, in the sense that we observe schools with the same value-added and enrolments, but with lower per student expenditures.

The efficiency results are also presented for each year of the sample separately. There is no discernible trend in average efficiency over time for both Primary and Secondary schools.

Table 3: DEA results						
	Overall	2012	2013	2014	2015	2016
<i>Primary schools</i>						
Efficiency score	0.77	0.77	0.74	0.76	0.80	0.75
<i>Secondary schools</i>						
Efficiency score	0.71	0.68	0.71	0.74	0.73	0.69

Within schools, efficiency scores are relative stable over the sample period, as indicated by the correlation between the contemporaneous efficiency score and its lag. Using this measure of stability, the efficiency scores of Primary schools appear more stable ($\rho = 0.75$) than Secondary schools ($\rho = 0.64$). However, this stability is likely driven to some extent by persistence in expenditure levels over time, rather than the value-added scores.

As a corollary, this analysis highlights the role of non-discretionary inputs in the measurement of the relative efficiency of schools. Figure 3 presents the distribution of efficiency scores, comparing our main results to an alternative model that does not include enrolments as a non-discretionary input.¹¹ Efficiency scores for Primary schools in 2016 are presented for expositional clarity, although this pattern persists across all years and for Secondary schools.

The key result is that scale effects are present: smaller schools spend more, on average, to achieve the same benchmark value-added scores. This result is consistent with findings in Blackburn et al. (2014), who find that average efficiency is increasing in school size for both Primary and Secondary schools. Further, not controlling for the effect of school size on cost, separate from the role of size in academic achievement (which is controlled for when estimating school value-added score), decreases average efficiency by up to 5 percentage points. Base funding estimates from an analysis that does not account for scale in the cost of production may mis-attribute inefficiency to scale. This would imply an inefficient redistribution of funds away from smaller schools.

¹¹To preserve anonymity of schools, enrolments have been re-scaled to lie between zero and one.

Figure 3: Effect of allowing for enrolments as an environmental input, Primary schools (2016)

(iii) Calibrating a base funding formula

An estimate of school level base funding requirements, given by Equation 4 above, is then used to calibrate a formula that summarises the level of resources required to maintain the current distribution of value-added. We emphasise here that this is primarily an exercise of calibration not statistical estimation (although we use the terms interchangeably), and is intended to illustrate the method developed above in the design of a school funding model. We begin with a graphical inspection of the distribution of efficient expenditure levels. Figure (iii) plots the relationship between efficient expenditure levels and school size (enrolments).¹² The green line plots a non-parametric estimate of this relationship.

¹²The efficient expenditure level is re-scaled to be bound by zero and one to preserve anonymity of the sample schools.

Figure 4: Efficient expenditure estimates, Primary schools (2012-16)

Based on this examination, our hypothesised expression for allocating base funding is to provide schools a large fixed amount and a constant marginal amount per students. That is, we consider the following specification,

$$\underline{E}_j = \beta_0 + \beta_1 \frac{1}{n_j} \quad (5)$$

where parameters β_0 and β_1 are to be estimated based on the calculated funding requirements.

A detailed exposition of this model and results is left for future research – however, initial testing suggests this model performs relatively well, marginally more so for secondary schools ($R^2 = 0.74$) than for primary schools ($R^2 = 0.67$)

It is notable from this analysis that measures of educational advantage (such as SES) do not appear to be highly correlated with efficiency to achieve value-added. This occurs, in part, because these factors are (by definition) uncorrelated with our measure of effectiveness (value-added) and therefore would only be expected to influence efficiency through their effect on expenditure (that is, if advantaged/disadvantaged schools were found to systematically spend more/less to achieve the same level of value-added). Such a phenomenon is not apparent from these results, in large part because there is no observable correlation relationship between expenditure levels and school

value-added (a result consistent with evidence from the international literature).

This outcome does not suggest that resourcing based on factors of educational need is not required to achieve a given level of school performance. Rather, because the measure of performance used here is intake adjusted, schools can be compared on an equal basis when considering their ability to add value to student learning growth. Importantly, the concept of needs-based funding would imply that students experiencing educational disadvantage require additional funding to overcome the effects of these characteristics on their learning outcomes. That is, additional needs-based loadings would be applied in addition to this base funding requirement to support schools in breaking the link between these factors of disadvantage and learning outcomes. The derivation and application of such loadings, in addition to the base allocation developed here, is an important topic, but is nonetheless left for future research.

V Discussion

(i) An improved approach

The extent to which these results are more efficacious is difficult to test empirically, as it is not possible to define the true or most accurate base level of funding to be used in a SRS formula. Nonetheless, we contend that the new approach developed here is more appropriate and accurate because it overcomes two of the most notable limitations of the Gonski SRS reference school approach – namely that it:

1. incorporates a measure of school **effectiveness** that provides a more *accurate and appropriate* indication of relative school performance, through the use of a value-add methodology; and
2. adopts an approach to measuring **efficiency** that is more *robust*, by estimating the frontier of school efficiency towards achieving value-add through the use of DEA.

As a result of these improvements, we contend that the results from the use of this new methodology will be more *reliable* than those produced using the Gonski SRS reference school approach.

This is because this methodology does not rely on arbitrarily defined and discrete reference groups or standards for the purpose of measuring school effectiveness and efficiency.

(ii) Limitations and areas for further research

While we contend that this new methodology represents a significant step forwards from the original Gonski SRS reference school approach, there remain a number of significant limitations to the analysis and its applicability to the development of school funding formulae.

The most significant limitation arises from the empirical observation that there is no observable (conditional) correlation between per-student expenditure and outcomes across schools within the schooling system considered in this study, combined with the fact that the application of DEA to the measurement of efficiency does not allow for statistical inferences to be made regarding the relationship between funding and outcomes.

The implication of this observation is that we cannot necessarily infer that the efficient funding requirement derived from this analysis is sufficient to drive a benchmark level of performance for all schools in a causal sense. To support this inference, further analysis is required to understand why some schools are more or less efficient than others, and the extent to which the efficiency frontier represents a valid benchmark against which all schools may be assessed (that is, where the drivers of efficiency are within the control of a school – such as the extent of effective resource management). Recent research by Cobb-Clark & Jha (2016) suggests that this may be the case for some schools, but further research is required to understand this phenomenon more fully.

One of the reasons why there appears to be no correlation between expenditure and outcomes is that our measure of school performance (value-added) is not in itself a significant predictor of variation in student outcomes. That is, the estimated relative effects of schools on student outcomes do not vary greatly between schools, and the majority of measures of value-added are statistically indistinguishable from each other. As previously discussed, this finding is consistent with literature from Australia and internationally, and reinforces the value of gaining further insight into how variations of practice *within schools* affects students' learning outcomes.

Further, existing expenditure levels are not randomly distributed among schools in the schooling

system used in this analysis. Rather, funding levels (and therefore expenditure levels) are a function of existing school funding models, which already account for some variation in educational need and efficient costs across schools. This means that there are limitations to the extent to which randomly occurring variation in levels of school expenditure can be used to support inference of the impact of funding on outcomes. In this regard, further research would benefit from combining outcomes and expenditure data across longer time periods and, potentially, for multiple education systems.

Finally, this study has only considered the methodological approach to determining the base funding component of a needs-based funding formula (which would also include additional loadings for educational disadvantage). Building upon this analysis, further research may consider the interaction, both conceptually and empirically, between the development of the base funding requirement and loadings of a SRS formula, and would attempt to develop a clearer understanding of how funding can be used to break the link between factors of educational disadvantage and students' learning outcomes.

References

- Australian Curriculum, Assessment and Reporting Authority (2016), NAPLAN Achievement in Reading, Writing, Language Conventions and Numeracy: National Report for 2016, Technical report, ACARA.
- Banker, R. & Morey, R. (1986), 'Efficiency Analysis for Exogenously Fixed Inputs and Outputs', *Operations Research* **34**, 513–521.
- Blackburn, V., Brennan, S. & Ruggiero, J. (2014), 'Measuring Efficiency in Australian Schools: A Preliminary Analysis', *Socio-Economic Planning Sciences* **48**, 4–9.
- Chakraborty, K. & Harper, R. (2017), 'Measuring the Impact of Socio-Economic Factors on School Efficiency in Australia', *Atlanta Economic Journal* **45**, 163–179.
- Chetty, R., Friedman, J. & Rockoff, J. (2014), 'Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates', *The American Economic Review* **104**(9), 2593–2632.
- Cobb-Clark, D. & Jha, N. (2016), 'Educational Achievement and the Allocation of School Resources', *The Australian Economic Review* **49**(3), 251–271.
- Gonski, D. (2011), *Review of Funding for Schooling — Final Report*, Australian Government.
- Hanushek, E., Link, S. & Woessmann, L. (2013), 'Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA', *Journal of Development Economics* **104**, 212–232.
- Justman, M. & Ryan, C. (2013), 'What's Wrong with the Gonski Report: Funding Reform and Student Achievement?', *Melbourne Institute Policy Briefs Series* **13**(2).
- Mante, B. (2001), 'Measuring the Performance of State Secondary Schools in Victoria: An Application of Data Envelopment Analysis', *Education Research and Performance* **28**(1), 105–132.
- Muñiz, M., Paradi, J., Ruggiero, J. & Yang, Z. (2006), 'Evaluating Alternative DEA Models used to Control for Non-Discretionary Inputs', *Computers & Operations Research* **33**, 1173–1183.

Ruggerio, J. (2006), 'Measurement Error, Education Production and Data Envelopment Analysis',
Economics of Education Review **25**(3), 327–333.

Author Manuscript

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Gort, W;King, R;Weiss, A

Title:

Measuring School Efficiency and Effectiveness to Inform the Design of a School Funding Formula

Date:

2019-06-01

Citation:

Gort, W., King, R. & Weiss, A. (2019). Measuring School Efficiency and Effectiveness to Inform the Design of a School Funding Formula. The Economic Record, 95 (S1), pp.65-78.
<https://doi.org/10.1111/1475-4932.12470>.

Persistent Link:

<http://hdl.handle.net/11343/286853>