Oxford Review of Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/core20

Good schools or good students? The importance of selectivity for school rankings

Aedín Doris, Donal O'Neill & Olive Sweetman

To cite this article: Aedín Doris, Donal O'Neill & Olive Sweetman (2022) Good schools or good students? The importance of selectivity for school rankings, Oxford Review of Education, 48:6, 804-826, DOI: 10.1080/03054985.2022.2034611

To link to this article: https://doi.org/10.1080/03054985.2022.2034611

Published online: 18 Feb 2022.

OXFORD REVIEW OF EDUCATION
2022, VOL. 48, NO. 6, 804–826
https://doi.org/10.1080/03054985.2022.2034611

Good schools or good students? The importance of selectivity for school rankings

Aedín Doris[a], Donal O'Neill[b,b] and Olive Sweetman[a]

[a]Department of Economics, Maynooth University, Maynooth, Ireland; [b]Department of Economics, IZA, Bonn, Germany

ABSTRACT
This paper uses a rich set of student background characteristics to estimate the value added of second-level schools in Ireland. We show that there is a considerable degree of reranking of schools when we move from analysing raw outcomes to value added; in many cases the best performing schools in raw terms are not the best in value-added terms. We show that, contrary to popular perception, fee-paying schools do not add higher value than other schools. A simulation exercise suggests that if parents chose the best value-added school from among the set of feasible schools, then this reallocation of students has the potential to increase academic achievement substantially.

KEYWORDS
School value added; school choice

1. Introduction

Choosing a good school for their child is a key decision for parents. In making their choice, parents often rely on outcomes such as the average exam performance of students in a school. However, when students non-randomly select into schools, such outcomes partly measure the prior ability or family background of the student body rather than the contribution of the school. In some countries, governments provide information on the value added of schools, taking account of compositional differences. However, others do not. For example, while the Irish government publishes inspection reports on individual schools, these are qualitative and there is no information on school rankings, either in terms of raw examination results or value added. In the absence of other information, newspapers produce *ad hoc* league tables based on progression to university. Despite their acknowledged shortcomings, these league tables are believed to affect parents' choices. In this paper, we provide estimates of the value added of Irish second-level schools based on the results of a state examination taken by all students at age 15 and examine the implications of this for school choice.

The data we use allow us to control for a rich set of background characteristics, including information on non-cognitive skills, parental education, family structure, and prior achievement. The choice of control variables in a value-added model is the

CONTACT Donal O'Neill donal.oneill@mu.ie Department of Economics, Maynooth University, Rhetoric House, Maynooth, Co. Kildare, Ireland

focus of the literature on contextual and non-contextual value-added models. While non-contextual value-added (NCVA) models adjust only for prior attainment, contextual value-added (CVA) models additionally control for a range of other socioeconomic and demographic characteristics. Researchers further distinguish between pupil-level contextual variables such as age, gender, ethnicity and socioeconomic status and school-level contextual variables such as area deprivation indicators, school ethnic mix or average prior achievement. These latter controls measure factors that operate at the school level but are beyond the control of teachers or the school. Raudenbush and Willms (1995) argue that the choice of controls depends on the purpose of the value-added exercise; school-level factors should be included if the purpose is to hold schools accountable, whereas if the purpose is to inform parents on school choice, controlling only for student-level characteristics is preferred.

The Irish case is interesting because of the notable absence of information available to parents when choosing second-level schools. In contrast, Wilson et al. (2006) and Leckie and Goldstein (2017) document the detailed information available to parents in England when choosing a school. Since the early 1990s school league tables of student progress in state-funded secondary schools have been published. The measure of progress has changed over time. From 2006–2011 school league tables were constructed on the basis of CVA models, but in 2011 this was abandoned due to concerns that CVA encourages schools to set lower targets for disadvantaged students than for advantaged students with the same prior achievement. However, this view is challenged by Leckie and Goldstein (2017, 2019) who argue in favour of a CVA approach and conclude that accountability systems that ignore student background reward and punish the wrong schools (see also Dearden et al., 2011). The current system used in England, an NCVA approach called Progress 8, is evaluated in Prior et al. (2021). A student's Progress 8 score is calculated as the difference between their attainment score at the end of KS4 (typically at age 16) and the average attainment of all pupils nationally with the same KS2 score (typically at age 11). The average scores of all schools are published online. In contrast, parents in Ireland are given almost no quantitative information on schools and must rely on league tables of schools' higher education progression rates published in national newspapers, constructed using information obtained from Higher Education institutions on the schools of origin of their first-year students. They are known to have serious limitations. Consequently, policies that improve information regarding school effectiveness are particularly relevant in Ireland.

We use a CVA approach to compare value added in Irish schools. While there are several papers explaining differences in academic outcomes across Irish schools (Cullinan et al., 2021; Pfeffermann & Landsman, 2011; Sofroniou et al., 2000; Weir & Kavanagh, 2018), measuring value added is not the focus of these papers. An exception is Smyth (1999), who estimates value added across a sample of Irish schools, However, prior achievement is measured at a time when students have already spent almost three years in the school. As noted by Smyth, to the extent that this measure is capturing some school effects, this approach will lead to an underestimate of school value added. In our data, prior achievement is measured before the students enter secondary school. We find that the variance of value-added scores across schools is substantially lower than the variance of raw scores. Moreover, the ranking using value added differs substantially from that using raw examination scores.

Having obtained estimates of value added for each school, we then look at the factors associated with a school's rank in raw scores and value added. We find that although school type is important in determining raw rank, it is unimportant for value-added rank. We also find that there are very few observable school characteristics that are correlated with value added, an exception being that schools with female principals have higher value added. Finally, our analysis suggests that if parents chose the best value-added school from among the set of feasible schools, then this reallocation of students could potentially increase academic achievement substantially.

2. Irish school system

There are three main categories of schools in Ireland: voluntary secondary schools; vocational schools and community colleges; and community and comprehensive schools. About 50% of schools are voluntary secondary schools and they account for almost 60% of students. These are schools that were established by religious orders and so are all denominational. Traditionally, they were single-sex schools, but over the years about 40% have become mixed, often due to the amalgamation of the boys' and girls' school in an area.

About 46% of schools are vocational schools or community colleges, almost all of which are mixed. These schools, which cater for about a quarter of children, are run by area-based Education and Training Boards (ETBs) and are non-denominational. They were traditionally oriented towards technical and manual rather than academic education but that changed since the 1970s and they now provide a full range of academic subjects (Hannan & Boyle, 1987). Nevertheless, the perception that academically inclined and/or middle-class children attend secondary schools rather than vocational schools and community colleges persists to some extent. The remaining 2% of schools are comprehensive and community schools, which also tend to offer a wider range of subjects than secondary schools. They are denominational and mixed.

About 7% of Irish schools are fee-paying, which although not eligible for government capital funding, have the unusual feature that teachers' salaries are predominantly state funded.[1] For this reason, fees are relatively low by international standards, ranging from about €3,000 to €8,000 per annum.[2]

While there is variation in management structure across school types, individual schools have limited autonomy. For all schools, the Department of Education and Skills sets curricula, regulates the management, funding and staffing of schools, and centrally negotiates teachers' salary scales (Darmody & Smyth, 2013). However, there is no centralised system of allocation of children to schools. Schools set their own criteria, which can be based on location, religious denomination, or whether the child's siblings or parents attended the school; academic performance may not be used as a criterion for admission.

To address educational disadvantage, the government operates a scheme designed to support schools with a high proportion of students from disadvantaged backgrounds, the DEIS (Delivering Equality of Opportunity In Schools) scheme. DEIS schools benefit from a slightly lower pupil-teacher ratio. About 27% of second-level schools are in the DEIS scheme.

As well as the fee income received by fee-paying schools, other schools can solicit 'voluntary contributions' from students' parents, with varying amounts requested and varying degrees of persuasion applied. Thus, there may be differences in the resources available to schools.

Students typically enter second-level education at age 12 and all begin the three-year Junior Cycle. All students must study English, Irish, Maths and CSPE (Civic, Social & Political Education). Most students also take a foreign language, History, Geography and Science, and typically at least two other optional subjects, including everything from Ancient Greek to Woodwork. At the end of the Junior Cycle, at about age 15, they sit the Junior Certificate (JC), which entails a written state-wide examination in each of the 10–13 subjects studied. They then enter the Senior Cycle, which takes a further two or three years. At the end of the Senior Cycle, students sit the Leaving Certificate (LC); the grades on this examination determine the higher education programme to which students are admitted. Schools are not involved in assessing their own students; both the JC and the LC are graded centrally following strict national guidelines.

Due to data availability, we use the JC examination scores in our analysis. Although this is a low-stakes exam compared to the LC, many students and parents take it seriously as it can determine what stream or class a student enters at Senior Cycle. The JC has the advantage over the LC of being available for students prior to the compulsory school leaving age. Smyth (1999), finds that performance in the JC is highly predictive of performance in the LC, with a correlation of 0.8 between the two

3. Methodology

There are a number of approaches to estimating the effectiveness of schools (Eyles et al., 2016). A common approach to estimating value added, which we adopt, uses a regression to control for prior student achievement and other differences in observed characteristics across schools (Belfield et al., 2018; Cunha & Miller, 2014; Dearden et al., 2011; Ehlert et al., 2014; Ladd & Walsh, 2002). We estimate the following regression model:

$$Y_i = \beta_0 + \beta_1 A_i + \beta_2 X_i + \sum_{s=1}^{S} \mu_s D_{is} + \varepsilon_i \quad (1)$$

where Y_i is the outcome of student i, A_i is prior achievement, X_i is a vector of student-level characteristics that are important in determining outcomes and D_{is} are a series of dummy variables taking the value of one if student i is in school s and zero otherwise. μ_s therefore measures the individual contribution of the school after student-level background factors and prior achievement have been taken into account.

The school effect, μ_s, can be interpreted as the value-added score but because this will not be on the same scale as the raw score, researchers often calculate a predicted score (Meyer, 1997):

$$\widehat{Y_s} = b_0 + b_1\bar{A} + b_2\bar{X} + \widehat{\mu_s}$$

where $\widehat{Y_s}$ is the predicted score for a given school s with values of A and X set to some benchmark levels, $\bar{A}$ and $\bar{X}$.

An issue to consider when estimating value-added models is the fact that test scores are bounded, which may give rise to ceiling effects. The concern is that if advantaged schools can only move students from one test ceiling to another then value-added measures may be biased against these schools. However, Koedel and Betts (2010) show that over a wide range of test-score-ceiling severity, value-added estimates (of teachers) are only negligibly influenced by ceiling effects. Only when over half of students are affected by the ceiling do they find significant biases in value-added estimates. In our sample the average percentage of students scoring the maximum within a school is 5.17%. The low percentage of students attaining the maximum grade suggests that ceiling effects are not an issue for our analysis.

A further issue when estimating value-added models is that some schools are small. The Bayesian shrinkage estimator can be used to adjust for small numbers of observations within a school (e.g. Jacob & Lefgren, 2008; Kane & Staiger, 2002), although in recent research Guarino et al. (2015) urge caution when using this estimator. The choice of estimator makes very little difference to our results (see Appendix A).

4. Data

The data used in this paper are drawn from the Growing Up in Ireland (GUI) study, which tracks the development of a cohort of children born between November 1997 and October 1998. We use data from the first three waves when the children were aged 9, 13 and 17/18. The Wave 1 sample was generated through the national primary school system, with children from 910 randomly selected schools participating in the study.[3]

By Wave 2, the study children had moved into second-level schools. As a result of the relatively free school choice, 627 (of about 720) second-level schools were included in the second wave. By Wave 3, the students had all sat their JC examinations. We use self-reported scores from the JC to construct the outcome variable. Following Sofroniou et al. (2000), we convert the grades for each JC subject into a numerical value ranging between 1 and 12 for each subject (see Appendix B for more detail). To construct our overall score, we add the scores from each student's seven best subjects.[4]

Prior achievement is measured at Wave 1, using reading and maths 'Drumcondra' tests, which were administered by GUI fieldworkers at the child's primary school. Drumcondra tests are curriculum-based standardised tests that are widely used in Ireland.

In this paper, we adopt a CVA approach by controlling for a range of additional student characteristics. The GUI has a rich set of family background variables, which allows us to control for family income, parental education and family structure as well as the gender of the students. In addition, the GUI constructs several measures of non-cognitive skills of the student using the Strengths and Difficulties Questionnaire (SDQ) to assess the student's socio-emotional development and behaviour. We use these SDQ scores in estimating our value-added models. We are unaware of any previous study that has controlled for non-cognitive skills when measuring the value added of schools.

We follow Raudenbush and Willms (1995) and omit school-level variables from the value-added model; instead we consider their importance when looking at the determinants of value-added.

We exclude a very small number of schools that solely cater for students with special needs or for whom school type was unavailable. The only additional restriction we impose on our sample is to exclude schools with fewer than five study children. We do this in order to balance the precision of the value-added estimates against the number of schools included in the analysis. Although the minimum is five, the average number of sampled students in a school in our analysis is 11.8 and the maximum is 48. As mentioned earlier, when we examine the sensitivity of our results to the fact that some schools have small sample sizes using a Bayes shrinkage estimator, we find very similar results, indicating that our results are robust to corrections for small samples, as reported in Appendix A.

Table 1 provides summary statistics for the student-level variables used in our analysis. The first column presents statistics for the full Wave 1 sample of 8,568 individuals. Sample attrition between Waves 1 and 3 reduces the potential sample of individuals to 6,216; statistics on this group are shown in column 2. For both columns 1 and 2, the number of observations available for each variable is shown in italics. Missing data on covariates reduce the sample further to 5,123, while our restrictions reduce it to 4,577. Summary statistics for this final sample are provided in column 3. There is some evidence that those in our final sample have higher prior achievement scores and perform better on the non-cognitive tests than the initial sample. However, for the other variables the differences across the three columns are small.

Table 2 reports summary statistics at the school level. The number of schools represented in the dataset falls from 627 in Wave 2 to 494 in Wave 3 because of sample attrition. Following our restrictions, we construct value-added estimates for 388 of these

Table 1. Individual summary statistics.

	Wave 1	All Waves	Final Sample
	Mean *N*	Mean *N*	Mean (Standard Deviation)
JC Overall Score		72.65	73.24
		5888	(7.43)
Reading Score – age 9	0.02	0.05	0.10
	8356	*5917*	(0.99)
Maths score – age 9	−0.75	−0.71	−0.67
	8449	5977	(0.92)
SDQ Score (Wave 2)	-	7.04	6.5
		6038	(4.94)
Attention Span – medium	0.34	0.33	0.33
	8226	*5812*	(0.47)
Attention Span – high	0.52	0.53	0.56
	8226	*5812*	(0.50)
Male	0.51	0.51	0.50
	8568	*6039*	(0.50)
Irish	0.91	0.92	0.93
	8560	*6032*	(0.25)
Family Equivalised Income (Wave 2)	-	16,121	16,662
		5610	(10,352)
Parent Degree or Higher (Wave 2)	-	0.30	0.33
		6017	(0.47)
Father Not Present in Household (Wave 2)	-	0.25	0.23
		5986	(0.42)
			N 4577

SDQ score refers to the Strengths and Difficulties Questionnaire.

Table 2. School summary statistics.

	Mean (Standard Deviation)
Fee-Paying	0.09
	(0.28)
Non-DEIS Secondary	0.51
	(0.50)
Non-DEIS Vocational	0.28
	(0.45)
DEIS	0.13
	(0.34)
Single-Sex Boys	0.20
	(0.40)
Single-Sex Girls	0.24
	(0.43)
Medium Size (400–700)	0.52
	(0.50)
Large Size (700+)	0.29
	(0.45)
Streaming	0.18
	(0.39)
Class Tutors	0.98
	(0.14)
Student Mentors	0.86
	(0.35)
Study Skills	0.77
	(0.42)
Female Principal	0.39
	(0.49)
Age Principal (40–49)	0.39
	(0.49)
Age Principal (50–59)	0.45
	(0.50)
Age Principal (60+)	0.09
	(0.29)
Experience of Principal	7.02
	(6.03)
Emotional Problems	0.21
	(0.41)
Average Daily Attendance	89.16
	(16.18)
Local Unemployment Rate	4.87
	(1.64)
N	337

DEIS refers to schools with a high proportion of educationally disadvantaged students.

schools, corresponding to approximately 54% of all second-level schools in Ireland. However, when reporting summary statistics, we restrict the sample to the 337 schools with complete information on the variables determining value added. These variables describe the characteristics and practices of each school, including the experience level, age and gender of the principal, whether the school streams students based on ability and whether the school offers study skills programmes, student mentors and class tutors. In addition, we use the total number of students enrolled in the school, the daily attendance rate and the proportion of students with emotional problems. We also use the average unemployment rate for the school based on the students' detailed local area.

Of particular interest in our analysis is school type, which has four categories: fee-paying, non-DEIS secondary, non-DEIS vocational/community and DEIS. From Table 2, we see that about 9% of schools are fee-paying and 28% are non-DEIS vocational/community, both figures that are consistent with the national distribution. However, at 13%, DEIS schools are substantially underrepresented in our sample, given that they account for 27% of all schools nationally. This is partly because DEIS schools are underrepresented in the GUI dataset and partly because these schools are typically smaller and therefore less likely to satisfy our inclusion criteria. By the same reasoning, non-fee paying non-DEIS secondary schools are overrepresented in our sample because they tend to be larger. When conducting our analysis, we estimate all the models on both the full sample including DEIS schools and on the smaller sample excluding DEIS schools, to see if their underrepresentation affects our findings. We find no evidence of this.

5. Results

As a preliminary analysis, we assess the relative importance of within- and between-school variation in JC scores. The total variance of scores across all students is 55.25, whereas when we assign to each student the school average test score, the variance is 12.84. Therefore 23.3% of the total variance in overall JC scores is accounted for by between-school variation. Smyth (1999) obtains a similar result (22%) in an Irish data set with fewer schools but more observations within each school. Previous international research also typically finds within-school variation accounting for the majority of the variance in test scores (OECD, 2004).

5.1. Measuring value added

Figure 1 presents the average JC score in each school, together with its 95% confidence interval, ranking the schools from lowest to highest average score. Although it is difficult to distinguish between schools in the middle of the rankings, we can distinguish between

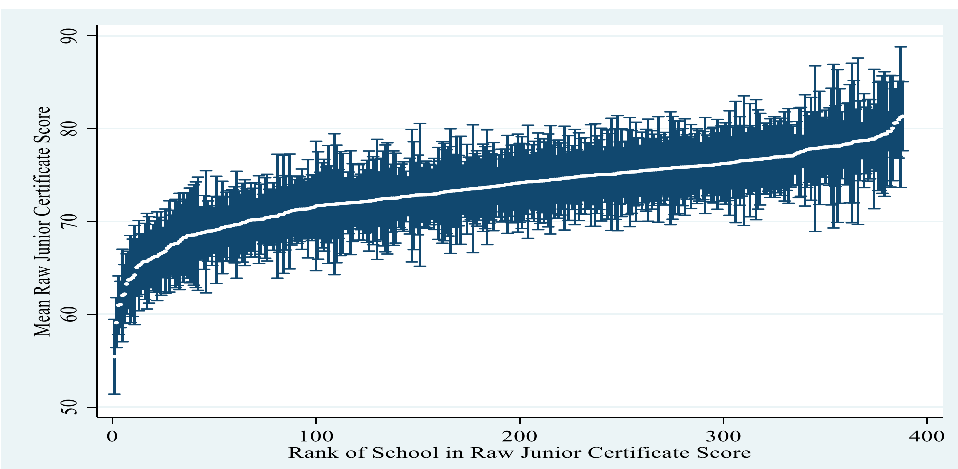

Figure 1. School mean raw JC scores.

high- and low-performing schools. The bottom 15% of schools, whose upper bound is approximately 74, can be distinguished from the top 15%, whose lower bound exceeds 74.

However, if children with higher prior achievement or from a higher socioeconomic background cluster in particular schools, then it is difficult to distinguish good schools from schools with good students. To examine this selectivity issue, we group schools into quartiles based on average JC scores and examine the distribution of student characteristics for each quartile. For prior maths and reading scores, equivalised household income and SDQ scores, the results are plotted in Figure 2. For all of these variables, the distributions for schools in the top quartile of JC scores are furthest to the right, followed in order by each of the other quartiles. The same pattern applies for the discrete variables used. The proportion of parents with a university education is 33% overall, but 48% (21%) for students attending schools in the top (bottom) quartile. Likewise, the proportion of students living in households without their father is 23% overall, but 17% (31%) for students in schools in the top (bottom) quartile. These clear differences across schools illustrate the importance of controlling for compositional differences when assessing school effectiveness.

The results from the value-added model are given in Table 3. All the estimates are statistically significant with the expected signs. When interpreting the magnitude of these coefficients, it is useful to recall that the standard deviation of JC scores in our sample is approximately 7.4. The point estimates on prior achievement indicate that a one unit (one standard deviation) increase in a student's prior reading score is associated with a 0.28

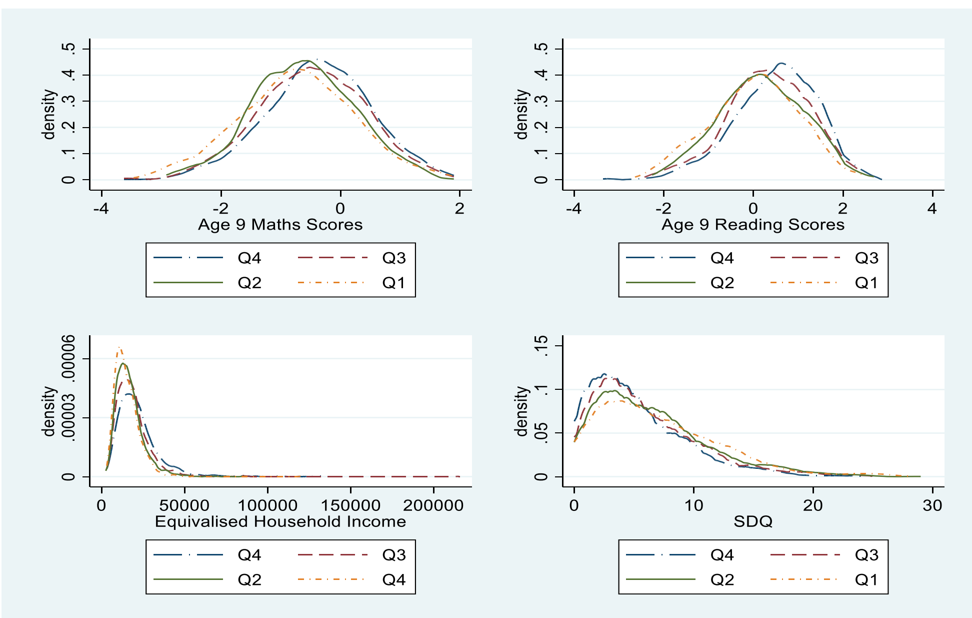

Figure 2. Distribution of characteristics by school raw JC ranking (Quartile 1 to Quartile 4).

Table 3. Value added regressions: dependent variable is JC score.

Reading Score (age 9)	2.09***
	(0.11)
Maths score (age 9)	1.44***
	(0.12)
SDQ Score	−0.25***
	(0.02)
Attention Span – medium	0.78***
	(0.29)
Attention Span – high	2.61***
	(0.29)
Male	−0.44**
	(0.23)
Irish	−0.58*
	(0.33)
Family Equivalised Income/1000	0.05***
	(0.009)
Parent Degree or Higher	1.59***
	(0.19)
Father Not Present in Household	−1.33***
	(0.20)
N	4577

School fixed effects included. Standard errors in parentheses. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

standard deviation increase in their JC score, while a one unit increase in the prior maths score increases the JC score by 0.20 standard deviations. The other results in the table can be similarly interpreted.

We use the fixed effects estimated from the above model in calculating the value-added measure $\widehat{Y_s}$. To put the value-added scores on a similar scale to the raw scores, we estimate $\widehat{Y_s}$ for boys with average values of prior achievement scores, equivalised income and SDQ scores, with a high level of persistence and whose parents are Irish, have a university degree and are both present in the household. However, these baseline choices are simply normalisations that do not affect our analysis.

5.2. *Model validity*

In this section we consider the validity of our value-added estimation. While we include a rich set of control variables in our model, these may still not adequately control for selection. Although we do not have access to experimental variation to assess the extent of any remaining selection bias (e.g. Angrist et al., 2017), we can examine the validity of our approach by 'checking for balance' in characteristics not used in the value-added model. The intuition behind the test is that if high value-added schools are found to 'affect' predetermined student characteristics then this is almost certainly an indication of remaining selection bias. While there is ongoing debate on the use of balance tests of value-added models (Chetty et al., 2016), they are still widely used.

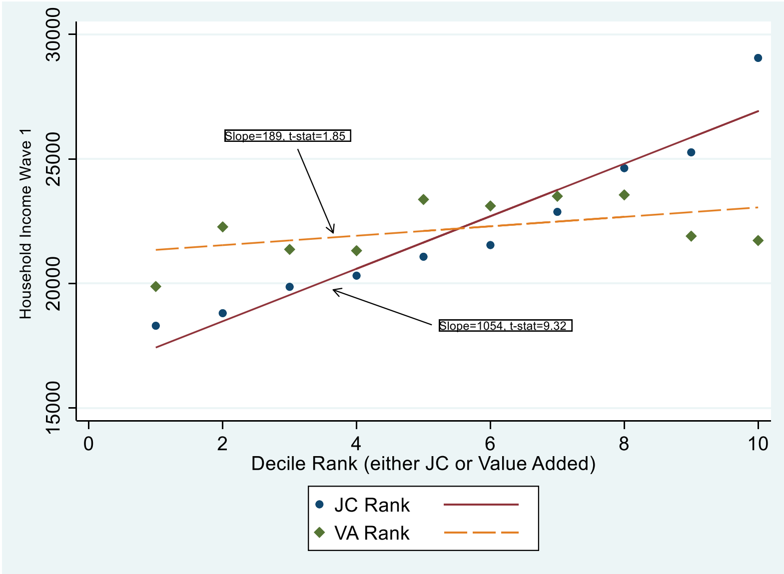

Figure 3. Test for balance in Wave 1 household income.

The balance test that we apply uses Wave 1 family income as a 'held out' variable. Parental income from Wave 2 (when the children are aged 13) is included in our value-added specification, but parental income from Wave 1 (when aged 9) is not. The correlation between parental income at Waves 1 and 2 is 0.56; the fact that it is substantially less than one leaves open the possibility that Wave 1 income may capture some characteristics that determine selection into schools. If our existing controls are effectively capturing selection however, we should see no relationship between value added and Wave 1 parental income.

To implement this test, we group schools into deciles based on estimated school value added and calculate average Wave 1 parental income in each decile. For comparison, we also construct average Wave 1 parental income by decile of raw JC scores. Both of these are plotted in Figure 3. Average parental incomes across raw JC deciles are represented by the circles, while the averages across value-added deciles are represented by the diamonds. The solid and dashed lines represent the associated fitted least squares lines.

The clear positive relationship between raw JC scores and Wave 1 income confirms the selection documented in Figure 2; students in schools that perform well in terms of raw JC results are more likely to come from high income families. In contrast, there is no evidence of a relationship when schools are ranked based on their value-added scores. This substantial improvement in balance in Wave 1 income when moving from raw scores to value-added scores supports our approach.

5.3. Value added across schools

Figure 4 shows the estimated value-added score in each school, ranked from lowest to highest. The graph in Figure 4 is notably flatter than that for raw scores in Figure 1. To verify this, we fit a non-parametric kernel-based local-linear regression through the data

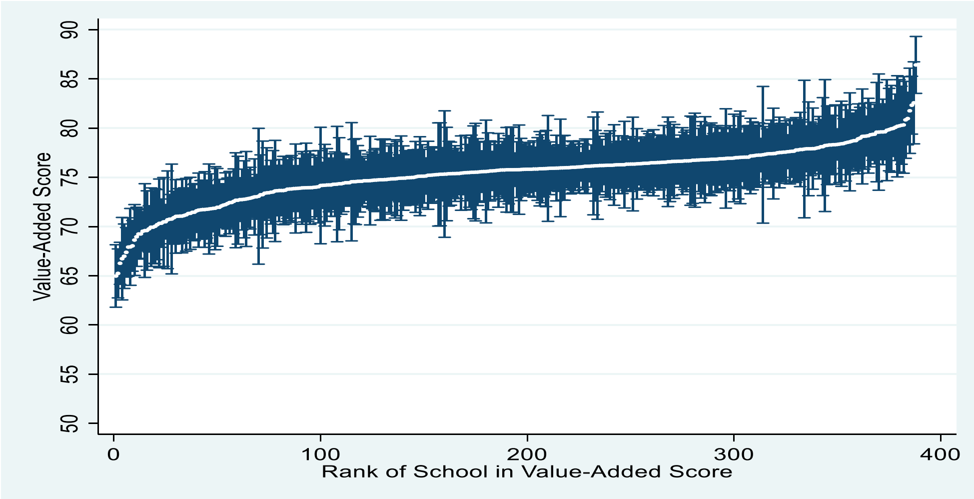

Figure 4. School value-added scores.

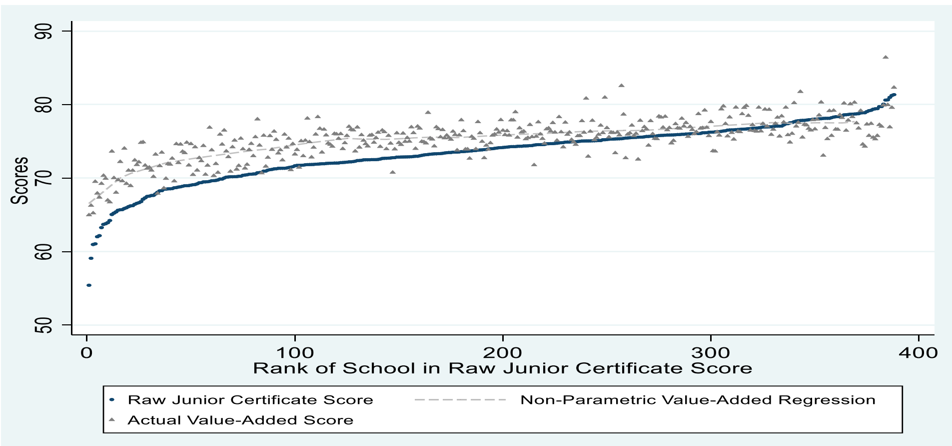

Figure 5. Raw and value-added scores ranked by raw scores.

given in Figures 1 and 4. The average estimated derivative of this regression line for Figure 4 is half that of Figure 1, with the difference in slopes particularly pronounced at the lower end. This reduction in slope translates into a 28% reduction in the standard deviation of value-added scores relative to raw JC scores.

Despite the reduced variance, there is still evidence in Figure 4 of high and low value-added schools. As with raw scores, we can distinguish between the bottom and top 15% of schools, based on a value-added threshold of approximately 75. However, schools that perform well in raw JC scores (Figure 1) are not necessarily the same schools that perform well in value-added terms (Figure 4). To examine this, Figure 5 plots both a school's value-added and its average raw score against its ranking in raw scores. The value-added scores are given by the triangles, while the dark dots indicate raw performance. In general, schools that have the lowest raw scores perform better in value-added terms and the top-

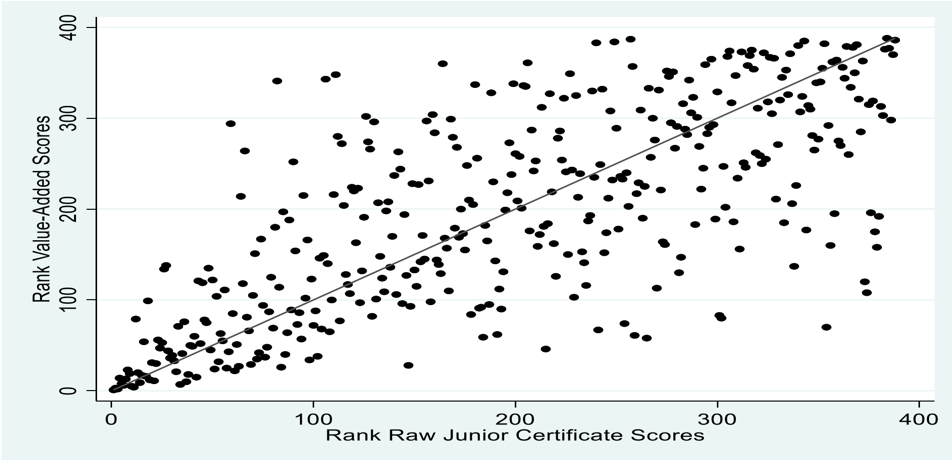

Figure 6. Scatterplot of raw rank and value-added rank.

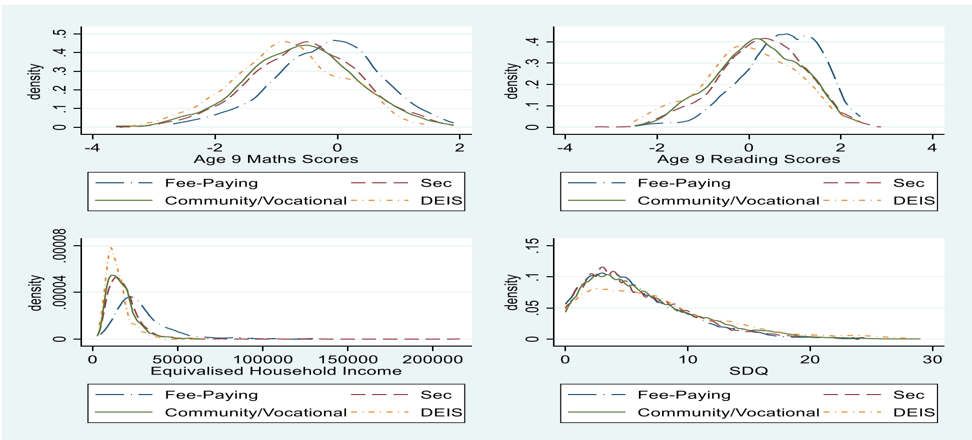

Figure 7. Distribution of characteristics by school type.

ranking schools perform worse. To illustrate this, we add a non-parametric regression of value added on raw rank (the light dashed line), which is clearly flatter than the line of dark dots.

These graphs indicate a substantial reranking of schools when we switch from raw scores to value-added scores. Figure 6 illustrates this reranking using a scatterplot of the raw and value-added ranks. Schools above the 45-degree line improve their ranking when selection is accounted for, while those below disimprove. The rank correlation between raw and value-added scores is 0.72. The Kendall tau rank distance between the two ranking lists is 0.23, implying that 23% of all pair comparisons differ in ordering between the raw and value-added measures.

Anecdotal evidence suggests that school choice is often based on school type (fee-paying, secondary, community/vocational and DEIS schools). To examine the extent to which student characteristics vary by school type, we plot the prior maths and reading

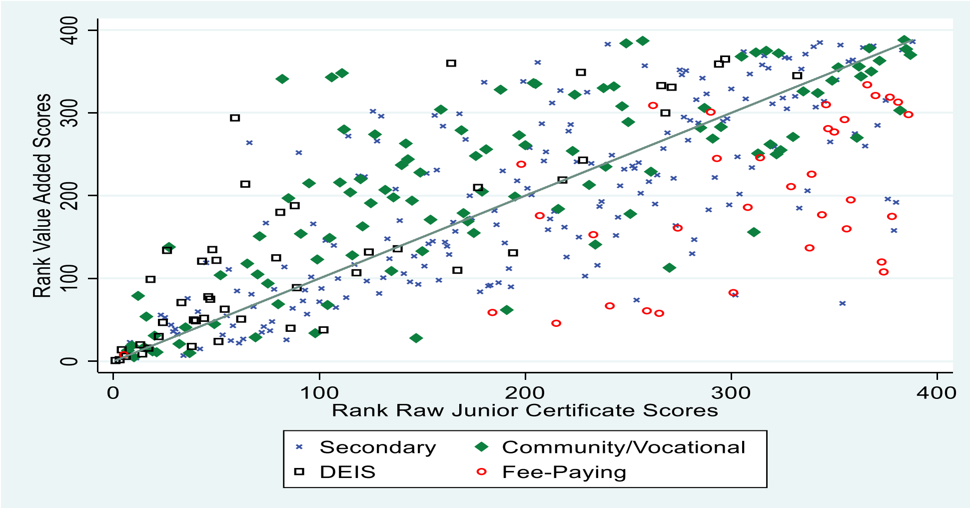

Figure 8. Scatterplot of raw rank and value-added rank by school type.

scores, household income and non-cognitive scores by school type. These are given in Figure 7, which show clear differences across school types in terms of student intake. In order to consider whether the reranking discussed above depends on school type, Figure 8 reproduces Figure 6, including identifiers for school type. As fee-paying schools are predominantly located below the diagonal line and DEIS schools above, it is apparent that the reranking of schools is not independent of school type. However, it is also evident that there is a lot of variation in value-added performance within each school type.

5.4. A comparison of a non-contextual and contextual value-added approach

In this section we examine whether the additional student-level covariates that we include in our CVA matters for the overall ranking of schools. As well as prior achievement measures, our model includes family income, nationality, gender, non-cognitive skills, parental education and family structure. To examine the impact of including these additional background variables, we compare our results with those obtained from a restricted version of Equation (1) using only prior performance. This represents the NCVA model discussed in previous work (Timmermans et al., 2011).

The correlation in rankings between our NCVA and CVA models is 0.89. This is similar to findings from other countries. Leckie and Goldstein (2019) find a correlation of 0.89 between their Progress 8 and Adjusted Progress 8 ranks in English schools; Timmermans et al. (2011) find a correlation of 0.82 between CVA and NCVA models in Dutch schools; and Marks (2021) finds correlations in the range of 0.64 to 0.84 in Australian secondary schools.

Despite the relatively high correlations, the differences between the models can have large implications for schools. This is also evident in our results. Figure 9 plots a school's rank in NCVA against its rank in CVA, distinguishing by school types. Schools below the diagonal have lower rank when the additional student-level variables are included in the

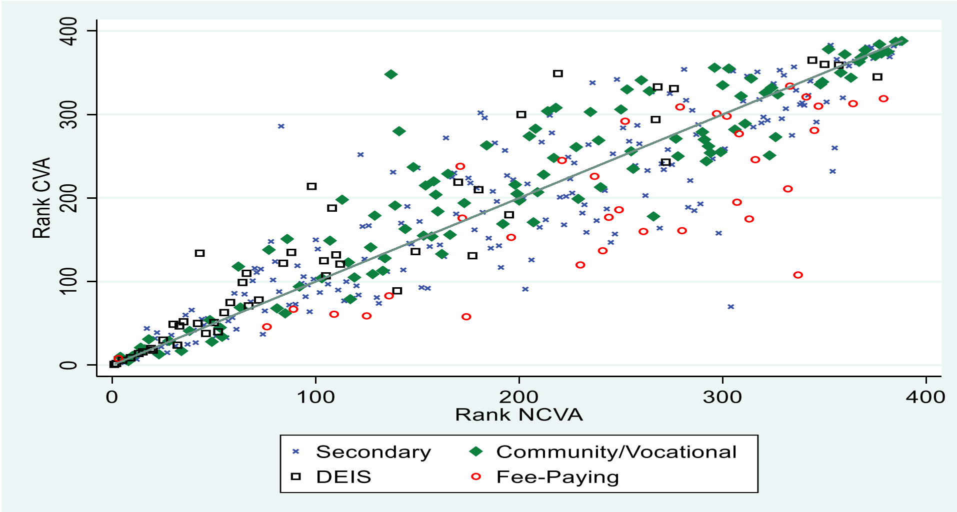

Figure 9. Scatterplot of CVA v NCVA rankings by school type.

value-added model. Changing from an NCVA to a CVA approach would lead to 25% of schools in our sample moving up or down the school rankings by 50 or more ranks, with 6.4% moving over 100 ranks. Furthermore, the change in rankings when moving from NCVA to CVA varies systematically with school type. Figure 9 shows that controlling only for prior attainment favours fee-paying schools. The average ranking of fee-paying schools fell from the 98th percentile in the raw rankings to the 80th percentile in the NCVA rankings and fell further to the 63rd percentile in the CVA rankings.

When we apply the balance test discussed in Section 2.2 to the NCVA model, we reject the null hypothesis of no relationship between Wave 1 parental income and value added, with a p-value close to zero. This provides evidence that prior achievement is not sufficient to remove all the selection bias in the raw scores and illustrate the importance of the CVA approach.

5.5. Relationship between school characteristics and value added

In this section, we examine the school characteristics that are related to estimated school value added, with results shown in Table 4. We regress both the school average raw JC scores (column 1) and the school value-added estimates (column 2) on the school-level variables discussed in Section 4. The results in column 1 show that on average, fee-paying schools are ranked about 89 places ahead of secondary schools and about 163 places above DEIS schools; school type clearly matters for raw scores. Schools with a female principal tend to be ranked higher (by about 29 places), while schools with older principals tend to be ranked lower. However, conditional on age, experience as principal in the school is positively associated with raw JC scores. Furthermore, schools with students from high unemployment areas rank significantly lower. None of the other regressors included in column 1 has a significant effect on raw scores.

Table 4. Determinants of school raw rank and value-added rank.

	(1)	(2)
	JC Raw Rank	JC Value-Added Rank
Constant	376.90***	216.05***
	(62.46)	(70.20)
Non-DEIS Secondary	−89.41***	14.53
	(20.32)	(22.83)
Non-DEIS Community/Comprehensive	−106.24***	36.73
	(23.59)	(26.51)
DEIS	−162.93***	−30.68
	(26.90)	(30.23)
Single-Sex Boys	0.55	0.07
	(17.19)	(19.32)
Single-Sex Girls	12.89	−3.46
	(16.85)	(18.94)
Medium Size (400–700)	−21.62	−20.45
	(14.31)	(16.09)
Large Size (700+)	3.56	−25.56
	(16.05)	(18.05)
Streaming	−17.89	−14.79
	(13.93)	(15.66)
Class Tutors	12.58	−37.88
	(37.55)	(42.20)
Student Mentors	−13.15	18.00
	(−15.20)	(17.09)
Study Skills	−8.64	3.06
	(12.90)	(14.50)
Female Principal	29.07**	39.25***
	(12.59)	(14.15)
Age of Principal (40–49)	−26.31	2.74
	(22.19)	(24.95)
Age of Principal (50–59)	−35.93	−6.69
	(22.52)	(25.31)
Age of Principal (60+)	−55.50*	−16.62
	(29.11)	(32.72)
Experience of Principal	1.79*	1.97*
	(0.99)	(1.12)
Emotional Problems	−18.26	−5.23
	(13.76)	(15.46)
Average Daily Attendance	0.35	0.58
	(0.33)	(0.38)
Local Unemployment Rate	−18.33***	−14.17***
	(3.44)	(3.87)
N Schools	337	337

Standard errors in parentheses. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Looking at the characteristics associated with value added in column 2, the dominance of fee-paying schools is no longer evident. While some fee-paying schools continue to rank highly on measures of value added (see Figure 8), fee-paying schools as a group perform no better than other second-level schools. These findings are in keeping with the reranking results reported earlier and with the work of Cullinan et al. (2021), who found no effect of attending a fee-paying school in Ireland when family background is controlled for. In contrast, research on England by Malacova (2007) and Henderson et al. (2020) found a modest statistically positive effect even after controlling for background characteristics and prior achievement. The difference between the results for Ireland and England possibly reflects the much higher resources and greater autonomy afforded to fee-paying schools in the English education system.

Looking at the other variables, schools with students from low unemployment areas have a higher value-added score, possibly due to local peer effects or fewer resources in these schools. Schools with female principals also have higher value added, possibly due to more effective leadership styles, evidence of which is reported in Bloom et al. (2015) and Hallinger et al. (2016). Alternatively, the positive association may reflect a selection effect, rather than a causal relationship. If only particularly able women go forward for principalships or more progressive schools hire female principals, then we would observe this positive effect even in the absence of a causal relationship.[5]

Although we find differences in value added across schools, it is striking how few of the included variables are significant in determining a school's value added. This is consistent with recent work by Masci et al. (2018), who consider information about school management, teaching staff, assessment and streaming in their analysis. They find no relationship between these measures and school value added. This suggests that the factors associated with value added are idiosyncratic and may concern good management practices that are difficult to measure.

5.6. *The impact of school choice*

In this section, we examine the extent to which school choice matters to a child's performance, by carrying out a counterfactual simulation where we reassign each student to the best value-added school accessible to that student. Because there is no information on the address of either the child or the school, we identify the second-level schools accessible to student *i* as the set of schools attended by any classmate of student *i* in their primary school. For some students (9.6% of our sample) there is effectively no choice, as all the sample children in their primary school attend the same second-level school. However, for the remaining children, there is choice: from any given primary school in Wave 1, children attend up to 16 second-level schools in Wave 2, with a median of 5. For each child, we rank each accessible school by value-added score and calculate the child's predicted score for each of these schools using his/her own characteristics. We then examine the effect on each child's score of moving them to the best available school. In doing this, we take into account that boys may not attend single-sex girls' schools and *vice versa*; we also assume that any student not currently attending a fee-paying school cannot afford to do so, meaning that only students already attending fee-paying schools are assigned to schools in this category. However, we do assume that all pupils attending a given school benefit to the same extent, so schools cannot be differentially effective for different students (Strand, 2010). Our sample size is too small to model the interactions that would be needed to account for heterogeneous student effects. Although our reassignment exercise abstracts from differential effects, peer effects and reallocation costs, this thought experiment provides a useful indication of the potential effects of policy decisions based on value added.

Our results indicate that the effect of each child attending the best available school is to raise average scores by 1.11 points or 0.22 standard deviations. This effect is substantial when placed in the context of the value-added results presented earlier; reallocation to the best available school is worth 70% of the effect of having a parent with a higher education, or €22,000 extra equivalised family income. Our estimate is also consistent with international research. Angrist et al. (2017) find that a simulation that involved closing the lowest-ranked district school in the United States and sending the students to schools with average value added boosts achievement by 0.37 standard deviations, while Schiltz et al. (2018) find that the same policy in Italy raises achievement by 0.16 standard deviations. Our results indicate that enabling parents to choose on the basis of value added could have substantial effects. To the extent that academic success is one of the criteria that parents consider when choosing schools, the results suggest that information on value added could greatly benefit their decision making.

6. Conclusion

When students non-randomly select into schools according to prior achievement or socioeconomic factors, it can be difficult for parents to identify good schools. This is especially true in countries, like Ireland, where the information available to parents is very limited. In this paper we use a rich data set with detailed information on prior achievement, non-cognitive skills, family income, parental education and family structure. We show that these characteristics are strongly correlated with academic achievement and vary substantially across schools. As a result, controlling for selection substantially changes the rankings of schools.

School type is important in determining raw ranks, with fee-paying schools over-represented among the highest performing schools in raw terms. However, this is not true once we account for compositional differences. Indeed, our analysis suggests no significant differences between school types in value-added terms. Schools with female principals rank highly in terms of both raw and value-added performance. However, in general the factors determining value added appear to be related to unobserved individual school practices, rather than systematic school policies. Identifying these school practices could have important policy implications.

Our results suggest that reallocating students to schools based on value added has the potential to substantially boost student outcomes. Under our simplifying assumptions, the simulated effect of all children attending their best available school is almost as important as having a parent with a university degree. Providing information on the value added of schools could therefore be a low-cost way to assist parents in choosing a school.

Notes

1. Until 2009, teachers at fee-paying schools were funded at the same pupil-teacher ratio as other schools, but since then, the pupil-teacher ratio at fee-paying schools has been reduced so that it is about 20% lower.
2. For details see https://www.irishtimes.com/news/education/the-rise-of-private-schools-there-s-a-lot-more-money-around-1.3338945.
3. For more details see https://www.growingup.ie/data-documentation/.
4. We also conducted the analysis using a broader set (best 10) and a narrower set (only English and Maths, which are taken by all students) and found very similar results.
5. We interacted the gender of the principal with the single-sex status of the school and found no significant interaction effects.

Acknowledgments

We thank Greg Connor, John Cullinan, Kevin Denny, Darragh Flannery and Irene Mosca for helpful comments.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This paper uses RMF data provided by the Central Statistics Office (CSO). Growing Up in Ireland (GUI) is funded by the Department of Children and Youth Affairs (DCYA). It is managed by DCYA in association with the CSO. Neither the CSO nor DCYA take any responsibility for the views expressed or the outputs generated from these analyses.

Notes on contributors

Aedín Doris is a Lecturer in the Department of Economics, Maynooth University and obtained her PhD at the European University Institute in Florence, Italy. She is an applied labour economist doing research mainly in the areas of earnings inequality, the economics of education and labour market public policy. Her research addresses policy-relevant questions using empirical analysis, typically of Irish data. She has research interests in the determinants of earnings and earnings inequality; educational inequality; gender inequality in the labour market; the labour market effects of benefit systems; and higher education funding and access. She has published in leading international journals such as the Journal of Human Resources, Economica, the Economics of Education Review, the Journal of Economic Inequality and the IZA Journal of European Labor Studies. She is Managing Editor of the Economic and Social Review, Ireland's primary Economics journal. She has been a member of the government's Pensions Commission and is a member of the External Advisory Group for IGEES (the Irish Government Economic and Evaluation Service).

Donal O'Neill is a professor of Economics at Maynooth University, a Research Fellow at the IZA Institute for the Study of Labor, Bonn and past President of of the Irish Economic Association. He obtained his Ph.D from the University of Iowa and in 2000 was awarded the Barrington Medal for Economics in recognition of his contribution to research in the Social Sciences in Ireland. His research focuses on the areas of labour economics, the economics of education and econometrics. He has published extensively in leading international economics and statistics journals including the American Statistician, the European Journal of Health Economics, Economic Journal, Journal of Economic Inequality, Journal of Human Resources, Journal of Political Economy, Journal of Labor Economics, Labour Economics, Oxford Economic Papers, the Oxford Bulletin of Economics and Statistics, the Economics of Education Review and the IZA Journal of Labor Economics. He served as labour market expert on the inaugural Low Pay Commission of Ireland, a statutory body whose principal function is, on an annual basis, to examine and make recommendations to government on the appropriate level of the national minimum wage and related matters.

Olive Sweetman is a lecturer in the Department of Economics at Maynooth University. She obtained her Ph.D from the University of Iowa. Her main interests are in the areas of labour economics and the economics of education. She has published articles examining the gender and racial wage gap, single-sex education, inequality and intergenerational mobility. Her research has been published in leading international economics journals including the Journal of Economic Inequality, Journal of Human Resources, Labour Economics, the Oxford Bulletin of Economics and Statistics, the Economics of Education Review and the IZA Journal of Labor Economics. She has served as a member of the board of Combat Poverty and as editor of the Economic and Social Review.

ORCID

Donal O'Neill http://orcid.org/0000-0001-5516-2603

References

Angrist, J., Hull, P., Pathak, P., & Walters, C. (2017). Leveraging lotteries for school value-added: Testing and estimation. *Quarterly Journal of Economics*, *132*(2), 871–919. https://doi.org/10.1093/qje/qjx001

Belfield, C., Britton, J., Buscha, F., Dearden, L., Dickson, M., van der Erve, L., Sibieta, L., Vignoles, A., Walker, I., & Zhu, Y. (2018). *The relative labour market returns to different degrees* (IFS Research Report). Department of Education, U.K.

Bloom, N., Lemos, R., Sadun, R., & Van Reenen, J. (2015). Does management matter in schools? *The Economic Journal*, *125*(584), 647–674. https://doi.org/10.1111/ecoj.12267

Chetty, R., Friedman, J., & Rockoff, J. (2016). Using lagged outcomes to evaluate bias in value-added models [Article|Proceedings paper]. *American Economic Review*, *106*(5), 393–399. https://doi.org/10.1257/aer.p20161081

Cullinan, J., Denny, K., & Flannery, D. (2021). A distributional analysis of upper secondary school performance. *Empirical Economics*, *60*, 1085–1113 https://doi.org/10.1007/s00181-019-01756-8 .

Cunha, J., & Miller, T. (2014). Measuring value-added in higher education: Possibilities and limitations in the use of administrative data. *Economics of Education Review*, *42*, 64–77. https://doi.org/10.1016/j.econedurev.2014.06.001

Darmody, M., & Smyth, E. (2013). (ESRI Research Paper No. 34 (Dublin: Economic and Social Research Institute)).

Dearden, L., Miranda, A., & Rabe-Hesketh, S. (2011). Measuring school value added with administrative data: The problem of missing variables. *Fiscal Studies*, *32*(2), 263–278. https://doi.org/10.1111/j.1475-5890.2011.00136.x

Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. J. (2014). The sensitivity of value-added estimates to specification adjustments: Evidence from school- and teacher-level models in Missouri. *Statistics and Public Policy*, *1*(1), 19–27. https://doi.org/10.1080/2330443X.2013.856152

Eyles, A., Hupkau, C., & Machin, S. (2016). Academies, charter and free schools: Do new school types deliver better outcomes? *Economic Policy*, *31*(87), 453–501. https://doi.org/10.1093/epolic/eiw006

Guarino, C., Maxfield, M., Reckase, M., Thompson, P., & Wooldridge, J. (2015). An evaluation of empirical Bayes's estimation of value-added teacher performance measures. *Journal of Educational and Behavioral Statistics*, *40*(2), 190–222. https://doi.org/10.3102/1076998615574771

Hallinger, P., Dongyu, L., & Wang, W.-C. (2016). Gender differences in instructional leadership: A meta-analytic review of studies using the principal instructional management rating scale. *Educational Adminstration Quarterly*, *52*(4), 567–601. https://journals.sagepub.com/doi/10.1177/0013161X16638430

Hannan, D., & Boyle, M. (1987). (ESRI Working Paper No. 136 (Dublin: Economic and Social Research Institute)).

Henderson, M., Anders, J., Green, F., & Henseke, G. (2020). Private schooling, subject choice, upper secondary attainment and progression to university. *Oxford Review of Education*, *46*(3), 295–312. https://doi.org/10.1080/03054985.2019.1669551

Jacob, B., & Lefgren, L. (2008). Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of Labor Economics*, *26*(1), 101–136. https://doi.org/10.1086/522974

Kane, T., & Staiger, D. (2002). The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives*, *16*(4), 91–114. https://doi.org/10.1257/089533002320950993

Koedel, C., & Betts, J. (2010). Value added to what? How a ceiling in the testing instrument influences value-added estimation. *Education Finance and Policy*, *5*(1), 54–81. https://doi.org/10.1162/edfp.2009.5.1.5104

Ladd, H. F., & Walsh, R. P. (2002). Implementing value-added measures of school effectiveness: getting the incentives right. *Economics of Education Review*, *21*(1), 1–17. https://doi.org/10.1016/S0272-7757(00)00039-X

Leckie, G., & Goldstein, H. (2017). The evolution of school league tables in England 1992–2016: 'Contextual Value-Added', 'Expected Progress' and 'Progress 8'. *British Educational Research Journal*, *43*(2), 193–212. https://doi.org/10.1002/berj.3264

Leckie, G., & Goldstein, H. (2019). The importance of adjusting for pupil background in school value-added models: A study of Progress 8 and school accountability in England. *British Educational Research Journal*, *45*(3), 518–537. https://doi.org/10.1002/berj.3511

Malacova, E. (2007). Effect of single-sex education on progress in GCSE. *Oxford Review of Education*, *33*(2), 233–259. https://doi.org/10.1080/03054980701324610

Marks, G. N. (2021). Should value-added school effects models include student- and school-level covariates? Evidence from Australian population assessment data. *British Educational Research Journal*, *47*(1), 181–204. https://doi.org/10.1002/berj.3684

Masci, C., Johnes, G., & Agasisti, T. (2018). Student and school performance across countries: A machine learning approach. *European Journal of Operational Research*, *269*(3), 1072–1085. https://doi.org/10.1016/j.ejor.2018.02.031

Meyer, R. (1997). Value-added indicators of school performance: A primer. *Economics of Education Review*, *16*(3), 283–301. https://doi.org/10.1016/S0272-7757(96)00081-7

OECD. (2004). *Learning for tomorrow's world: First results from PISA 2003*. https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/34002216.pdf

Pfeffermann, D., & Landsman, V. (2011). Are private schools better than public schools? Appraisal for Ireland by methods for observational studies. *The Annals of Applied Statistics*, *5*(3), 1726–1751. https://doi.org/10.1214/11-AOAS456

Prior, L., Jerrim, J., Thomson, D., & Leckie, G. (2021). A review and evaluation of secondary school accountability in England: Statistical strengths, weaknesses and challenges for 'Progress 8' raised by COVID-19. *Review of Education*, *9*(3), e3299. https://doi.org/10.1002/rev3.3299

Raudenbush, S. W., & Willms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, *20*(4), 307–335. https://doi.org/10.3102/10769986020004307

Schiltz, F., Sestito, P., Agasisti, T., & De Witte, K. (2018). The added value of more accurate predictions for school rankings. *Economics of Education Review*, *67*, 207–215. https://doi.org/10.1016/j.econedurev.2018.10.011

Smyth, Emer 1999 Do Schools Differ? (Dublin: Oaktree Press)

Sofroniou, N., Shiel, G., & Cosgrove, J. (2000). Explaining performance on the JC examination: A multilevel approach. *The Irish Journal of Education/Iris Eireannach an Oideachais*, *31*, 25–49. http://www.jstor.org/stable/30077469

Strand, S. (2010). Do some schools narrow the gap? Differential school effectiveness by ethnicity, gender, poverty, and prior achievement. *School Effectiveness and School Improvement*, *21*(3), 289–314. https://doi.org/10.1080/09243451003732651

Timmermans, A. C., Doolaard, S., & De Wolf, I. (2011). Conceptual and empirical differences among various value-added models for accountability. *School Effectiveness and School Improvement*, *22*(4), 393–413. https://doi.org/10.1080/09243453.2011.590704

Weir, S., & Kavanagh, L. (2018). *The evaluation of DEIS at post-primary level: Closing the achievement and attainment gaps*. Educational Research Centre. https://www.education.ie/en/Schools-Colleges/Services/DEIS-Delivering-Equality-of-Opportunity-in-Schools-/the-evaluation-of-deis-at-post-primary-level-closing-the-achievement-and-attainment-gaps.pdf

Wilson, D., Croxson, B., & Atkinson, A. (2006). What gets measured gets done. *Policy Studies*, *27*(2), 153–171. https://doi.org/10.1080/01442870600637995

Appendix A. Empirical Bayes (EB) estimates of value added

Figure A1 plots the rankings of our reported measure against rankings based on the EB estimator. The correlation between the reported estimates and the EB estimates is 0.98. The Kendall tau rank distance between the two ranking lists is 0.03. This implies that only 3% of all pair comparisons differ in ordering between the two value-added measures.

Furthermore, the use of the EB estimates in the subsequent analysis of the determinants of value added did not change our findings. Based on this we are confident that the results reported in the main body of the paper are robust to corrections for noise and small samples.

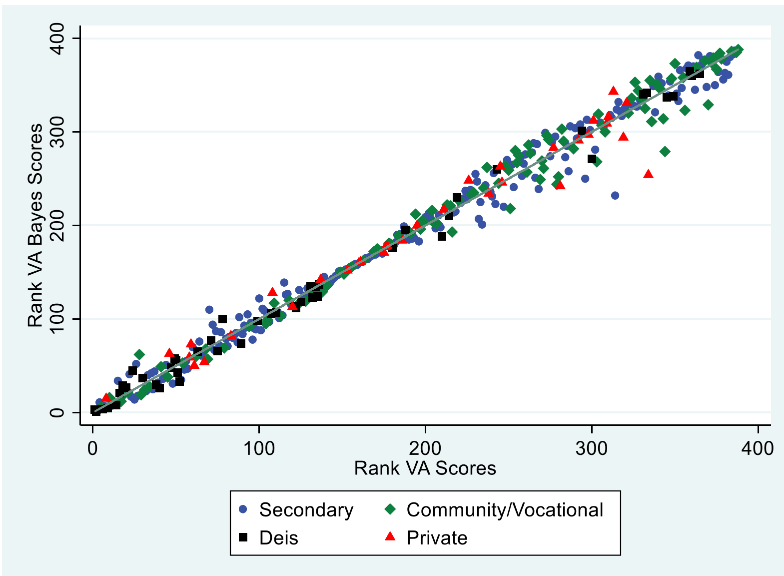

Figure A1. Scatterplot of unadjusted and empirical Bayes value-added ranks by school type.

Appendix B. Conversion of JC grades to individual performance scores

The conversion of JC letter grades to a numerical scale was constructed using the following conversion rates (see also Sofroniou et al., 2000).

Higher	Ordinary	Foundation	Score
A			12
B			11
C			10
D	A		9
E	B		8
F	C		7
	D	A	6
	E	B	5
	F	C	4
		D	3
		E	2
		F	1