

School accountability and fairness: Does 'Progress 8' encourage schools to work more equitably?

Executive Summary

Simon Burgess, University of Bristol

Dave Thomson, FFT Education Datalab

This work has been funded by the Nuffield Foundation, but the views expressed are those of the authors and not necessarily the Foundation. Visit www.nuffieldfoundation.org.

This research was conducted by the authors between 2018 and 2020. For further enquiries please contact educationdatalab@fft.org.uk.

Background

School accountability in the form of published “league tables” of performance indicators and routine inspection have been a feature of the education system in England since the early 1990s. This creates incentives, both positive and negative, to which schools respond. For example, changes to the list of qualifications that count towards a school’s published performance indicators will change the nature of qualifications offered by schools.

In 2013, the UK Government announced an important reform to the accountability framework for state-funded schools in England by publishing a new performance indicator: Progress 8. This change was notable as the previous framework was dominated by a threshold measure of pupil attainment, whether or not a pupil achieved 5 or more General Certificates of Secondary Education (GCSEs) at grades A*-C including English and maths (5ACEM).

The reform attempted to address two concerns. Firstly, that the 5ACEM indicator encouraged schools to focus on a particular set of pupils, namely those that were on the grade C/D borderline. Secondly, it was recognised that published statistics on secondary schools’ attainment did not account for differences in attainment on entry to school.

Progress 8 is a value-added measure, summarising attainment at the end of compulsory secondary education (usually at age 16), controlling for attainment at the end of primary education (usually at age 11). Unlike the previous regime, it offered no particular incentive for schools to focus on a narrow segment of the pupil population. Not only that, it was felt to offer a fairer way of comparing the performance of schools. Although measures of value-added had been published for many years, they had not been conferred with the same level of precedence as Progress 8.

Research aims and methodology

We aimed to understand whether the introduction of Progress 8 encouraged schools to work more equitably. By this, we mean whether we find evidence of schools focusing their

efforts less on pupils at the C/D borderline and instead spread their effort more evenly across the full range of attainment.

Administrative data on the attainment of all pupils in state-funded schools in England is available dating back to 2001/02. We attempt to use such data to estimate the effect of the introduction of Progress 8 on the attainment of above-borderline and below-borderline pupils relative to borderline pupils.

However, there are a number of methodological issues to overcome. Firstly, we do not know which pupils schools considered to be at the C/D borderline. Secondly, Progress 8 was just one of a series of policy changes that have occurred to the secondary school accountability framework in England since 2010. Thirdly, improvements in examination results from year to year are controlled by the regulator for qualifications, Ofqual, by a method known as comparable outcomes.

Our analytical approach is shaped by these issues. We therefore use six years of pupil-level data covering all state-funded school pupils in England. This starts in 2011/12, the first year that the Comparable Outcomes policy was applied to GCSE results in English and maths. Progress 8 was first published in 2015/16. This means we have data for four years prior to its introduction and two years since its introduction. We also track changes in pupil attainment using a number of indicators that are reasonably stable in definition and coverage over the period we observe. We also observe changes in the types of qualifications pupils entered.

We adopt a difference-in-difference (d-in-d) approach to isolate the causal effect of the reform on a number of pupil outcomes. That is, we compare outcomes between two groups of pupils ("difference") before and after the policy change ("difference in difference"). We examine, following the reform, changes in the outcomes for the different groups of pupils that theory suggests will be differentially affected. We use a flexible approach to modelling that controls for a wide range of pupil characteristics, and for school factors and other policy shocks.

Key findings

Our results are consistent with the view that some schools had reacted to the previous regime of high implicit incentives for the exam results of students at the GCSE grade C/D borderline. Once that incentive was removed, the borderline group appeared to make less relative progress compared to other groups. The effects are small but not trivial: our headline findings show a post-reform gain of 0.01 standard deviations (SD) in GCSE English and maths for the above-borderline group and 0.06SD for the below-borderline group. This latter effect on GCSE attainment is the same size as that arising from a 1SD increase in school expenditure (Jenkins et al, 2005).

We are, however, cautious in presenting these results noting the issue of trends subsequent to announcement but before implementation. We judge the results to be supportive of the hypothesis but not clinching. We are also aware that the results may have been sensitive to some of our modelling choices. Therefore, we show the effects of making different choices. Our tests of robustness show that these different choices make little substantive difference.

The results also have a bearing on the test score gap between disadvantaged pupils and their peers. Because disadvantaged pupils are disproportionately likely to be in the below-borderline group, and so are more likely to gain from the reform, our findings show a slight post-reform improvement of 0.01 SD.

The change to the accountability framework was far-reaching and had other implications beyond simply test scores. The machinery of school accountability also incentivises schools to enter pupils for particular qualifications. Just as the Government response to the Wolf Review had done two years earlier, the introduction of Progress 8 led to large changes in the types of qualification for which pupils were entered as schools increasingly began to fill the eight qualification 'slots' available in the new accountability measures. In most cases, this was a result of switching away from other types of qualification that were not eligible for inclusion in the accountability measures.

Changes in school behaviour

Schools' responses to the introduction of Progress 8 were varied, although it is difficult to disentangle specific responses to Progress 8 from responses to other changes that happened around the same time, such as reforms to GCSEs. We surveyed over 400 school leaders and teachers in England to find out more about how they responded to the introduction of Progress 8. The results suggest a general shift away from running intervention sessions aimed specifically at borderline pupils towards pupils judged to be falling behind.

Policy Implications

This analysis has a number of implications for policy-makers.

First, and bearing in mind the caveats noted in the report, our results suggest that the introduction of Progress 8 had the intended effect of shifting schools' focus away from students who were borderline to the previous accountability threshold. In that sense, the policy had the intended effect of making schools work more equitably.

Second, this reinforces the view that accountability measures are an effective policy tool. They do not impinge directly on schools' operational autonomy, unlike explicit Ministerial directives, but they do adjust the incentive structure that schools face. This research shows that this can be effective in changing behaviour. The setting, and occasional re-setting, of the accountability framework seems an appropriate role for Government – it is the practical expression of its view of what society deems valuable in education, of what schools 'ought' to do. Problems arise if the framework is changed very frequently so that schools do not have a stable environment for planning.

Problems can also arise if different parts of schools' incentives pull in different directions, and this is the third and final policy message from this study. The previous accountability regime was based on the threshold of achieving 5 or more GCSEs at grades A*-C including English and maths (5ACEM). Schools were strongly incentivised to maximise the number of their pupils that achieved this. This drive meshed well with the goal of the typical pupil because for her, passing that threshold was the key to accessing higher or further education and to the job market. Schools could allocate their resources knowing that the goal of doing well by their pupils and the goal of doing well on the performance metrics were reasonably well aligned. In the new regime, currently, that remains true for pupils but less so for schools. Access to further education and to jobs is still to an extent dominated by achieving at least grade C (now grade 4) passes in GCSE English and maths, and no attention is paid to the achievement of pupils in Progress 8 terms by employers. This may mean that schools are partially conflicted, and that a goal for the school of keeping the 5ACEM "pass rate" high is still important to them. This in turn may partly explain why the impact of the reform on test scores was rather modest. It may be that the labour market and higher education admissions

departments will respond and place more emphasis on Progress 8 scores, or it may be that these two goals for schools will remain in tension.