

Impact case study (REF3b)

Institution: Institute of Education
Unit of Assessment: 25
Title of case study: Pupil performance tables: finding fairer measures
1. Summary of the impact (indicative maximum 100 words)

Educational performance tables – some comparing countries as well as schools – have come to assume great importance. They now influence not only parents' school choices but some national education policies. Tables can, however, mislead as well as enlighten. The three studies featured here demonstrate this and help to ensure that the public will be better informed in future. Two played a key role in convincing the government that it should revise England's school performance tables. The third gave civil servants and politicians good reason to be more circumspect about how they publicly interpret international pupil performance data.

2. Underpinning research (indicative maximum 500 words)
--

The three studies discussed in this case study develop the long line of work on education performance indicators that was started by Harvey Goldstein, the IOE's Professor of Statistical Methods from 1977 to 2005. The quality and impact of these studies also reflect 20 years of IOE involvement in the economics of education and its application to policy and practice.

Study 1 was carried out between 2010 and 2012 by Dr Rebecca Allen, then senior lecturer (now Reader) in the Economics of Education at the IOE, and Professor Simon Burgess, University of Bristol.

Key findings: Two thirds of children attending the highest-performing school available to them in 2003 – rather than picking randomly – benefited from their choice – see references **R1 and R2**. Perhaps surprisingly, raw outcome measures were found to be more useful than tables showing the 'contextual value added' (CVA).

Research methods: The study analysed the consequences of the secondary school choices of more than 500,000 pupils who transferred to more than 3,000 schools in 2004, completing compulsory education in 2009. Pupil characteristics data were linked to attainment at 7, 11 and 16 and each child was compared with similar local pupils who chose different schools. This showed whether picking top local schools paid dividends. Allen did nearly all of the technical work that this study entailed. She developed the specific method of calculating local school choice sets and predicting a child's likely exam performance in all local schools using a regression-based approach. She also conducted all the data analysis.

Study 2 was undertaken by Professors Lorraine Dearden, John Micklewright and Anna Vignoles in 2010-11 (Vignoles moved to Cambridge University in 2012).

Key findings: Between 25% and 40% of schools in England are differentially effective for pupils of differing prior ability – i.e. either good for low-attaining pupils but not for high-attaining children or vice versa (**R3**). An even larger proportion of children is affected as bigger schools are more likely to be differentially effective. Having only a single measure of school performance is therefore often misleading. The study also found that schools that are differentially effective tend to outperform those that are not. This is because the 'added value' they offer is generally at or above the average for *all* prior attainment groups – even though they may be especially effective for just one group.

Methods: The researchers analysed the performance of 1,116,982 state school children who were aged 16 by the end of the 2006-7 and 2007-8 academic years. Pupil performance in key stage 2 (age 11) English and mathematics tests was compared with their GCSE scores.

Study 3 was conducted in 2011 by Dr John Jerrim, a lecturer in Economics and Social Statistics.

Key finding: Contrary to common belief, there was no hard evidence of any decline in the performance of English secondary school pupils in international mathematics tests (**R4**). This was highly significant as the coalition government had justified its radical school reforms by pointing to England's apparently plummeting performance in the three-yearly Programme for International

Impact case study (REF3b)

Student Assessment (PISA) tests set by the Organisation for Economic Co-operation and Development (OECD).

Methods: Jerrim looked at English pupils' scores in international tests since 1999. He noted that another major survey, the Trends in International Mathematics and Science Study (TIMSS), found that – contrary to PISA – the scores of England's 13 and 14-year-olds rose in comparison with other nations between 1999 and 2007. He calculated that the disagreement between PISA and TIMSS on England's mathematics performance was much bigger than for any other country and conducted statistical analysis to check that this difference was not being driven by sampling variation. He then identified several possible reasons for this discrepancy which no previous study had highlighted. For example, one third of English pupils tested by PISA in 2000 and 2003 were in Year 10 and two thirds in Year 11. However, in 2006 and 2009 almost all English pupils were in Year 11. This change might be expected to boost England's rank position but this effect was confounded by changes in the date during the school year that the pupils sat the test, which also affected the results.

3. References to the research (indicative maximum of six references)

- R1: Allen, R. & Burgess, S. (2013) Evaluating the provision of school performance information for school choice, *Economics of Education Review*, 34,175-190.
- R2: Allen, R. & Burgess, S. (2011) Can school league tables help parents choose schools? *Fiscal Studies*, 32(2), 245-261.
- R3: Dearden, L., Micklewright, J. & Vignoles, A. (2011) The effectiveness of English secondary schools for pupils of different ability levels, *Fiscal Studies*, 32(2), 225-244.
- R4: Jerrim, J. (2013) The reliability of trends over time in international education test scores: Is the performance of England's secondary school pupils really in relative decline? *Journal of Social Policy*, 42(2) 259-279.

Indicators of quality:

- IQ1: In 2009, Allen received the British Educational Research Association's best PhD dissertation of the year award for a closely related study analysing the consequences of allowing parents to choose their children's school. She also received a Future Research Leaders award from the Economic and Social Research Council in 2012, partly on the basis of the quality of her research on performance tables with Burgess.
- IQ2: The work conducted by Dearden and her colleagues was funded by the ESRC under the prestigious National Centre for Research Methods programme.
- IQ3: Although Jerrim's study could be seen as having unwelcome findings for the DfE, he and his colleagues, including Micklewright, Vignoles and Allen, later won the government contract to write the national reports on the PISA 2015 survey for England, Wales and Northern Ireland¹. The PISA analysis described here was also largely responsible for Jerrim winning two coveted ESRC awards in 2013 – the Outstanding Early Career Impact prize and a Future Research Leaders award.

Grants:

- G1: The research by Dearden et al was one of the National Council for Research Methods ADMIN (Administrative Data - Methods, Inference and Network) node projects funded by an ESRC grant of £1,609,189 (July 1, 2008 to August 31, 2011). Dearden was the grant-holder.
- G2: Allen's research was financed via the DfE-funded Centre for Understanding Behaviour Change (CUBeC). IOE received just under £40,000 to cover her salary costs between March 2010 and December 2013.
- G3: Jerrim's work was funded by an ESRC post-doctoral fellowship.

¹ The work on the data collection, but not their analysis, is in partnership with RM Education, who lead this part of the project, and World Class Arena.

Impact case study (REF3b)

4. Details of the impact (indicative maximum 750 words)

Principal beneficiaries: Studies 1 and 2 have benefited the Government, parents and schoolchildren while Study 3 has been of value to society as a whole by encouraging policy-makers to be more careful about the conclusions they draw from international comparative studies.

Dates of benefit: Studies 1 and 2: early January 2011 onwards; Study 3: from December 2011.

Reach and significance: As the evidence below shows, all three studies have had both an instrumental impact² (influencing policy) and a conceptual one (enhancing general understanding and informing debate). Each, in different ways, has ensured that teachers, parents and the public should be better briefed about the measurement, representation and substance of educational performance in future. Parents in England now receive more useful information on which to base school choices as a result of Studies 1 and 2. Children are therefore likely to do better academically - because more families should make more appropriate school choices. Jerrim's study has made it less likely that misleading evidence will be used to justify school reforms. He has also made the public more aware of the dangers of international comparisons.

Studies 1 and 2:

Allen and Burgess began discussing their proposed amendments to England's performance tables at a DCSF meeting in February 2010 attended by the Head of Research. Vignoles also gave a presentation to the then shadow education secretary, Michael Gove, in early 2010, highlighting the ADMIN research on how best to measure school effectiveness and the problems with CVA. She and her colleagues had meetings with government officials and Vignoles made presentations at the 2010 DCSF Research Conference and the 2011 Ministerial Seminar hosted by universities minister David Willetts. The then IOE director, Geoff Whitty, briefed Graham Stuart MP, chair of the House of Commons Education Select Committee, on the ADMIN research in July 2011. Mr Stuart then asked Allen and Burgess to present their proposals to him in Westminster on November 10, 2011. Allen has since had several meetings with Tim Leunig, senior policy adviser to ministers at the Department for Education (DfE) – see impact source **S1**.

Impact on tables: The research by Allen and Burgess contributed to the DfE decision to remove the CVA measure from performance tables from January 2011. Lord Hill of Oareford, Parliamentary Under Secretary of State for Education, confirmed this in the House of Lords on January 19, 2011. Asked by Lord Hunt of Kings Heath what academic evidence the government had relied on in deciding to end the use of CVA, he replied: "Research conducted by Allen and Burgess in 2010...found that CVA is a less strong predictor of how well a child will do academically than raw attainment measures" (**S2**). The studies by Allen and Burgess, and Dearden, Vignoles and Micklewright, also helped to convince the DfE that it should change the tables again in January 2012. This time it decided to offer parents information about the GCSE performance of children of similar ability to their own - in all schools in their local area. Both teams of researchers had recommended this change. The DfE agreed to divide children into three groups – low, middle and high attainers – and indicate how they had progressed. One advantage of this approach is that it militates against schools focusing disproportionately on pupils at the GCSE C/D threshold. The tables are still not exactly as the IOE researchers would wish – for example, each band includes a wider ability range than they recommend – but it is an important step in the right direction. Furthermore, senior politicians are continuing to seek their advice. Allen was invited to take part in a closed roundtable discussion on school accountability with the Deputy Prime Minister, Nick Clegg, on April 16, 2013. She and Dearden also advised the Education Select Committee on performance tables on the same day. On July 2, 2013, Allen again briefed Graham Stuart as he prepared to discuss performance tables with Michael Gove. She pointed to problems with the calculation of progress measures for the KS4 tables, the potentially distorting effects of the new 'best 8' GCSEs measure, and the case for and against threshold measures such as 5 A-C grades at GCSE. Allen was also asked to provide the DfE with technical advice (in June 2013) on the new value added measures to be used in the 2013 performance tables. This advice included recommendations for the most appropriate statistical model for predicting pupil GCSE performance.

² Using Evidence: How Research can Inform Public Services (Nutley, S., Walter, I., Davis, H. 2007)

Impact case study (REF3b)

Wider influence: Audiences beyond Westminster and academia have been reached via the very accessible blogs that Allen (www.rebeccaallen.co.uk) has written and by the media coverage that the ADMIN node work has attracted (**S3**).

Study 3:

Jerrim issued a press release about his study that triggered substantial media coverage in December 2011. Stephen Twigg, the then Shadow Education Secretary, told the BBC: “This report demonstrates that the claims that pupils in England have been sliding down the international performance tables are unfounded” (**S4**). Headteachers’ leader Brian Lightman said that the study had shown that international comparisons should come with health warnings. The DfE, however, continued to insist that PISA highlighted the scale of the decline in pupil performance. The following September the Chief Inspector, Sir Michael Wilshaw, also reasserted that England had fallen from 7th to 28th in the PISA mathematics table between 2000 and 2009.

Dilnot investigates: The Wilshaw statement prompted former schools minister David Miliband to complain to Sir Andrew Dilnot, chair of the UK Statistics Authority. Sir Andrew investigated and sent Mr Miliband an open letter supporting his complaint in October, 2012 (**S5**). He said that the OECD had confirmed that gaps in some years’ PISA data made accurate comparisons difficult. Sir Andrew said he had also noted Dr Jerrim’s study which reported that 1) there were problems with identifying change over time using PISA data for England 2) conclusions should not be based on this resource alone, and 3) other evidence, including TIMSS, contradicts PISA findings (**S6**). He (Dilnot) was concerned that PISA rankings had been used in a DfE press release in December 2010 without the necessary “detailed advice or caveats”. He said he would discuss this issue with the Department and copy his letter to the Chief Inspector, the National Statistician, and to the Heads of Profession for Statistics at the DfE and Ofsted. Sir Andrew later had a high-level meeting at the DfE and there is evidence that his advice has been heeded. In December 2010, Michael Gove, the Education Secretary, told *TES* readers that “PISA 2009 shows that thoroughgoing reform of our schools is urgently necessary” (**S7**). But by May 2013, when he appeared before the Education Select Committee, Mr Gove had modified his claims. Under questioning, he acknowledged that cross-country comparisons using PISA data were “complex” (**S8**). It is therefore evident that Jerrim’s study, which was also referred to in a *Guardian* leader in September 2012 (**S9**), has helped to shape this important debate.

5. Sources to corroborate the impact³ (indicative maximum of 10 references)

S1: Tim Leunig, Senior Policy Adviser, DfE

S2: House of Lords statement by Lord Hill of Oareford (*Hansard*: HL Debate, January 19, 2011, c50W) <http://www.theyworkforyou.com/wrans/?id=2011-01-19b.50.7&s=Allen+and+Burgess#g51.0>

S3: “UK school league tables in doubt”, *Financial Times*, August 15, 2011.

S4: Coughlan, S., “No evidence for England’s schools falling behind”, BBC News, December 7, 2011 <http://www.bbc.co.uk/news/education-16054654>

S5: Letter from Sir Andrew Dilnot, chair of the UK Statistics Authority, to David Miliband MP, October 3, 2012 <http://www.statisticsauthority.gov.uk/reports---correspondence/correspondence/index.html>

S6: Testimonial from Sir Andrew Dilnot confirming the importance of Jerrim’s research (IOE can provide).

S7: *TES*, December 21, 2010 <http://www.tes.co.uk/article.aspx?storycode=6066185>

S8. House of Commons Education Select Committee, Oral evidence, May 15, 2013 <http://www.publications.parliament.uk/pa/cm201314/cmselect/cmeduc/uc65-i/uc6501.htm>

S9: “Overhauling exams: lessons in nostalgia”, the *Guardian*, September 17, 2012

³ All web links accessed 11/10/13