Planning v competition in education: outcomes and efficiency

Geraint Johnes, Lancaster University Management School

https://doi.org/10.18573/wer.233 Accepted: 03/12/18

Introduction
Governments have a responsibility to secure the best return on their investments in services. This being the case, the institutions that are involved in organising the supply of these services are a topic of interest and concern. In the case of educational services in Wales, particular concern surrounds the performance of school pupils as measured by international tests, in which the recent record has been disappointing. This paper examines this record and evaluates reform measures that have been put in place.

The remainder of the paper is structured as follows. In the next section, I look at some data that allows comparison of educational outcomes in Wales and England. This is followed by a broader discussion about the system of education. The final section draws together the main ideas of the paper in a conclusion.

Data analysis
The data used here come from the Programme for International Student Assessment (PISA) undertaken by the Organisation for Economic Cooperation and Development (OECD). Every three years, samples of 15 year old pupils from schools in OECD countries and elsewhere undertake tests in mathematics, reading and science. More than 80 countries have participated in PISA; in the case of the United Kingdom, the data are collected by the National Foundation for Educational Research (NFER) and are enhanced by the presence of separately identifiable samples from each of England, Wales, Scotland and Northern Ireland. The latest available data, used here, are for 2015. The microdata include information about pupil and school characteristics as well as about pupil performance on the tests, and are made freely available to researchers on the OECD website.

The PISA report for 2015 provides data on average scores achieved on tests in mathematics, reading and science that are in each case higher in England (493, 500 and 512 respectively) than in Wales (478, 477 and 485 respectively). These mean values are based on samples of pupils, of course, but it should be noted that the England and Wales 95% confidence intervals do not overlap in the case of any of the three subject areas. This follows performance in earlier rounds of PISA testing that likewise indicated lower scores in Wales than elsewhere in the United Kingdom. This being the case, the media has shown interest and the results have been the subject of debate both in the Welsh Assembly and the House of Commons. Various reforms have been introduced in Wales over recent years, and while it is hoped that these will have positive impact it is likely that their effects on pupil performance are only now beginning to be realised (Reynolds, 2016).
Average scores based on outputs (test results) can provide a misleading picture of the educational process, however. The literature on educational production functions (Hanushek and Woessmann, 2011a, 2011b; Glewwe et al., 2013) provides a wealth of information about how various characteristics of pupils, schools and systems influence performance. The differences that we observe between educational output as measured by PISA performance across England and Wales are likely due, in part at least, to differences in inputs – including, for example, socio-economic composition.

To interrogate this further, I investigate the PISA data for England and Wales using methods drawn from the literature on data envelopment analysis (DEA). This method was developed by Charnes et al. (1978) following earlier work by Farrell (1957), and uses linear programming techniques to identify, separately for each decision-making unit in the data, a frontier against which can be evaluated the efficiency with which the unit maps inputs onto outputs.

Specifically in the present context, I use the DEA approach of meta-frontiers pioneered by Charnes et al. (1981), Portela and Thanassoulis (2001) and Rao et al. (2003) and subsequently applied in the context of education by Johnes (2006), Waldo (2007), De Witte et al. (2010), and Thieme et al. (2013). This approach involves assessing the efficiency with which each pupil converts inputs (characteristics) into outputs (educational performance) relative, first, to the frontier within their own country, and secondly to the frontier across countries. Thus it is possible to analyse, for students of different types, how efficient is the system of education in each country.

Using linear programming methods, the DEA can be extended to a situation in which there are multiple inputs and multiple outputs. The analysis reported in this paper have been obtained using Limdep software.

I use, as inputs into the production process, the pupil’s age (by month), the number of books they have at home, weekly hours spent in class for maths, language (reading) and science respectively, hours spent studying per week out of class, and measures of economic, social and cultural status (ESCS), household possessions, and wealth. As outputs, I use the pupils’ (‘plausible value’) scores on the PISA tests for maths, reading and science respectively. Thus the DEA model has 9 inputs and 3 outputs. I employ the variable returns to scale method with output orientation, developed by Banker et al. (1984).

Table 1 reports, in the first row, the average efficiency scores obtained by pupils in England and Wales respectively, relative to best performance within the relevant country. These averages are fairly high, indicating that pupils are generally quite successful in converting their endowments (characteristics) into positive educational outcomes. The score of 0.8338 for Wales is slightly higher than the 0.8174 for England, and this likely reflects a lower diversity of experience in Wales than in England.

The second row of the table shows how the frontier defined by the ‘efficient’ pupils in Wales compares with that which applies to England. For the typical student, the frontier in Wales is lower than that in England.

<table>
<thead>
<tr>
<th>Table 1 Average value of pupil-level efficiencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>England</td>
</tr>
<tr>
<td>Pupil within country</td>
</tr>
<tr>
<td>Country within all countries</td>
</tr>
</tbody>
</table>

Note: Author’s calculations based on the PISA microdata for England and Wales, obtained from http://www.oecd.org/pisa/data/2015database/. Numbers of observations are 1693 pupils for Wales, 2672 for England.
But the fact that the score for England is less than unity indicates that there are some pupils whose characteristics are such that, for them, the frontier in Wales dominates that in England. It is not, therefore, unambiguously the case that the English system is performing better (or worse) than the Welsh. Put simply, it depends who you are.

It is possible to construct similar averages to those reported in Table 1 for subsets of the sample of pupils. This is done in Table 2, where the average scores are reported for those pupils with above average, below average and bottom decile wealth, and those with above average, below average and bottom decile ESCS. In all cases but one, the ‘country within all countries’ score is lower for Wales than for England. The exception is that of pupils in the bottom decile of the ESCS. For these pupils, frontier performance in Wales dominates that observed in England. Interestingly, this corresponds with the finding of Pont et al. (2017) that ‘a student’s socio-economic background in Wales has less impact on their performance than for students in other parts of the United Kingdom’.

Two important conclusions may be drawn from this analysis. First, once allowance is made for the different characteristics of pupils, spatial variation of performance in standardised tests may not be as great as at first appears. This echoes the findings of Rees and Taylor (2014), who find that the gap in measured performance between Welsh and English schools is much reduced when they undertake a comparison of schools in Wales with those in the English local education authorities whose characteristics most closely match those of Wales. Secondly, the shape of the production frontier likely differs across space so that it is not necessarily the case that all pupils are best served by the technology

<table>
<thead>
<tr>
<th>Table 2 Average value of pupil-level efficiencies, disaggregated by characteristics</th>
<th>England</th>
<th>Wales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wealth average or above</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pupil within country</td>
<td>0.8016</td>
<td>0.8467</td>
</tr>
<tr>
<td>Country within all countries</td>
<td>0.9926</td>
<td>0.9906</td>
</tr>
<tr>
<td>Wealth below average</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pupil within country</td>
<td>0.8467</td>
<td>0.8696</td>
</tr>
<tr>
<td>Country within all countries</td>
<td>0.9906</td>
<td>0.9194</td>
</tr>
<tr>
<td>Wealth below bottom decile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pupil within country</td>
<td>0.9604</td>
<td>0.9669</td>
</tr>
<tr>
<td>Country within all countries</td>
<td>0.9954</td>
<td>0.9588</td>
</tr>
<tr>
<td>ESCS average or above</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pupil within country</td>
<td>0.8108</td>
<td>0.8244</td>
</tr>
<tr>
<td>Country within all countries</td>
<td>0.9941</td>
<td>0.9215</td>
</tr>
<tr>
<td>ESCS below average</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pupil within country</td>
<td>0.8340</td>
<td>0.8525</td>
</tr>
<tr>
<td>Country within all countries</td>
<td>0.9865</td>
<td>0.9259</td>
</tr>
<tr>
<td>ESCS below bottom decile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pupil within country</td>
<td>0.8150</td>
<td>0.8318</td>
</tr>
<tr>
<td>Country within all countries</td>
<td>0.9919</td>
<td>0.9936</td>
</tr>
</tbody>
</table>
underpinning the education system in one locale rather than another. This recalls the finding of Hanushek et al. (2011) that school autonomy is beneficial in developed systems of education, but less so elsewhere – there appear to be horses for courses. This is particularly significant in the context of Wales, given the political commitment, arguably more pronounced there than in England, to equality and social justice (Power, 2016).

**Discussion**

The third Thatcher government introduced a wide range of policies aimed at transforming the provision of public services by way of the introduction of quasi-markets (Le Grand, 1991). The new arrangements separated out the funding of services by the state from their provision. Providers – whether in the public or private sector – would compete with each other to provide services, and would be remunerated on the basis of the amount of custom they attracted. In education, for example, the 1988 Education Reform Act replaced block grant funding by formula funding where monies followed pupils and, crucially, where pupils (and their parents) were given new freedoms to choose the school that they attended. At the same time, greater managerial autonomy was granted to schools, which for the first time could opt out of local authority control. In some areas, such as the provision of nursery education, similar principles were applied but in a yet more vivid fashion – through the introduction of vouchers (Sparkes and West, 1998).

Much early evidence on the impact of quasi-market reforms on school performance in England is consistent with the view that competition enhances performance (Levačić, 1994; Bradley et al., 2000; Bradley et al., 2001). More recent evidence provided by Burgess et al. (2013) is particularly interesting because it uses the abolition of school rankings in Wales as a natural experiment with which to test the hypothesis that publication of such rankings improves school effectiveness. The evidence supports the hypothesis, albeit in a way that varies substantially across the distribution of schools. This is an important point, not least because of the emphasis placed in Wales on equity as a desired outcome from the education system; hence not only average effects of policy but the impact of policy throughout the distribution of experience should be considered. Indeed, in the Burgess et al. study, there is no evidence of an effect on performance of the introduction of quasi-market mechanisms on the top quartile of schools. Goldstein and Leckie (2016) confirm that the reform was followed by a relative decline in average achievement in schools in Wales, though they argue that this may be as much due to pupils’ lack of practice in high-stakes tests as to competitive effects. Likewise West and Pennell (1997) document improvements due to the greater freedom to choose schools, but (in contrast to Burgess et al.) argue that the greatest benefits were realised by those drawn from the higher end of the income distribution.

Yet the fundamental assumption underpinning competition in the provision of school places deserves some interrogation. Schools are not shareholder-owned firms in which management faces an imperative to maximise profit, and the freedom of entry and exit essential to a truly competitive climate, is absent. Moreover, until the advent, in England but not in Wales, of academies in the early years of this century, adherence to a national curriculum further limited competition. Schools in reality face a plethora of incentives and respond to a multiplicity of social needs. They offer their customer base an experience good – one where the quality of the match between a school and pupil can only be assessed over time and through the commitment of engagement. Schooling takes time, school quality is not static, and choice made on the basis of necessarily retrospective data may not be optimal ex post. Furthermore, schooling is a club good – one that is
provided simultaneously, on a non-rival basis, to many beneficiaries; individuals cannot make decisions about their schooling without this impacting on others, and in such a context markets cannot operate perfectly.

Following the establishment of the Welsh Government in 1999, the approach to policy in the sphere of education has certainly differed markedly between Wales and England. Power (2016) notes the tendency for government to be viewed in England as part of the problem, while in Wales it is more usually seen as part of the solution. Relatively few schools in Wales have become grant maintained, so that local authorities retain a strong voice here in the provision of education. Likewise, the curriculum, or at least the way in which it is delivered, has developed in ways that are more progressive in Wales than in England. Critically, there is less standardised testing in Wales, where league tables have not been produced since 2001. This means that the principle of school choice does not work in the same way in Wales as in England – parents do not have the same information sets. It is in any event not clear that school choice can work the same way across the two countries; in mid-2016, England had a population density of 424 per km², while the corresponding figure for Wales was just 150 per km².

As noted by Swaffield (2017), the relatively poor ranking achieved by Wales has resulted in a ‘PISA shock’, with politicians responding with the plea that we should ‘never waste a crisis’. Hopkins (2016) has identified a number of issues at systemic level that have prevented the Welsh system from achieving to its potential. In particular, he notes that ‘accountability systems are still relatively crude and not linked to increased performance, there is a lack of a secure pedagogy that reliably leads to enhanced student learning, the standards of entrants to the teaching profession are lower than in equivalent systems and the architecture for sustained self-improvement is missing’. He also notes a mismatch between the needs of the system and the solutions proposed by politicians, the latter tending to be more bureaucratic than effective – a criticism also noted by Grigg (2016).

Recent reforms in Wales have aimed to tackle these shortcomings, starting out from the 20 point plan outlined by Andrews (2011). This focuses on four areas (Pont et al., 2017). The first concerns the quality of teaching provision. This follows criticism by Estyn (2016) of the standard of teaching in many secondary schools. This is being addressed through Teach First, and through the various recommendations of Furlong (2015). These include a tightening of requirements for newly qualified teachers, and reform in the provision of initial teacher education through a process of competitive tendering. Beyond initial training, continuing professional development has been enhanced through the New Deal for the Education Workforce (Lewis, 2015), providing through the Education Workforce Council a professional learning passport, encouraging collaboration (between schools and between individual teachers) and reflection. Evaluation of the early progress on these initiatives will be crucial.

Secondly, the Donaldson (2015) report has highlighted the need for curriculum reform in Wales – and cites PISA outcomes as a major source of concern. Noting evolving employer needs, this emphasises learning skills, creativity, ethical behaviour and a commitment to society. The reform promotes a holistic curriculum in six areas of learning: arts; humanities; languages; literacy; mathematics; health and wellbeing; and science. How the change programme advocated in this review translates into practice will be crucial in determining the success of the reform, and evaluation will also be crucial in this area.

Thirdly, the emphasis on equity in education is particularly strong in Wales
(Welsh Government, 2016), and this distinguishes the system in this part of the UK from that in, say, England. The DEA results reported earlier confirm this, in showing that the system in Wales serves those in the bottom ESCS decile better than does that of England. Nevertheless, questions remain about the extent to which the oversight that the Welsh government has on resource allocation allows good decision making. Local authorities, working alongside regional consortia set up following the National Model for Regional Working (Welsh Government, 2014), devise funding formulae for the allocation of resource to schools. As has been the case in England, there is an open question about how equitable this allocation might be, and about the extent to which local decision making results in a ‘postcode lottery’. This relates to the tension between the desirability of equity and the principle of subsidiarity – that experts at local level are best placed to assess the local situation. An interesting literature on subsidiarity appeals to property rights and shows that local authorities likely underinvest in circumstances where there are spillovers in the returns to the investment (Lülfesmann, 2002). Specifically in the context of education, the optimal locus of decision making has been the subject of research by Johnes (1995) and De Witte and Schiltz (2016), the latter of which in particular indicates the presence of substantial economies of scale – that is, that decisions are better made at higher levels.

Fourthly, as has been noted by Bloom et al. (2015) and McCormack et al. (2014), leadership is a key determinant of the success of educational institutions. School leadership remains an area that is of concern to the OECD (Pont et al., 2017), although the National Academy of Educational Leadership launched in 2018 offers promise – but again will need to be evaluated.

A common thread running through these areas of activity is the recognition that reform is needed in order to improve the experience of young people as they progress through education in Wales. While PISA results have been a driver, they should not be (and have not been) the only driver. While recognising the mobility of labour within the UK and beyond, the development of the education system in Wales needs to be cognisant of the needs of Wales itself, and specifically of (existing and prospective) employers in Wales. The distance between employers and education has been a challenge, not only within Wales but more widely in the UK, and opportunities to shrink this distance need to be grasped.

Conclusions
The gap in average performance between pupils in Wales and those in England has justifiably caused concern. There are hazards in taking such measures at face value, but the prompt provided by PISA results to examine and, where appropriate, reform provision in Wales is welcome. The reform measures put in place heretofore have promise, but will need systematic and rigorous evaluation. Moreover, it should be recognised that performance in international tests should represent only one of the drivers of this reform. The Welsh commitment to equity should also (continue to) condition the policy response. Likewise the future needs of business need to be reflected. In this last respect, much remains to be achieved.

Endnotes
3. For example: [http://www.walesonline.co.uk/news/education/full-pisa-results-2016-show-12278551](http://www.walesonline.co.uk/news/education/full-pisa-results-2016-show-12278551)


5. There are other applications of DEA in the context of education, but not using the meta-frontier approach.

6. Where, as in the case of ESCS, household possessions and wealth, the measures provided in the PISA database contain negative values, I have added 10 to the measure to ensure that all values are positive; this is a requirement of the DEA model. The measures reported in the PISA data for these variables are based on standardised variables – and in the case of ESCS on the first principal component of a group of standardised variables. See [http://www.oecd.org/pisa/sitedocument/PISA-2015-Technical-Report-Chapter-16-Procedures-and-Construct-Validation-of-Context-Questionnaire-Data.pdf](http://www.oecd.org/pisa/sitedocument/PISA-2015-Technical-Report-Chapter-16-Procedures-and-Construct-Validation-of-Context-Questionnaire-Data.pdf).

7. Coase (1960) suggests that if property rights are well defined and externalities can be traded, market failure can be precluded. These conditions are not met in the classroom.


9. See [https://seneddresearch.blog/2018/05/17/national-academy-for-educational-leadership/](https://seneddresearch.blog/2018/05/17/national-academy-for-educational-leadership/)

References

Agasisti, T. and Johnes, G. (2009). Beyond frontiers: comparing the efficiency of higher education decision making units across more than one country. Education Economics, 17, 59-79. [https://doi.org/10.1080/09645290701523291](https://doi.org/10.1080/09645290701523291)


