

Department of Quantitative Social Science

Administrative Data and Economic Policy Evaluation

Lorraine Dearden

DoQSS Working Paper No. 10-14
July 2010



DISCLAIMER

Any opinions expressed here are those of the author(s) and not those of the Institute of Education. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

DoQSS Workings Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

DEPARTMENT OF QUANTITATIVE SOCIAL SCIENCE. INSTITUTE OF
EDUCATION, UNIVERSITY OF LONDON. 20 BEDFORD WAY, LONDON
WC1H 0AL, UK.

Administrative Data and Economic Policy Evaluation

Lorraine Dearden^{*†}

Abstract. This paper looks at the strengths and weaknesses of using administrative data for economic policy evaluation. It does this by looking at how school administrative data has been used to assess school effectiveness and the impact of month of birth on educational outcomes with varying degrees of success. It concludes that if there is some natural experiment in the way the education is delivered or an education initiative is introduced, then schools' administrative data offers the opportunity of answering questions of extreme policy interest in a robust way – even without rich background information on the students and their families.

JEL classification: C52, I20.

Keywords: administrative data, evaluation methods, school league tables, month of birth, natural experiment.

^{*}Institute for Fiscal Studies, 7 Ridgmount Street, London, WC1E 7AE; Institute of Education, University of London, 20 Bedford Way, London WC1H 0AL, UK. E-mail: L.Dearden@ioe.ac.uk

[†]**Acknowledgements.** This paper is based on elements on my keynote talk at the Australian Economists' Conference held in Adelaide in September 2009. I would like to thank conference participants, in particular Deborah Cobb-Clarke, Chris Ryan and Andrew Leigh for constructive comments. I would also like to thank Claire Crawford, Costas Meghir, John Micklewright, Alfonso Miranda, Sophia Rabe-Hesketh, Marcello Sartarelli and Anna Vignoles for allowing me to draw on research I have conducted with them in my talk and this paper. This research was supported by ESRC grant RES-576-25-0014 under the ADMIN node of the National Centre for Research Methods at the Institute of Education

1. Introduction

Administrative data is increasingly being used in a large number of countries to evaluate policy initiatives and to measure effectiveness of public service provision. In the coming years, in a climate of global public sector fiscal constraint, it is highly likely that funding of large national longitudinal surveys or indeed national censuses will come under close scrutiny and the pressure to use existing administrative data to supplement or even replace expensive surveys is likely to increase. It is therefore useful to consider the strengths and weaknesses of administrative data for economic policy evaluation.

In this paper I am going to focus on some examples of using *school* administrative data in evaluating education policy in order to illustrate the issues one must consider when using administrative data for economic policy evaluation. This is a topical issue in Australia at the moment where the new National Assessment Program, Literacy and Numeracy (NAPLAN) tests are now effectively being used to show how each school in Australia is performing relative to the national average and ‘statistically similar’ schools¹. In England, school league tables have been part of the education scene since the mid 1990s².

But can we use this school administrative data to really tell us how effective a school is? And does publishing school results in the way done in Australia and England really inform parents or the government about whether a school is ‘good’ or ‘failing’ and/or help parents make schooling choices for their child? This is the question I look at in section 2 of the paper.

The problem is that administrative data, like NAPLAN and the National Pupil Database (NPD) in England, typically do not have the extensive background on students which one would typically observe in survey data. Measures of family characteristics that are well known to influence school choice and educational outcomes, such as family structure and composition as well as parental education and socio-economic status, are typically not measured. This raises the strong possibility that analysis using this administrative data is not measuring the true causal impact of schools and/or other educational inputs or policies, but the effects of these unobservable factors which are correlated with the school and/or

¹ See <http://www.myschool.edu.au/> which is based on results from the NAPLAN tests.

² See <http://www.education.gov.uk/performance/tables/>.

input and/or policy and children's school outcomes. This means one has to be very careful when using administrative data to inform public policy because of the likely endogeneity of the variable of interest – a common problem in empirical economics. This is particularly true when considering questions like school performance or indeed differences in school performance by areas such as local authority in England or State in Australia.

However, the individual administrative data on which these tables are based, can be used by researchers to look at a number of other important educational issues, particularly when there is some sort of natural experiment that can be exploited. My second example in section 3 of this paper looks at one such example in England- the impact of month of birth and length of schooling on educational outcomes in England. This work exploits a regression discontinuity in the age at which children start school as well as varying admissions arrangements that occur across local authorities in England. It shows that children who are youngest in the year perform much worse than children who are the oldest in their academic year, and this has implications past the end of compulsory schooling.

2. Measuring School Effectiveness

School league tables have been available in England since the mid 1990s. The main drivers behind the introduction of the published League Tables was to identify 'failing' schools, to set incentives for schools to improve performance by effectively naming and shaming, and to aid parents in school choice. However, it was soon apparent that differences across schools did not necessarily represent differences in school quality (see Goldstein and Spiegelhalter (1996) and this has seen a raft of different measures of school effectiveness being introduced in England (see Ray (2006)).

So what are the problems with League Tables? Firstly, school effectiveness measures that just compare average outcomes in schools (as is currently done in Australia and is done in the England with the headline results), do not take into account *baseline differences in ability and backgrounds* of children attending different types of schools and therefore are not a good measure of *school* effectiveness. In the English context it is easy to show that most of the variation in outcomes at 16 can be explained by factors other than the school which the child is attending, the most important being outcomes at the end of primary school (see Chowdry et. al 2010, and Goodman and Greig (2010) for example). So in

order to use administrative data to measure school effectiveness, it is essential to have reliable and consistent baseline measures of student ability and background as a child enters their school. This is not possible with the current Australian testing regime with tests held in May towards the middle of Years 3, 5, 7 and 9). In England, the Year 6 tests at the end of primary school can form a baseline for tests at age 16, but information on family background is very limited.

There has been a lot of literature on how to measure value added in schools (see for example Ladd and Walsh (2000), Meyer (1996, 1997)) but even these more sophisticated metrics that measure value added by taking into account baseline differences in the characteristics of children again cloud differences in effectiveness across different types of students within schools (see Dearden, Micklewright and Vignoles (2010) for example) and suffer from other problems like missing covariates (e.g. parental education) and right and left censoring of baseline test scores (see Dearden, Miranda and Rabe-Hesketh(2010)). Also, because these value added models generally involve statistical modelling, there are issues of whether differences in school effectiveness are in fact statistically significant and stable over time (see Goldstein and Leckie (2009)). Further, the reporting of value added in a transparent and understandable way is difficult and there is evidence that parents, despite these alternative measures, tend to focus on average outcomes of schools (see Hansen and Machin (2010) for example). Finally all of these measures just come up with one mean measure for every school, not a school's effectiveness across the entire distribution of its student intake. Dearden, Micklewright and Vignoles (2010) show that around 40 per cent of secondary schools in England are differentially effective across the prior ability distribution so in these cases average measures of performance or value added do little to inform parents whether the school is effective for a particular child.

But coming up with understandable and transparent ways of constructing these league tables that are informative and a true reflection of school effectiveness is very difficult. Dearden, Micklewright and Vignoles (2010) attempt to come up with some simple, transparent solutions in the English context for secondary schools. But to do this well, you need rich baseline data as the child first enters primary and secondary school, including baseline ability and family background variables. Currently this rich baseline data does not exist in either the Australian or English Schools' administrative data used to construct these tables. And whilst that is the case, school league tables will continue to be misused and abused.

3. Impact of month of birth on educational attainment

In England and Australia there are different conventions and rules concerning when children start school. In any education system there will always be age differences within a school year. But if the education and testing system works well, then the month in which you are born should not have any long lasting impact on long term educational outcomes.

But there is clear evidence in England that this is not the case and children who are the oldest in a school year (September born children in England) do significantly better in academic outcomes than their younger colleagues in the same year group. These differences persist past the end of compulsory schooling and even impact on Higher Education participation (see Crawford, Dearden and Meghir (2010)).

In our research we exploit regional variation in the length of schooling/age of starting school for younger children in an academic year. In some English local authorities children born later in the academic year have one or two less terms of schooling than their older peers. We find that the month of birth differences in outcomes we observe are almost entirely due to differences in the age at which children sit the tests rather than length of schooling/age of starting school effects. We find that starting school earlier/having more terms of school is marginally better for younger children at the ages of 7 and 11 but any differences have disappeared by the age of 14.

Ideally the education system needs to create a level playing field for all children *regardless of date of birth*. But within a school year, some children are always going to be younger than others. This points to the need to age normalise tests/assessments so that young children are compared only to their similar aged peers and not incorrectly labelled as low ability. The English example shows that doing nothing is not a solution.

Understanding the nature and extent of the month of birth effects (and other education initiatives/policies) in England wouldn't have been possible without having easy access to the *individual* data underlying school's administrative data in England. Australia needs to follow the English lead by giving researchers access to the *individual* NAPLAN data to look at policy concerns beyond comparing schools.

If we go back to the month of birth question, in Australian almost every state has a different rule/convention regarding age of starting school which means the national individual NAPLAN data could be used to look at the impact of different starting ages and length of schooling on NAPLAN test results in years 3, 5, 7 and 9. Indeed it is almost certain that some of the explanation for State differences in test results will be due to the different ages and length of schooling of children sitting these tests. In some states, like Queensland, there has also been a

switch from 12 to 13 years of schooling and a switch in school starting age in 2007. This means when analysing Australian data you could separate not only age of sitting test effects, but length of schooling versus age of starting school effects. This analysis may actually be able to inform policy makers of the best system for children and may provide a strong evidence base to move to a common admissions system Australia wide.

Given there are effectively 8 different education systems in Australia, the potential questions that the data could be reliably used to answer, are likely to be numerous. But researchers need to have access to the Australia wide individual data. The UK government's decision to allow widespread access to English schools' administrative data is something the Australian governments should follow.

4. Conclusions

Administrative data is increasingly being used by empirical economic researchers. But one has to be careful using these data as typically it is not as rich as survey data. Using schools administrative data to measure school effectiveness, for example, is fraught with problems and it is highly likely that these measures are reflecting differences in prior academic ability and family background rather than school effectiveness.

That is *not* to say that we should abandon nationally consistent testing of children at key stages of their schooling. Schools administrative data based on these tests can be used to answer a large number of questions if the individual data is made available to researchers. Indeed, if there is some natural experiment in the way the education is delivered or an education initiative is introduced, then schools' administrative data offers the opportunity of answering questions of extreme policy interest in a robust way – even without rich background information on the students and their families. In Australia, with eight different education systems, the possibility of finding such natural experiments is probably higher than in most countries.

References

Chowdry, H., Crawford, C., Dearden, L., Goodman, A. and Vignoles, A. (2010), "Widening participation in higher education: analysis using linked administrative data", IFS Working Paper W10/04, May (<http://www.ifs.org.uk/publications/4951>).

- Crawford, C., Dearden, L. and Meghir, C (2010), "When you are born matters: The Impact of Date of Birth on Educational Outcomes in England", IFS Working Paper W10/06, May (<http://www.ifs.org.uk/publications/4866>).
- Dearden, L., Miranda A. and Rabe-Hesketh, S. (2010)), 'Contextualised Value Added: How important are missing covariates for school rankings?', Department of Quantitative Social Science Working Paper, Institute of Education.
- Dearden, L., Micklewright, J., and Vignoles, A. (2010), 'Measuring School Effectiveness', Department of Quantitative Social Science Working Paper, Institute of Education.
- Goldstein, H. and Spiegelhalter, D. J. (1996) 'League tables and their limitations: statistical issues in comparisons of institutional performance'. *Journal of the Royal Statistical Society: Series A*, 159, 385-443.
- Goodman, A. and Greig, P. (2010), 'Poorer children's educational attainment: how important are attitudes and behaviour', Joseph Rowntree Foundation Report, March 2010 (<http://www.jrf.org.uk/sites/files/jrf/poorer-children-education-full.pdf>).
- Hansen, K. And Machin, S. (2010), 'Demand for School Quality in the Early Years', Department of Quantitative Social Science Working Paper, Institute of Education.
- Ladd and Walsh (2000) 'Implementing value-added measures of school effectiveness: getting the incentives right', *Economics of Education Review*, vol. 2 part 1 pp. 1-17.
- Leckie, G. and Goldstein, H. (2009). 'The limitations of using school league tables to inform school choice', *Journal of the Royal Statistical Society: Series A*. vol. 127 part 4, pp835-52.
- Meyer, R. (1996). Value-added indicators of school performance. In E. A. Hanushek, & D. W. Jorgenson, *Improving America's schools: the role of incentives*. Washington, DC:National Academy Press.
- Meyer, R. (1997) 'Value-Added Indicators of School Performance: A Primer', *Economics of Education Review*, Vol. 16, No. 3, pp. 283-301.
- Ray, A. (2006) School Value Added Measures in England. Paper for the OECD Project on the Development of Value-Added Models in Education Systems. London, Department for Education and Skills <http://www.dcsf.gov.uk/research/data/uploadfiles/RW85.pdf>.