

IZA DP No. 1005

The Literacy Hour

Stephen Machin
Sandra McNally

January 2004

The Literacy Hour

Stephen Machin

*University College London,
London School of Economics and IZA Bonn*

Sandra McNally

London School of Economics and IZA Bonn

Discussion Paper No. 1005
January 2004

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

Email: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion 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 on the IZA website (www.iza.org) or directly from the author.

1. Introduction

Literacy matters. Many people in different countries fail to reach even basic levels of literacy. This severely hampers their individual circumstances and lowers national productivity. For example, in the UK, the Moser report (DfEE, 1999) identifies one in five adults as not being functionally literate. How can we ensure that future generations of adults do not suffer from such problems? One way is by trying to provide a means whereby literacy levels and by association overall pupil attainment can be raised through government policy.

More generally, there is a substantial body of research investigating how best to raise pupil attainment. The economic literature primarily addresses the impact of changing school resources such as class size, teacher quality and measures of expenditure (for example, Angrist and Lavy, 1999; Card and Krueger, 1992; Hanushek, 1997, 2003). Educationalists address many of the same issues (as reviewed by Sammons, 1999, or Teddlie and Reynolds, 2000), but also investigate the efficacy of what is taught in particular subjects (for example, see Stainthorp, 1996, for a review of evidence on what children need to learn to become skilled readers) and how this is put across in the classroom in terms of such issues as time on task, frequency of monitoring pupils' work and grouping arrangements (e.g. Scheerens, 2000; Creemers, 1994).¹

Some of this research appears to have influenced public policy. For example, with regard to reading, US research findings have been used as a rationale to underlie the

¹ One of the few papers in the economic literature addressing the potential effectiveness of how subjects are taught is Angrist and Lavy's (2002) paper addressing the role of Computer Aided Instruction. Another exception is the paper by Glewwe et al. (2000), which investigates the impact of flip charts in Kenyan classrooms.

provides a practical structure of time and class management for a daily literacy hour.⁴ Training material is also provided for teachers (i.e. course booklets, videotapes etc. to support three in-service training days and some further training in after-school sessions, in the first year of implementation, with continued funded training offered in subsequent years). The policy, which was introduced to all primary schools at the start of the 1998/99 school year and is still ongoing, was preceded by a project that was implemented in about 400 schools two years earlier, during the 1996/97 and 1997/98 school years - the National Literacy Project (NLP).⁵ From our perspective, the interesting feature of the NLP is that it resulted in some children being exposed to up to two years of the literacy hour in a period when other children were not.

In the United States, ‘Success for All’ has been described as a more intensive way of delivering the aims of the National Literacy Strategy (Hopkins, 1999). It has been suggested that this US program is, arguably, the most widely implemented, widely researched and widely critiqued educational program in the US (Borman and Hewes, 2002). However, like the National Literacy Project, it does not appear to have been evaluated by economists. Nor has it yet found its way into the economic literature on what can raise pupil attainment. On account of this, our assessment and evaluation of the National Literacy Project takes on an added significance - it potentially sheds more light on what can raise attainment in debates within the economics of education. Moreover, the NLP provides a better testing ground than the ‘Success for All’ program since the latter is

⁴ In this paper, we refer to these set of measures as ‘the literacy hour’.

⁵ The National Literacy Project was not explicitly viewed as a pilot for the National Literacy Strategy at the time of its introduction. However on the recommendation of the Literacy Task Force, the approaches were incorporated into the National Literacy Strategy, which all schools were encouraged to adopt, with additional funding and support, from 1998.

example, in the school year before the NLP was introduced, 49 percent of 11 year old boys achieved the government target of level 4 or above in Key Stage 2 (KS2) English, whereas 64 percent of girls achieved this target (see Table 1). There are many reasons for why this gender gap exists, and there is a very large literature on gender gaps in education - again mostly within the education field (Maynard, 2002; White, 1996; Millard, 1996), with some more recent work in economics (Jacob, 2002). One very relevant issue is whether the highly structured nature of the literacy hour had the potential to make inroads into this gender gap. If one thinks boys have a greater problem with concentration and focus relative to girls, it may well be that the NLP could differentially benefit boys. We investigate this hypothesis towards the end of the paper.

The structure of the rest of the paper is as follows. In Section 2, we discuss the National Literacy Strategy in English schools, and give details about the National Literacy Project. In Section 3, we describe the data used in our empirical analysis. Section 4 provides estimates of the impact of the NLP, both on primary school attainment in 1997 and 1998 and on secondary school attainment five years later. In Section 5, we address the question as to whether NLP had a differential impact by gender. Then in Section 6, we estimate the relationship between higher literacy skills and labour market earnings and use this to convert our estimate of the NLP impact to a monetary measure. We compare the present value of the discounted benefits per pupil to the costs and thus assess the cost-effectiveness of the NLP policy. Conclusions are presented in Section 7.

test scores for reading and writing). One sees a gradual increase in this measure over time. By 2002, the percentage achieving level 4 or above had risen to 75. Although sizable gender gaps are evident, it is interesting to note for the later analysis that between 1996 and 2002 boys improved their relative position, with a rise of 21 percentage points over this time period compared to a 15 percentage point rise for girls.

The same pattern is evident for reading. We only have statistics here from 1997 onwards. The percentage achieving at least level 4 increased from 67 to 80 percent between 1997 and 2002. Once again, although sizeable gender gaps are present at each point in time, over this period boys experienced an increase of 14 percentage points compared to an increase of 12 percentage points for girls.

The lower Panel of Table 1 provides statistics on secondary school performance in English. This shows the percentage of children who achieved a Grade C or higher in GCSE English (GCSE is the national examination system, the General Certificate of Secondary Education). Between 1996 and 2002, this measure increased from 49 to 56 percent of pupils. Again there are very large gender differences, which seem to widen out in secondary schools as compared to primary schools (Machin and McNally, 2003), although there is a small improvement in the relative position of boys over time.

3. The Data and Modelling Approach

Key Stage Attainment Data

The empirical analysis is based on administrative records of pupil-level attainment and on school-level data. The former consists of very detailed attainment information from when students were 11 and 16 years of age. At age 11, all pupils in England are tested at the

end of what is known as Key Stage 2 (i.e. the phase of education from 7-11 years, as stipulated in the National Curriculum) whereas exams at the end of Key Stage 4 (i.e. GCSE or GNVQ exams) mark the end of a pupil's compulsory education at age 16.¹⁰

The first available year of national Key Stage 2 data for pupils is in 1996, which corresponds to the school year before the National Literacy Project was introduced. We also use the national Key Stage 2 data in the 'policy on' school years of 1997 and 1998, before the National Literacy Strategy was introduced nationwide.

The National Literacy Project was introduced in two cohorts, in 1997 for the first cohort and 1998 for the second cohort. Pupils within the first cohort finished their compulsory education in 2002. Hence, to evaluate the impact of the NLP on attainment at secondary school, we match up pupil records from 1996 (i.e. pre-policy) and 1997 to GCSE/GNVQ¹¹ attainment data in 2001 and 2002 respectively.¹²

The pupil-level administrative files have detailed information on attainment, the gender of the student and codes for the primary and secondary schools attended. This latter information allows the files to be matched up to national school-level data available in the School Performance Tables and the LEA and School Information Service (LEASIS). Available information includes measures of school outcomes (results, absences), inputs (e.g. pupil-teacher ratios), disadvantage (e.g. the percentage of students eligible for Free School Meals or those identified as having Special Educational Needs) and other school characteristics (e.g. school type).¹³

¹⁰ The other two stages, not considered here, are Key Stage 1 (at age 7) and Key Stage 3 (at age 14).

¹¹ GCSE is the General Certificate of Secondary Education and GNVQ denotes General National Vocational Qualifications, the pupil qualifications taken at the end of secondary school for pupils following the academic (GCSE) or vocational (GNVQ) route.

¹² At the time of writing we do not yet have access to GCSE attainment data for 2003 and so can only look at the first cohort of children exposed to the NLP.

¹³ A full list of the detailed set of variables used as controls is provided in the notes to Tables 3 and 4.

With regard to outcome measures, we concentrate on two measures of attainment at the end of primary school: the percentile reading score and the percentage of students attaining level 4 or above in Key Stage 2 English. Reading is one of several tests students take as part of their Key Stage 2 assessment. Since the marking scheme can change over time, we convert raw scores to percentile scores. The scores of various tests are aggregated and then converted to an overall ‘level’ (which is in a range of 2-6). The key indicator of policy interest is the percentage of students attaining level 4 and above at age 11, which is the standard deemed to be appropriate at this age.¹⁴

To measure attainment at the end of secondary school, we consider the probability of obtaining a particular grade or higher for English (i.e. ranging from a low of grade G to the highest grade A*). We also consider the main overall indicator of attainment at age 16 - whether a student obtains 5 or more A*-C Grades at GCSE/GNVQ, as well as the GCSE/GNVQ points score.

Specifics of the NLP and Definition of Control Schools

As noted above, the NLP was introduced in the school year 1996/97 for the first cohort of schools and 1997/98 for the second cohort. Of about 400 schools in the project (about half in each cohort), 80 per cent were located in the inner cities – within several Local Education Authorities in London and also Sandwell, Liverpool, Manchester, Sheffield, Newcastle and Bristol. Within England these inner city areas are characterised by problems of economic and social disadvantage. Collectively, about 40 per cent of primary schools within these LEAs were involved in the NLP. The remaining NLP

¹⁴ The National Literacy Taskforce (which developed the National Literacy Strategy) advised that by 2002, 80% of 11 year olds should reach Level 4 in Key Stage 2 English. In fact, as the numbers in Table 1 show, this target has not been met.

schools were in eight LEAs in three counties (Hampshire, Essex and Norfolk).¹⁵ Only about 7 per cent of primary schools in these counties were involved in the NLP.

Local Education Authorities (LEAs) were involved in the administration of the NLP within the treatment schools. The Literacy Task Force (1997b) identified LEAs as the ‘first link in the chain’ of implementation, being expected to provide a strong lead to all their schools, profiling literacy as a priority for the LEA with publicity, producing visible targets and so on.

This matters a lot for the definition of control areas. It seems possible that all primary schools within an LEA taking part in the NLP may have (at least indirectly) benefited from the initiative – and not just those schools taking part in the project. For example, the NLP policy was well advertised within LEAs. Therefore, it is highly questionable whether ‘non-NLP’ schools within NLP LEAs should be used as controls for the treatment schools.

However, there is also a good case for not using all other primary schools in the country as a control group: there are over 14,000 in England, and these are likely to be heterogeneous in ways we cannot account for with available data.¹⁶ Thus, we adopt the following strategy. We first identify LEAs that are in areas geographically adjacent to LEAs involved in the NLP. Then, if multiple non-NLP areas are present, we choose those with the closest educational performance indicator in the pre-policy period.

This approach has a lot of similarities to the approach adopted in evaluations of recent UK area-level initiatives (such as the Educational Maintenance Allowance in

¹⁵ The precise LEAs are: Hampshire, Portsmouth, Southampton, Essex, Southend-on-Sea, Thurrock, Isle of Wight, Norfolk.

¹⁶ However, we do have very detailed school-level controls. Thus, as a robustness check, we later estimate regressions using the full sample of schools.

It is interesting to place this finding in the context of the national roll-out during the school year 1998/99 and with the national figures given in Table 1. It is evident that the gender gap in primary school reading and English attainment has reduced in recent years. The results we report here are entirely consistent with the literacy hour playing an important role.

Gender Differences at Secondary School

The lower panel of Table 6 reports gender differences in GCSE English, with and without KS2 controls. The picture here is patchier, but when one does not control for KS2 performance, NLP effects for boys are mostly higher than for girls (though even this differs for the difference-in-difference and matched estimates). All in all, at GCSE level, the evidence for (statistically different) gender gaps is not strong. Moreover, anything that is there disappears after inclusion of the KS2 controls, suggesting that the literacy hour impact on gender differences occurred more at primary level - namely at the time when the extra exposure to the literacy hour occurred.

6. Measuring Economic Costs and Benefits

This analysis has shown a significant impact of the literacy hour on reading and on English attainment. The question remains as to whether the policy was cost effective. Hence we now compare the per pupil costs of the policy with the economic benefits, as reflected in predicted labour market earnings.

The planned cost of the NLP was £12.5 million over 5 years. The main costs were 14 local centres (each costing about £25,000 per year) and literacy consultants in each participating Local Education Authority (about £27,000 per year for each consultant).

Schools also received some funding for teacher training and resources, which was broadly the same for each school (though some account was taken of the pupil-teacher ratio). However, since the national roll out took place two years after the NLP was introduced, only the first two years are relevant. The total cost per annum was thus £2.5 million (or about £2.8 million in 2001 prices). We observe the number of students affected from pupil numbers in the schools within Cohorts 1 and 2 in 1997 and 1998 (i.e. 222,261 pupils in aggregate).¹⁹ Hence the cost per pupil is £25.52 per annum.

To estimate benefits of the policy, we first convert the impact of the policy on reading scores (i.e. 2.63 percentiles, as shown in Table 3, column 3) to an equivalent estimate in terms of standard deviations. This is calculated as 0.091 standard deviations. Secondly, we estimate the impact of reading scores on future labour market earnings using the British Cohort Study. This is a panel survey of all those living in Great Britain who were born between 5th and 11th April 1970. We regress the log of labour market earnings (at age 30, in 2000) on age 10 percentile reading scores (from 1980).²⁰ Results are shown in Table 7. We show three specifications. In column (1), controls are included for gender and region; in (2) we add controls for family background; and in (3) we include dummy variables for the participant's highest educational qualification achieved by age 30. Since the latter variable is likely to partly capture the effect of the reading score, the effect of reading on labour market earnings in column (3) should be considered as a lower bound estimate (or even an under-estimate).

¹⁹ This includes infant schools.

²⁰ The reading test is a shortened version of the *Edinburgh Reading Test*, which is a test of word recognition. It examines vocabulary, syntax, sequencing, comprehension and retention (see Godfrey Thompson Unit, University of Edinburgh, 1978, or Plake and Impara, 2001, for more details).

The estimates are .54, .42 and .20 in each of the respective columns and are always statistically significant, showing a higher standard of reading at age 10 to be associated with higher earnings at age 30. The earnings impact of a .091 increase in the standard deviation (sd) of reading percentiles is then calculated as $.091 * sd * [\exp(\text{reading score coefficient}/100) - 1]$. This amounts to an annual sum of £196.32, £154.23 and £75.40 for each specification. Assuming that labour market participation occurs between the age of 20 and 65, and using a discount rate of 3 per cent, the corresponding present discounted value of the cumulative effect of the literacy hour is estimated as £5,476, £4302 and £2103.

There are, as with all calculations like this, certain issues that may make one question these economic benefits. One clear example in this case is that the beneficiaries of the NLP literacy hour tended to be children from less well performing schools. On average these children are likely to be located further down the reading score distribution, yet the earnings effect from age 10 reading is assumed linear in the regressions in Table 7. We have therefore considered economic benefits from parametric models where we allow separate effects on earnings for the top and bottom half of the reading score distribution and from non-parametric regressions (using the Nadaraya-Watson estimator). In both cases the linearity assumption seems reasonable: the bottom half effect for the column (2) specification in Table 7 is estimated to be .458 (standard error = .100) as compared to the top half effect of .435 (.084); a non-parametric regression is shown in Appendix Figure A4.

Whichever way one looks at it, the benefits of the literacy hour seem to be sizeable and the costs are much smaller. Even if we take the smallest impact estimate

from our analysis (the 1.72 percentile improvement in column (4) of Table 3, which corresponds to a .06 standard deviation increase), the economic benefits are measured in the range of £1375 to £3581.

The benefits of the policy (a .06 to .09 standard deviation in reading scores) seem comparable to more expensive programs like reducing class size. For example, Krueger and Whitmore (2001) found a class size reduction in the STAR program to lead to an increase in test scores of .13 standard deviations. Rivkin, Hanushek and Kain (2002) suggest that having a teacher at the higher end of the quality distribution raises student achievement by at least .11 standard deviations. However, the costs of such programs are more substantive. They are simply not comparable to this apparent low-cost literacy hour program of changing teacher practice. Of course, the financial costs of the program may not reflect its true resource cost. For example, it might be argued that the measured costs do not reflect any extra effort the teacher might have to put in to learning and implementing the new teaching method. On the other hand, such effort may be fully accounted for in the cost of training. Furthermore, there are reports of a very positive response by teachers (e.g. Fisher and Lewis, 1999, or Smith and Whitely, 2000), who find the learning objectives and structure of the literacy hour to provide a clear focus for what they teach.

One might also argue that the literacy hour takes teaching effort and resources away from other subjects and that this indirect cost effect (via substitution) should be taken account of in a cost-benefit calculation. However, it seems likely that literacy was being taught in some form before the policy, for a commensurate time period. Therefore, the literacy hour represents a change in how reading and writing are taught, rather than

merely an increase in the time devoted to the subject. One might also suspect that the literacy hour could lead to positive spillovers due to complementarities between pupil subject areas and associated teacher practice. Firstly, since the ability to read and write are important generic skills, an improvement in how these skills are taught might lead to improved performance in other subjects. Secondly, the literacy hour might have caused teachers to re-evaluate their teaching methods in other subjects. It may be that aspects of the approach are transferable to the teaching of other subjects (e.g. the highly structured nature of lessons; the balance between group work and whole class activity). This is important in English primary schools because generally pupils within a particular year group are taught every subject by the same teacher.

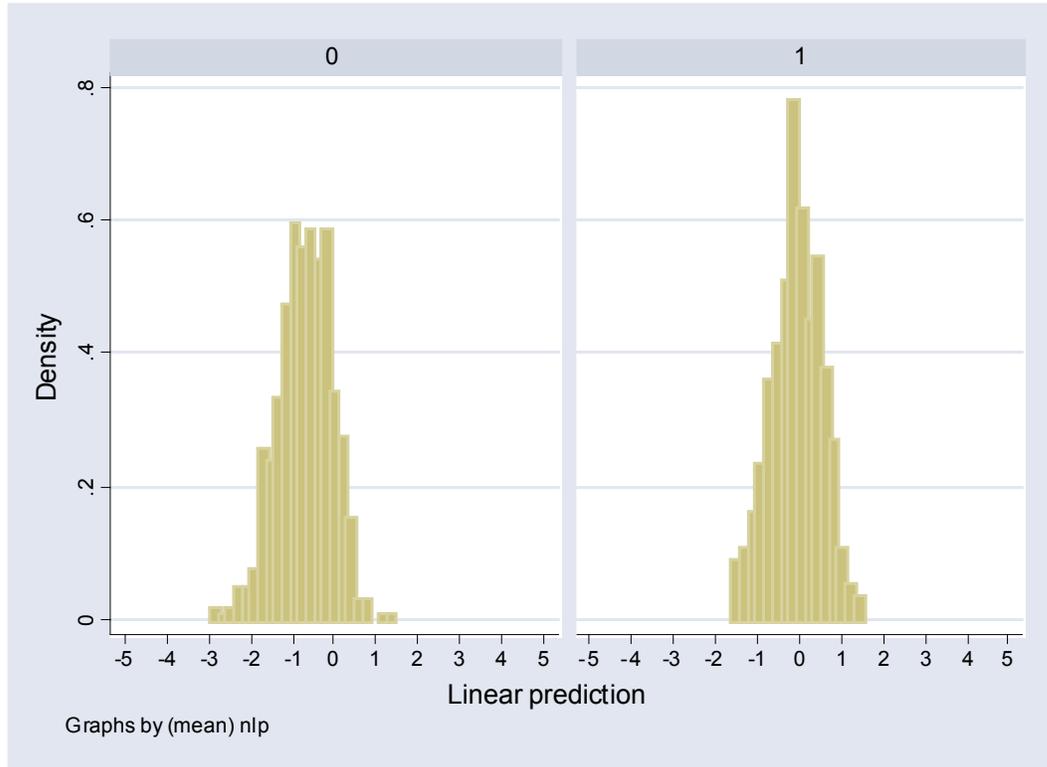
In Table 8 we consider the possible substitution/complementarity with other subjects and report a test of whether the literacy hour helped or hindered attainment in Mathematics. To do this, we simply re-estimate regressions reported in Table 3, where the dependent variables are now respectively, the percentile Mathematics score and whether the student obtains level 4 or above in Mathematics. The first three columns suggest a positive impact of the literacy hour. Interestingly, the magnitudes of the estimates are smaller, at around three-fifths of the impact on English (in Table 3). Including all controls and school fixed effects (column 3), NLP schools show higher scores of 1.53 percentiles and a .025 higher proportion achieving at least level 4. Also, when estimating these regressions using the matched sample of schools (column 5), these coefficients reduce to .838 and .014 respectively (about half of the effects for English seen in Table 3) and become statistically insignificant. Nonetheless, these results suggest, if anything, a complementary impact of the literacy hour on English and Mathematics

children affected by the NLP. Since significant gender gaps in English performance exist (in favour of girls), we consider whether the literacy hour had a differential impact by gender. We find some evidence that, at age 11, boys benefited more than girls. Finally, we show the benefits of the literacy hour to easily exceed the costs of the policy.

These findings are of considerable significance when placed into the wider education debate about what works best in schools for improving pupil performance. As suggested in the introduction, the research remit of economists has tended to be rather narrower than that of educationalists in this area. Our approach shows that one of the areas receiving much less attention from economists in the past, namely the content and organisation of what is taught, matters for attainment in English and reading. Indeed, as the effects we identify come from a government policy aimed at improving literacy - the literacy hour - the evidence we report suggests that public policy aimed at changing the content of teaching can significantly raise pupil attainment and can do so in a cost effective manner.

- Heckman, J., H. Ichimura and P. Todd (1997) Matching as an Econometric Evaluation Estimator, Review of Economic Studies, 65, 261-294.
- Herman, R., D. Aladjem, P. McMahon, E. Masem, I. Mulligan, A. O'Malley, S. Quinones, A. Reeve and D. Woodruff (1999) An educators' guide to schoolwide reform, Arlington, VA: Educational Research Service.
- Hopkins, D. (1999) Success for All – Part One, Literacy Today, 21.
- Jacob, B. (2002) Where the Boys Aren't: Non-cognitive Skills, Returns to Schooling and the Gender Gap in Higher Education, National Bureau of Economic Research Working Paper 8964.
- Krueger, A. and D. Whitmore (2001) The Effect of Attending a Small Class in the Early Grades on College-test Taking and Middle School Test Results: Evidence from Project Star, Economic Journal, 111, 34-63.
- Literacy Task Force (1997a) A Reading Revolution: How We Can Teach Every Child to Read Well, London: The Literacy Task Force c/o University of London: Institute of Education.
- Literacy Task Force (1997b) The Implementation of the National Literacy Strategy, London: Department for Education and Employment.
- Machin, S. and S. McNally (2003) Gender and Educational Attainment, Centre for the Economics of Education, unfinished draft.
- Maynard, T. (2002) Boys and Literacy: Exploring the Issues, RoutledgeFalmer: London.
- Millard, E. (1996) Differentially Literate: Boys, Girls and the Schooling of Literacy, Falmer Press: London.
- Office for Standards in Education (1996) 'The Teaching of Reading in 45 Inner London Primary Schools', A report by Her Majesty's Inspectors in collaboration with the LEAs of Islington, Southwark and Tower Hamlets, London: Ofsted.
- Office for Standards in Education (1998), The National Literacy Project: An HMI Evaluation, London: Ofsted.
- Office for Standards in Education (1999) 'Primary Education 1994-1998: A Review of Primary Schools in England', London: Ofsted.
- Plake, B. and J. Impara (Eds.). (2001) The Fourteenth Mental Measurements Yearbook, Lincoln, NE: Bureau of Mental Measurements.

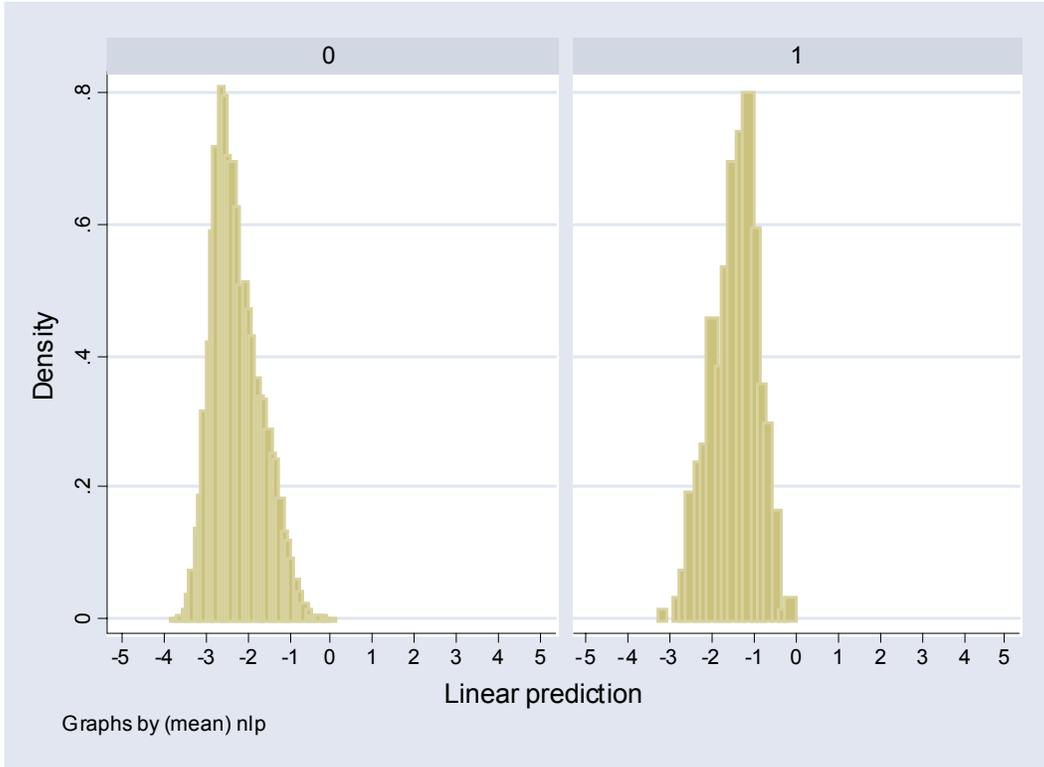
Figure A1:
Propensity Scores for NLP and non-NLP Schools,
For Table 3 Specifications



Selected for 'matching' specification:

Schools with a predicted linear index of the propensity score between -1.6 and 0.8.

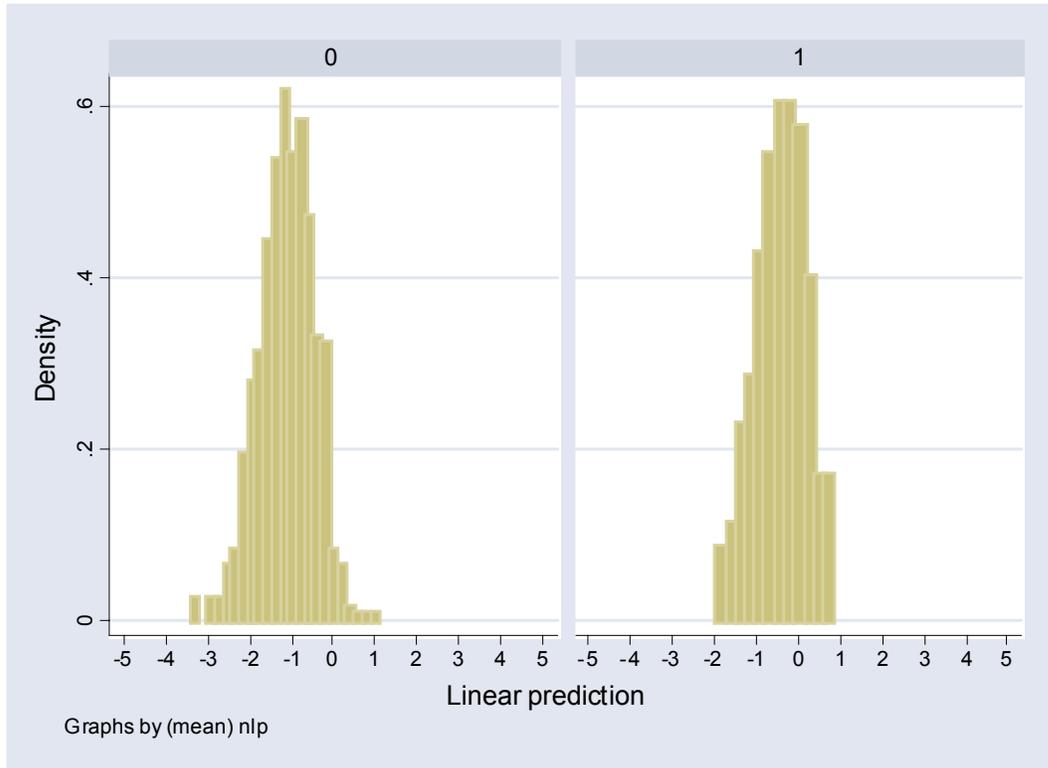
**Figure A2:
Propensity Scores for NLP and non-NLP Schools,
for Table 4 Specifications**



Selected for ‘matching’ specification:

Schools with a predicted linear index of the propensity score between -2.9 and -0.43.

**Figure A3:
Propensity Scores for NLP and non-NLP Schools,
For Table 5 Specifications**



Selected for ‘matching’ specification:
Schools with a predicted linear index of the propensity score between -2 and 0.6.

