Early interventions and children's educational attainment. Evaluating the impact of free part-time pre-school education for 3 year olds in England

J. Blanden[*], E. Del Bono[**], K. Hansen[†], S. McNally[*], and B. Rabe[**]

[*]University of Surrey and CEP, London
[**]ISER, University of Essex
[†]IoE, University of London

December 17, 2014

Abstract

In this paper we analyse whether entitlement to a free part-time place in a nursery or other registered setting at 3 years of age affects educational attainment in primary school. Our identification strategy exploits date-of-birth discontinuities in eligibility for free pre-school. In England, children are entitled to a free part-time pre-school place the term after they turn 3, while they usually start school in the month of September of the academic year in which they turn 5. This implies that some children born just a few days apart are entitled to different amounts of pre-school education (up to 3 months) while starting school at the same time and within the same cohort. Using a regression discontinuity approach and analysing administrative data on all state school pupils in England, we estimate that eligibility to free part-time pre-school education results in a small overall educational advantage for boys at age 5, but not for girls. Analysis by subgroups reveals that the effects of eligibility are unequally distributed across children from low and high income families, and contribute to increase the socio-economic gap for both boys and girls at age 5 and (to a lesser extent) at age 7.

Keywords: childcare, child outcomes, regression discontinuity
JEL codes: I22, I24, C21

Corresponding author: Emilia Del Bono, edelbono@essex.ac.uk. We thank the Department for Education for making available, under terms and conditions, data from the National Pupil Database. We would also like to thank seminar participants at the the University of Hamburg, the University of Essex and DIW (Berlin) for comments and suggestions. This work was funded by the ESRC under grant ES/K003232/1 and by the Nuffield Foundation.

1 Introduction

Understanding the impact of early childhood investments on children's outcomes is a question of considerable interest in social science. Neurological and psychological studies suggest that learning is easier in early childhood than later on in life (Shonkoff and Phillips, 2000), while economists point out that the returns to early investments are likely to be high simply because of the longer pay-off period (Becker, 1964). More recently, it has been argued that a child's developmental process might be characterised by "dynamic complementarities", where early learning makes subsequent learning more productive (Carneiro and Heckman, 2004). Some studies have also gone as far as to suggest that early investments are an effective way of reducing the cognitive gap between children from different socio-economic backgrounds (Knudsen et al., 2006).

Due to these arguments, many European governments have launched initiatives aimed at increasing the financial support for childcare, particularly for pre-school children. This support may consist of cash subsidies and tax breaks, which reduce the cost of childcare, or an expansion in the supply of places through direct provision or subsidies directed to private providers. As pre-school childcare is also thought to enable more parents to work (Brewer and Crawford, 2010; Fitzpatrick, 2010; Berlinski et al., 2011), many recent programmes implemented are *universal* rather than targeted at the most disadvantaged families. This makes these policies quite expensive. In the UK, for example, it is estimated that partially funded childcare places for all 4 and 3 year olds cost the government in excess of 2 billion per year (Department for Education, 2013).

In this paper we provide an evaluation of the latter policy, and in particular we ask whether part-time pre-school education for all 3 year olds can raise early educational attainment of children in England. We further ask whether there is any evidence to suggest that pre-school attendance has larger benefits for the most disadvantaged children, thus contributing to socio-economic mobility. Existing evidence from Merrell and Tymms (2011) on the performance of English children at the start of school shows no improvement in cognitive development since the early 2000s - when the provision for 4 year olds was already universal, while the provision for 3 year olds was still being rolled out - and no closing of the gap between children from different social backgrounds. However, to date there has be no formal evaluation of this policy.

The main problem with assessing the effects of pre-school childcare on later educational outcomes is that as long as childcare attendance is *voluntary* some parents are more likely to enroll their children than others, and these parents (or their children) may have unobservable characteristics which are correlated with education outcomes. Standard econometric methods are unlikely to account for this selection issue, and may lead to biased estimates of the effects of interest. Our paper overcomes this problem by exploiting strictly enforced date-of-birth rules which govern access to pre-school childcare in England and which imply that some children born just a few days apart are entitled to different amounts of pre-school education (up to 3 months) while starting school at the same time and within the same school-cohort.

In England all children enter primary school in the academic year in which they turn 5 (reception). While in the past many schools operated different intake policies, usually allowing younger children to start later in the year, in more recent years most schools have adopted a unique intake date in September. This implies that irrespective of their date of birth, all children within a school-cohort (going from September to August) start formal schooling at the same time (but at a different age). By contrast, eligibility to free part-time pre-school care changes discontinuously across the year; children born between 1st September and 31st December are entitled to claim their free hours as from the following January, children born between 1st January and 31st March as from April, while those born between 1st April and 31st August are allowed to claim their entitlement only from September of the following school year. This entitlement initially consisted of 12.5 hours of free pre-school childcare for 33 weeks of the year, this was later increased to 15 hours for 38 weeks, and by April 2004 it was universal for all 3 and 4 year olds.

Our analysis is based on the National Pupil Database (NPD), a longitudinal register with data on all pupils in state schools in England. We gained access to the precise date of birth for several cohorts of children who started school over the last few years. Our identification strategy is based on comparing the educational outcomes of children born a few weeks before and a few weeks after 31st December and 31st March cutoff dates. These two dates define eligibility for free part-time pre-school education, and do not imply a different starting school date.[1] Our outcomes are measured at age 5, at the end of reception, and at age 7 (Year 2). Since children are evaluated at the same point in time, and many studies show the existence of a statistically significant relationship between birth date and test scores (Crawford et al. 2010 and 2011), all our specifications control for a flexible function of date of birth.

[1]The third cut-off date for free pre-school eligibility, 31st August, creates a discontinuity in the amount of pre-school received but also separates children into different school cohorts and is therefore excluded from our analysis.

We find that eligibility to free part-time pre-school education results in a small overall educational advantage for boys at age 5, but no significant impact among girls. Analysis by subgroups shows that the effects of eligibility are unequally distributed across children from low and high income families, and contribute to increase the socio-economic gap for both boys and girls at age 5 and (to a lesser extent) at age 7. Our data does not contain information on actual pre-school enrollment or hours of childcare. This means that we cannot directly estimate the effect of pre-school education on children's attainment. As data from a recent cohort of English children (Millennium Cohort Study) indicates that most children (> 50%) receive formal childcare well before their entitlement date, and that the eligibility rule has a small impact on actual enrollment, our estimates could be interpreted as a lower bound of the effect of pre-school education on early attainment.

2 Related literature

The literature on the effects of early intervention programs and pre-school programs is very rich and extremely interesting. Many of the first studies were conducted in the US, and looked at the effects of small randomized early interventions such as the Perry Preschool project from Ypsilanti, Michigan, or the Abecedarian project from Chapel Hill, North Carolina. Both these projects are thought to have had positive effects on children's attainment during the school years, and there is evidence that these gains extended into adulthood. Non-randomized larger scale programs such as Head Start have also received a lot of attention. Here the general consensus is that the program improved outcomes in the short-run, but the long-term effects are more debated (see Blau and Currie, 2006, for a survey). All these programs have one distinctive feature in common, they are *targeted* at children from disadvantaged families.

There is a smaller, but rapidly growing, literature on large scale, publicly funded *universal* pre-school programs targeted at *all* children. Most of this literature focuses on short-term outcomes, with findings which are rather mixed. For example, Gormley and Gayer (2005) find large positive effect of a pre-K program implemented in Oklahoma on cognitive scores and language scores of children in the following year. Positive effects of child care on children's short-run outcomes are also found by Fitzpatrick (2008) in the US, Melhuish et al. (2008) in the UK and Berlinski et al. (2009) in Argentina. On the other hand, Baker et al. (2008) and Herbst and Tekin (2008) find negative effects of the introduction of subsidised universal childcare in Quebec and of subsidies for child care provided to working mothers in the US, respectively, on various measures of early attainment but in particular on aspects of social and emotional development. More recently, Datta Gupta and Simonsen (2010) find no effects of pre-school enrollment on outcomes at age 7 in Denmark. Long-term effects are more rarely analysed. Some evidence of positive impacts comes from Berlinski et al. (2008) for Argentina, and Cascio (2009) for the US, while Drange et al. (2012) find that a compulsory kindegarten program for 5-6 year olds had almost no impact on achievement at the end of compulsory schooling in Norway.

There is also evidence pointing out that the effects of these programs might be heterogeneous across the population, with some indication that disadvantaged groups stand to benefit more. So, for example Gormley and Gayer (2005) show that the most positive effects of the pre-K program in Oklahoma were felt by Hispanics and Blacks, while for the white majority the impacts were largely non-significant. Havnes and Mogstad (2011) find that an increase in publicly provided child-care in the 1970s in Norway had larger long-term benefits for children of low educated parents, while Dustmann et al. (2013) document large effects of free public child-care places on school readiness measures of children from immigrant families, with no effects on children of natives in Germany. However, we must note that in all these cases the treatment consists of the free provision of publicly provided *childcare places*, which are presumably rather homogeneous in terms of quality.

The contribution of this paper with respect to the existing literature is to evaluate whether the provision of a universal *subsidy* towards the cost of pre-school child-care has an impact on the early educational outcomes of all children in the UK, and whether the effects are different across different socio-economic groups. The key difference with respect to previous work in this area is that in our case the subsidy could be used to purchase a certain number of hours of child-care in settings of potentially different *quality*, it could be used to purchase additional hours (effects at the intensive margin), or it could be used simply as an unconditional cash benefit. To our knowledge there is only one other paper in the literature which asks a similar question. Black et al. (2012) analyse the effect of child care subsidies in Norway and find that these had little impacts on children's pre-school attendance and maternal employment, but statistically significant effects on academic performance in junior school, thus suggesting the relevance of income effects. In that study the receipt of the subsidy is based on an income threshold, and this implies that it is not possible to examine whether the impacts are different for disadvantaged children. Our study is based on a different identification strategy, where exogenous variation in entitlement to the subsidy is based on date of birth rules. We will thus be able to analyse whether the response to the policy differs by family income and therefore assess whether one of the stated objectives of this intervention - namely to reduce early socio-economic gaps in achievements - can be effectively reached.

3 Institutional background

In England the move towards *universal* pre-school education began in the late 1990s. Prior to this, Local Education Authorities (LEAs) had complete choice over whether they should offer nursery education to 3 and 4 year olds and how to target their funding. This could be done through nursery classes attached to schools, or more rarely, through stand-alone nursery schools. This type of 'maintained' provision was more common in urban areas with Labour controlled councils and was usually targeted at low-income households in receipt of means-tested or unemployment benefits.

In 1996 the Conservative administration put in place a scheme which entitled parents of 4 year old children to vouchers for nursery education. The idea was to make more places available by subsidizing private provision. In 1997 the incoming New Labour Government replaced nursery vouchers with an LEA-administered Nursery Education Grant (NEG). Compared to the previous voucher scheme, there were more conditions attached, primarily concerning the educational content of children's experiences. In 1999-2000 the NEG was extended to all 3 year olds in 65 pilot LEAs. By 2004 every LEA in England was in receipt of the grant, and all children in England were entitled to 12.5 hours of free child care per week for 33 weeks per year as from the term following their 3rd birthday. In 2010 the entitlement was extended to 15 hours per week and 38 weeks per year, while flexibility was also increased so that the free hours could be pooled over a minimum of two days.

As noted above, even before 1997 existing maintained provision was already accommodating disadvantaged 3 year olds in some areas. This type of provision remained fairly constant, while the main expansion occurred in the Private, Voluntary or Independent (PVI) sector. For instance, while the percentage of free places assigned to 3 year olds in the maintained sector increased from 36.2% in 1999 to 38.3% in 2007, PVI coverage went from 0% in 1999 to 49% in 2007. Interestingly, the increase in free places through the NEG did not generate a one-to-one increase in the total number of places; between 2000 and 2007 the total proportion of 3 year olds in childcare increased by just 13.4 percentage points.[2]. This indicates that a considerable proportion of PVI sector childcare was substituted for free PVI places, that is the 'place' was the same as before the introduction of the

[2]Data on all places is not available for 1999.

subsidy, with the only difference that funding was now from public rather than private funds. We exploit the expansion over the period between 1997 and 2007 in a companion paper (Blanden et al., 2014).

Although the policy entitles all children to the same number of hours, the type of early education experience the child will have will vary depending on where they take up their place. Maintained provision will usually be more restrictive in terms of hours available, often either five mornings or five afternoons, and usually will not extend outside school hours. Nursery schools and classes (which are to be found in the maintained sector) require that a qualified teacher is present, and have an adult-child ratio of 1:13. Requirements for qualifications are lower in PVI settings, but if there is no qualified teacher present then the ratio of adult per child is increased to 1:8 (Gambaro et al., 2013). There is also substantial variation within the PVI sector. Private and independent day nurseries focus on full-time care, so that often the entitlement acts only as a discount on fees, with few 15 hour places available. Voluntary pre-schools, which evolved from community play-groups, often offer care over more restricted hours. If maintained settings are not available in the area it is likely that those not already in childcare will take up their entitlement at a voluntary pre-school.

Although the structural aspects of provision vary, all providers are required to follow a common curriculum. Since 2008 this is the Early Years Foundation Stage, and prior to this, the (very similar) Curriculum Guidance to the Foundation Stage. The curriculum emphasises learning through play, ensuring that a range of stimulating activities are provided and that children's development across a range of areas is monitored and encouraged. Moreover, all settings are subject to inspection by the Government regulator OFSTED (Office for Standards in Education), although it should be noted that nursery classes and schools are assessed as part of the school assessment.

Given that a focus of our work is on how the impact of entitlement to pre-school education varies across children from different backgrounds, it will be very important to document whether and how the *quality* of provision differs according to the socio-economic characteristics of the family. The evidence in Gambaro et al. (2013) would seem to suggest that disadvantaged children are likely to have access to more qualified staff, but that they tend to be more segregated and in nurseries who receive lower quality scores in OFSTED inspections.

4 Empirical strategy

The parameter of interest

Empirically, we can formalize the relationship between pre-school education and educational attainment, measured for example by test scores, by the system of equations:

$$Y_i = \theta S_i + \Pi X_i + u_i \tag{1}$$

$$S_i = \Gamma X_i + \mu_i \tag{2}$$

where Y_i is the test score for child i, S_i is an indicator variable for whether the child attends pre-school, X_i is a vector of observed characteristics of the child, u_i captures the unobserved determinants of attainment, and μ_i captures the unobserved determinants of participation in pre-school education.

In order to obtain an unbiased estimate of the parameter of interest θ, it must be that $E[u_i\mu_i|X_i] = 0$. If this is not the case, and the unobserved determinants of whether a child attends pre-school and the unobserved determinants of educational attainment are correlated, standard analysis will yield biased estimates of the impact of pre-school on test scores.

The typical solution to this problem is to find a variable T_i which is correlated with pre-school attendance, or $E(S_iT_i|X_i) \neq 0$, but is not correlated with the unobserved determinants of test scores, $E[T_iu_i|X_i] = 0$. Our system will thus become:

$$Y_i = \theta S_i + \Pi X_i + u_i \tag{3}$$

$$S_i = \gamma T_i + \Gamma X_i + \nu_i \tag{4}$$

As we have discussed in section 2, in recent years in England access to free part-time pre-school education was based on strict date-of-birth rules. There is no reason to expect that the cut-offs which define eligibility are directly related to children's test scores, nor that they are systematically related to unobserved determinants of test scores. Therefore one variable which could be used to identify the effect of S_i on Y_i is an indicator variable which assumes value 1 when the child is born before a certain cut-off date $\bar{t}$ (and is therefore entitled to start at $\bar{t}+1$) and 0 otherwise.

This variable effectively defines eligibility to free part-time pre-school childcare over the next school term; we will call it T_i, and could write it as:

$$T_i = I\{t_i < \bar{t}\} \tag{5}$$

where t_i is the date of birth of the child.

However, in our data we do not observe whether a child attends or not pre-school, that is to say S_i is unknown to us. What we can estimate instead is a *reduced form* equation of test scores on eligibility to pre-school, T_i:

$$Y_i = \beta T_i + \Pi X_i + \epsilon_i \tag{6}$$

It is clear that (6) differs from (1) in various respects. First of all the parameter we estimate is not θ, the effect of pre-school attendance on test scores, but β, the effect of being eligible to attend pre-school in the following term on test scores. To the extent that there is not a perfect correlation between eligibility and attendance, β will represent an *intention to treat* effect. The second problem is that eligibility may determine whether a child (i) attends pre-school as opposed to stay at home with relatives or parents, but may also affect (ii) the number of hours of pre-school education a child is exposed to, (iii) the type of setting a child attends (if for example on becoming eligible a child is transferred from a maintained or public place to a place in the private sector), and possibly (iv) the disposable income of the family. The effect we can estimate using (6) is an *intention to treat effect* where the treatment is a combination of all the different channels through which eligibility may affect the outcome. Additional analysis will be needed in order to determine which channel is more likely to be correlated with eligibility before offering further insights into the economic relevance of the effects we estimate. In what follows we will simply refer to the estimated parameter β as the effect of eligibility on child educational outcome, stressing the fact that this is a parameter of interest *per se* if we are interested in the effect of the policy.

The identification strategy

There is one main empirical challenges to estimating (6) which we need to address. Our variable of interest, eligibility, is a function of date of birth. In a system like the English educational system, where children start school at the same point in time but at a different age and take their exams on the same day, date of birth determines age at test. As there is a well-documented positive relationship between age at test and test scores (Crawford et al., 2010; Leuven et al., 2010), simply introducing an indicator variable for eligibility status will not be sufficient to ensure the parameter we estimate is the effect of eligibility only. Eligible children will be on average older than non-eligible children and therefore will have (by virtue of the positive relationship between test scores and age at test) better outcomes.

Our identification strategy is based on a *Regression Discontinuity Design* (Thistlethwaite and Campbell, 1960; Imbens and Lemieux, 2008), i.e. on the idea that while the relationship between age at test and child outcomes is continuous in age, eligibility changes discontinuously at specific cut-off dates. So, for example, the test scores of a child born on 31st December should not be very different from the test scores of a child born on 1st January as their age is virtually the same, but since the child born on 31st December will receive one additional term of free part-time pre-school education with respect to the child born on 1st January if we compared these two groups of children we should be able to identify the effect of eligibility on the outcome. However, even with a very large dataset comparing children born over just 2 dates of birth will not yield sufficient statistical power, so we will have to include in our empirical analysis children born further away from the cut-off dates. This implies that we cannot assume that differences in age at test are irrelevant, but we will have to control for these differences by introducing in our equation a term which is a flexible function of date of birth.

In line with other studies which have adopted a similar identification strategy, and in particular with the work by Gormley and Gayer (2005), our main equation is specified as follows:

$$Y_i = \beta T_i + \delta f(t_i) + \Pi X_i + \epsilon_i \tag{7}$$

where $f(t_i)$ is alternatively specified as: (i) a linear function of date of birth, (ii) a quadratic function of date of birth, (iii) a linear function of date of birth where this function is allowed to have a different slope before and after the cutoff, and (iv) a quadratic function of date of birth where this function is allowed to have a different slope before and after the cutoff. In formulas:

$\mathrm{f}(\mathrm{t}_i) = t_i$

$f(t_i) = t_i + t_i^2$

$f(t_i) = |t_i - \bar{t}|$

$f(t_i) = |t_i - \bar{t}| + |t_i - \bar{t}|^2$.

Our data includes information on two different cut-off dates: 31st December and 31st March. The effect of eligibility might be different by cut-off, so in our analysis we will additionally include interaction terms between the eligibility indicator, T_i, and the March cut-off, C_i, as well as interactions between the cut-off and the trend in date of birth.

5 Data

Our analysis is based on data from the the National Pupil Database (NPD). This is an administrative dataset containing information on all children attending state (public) schools in England, and covers about 93% of all pupils in the country. The dataset combines individual data on attainment and a limited set of characteristics of the children which are recorded by the school. The dataset is longitudinal, in that it follows each child over the primary and secondary school years, and contains school and Local Education Authority (equivalent to districts in the US) identifiers.[3]

Educational attainment measures

We focus our analysis on children's educational attainment at age 5 and 7. This is partly because the policy was only recently implemented and later outcomes for the cohorts affected are not available, but also because it seems most appropriate to start looking for the effects of pre-school education on early measures of achievement rather than long-term outcomes.

Primary school in England starts with the reception year, which children usually begin at age 4 in the academic year in which they turn 5. From birth to the end of reception year, at age 5, the Early Years Foundation Stage sets standards for the learning, development and care of children in schools and pre-school settings. At the end of reception year children are assessed by their teacher according to the Foundation Stage Profile. This measures achievements of children against 13 assessment scales, each scale having a range between 1 to 9. The 13 assessment scales are grouped into six areas of learning: personal, social and emotional development; communication, language and literacy; problem solving, reasoning and numeracy; knowledge and understanding of the world; physical development, and creative development. Children who score between 1 and 3 points are

[3]There are around 150 Local Education Authorities which are responsible for setting admissions policies in England.

deemed to be working "towards" the Early Learning Goals (ELGs), children scoring between 4 to 8 points are working "within" the ELGs, a score of 6 or above is considered as working "securely within" the ELGs and a 9 point score is regarded as working "beyond" the ELGs. Moreover, there are 2 additional summary measures used by the Department for Education. Children with 78 or more points across the 13 assesment scales are considered to be developing well, and children with 78 or more points and at least 6 points in all assessment scales entering the areas of literacy and of social development are deemed to show a good level of development (Department for Education, 2012).

The outcome measures we use in our analysis of age 5 educational outcomes are inspired by the thresholds identified above, although our terminology differs somewhat. We use binary indicators assuming value 1 when the child scores: (i) 6 or above (*expected level*) in language and literacy (*literacy*), (ii) 6 or above in problem solving, reasoning and numeracy (*numeracy*), (iii) 6 or above in both *literacy* and *numeracy*, (iv) 6 or above in all six areas of learning, and (v) 6 or above in all six areas of learning and 9 in at least one area. [4]

At age 7, we analyse the results of the first test of attainment which the children take at the end of Year 2 (Key Stage 1).[5] The test consists in assessments in English, Mathematics and Science, where English has sub-assessments for Reading, Writing and Speaking and Listening. In line with previous work, we focus on the outcomes in Reading, Writing and Mathematics, as these have been shown to be the best predictors of long-term outcomes (Crawford and Cribb, 2013; Machin and McNally, 2008). Grading is by levels, with possible outcomes ranging from "W" (working towards) to levels 1 to 3, higher levels are rarely reached in Year 2. The *expected level* of attainment at age 7 is level 2B (for recent cohorts level 2 is subdivided into 2A, 2B and 2C), and children working at level 3 or higher are often considered to be highly able in that subject.

Our outcome measures at age 7 consist of a set of binary indicators assuming value 1 when the child: (i) achieves level 2B in reading, (ii) achieves level 2B in writing, (iii) achieves level 2B in mathematics, (iv) achieves level 2B in reading, writing and mathematics, and (v) achieves level 2B in all these subjects and level 3 in at least one subject.

---

[4] We also use standardized point scores as outcomes, where standardization is by gender and academic year.

[5] The period between Year 1 (the year after reception) and Year 11 is divided into four Key Stages. Pupils are assessed against the National Curriculum at the end of Year 2 (Key Stage 1), at the end of Year 6 (Key Stage 2), at the end of Year 9 (Key Stage 3) and at the end of Year 11 (Key Stage 4).

Although the construction of the output measures at age 7 resembles those used at age 5 (as we distinguish between achieving the expected level in a subject/area or going beyond it) we should note that the age 7 grading scales are very different from the grading scales used at age 5 and there is no relationship between the two. Most importantly, the age 7 tests are formal tests taken on a specific day and assessed by a teacher outside the school, while the age 5 performance measures are the result of the ongoing assessment of the class teacher. This implies that results at age 5 and age 7 will not be comparable, at least not in a quantitative sense.

Other observable variables

In the NPD we also observe a number of individual background variables. These are gender, eligibility for Free School Meals (FSM), ethnicity, whether the first language spoken at home is English, whether the child has been defined by the school as having Special Educational Needs, and the level of income deprivation in the area around the child's postcode of residence. As families entitled to free school lunches are families in receipt of means tested benefits and/or with one if not both jobless parents, this indicator will be used to distinguish low- from high-income families (Hobbs and Vignoles, 2009). The data also contains date of birth, and the date at which the child was enrolled in the school. We use this information to define school entry policies, as described below.

School entry policies

In England children start formal schooling at age 4, in the academic year in which they turn 5.[6] The school cohort consists of all children born between the month of September of one year and the month of August of the following year. Most children start school in September irrespective of their date of birth, but in some schools children born later in the school year are allowed (or even encouraged) to start school in the second or third term (i.e. in January or April, respectively). School entry policies can differ quite markedly across the country, and this variation in school entry age has been exploited in previous work which looked at the effect of month of birth (Crawford et al., 2010) or early schooling (Cornelissen et al., 2013) on educational outcomes.

Taking into account differences in school entry policies is very important in our analysis. Our identification strategy requires that children born either side of the cut-offs which define entitlement to free part-time pre-school education are the same in respect of all other observed and unobserved

[6]About 96% of children in England are educated through the school system, and it is very rare for a child to delay for one year or start earlier than age 4.

characteristics. There are three such cut-offs in England: 31st December, 31st March and 31st August. Children born either side of 31st August will automatically belong to different school cohorts, which implies that they will start school at different ages and at a different time, and that they will have a different *relative age* within the class (with children born in September being the oldest in the year, and children born in August being the youngest in the year). It is clear that 31st August will not be a valid discontinuity point for the purpose of our analysis. Also, in some schools children born in December have a different starting date with respect to children born in January, and in some other schools children born in March have a different starting date with respect to children born in April. Using information on date of birth and date of school enrollment we identify the schools where school entry differs either side of 31st December or 31st March. These schools are excluded from our analysis.

Estimation sample

In order to perform our analysis we need information on precise date of birth of the child. This information is usually restricted, but we obtained access to NPD data with date of birth of the child for several cohorts, including children starting school between the academic years 2006/07 and 2011/12 for age 5 (Foundation Stage Profile) outcomes, and children starting school between the academic years 2005/06 and 2009/10 for age 7 (Key Stage 1) outcomes. Because of the confidential nature of the data we obtained information on a subsample of all children within the year, and can only analyse data on children born up to 4 weeks before and after 31st December and 31st March cut-offs. This means that we do not have information on children born in the month of February or May, for example.

We apply the usual checks on the data and remove pupils with duplicate data entries or with missing data on school-identifiers, but keep children with missing information on some of the background and outcome variables. We also exclude "special schools" that exclusively cater for children with specific needs, for example because of physical disabilities or learning difficulties, as well as schools specifically focused on children with emotional and/or behavioural difficulties (less than 1% of pupils). As explained above, we exclude schools which operate a differential school entry intake for children born either side of 31st December or 31st March cut-offs (overall about 4% of pupils) and schools for which we were unable to define school entry policies by month of birth (about 10% of pupils).[7] The remaining sample contains 874,938 pupils at age 5 and 706,010 at age 7.

Appendix Table 1 shows descriptive statistics for child outcomes at ages 5 and 7. Here we report the proportions achieving the expected level, or going beyond the expected level, in various areas of assessment at age 5 and 7. We present these proportions separately for the whole sample, as well as by gender and free school meal status. As we would expect, girls outperform boys in all outcome measures at both ages, with the gap being generally smaller in numeracy and mathematics than in literacy, reading or writing. Even larger differences in attainment can be found between children eligible for free school meals and all the other children. When comparing levels of achievements across time we see that a lower proportion of children performs at the expected level at age 5 compared to age 7; for example the proportion who achieves the expected level in numeracy is 57% at age 5 while the proportion who achieves the expected level in mathematics is 76% at age 7. Similarly for the proportion of children scoring beyond the expected level; only 26% do so in relation to the Foundation Stage Profile criteria, while 33% are in this category at the end of Key Stage 1. However, other measures of achievement are more immediately comparable, as for example "achieving the expected level in literacy and numeracy at age 5" (57% of the sample) and "achieving the expected level in all 3 subjects (reading, writing and mathematics) at age 7" (59% of the sample).

6 Results

6.1 Preliminary analysis

As standard in analyses based on a regression discontinuity research design, we start by plotting the distribution of births (date of birth is our *running variable*) either side of the two cut-offs. This is in order to investigate whether the entitlement date had any effect on the day on which a child was born. One could think, for example, that parents who were aware of the importance of the eligibility rule (because they were well-informed or because they had an older child) might have timed the birth of their child so to receive more free part-time child care. If so, we would expect to see relatively more births in the days preceding the 31st December or the 31st March, and

[7]Our initial checks reveal that our analysis is robust to including/excluding the schools for which we cannot define a school entry policy cannot be defined.

fewer births in the first few days after these dates. As we can see from Figure 1a (31st December) and Figure 1b (31st March), this is not the case. While there is a clear weekly pattern in the number of births - with fewer of them occurring at weekends - and an influence of festivities (such as Christmas), there is no evidence that the number of births is higher in the days or weeks before the cut-off dates.[8]

Another important check we run is presented in Figure 2a and 2b, which provide visual evidence about the distribution of the individual characteristics we observe by day of birth of the child. Two important issues emerge from these figures. First, there is a clear indication that a disproportionate number of children from a non-white ethnic background, and non-English-speaking homes are born on 1st January. We suspect that in cases where the date of birth is not known (because for example a certificate of birth has not been provided) the 1st January date is used by the school in filling in their administrative forms. As children from immigrant families are more likely to be born in other countries, and official information is more likely to be missing in these cases, we observe a higher proportion of individuals from ethnic minorities with a 1st January date of birth. Since these children also have higher levels of deprivation, and usually lower attainment, by including them in the analysis we will be more likely to observe a positive effect of eligibility (born before the cut-off) on educational attainment. For this reason, all our regressions will exclude children born on 1st January and (for symmetry) 1st April.

The second problem highlighted by these figures is that, while there is no evidence of systematic differences either side of the cut-off by gender, FSM status, deprivation score, ethnicity, and non-English-speaking language status, there is some indication that children born before the cut-off date are more likely to be registered as Special Education Need (SEN) pupils (bottom panel of Figure 2b). This is confirmed by results from a regression analysis where we test for the presence of a discontinuity in observable characteristics either side of the cut-off according to the model presented in section 3. These results are shown in Appendix Table 2, and indicate that the effect of eligibility on the observed individual characteristics is usually not significantly different from zero, except for SEN status, where among boys eligibility results in a 0.7 percentage point increase (or an almost 6% increase over the mean) in the probability of being identified as a child with difficulties in learning or having social and behavioural problems.[9]

[8]We plot the figures by year of birth as aggregating over different years leads to a more confusing picture because of the uneven distribution of the days of the week and festivities around the cut-offs.

[9]The problem is less evident for girls, but since the percentage of girls identified as SEN is much lower than for boys, we think that in this case we do not have sufficient power to detect a significant effect.

The effect of eligibility on SEN status for boys could be a true or causal effect of the policy. Indeed, it is possible that by being exposed to pre-school education earlier than comparable children, eligible children develop more problematic behaviour and are therefore more likely to require a statement of Special Education Needs. On the other hand, it is also possible that eligible children are more likely to be identified as having learning or behavioural difficulties simply because they are observed by qualified staff for a longer period of time. Either way, SEN status appears to be affected by eligibility status, and as such it will be excluded from our set of controls in the regression analysis which follows.

6.2 Educational outcomes at age 5

As a first piece of evidence about the impact of eligibility on educational attainment at age 5, Figure 3 plots the proportion of boys achieving the expected level in both literacy and numeracy by date of birth (i.e., at least 6 points out of a maximum of 9). The data adjust for average differences across schools, but do not control for any of the observed characteristics of individuals, nor for academic year. The figure also plots two regression lines showing the estimated linear relationship between the outcome and the birth date separately for the period before and after the cut-offs in order to facilitate the visual interpretation of the results. We see that during the period of 4 weeks before and after the cut-off dates (31st December in the left panel and 31st March in the right panel) there is a clear negative relationship between the outcome and the date of birth. There is also some evidence of a small downward 'jump' in attainment after the discontinuity points, but as we have not plotted confidence intervals around the regression lines, it is not possible to see whether this positive effect of eligibility is significantly different from zero. Figure 4 reports similar graphs, now distinguishing between FSM and non FSM boys. Again, we see a clear negative relationship between the outcome and the test score in both cases, but only for the non-FSM group we see any evidence of a discontinuity in the outcome at the cut-offs.

For girls things appear to differ somewhat. Figure 5 reports the proportion of girls achieving the expected level in both literacy and numeracy by date of birth. Here it is impossible to detect any 'jump' in the outcome at the discontinuity points. The analysis by FSM status in Figure 6 seem to suggest a positive effect of eligibility for the FSM group at the March cut-off, and a change in the slope of the relationship between the outcome and date of birth after the December cut-off.

The graphical analysis reveals some important features of the data which will guide the specification of our regression models. First, we see evidence of a mostly *linear* relationship between the outcome and the birth date. Although we will consider also quadratic specifications, there is no suggestion that significant non-linearities are at work within the 4-week window before and after each cut-off. Second, we see that the relationship between the outcome and birth date has in some cases a different slope at the two sides of the cut-off point, and it will therefore be important to consider specifications which allow for this. Third, the graphs show that the effect can be different at the December vs. the March cut-off points, so we will have to formally test for an interaction effect between eligibility and cut-off date. Finally, while overall effects might be zero, significant impacts could be seen in some subgroups.

Next we turn to regression analysis. For this we will use a more restricted data window, spanning 2 weeks before and after the cut-off dates. This will allow us to use the available data to maximum effect, as we will conduct also some placebo tests. To maximise efficiency, we will pull the data around the December and the March cut-offs, using a dummy to account for the fact that average attainment is lower on average for the children born around the March cut-off than for the children born around the December cut-off. We run four main specifications, as outlined in section 3 we take into account date of birth effects using: (i) a linear term (specification 1), (ii) a quadratic term (specification 2), (iii) a linear term which is allowed to change at the cut-off point (specification 3), and finally (iv) a quadratic term which is allowed to change at the cut-off point. This latter specification is likely to be too demanding as our data window consists of only 14 days either side of the discontinuity point, but it is still important to test for the presence of various forms of non-linearities. All our regressions also control for a full set of dummies for year, free school meal status, ethnicity, language spoken at home, and area of deprivation deciles. We account for school-level fixed effects and cluster the standard errors by date of birth.

Table 1 reports the main results for boys and girls separately. We show here the findings obtained from a model in which the effect of eligibility is not interacted with the cut-off point (i.e., it is the same at December as well as March) as initial tests indicate that the interaction between eligibility and cut-off is not significantly different from zero.[10] In line with the graphical analysis, we see some positive effects of eligibility for boys, but no significant impacts for girls. The effects

[10] As the effects we report are small it is possible that the effect of eligibility changes with the cut-off point, but we are unable to detect a statistically significant difference due to limited power. Future analysis will investigate this issue more accurately.

are rather small, however. Being born in the two weeks before the cut-off point results in a 1 percentage point change in the probability of achieving the expected level in numeracy at age 5 (linear specification). This is a change of about 1.4% with respect to the mean. Slightly larger effects (in relation to the mean) are to be found when looking at other outcomes. For example, achieving the expected level in all 6 areas of assessment results in 0.9 percentage points change, which is a 2% increase in relation to the mean, and going beyond the expected level results in a 1.1 percentage points change, which is equivalent to a 5% increase in relation to the mean. The results are not very sensitive to changes in the functional form for date of birth, although there is some evidence that the most flexible model (specification 4) is not always able to capture the effect of interest.

Appendix Table 3 reports a parallel analysis conducted on the standardized test scores. As for the main analysis, we see some positive effects of eligibility only on boys. The size of the effects is again small, in the range of 0.013 to 0.015 points of a standard deviation for most outcomes. Interestingly, we see effects of roughly the same size in many different areas, apart from physical development. Also rather interestingly, we see positive (although not precisely estimated) effects in the social and personal development area. This is to be seen in contrast with the results for SEN status, which on the contrary suggest negative effects of eligibility on behaviour (eligibility had a positive effect on the probability of being registered as SEN). This would seem to indicate that the effect of eligibility on SEN status is not a true effect, but rather a result of the fact that eligible children are observed by qualified educational carers for longer periods of time.

Table 2 and Table 3 report results disaggregated by FSM status for boys and girls, respectively. Here an interesting difference emerges. Similarly to what we saw in the graphical analysis, the positive effects of eligibility for boys are mainly to be found among the relatively more affluent children (non FSM group). Among the low-income/FSM boys the effects are not only never statistically significant at the conventional level, but change sign for different outcomes. For girls, we did not see any effect at the aggregate level, but the results in Table 3 suggest that there are in fact some negative effects of eligibility in the low-income/FSM group. The effects are relatively larger in size, about 3.4 percentage points for achieving the expected level in all 6 area of assessment at age 5, for example, which corresponds to an almost 8% decrease with respect to the mean. At this point in our analysis, it is unclear why we observe such differences in the effects of eligibility according to gender, but it is interesting to see that for both boys and girls the effects are such that the policy leads to an increase (rather than the intended reduction) in socio-economic inequalities in early achievements, and this is certainly a striking finding.

To conclude this section, we conduct a placebo test by testing for differences in outcomes 2 weeks either side of two arbitrary cut-off dates, which are set to 15th January and 15th April. As these dates do not separate children into different groups as far as their eligibility to free part-time pre-school childcare is concerned, we should not observe any difference across these arbitrary thresholds. As we can see from Table 4, this is indeed the case, and for both boys and girls we see that the effect of eligibility is never statistically significantly different from zero.[11]

6.3 Educational outcomes at age 7

It is important to ask whether the effects of eligibility on attainment we identify at age 5 carry on to later years or whether these effects are very short-lived. Thus, we extend our analysis to tests taken at age 7, which marks the end of the first Key Stage of the National Curriculum. As for the age 5 measures, we also consider here binary indicators for achieving the expected level (which is level 2B for these cohorts), or going beyond the expected level (achieving level 3 or above). The regression equations are specified as previously, and each of the four models shown allows for a different functional form in date of birth. The number of observations is in this case lower as we have access to information on 5 rather than 6 school cohorts.

As we can see in Table 5, there are no overall statistically significant effects of eligibility for boys or girls in any of the measures of attainment considered. By contrast, at age 5 we saw some small but positive and statistically significant effects for boys. It would therefore seem that the age 5 effects fade out rather quickly. But there is evidence this is not the case, at least not for all subgroups of the population. Tables 6 and 7 perform the same analysis on boys and girls separately according to our indicator of socio-economic status of the family, i.e. FSM status. Here we see that eligibility has negative effects on achievement in mathematics (expected level) for boys. These effects are about 3.5 percentage points, or a 6% decrease in relation to the mean. We see similar effects on girls' achievements in the same measure, although here the effect size is slightly smaller.

[11] There is only one exception, for the outcome "going beyond the expected level in at least one are of assessment" for girls we see that the most flexible specification shows a positive effect of eligibility on the outcome.

The results for girls at age 7 are in line with what we saw for age 5 in Table 3. That is, there is a clear indication that eligible girls from disadvantage families perform worse than comparable non-eligible girls, and this is a pattern which is persistent over the first years of primary education.By contrast, we saw that eligibility had positive effects on boys from non-disadvantaged families (non FSM) at age 5 (see Table 2), but negative effects on disadvantaged (FSM) boys at age 7 (Table 6). Our analysis on age 7 outcomes is still preliminary and these findings deserve further investigation. However, overall these results would seem to suggest that rather than reducing the socio-economic attainment gap, the provision of free part-time pre-school education in England might have increased it.

7 Conclusions

In this paper we carry out an evaluation of the introduction of a universal subsidy towards the costs of pre-school education for 3 year old children which came into effect in England in the early 2000s. The stated objective of this policy was to achieve an overall increase children's attainment and at the same time reduce early socio-economic differences in outcomes. Our identification strategy exploits the fact that eligibility to the subsidy was based on date-of-birth rules which imply that children born just a few days apart are entitled to a different amount (up to 3 months) of pre-school education. Using a regression discontinuity design, we compare the outcomes of children born either side of the cut-off points for eligibility (31st December and 31st March) while controlling for a flexible specification of date of birth in order to take into account age at test effects.

Our findings are still preliminary, but so far they reveal that eligibility to free part-time pre-school education had small effects on teachers' evaluations at age 5 (Foundation Stage Profile scores) and no statistically significant impacts on test scores in reading, writing, and mathematics at age 7 (Key Stage 1). However, these overall effects however mask significant differences across socio-economic groups. In particular, we find that eligibility had positive effects on boys from more advantaged backgrounds (non-FSM) at age 5, and negative effects on boys from disadvantaged backgrounds (FSM) at age 7. For girls, we see negative effects of eligibility among disadvantaged families at both age 5 and age 7. It would thus appear that the effect of the policy was to ultimately increase socio-economic inequalities in achievements, with small or no effects on the general population.

An interesting question we still need to answer completely is whether the impact of eligibility can help us identify other parameters of interest, and in particular whether this study could tell us something about the effect of "attending pre-school" on children's early educational attainment. The answer to this question depends on *how* eligibility affects children outcomes; whether it is by increasing attendance, hours, by inducing changes in the quality of pre-school education or simply through income effects. Preliminary analysis on a cohort of children born in 2000/2001 (the Millennium Cohort Study) shows that eligibility had modest effects on attendance; most children are in some form of formal childcare ($> 50\%$) before they turn 3, for those born September-December attendance increased only by 10% in January (the first eligible month), while for those born January-March it increased by a mere 5% in April. Assuming that the main channel through which eligibility to free part-time pre-school affects early educational outcomes is pre-school attendance, then our estimates could be interpreted as a lower bound of the effect of pre-school education on early attainment.

However, before we assess the economic significance of our findings, our analysis will have to consider *all* the channels through which eligibility may affect children outcomes. In particular we will have to examine in further detail whether eligibility had effects on hours of child care, as it is possible that the effects are non-linear in hours. Given that we observe mainly negative effects of eligibility on low-income children, it is unlikely that changes in the the quality of the setting or income effects could be a relevant part of the explanation. Nevertheless, the issue of *quality* deserves further attention, as it is possible that there is a significant interaction between exposure to child care and its quality. Research from the UK has suggested that only high-quality formal childcare can improve outcomes for children (Sylva et al., 2004). Our results would be consistent with this hypothesis if - as it has been documented elsewhere (Gambaro et al., 2013) - access to high-quality settings varies systematically with the socio-economic characteristics of the family.

References

Baker, M., Gruber, J. and K. Milligan (2008), "Universal Child Care, Maternal Labor Supply, and Family Well-Being", *Journal of Political Economy*, 116(4): 709-745.

Becker, G.S. (1964), *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* Chicago: University of Chicago Press.

Berlinski, S., S. Galiani and P. Gertler (2009), "The effect of pre-primary education on primary school performance", *Journal of Public Economics*, 93: 219-234.

Berlinski, S., S. Galiani and M. Manacorda (2008), "Giving children a better start: Preschool attendance and school-age profiles", *Journal of Public Economics*, 92(5-6): 1416-1440.

Berlinski, S., Galiani, S. and P. McEwan (2011), "Preschool and Maternal Labor Market Outcomes: Evidence from a Regression Discontinuity Design", *Economic Development and Cultural Change*, 59: 313-344.

Black, S.E., Devereux, P.J., Løken, K. and K.G. Salvanes (2012), "Care or Cash? The Effect of Child Care Subsidies on Student Performance", NBER Working Paper n. 18086.

Blanden, J., Del Bono, E., Hansen, K., McNally, S. and B. Rabe (2014), "Childcare and early child development. Evaluating the impact of universal part-time preschool education in England", University of Essex, *mimeo.*

Blau, D. and J. Currie (2006), "Pre-school, day care, and after school care: who is minding the kids?", in E. Hanushek and F. Welch (eds.) *Handbook of the Economics of Education.* Elsevier: North Holland.

Brewer, M. and C. Crawford (2010), "Starting school and leaving welfare: the impact of public education on lone parents' welfare receipt", IFS Working Paper W10/19.

Carneiro, P., and J. Heckman (2004), "Inequality in America: What Role for Human Capital Policies?", in *Human Capital Policy*, by J. Heckman and A.B. Krueger (eds.). MIT Press.

Cascio, E. (2009), "Do Investments in Universal Early Education Pay Off? Long-Term Effects of Introducing Kindergartens into Public Schools", NBER Working Paper n. 14951.

Cornelissen, T., Dustmann, C. and C. Trentini (2013), "Early School Exposure, Test Scores and Non-Cognitive Outcomes", University College London, *mimeo.*

Crawford, C. and J. Cribb (2013), "Reading and maths skills at age 10 and earnings in later life: a brief analysis using the British Cohort Study", Centre for Analysis of Youth Transitions Impact Study: REP03.

Crawford, C., Dearden, L. and C. Meghir (2010), "When you are born matters: the impact of month of birth on educational outcomes in England", Institute for Fiscal Studies Working Paper No. W10/06.

Crawford, C., Dearden, L. and E. Greaves (2011), "Does when you are born matter? The impact of month of birth on children's cognitive and non-cognitive skills in England", Institute for Fiscal Studies Briefing Note 122.

Department for Education (2012), *Early Years Foundation Stage Profile Results in England, 2011/12*, Statistical First Release SFR 23/2012.

Department for Education (2013), *More great childcare. Raising quality and giving parents more choice*, DFE-00002-2013.

Drange, N., Havnes, T. and A. Sandsør (2012), "Kindegarten for all: Long-run effects of a universal intervention", University of Oslo, *mimeo.*

Dustmann, C., Raute, A. and U. Schönberg (2013), "Does Universal Child Care Matter? Evidence from a Large Expansion in Pre-School Education", University College London, *mimeo.*

Fitzpatrick, M.D. (2010), "Preschoolers Enrolled and Mothers at Work? The Effects of Universal Prekindergarten", *Journal of Labor Economics*, 28: 51-84.

Fitzpatrick, M.D. (2008), "Starting School at Four: the Effect of Universal Pre-Kindergarten on Children's Academic Achievement," *B.E. Journal of Economic Analysis and Policy*, 8(1): 1-38.

Gambaro, L., Stewart, K. and J. Waldfogel (2013), "A question of quality: Do children from disadvantaged backgrounds receive lower quality education and care in England?", CASE Paper 171.

Gormley, W.T. and T. Gayer (2005), "Promoting School Readiness in Oklahoma", *Journal of Human Resources*, 40(3): 533-558.

Havnes, Y. and M. Mogstad (2011), "No Child Left Behind: Subsidized child care and children's long-run outcomes", *American Economic Journal: Economic Policy*, 3(May): 97-129.

Herbst, C.M. and E. Tekin (2008), "Child Care Subsidies and Child Development", NBER Working Paper n. 144174.

Hobbs, G. and a. Vignoles (2009), "Is children's free school meal 'eligibility' a good proxy for family income?", *British Educational Research Journal*, 36(4): 1469-3518.

Imbens, G.W. and T. Lemieux (2008), "Regression discontinuity designs: a guide to practice", *Journal of Econometrics* 142: 615-635.

Knudsen, E.I., Heckman, J.J., Cameron, J.L. and J.P.Shonkoff (2006), "Economic, neurobiological, and behavioral perspective on buildin America's future workforce", *Proceedings of the National Academy of Sciences*, 103(27): 10155-10162.

Leuven, E., M. Lindahl, H. Oosterbeek and D. Webbink (2010), "Expanding schooling opportunities for 4-year-olds", *Economics of Education Review*, 29: 319-328.

Machin, S. and S. McNally (2008), "The Literacy Hour", *Journal of Public Economics*, 92: 14411462.

Melhuish, E., Sylva, K., Sammons, P., Siraj-Blatchford, Taggart B., Phan, M.B. and A. Malin (2008): "The Early Years: Preschool Influences on Mathematics Achievement", *Science*, 321(5893): 1161-1162.

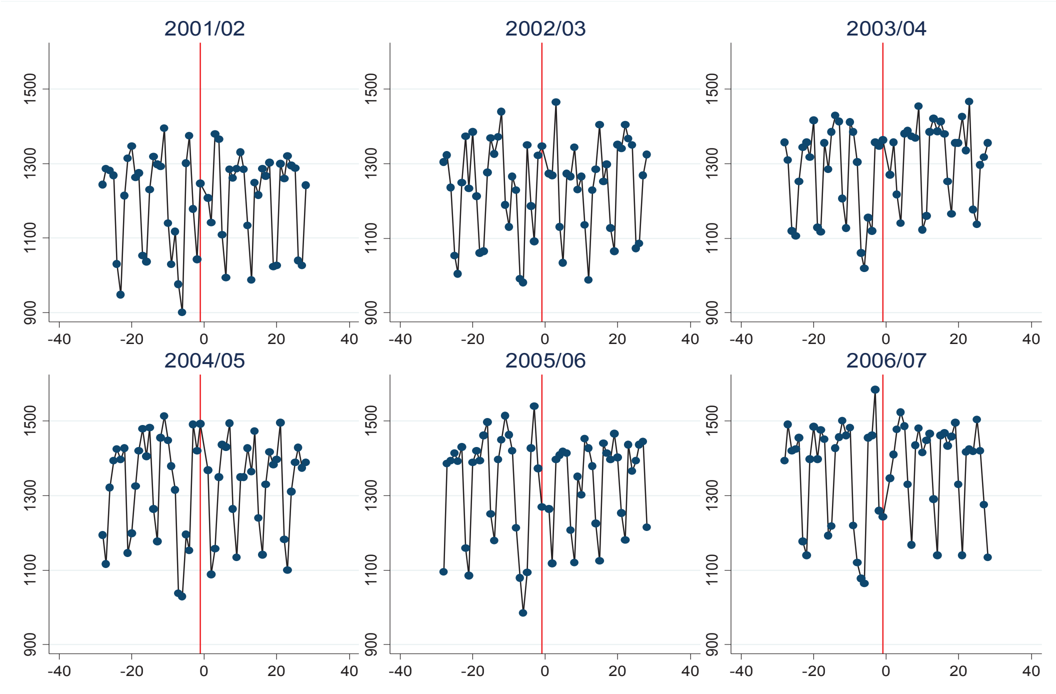
Merrell, C. and P. Tymms (2011), “ Changes in Children’s Cognitive Development at the Start of School in England 2001-2008”, *Oxford Review of Education*, 37(3): 333-345.

Thistlethwaite, D. and D. Campbell (1960), “Regression-Discontinuity Analysis: An alternative to the ex post facto experiment”, *Journal of Educational Psychology*, 51: 309317.

Shonkoff, J.P. and D.A. Philipps (2000), *From Neurons to Neighbourhoods: The Science of Early Childhood Development.* National Academy Press: Washington, D.C.

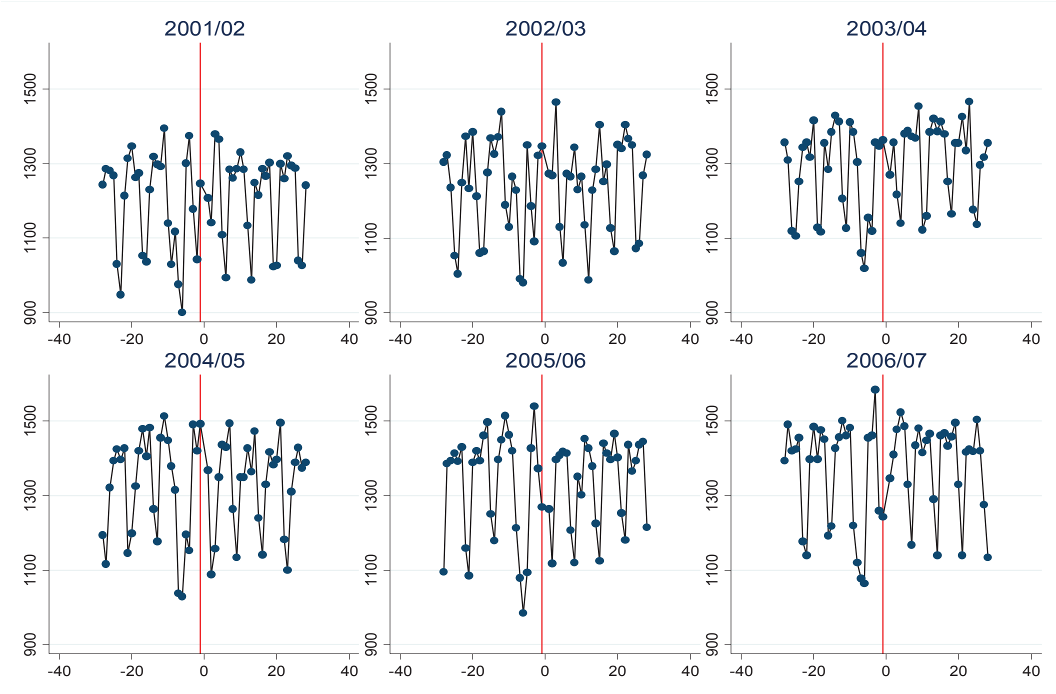
Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. and B. Taggart (2005), *Effective Pre-School Education. The Effective Provision of Pre-School Education (EPPE) Project: Final Report.* Institute of Education: London.

Figure 1a: Number of births before and after 31st December



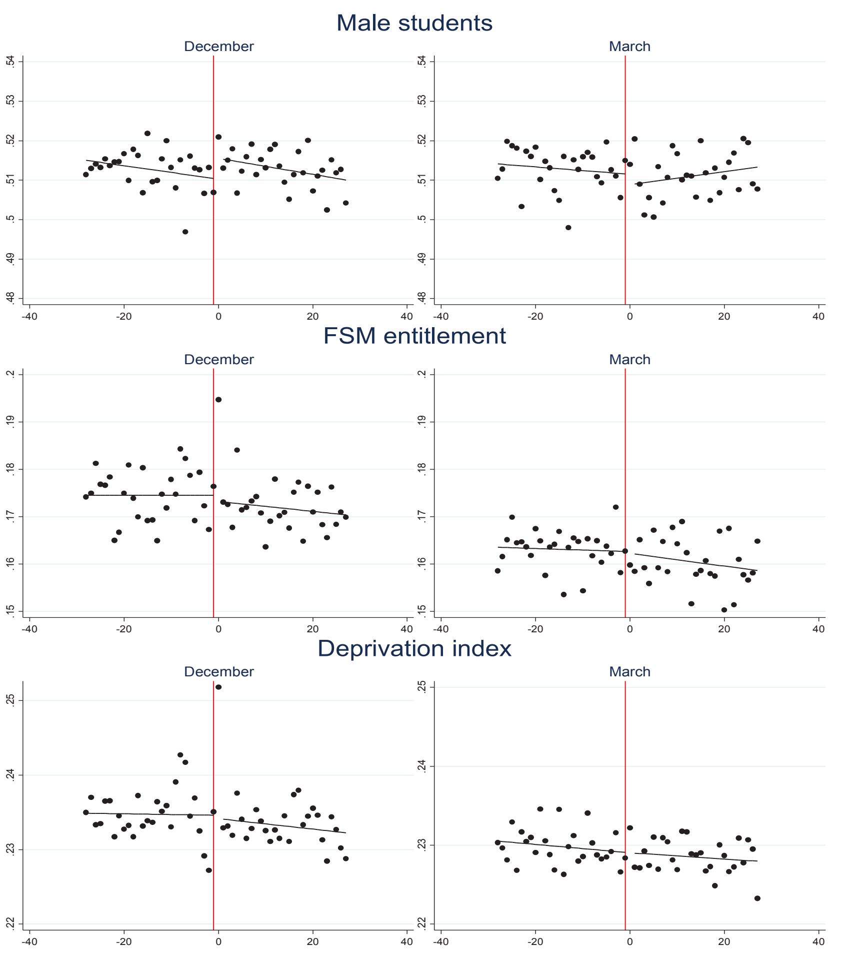
Notes: Each point represents the number of children born at a specific date over the 4 weeks before and 4 weeks after 31st December (vertical line). Data is plotted separately by year of birth.

Figure 1b: Number of births before and after 31st March



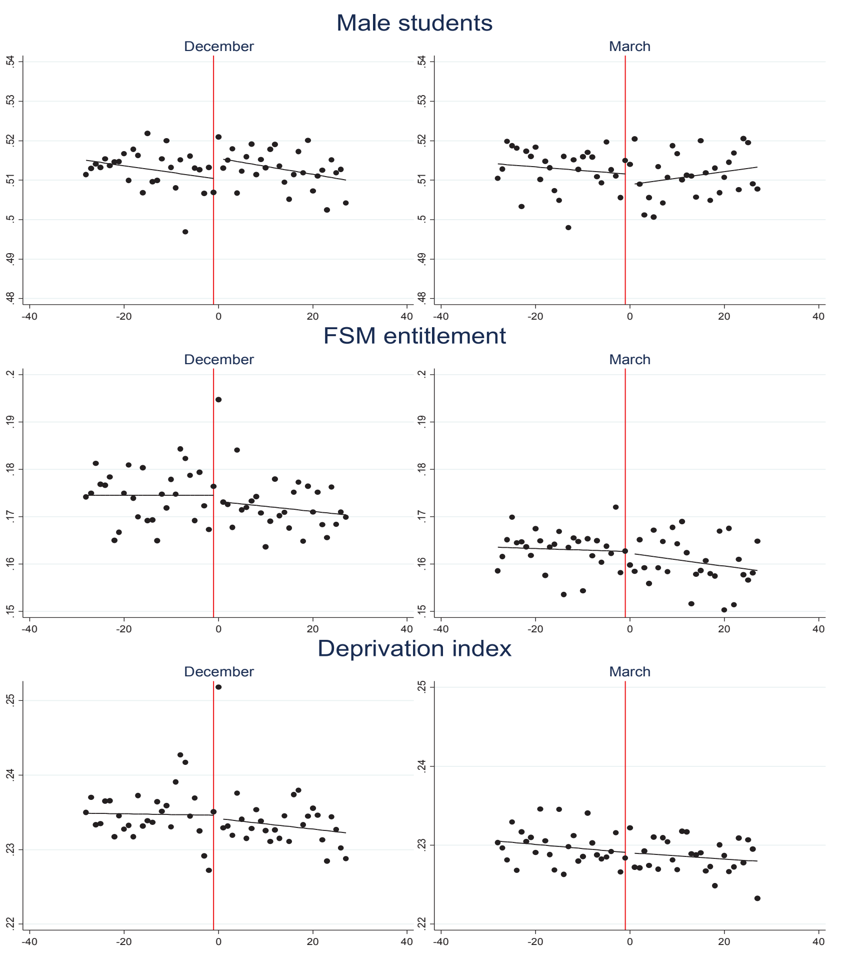
Notes: Each point represents the number of children born at a specific date over the 4 weeks before and 4 weeks after 31st December (vertical line). Data is plotted separately by year of birth.

Figure 2a: Observable characteristics by day of birth before and after cut-off dates



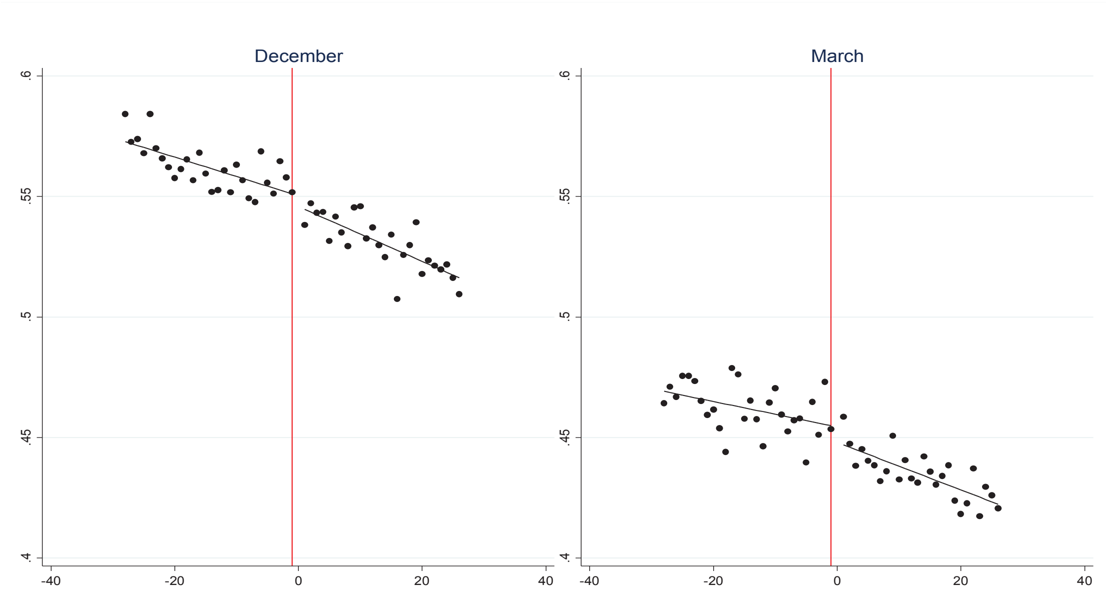
Notes: Each point represents the proportion of male children, children with FSM entitlement, and the mean value of the index of deprivation over the 4 weeks before and 4 weeks after 31st December (left panel) or 31st March (right panel). The solid lines are obtained by plotting the predicted values of a regression of the observable characteristic (e.g., male) on date of birth (in days) separately for the period before and after the cut-offs. The vertical red line represents 31st December and 31st March.

Figure 2b: Observable characteristics by day of birth before and after cut-off date



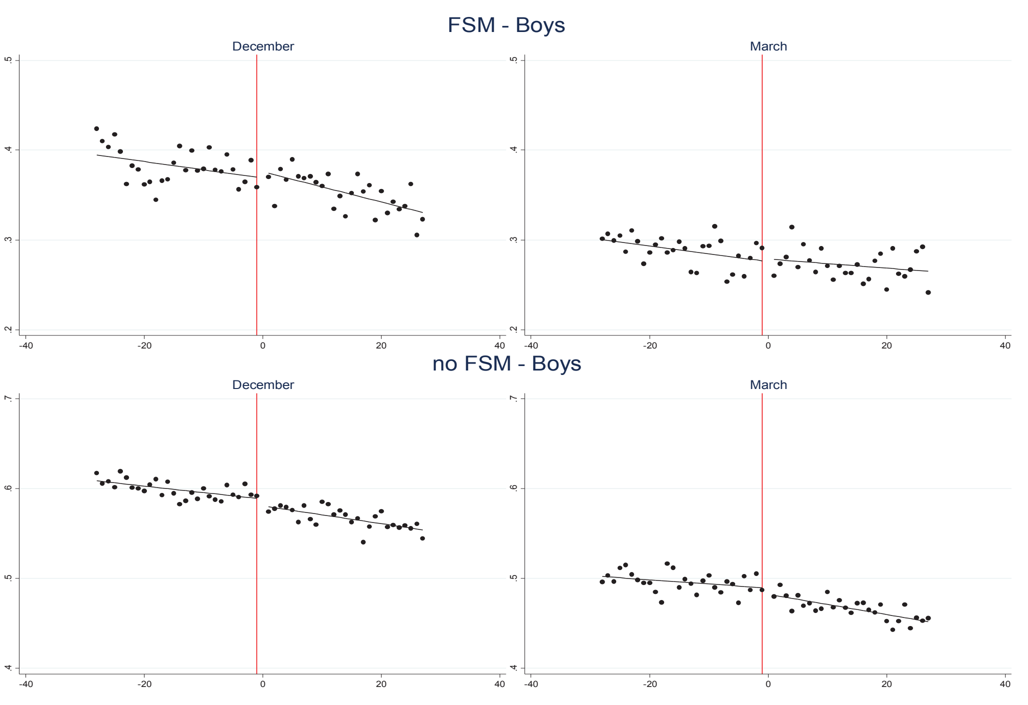
Notes: Each point represents the proportion of male children, children with FSM entitlement, and the mean value of the index of deprivation over the 4 weeks before and 4 weeks after 31st December (left panel) or 31st March (right panel). The solid lines are obtained by plotting the predicted values of a regression of the observable characteristic (e.g., male) on date of birth (in days) separately for the period before and after the cut-offs. The vertical red line represents 31st December and 31st March.

Figure 3: Effect of eligibility on at achieving the expected level in literacy and numeracy at age 5 - Boys



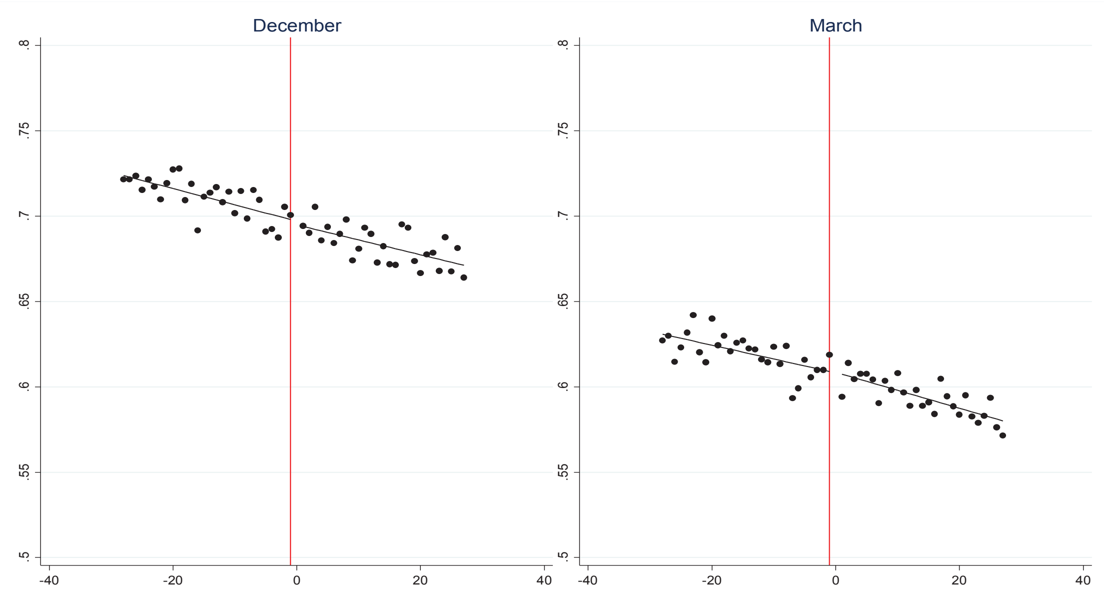
Notes: Each point represents the fraction of children achieving the expected level in literacy and numeracy over the 4 weeks before and 4 weeks after the cut-off points. The data account for unobserved school-fixed effects. The solid lines are obtained by plotting the predicted values of a regression of the outcome (fraction achieving the expected level) on date of birth (in days) separately for the period before and after the cut-offs. The vertical line represents 31st December (left panel) and 31st March (right panel), respectively.

Figure 4: Effect of eligibility on at achieving the expected level in literacy and numeracy at age 5 by FSM status - Boys



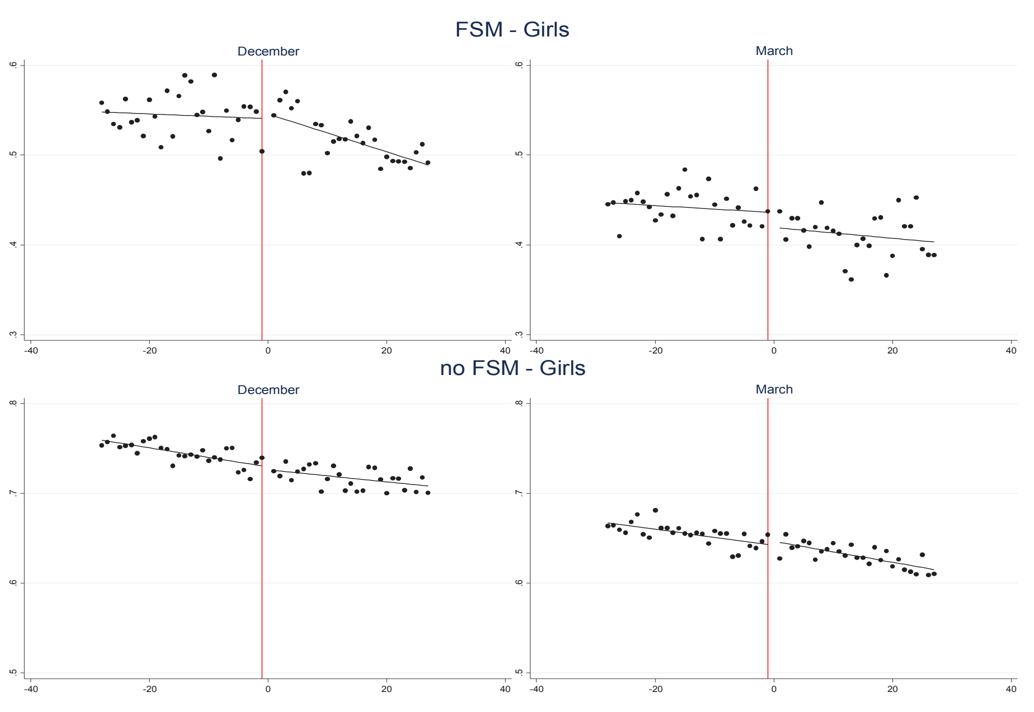
Notes: Each point represents the fraction of children achieving the expected level in literacy and numeracy over the 4 weeks before and 4 weeks after the cut-off points. The top panel reports data for children entitled to a free school meal (FSM), while the bottom panel reports data for children not entitled to a free school meal (no FSM). The data account for unobserved school-fixed effects. The solid lines are obtained by plotting the predicted values of a regression of the outcome (fraction achieving the expected level) on date of birth (in days) separately for the period before and after the cut-offs. The vertical line represents 31st December (left panel) and 31st March (right panel), respectively.

Figure 5: Effect of eligibility on at achieving the expected level in literacy and numeracy at age 5 - Girls



Notes: Each point represents the fraction of children achieving the expected level in literacy and numeracy over the 4 weeks before and 4 weeks after the cut-off points. The data account for unobserved school-fixed effects. The solid lines are obtained by plotting the predicted values of a regression of the outcome (fraction achieving the expected level) on date of birth (in days) separately for the period before and after the cut-offs. The vertical red line represents 31st December (left panel) and 31st March (right panel), respectively.

Figure 6: Effect of eligibility on at achieving the expected level in literacy and numeracy at age 5 by FSM status - Girls



Notes: Each point represents the fraction of children achieving the expected level in literacy and numeracy over the 4 weeks before and 4 weeks after the cut-off points. The top panel reports data for children entitled to a free school meal (FSM), while the bottom panel reports data for children not entitled to a free school meal (no FSM). The data account for unobserved school-fixed effects. The solid lines are obtained by plotting the predicted values of a regression of the outcome (fraction achieving the expected level) on date of birth (in days) separately for the period before and after the cut-offs. The vertical line represents 31st December (left panel) and 31st March (right panel), respectively.

Table 1: Effect of eligibility on educational outcomes at age 5

	Boys				Girls			
	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Achieving the expexted level in literacy at age 5								
Eligibility	0.008+	0.008+	0.007*	0.005	-0.002	-0.001	-0.001	0.013
	(0.004)	(0.004)	(0.003)	(0.006)	(0.004)	(0.004)	(0.006)	(0.009)
mean d.v.	0.52				0.69			
Achieving the expexted level in numeracy at age 5								
Eligibility	0.010*	0.010**	0.011**	0.006	-0.007+	-0.006+	-0.006	0.002
	(0.004)	(0.004)	(0.003)	(0.008)	(0.004)	(0.004)	(0.004)	(0.005)
mean d.v.	0.71				0.77			
Achieving the expexted level in literacy and numeracy at age 5								
Eligibility	0.011*	0.010*	0.010**	0.008	-0.005	-0.004	-0.004	0.007
	(0.004)	(0.004)	(0.003)	(0.006)	(0.004)	(0.004)	(0.005)	(0.007)
mean d.v.	0.5				0.65			
Achieving the expexted level in all 6 areas of assessment at age 5								
Eligibility	0.009*	0.008+	0.008+	0.001	-0.007	-0.005	-0.005	0.004
	(0.004)	(0.004)	(0.004)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)
mean d.v.	0.44				0.61			
Going beyond the expected level in at least one area of assessment at age 5								
Eligibility	0.011**	0.011**	0.010**	0.013*	0.002	0.003	0.002	-0.002
	(0.004)	(0.004)	(0.003)	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)
mean d.v.	0.22				0.3			
Observations	206,845				197,049			

Notes: Dependent variable is 0/1 for achieving the expected level (at least 6) in each of the areas considered or achieving the top score in at least one of the areas (going beyond). Estimation is by OLS. Coefficient shown is the effect of being born 2 weeks before 31st December or 31st March cut-off. Each specification includes a function of birth date and this is specified as follows: linear trend (1); quadratic trend (2); linear trend with effect allowed to change at cut-off point (3) and quadratic trend with effect allowed to change at cut-off point (4). Each specification also includes a dummy for the 31st March cut-off and full set of dummies for year, free school meal status, ethnicity, language spoken at home, and area of deprivation deciles. Estimation accounts for school-level fixed effects; standard errors are clustered by birth date. Symbols: ** significant at 1% level; * significant at 5% level; + significant at 10% level.

Table 2: Effect of eligibility on educational outcomes at age 5 by FSM status - Boys

	Boys - FSM				Boys - no FSM			
	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Achieving the expexted level in literacy at age 5								
Eligibility	-0.006	-0.006	-0.006	0.034+	0.012*	0.011*	0.011*	0.002
	(0.012)	(0.012)	(0.011)	(0.018)	(0.005)	(0.005)	(0.004)	(0.007)
mean d.v.	0.35				0.55			
Achieving the expexted level in numeracy at age 5								
Eligibility	0.020	0.021	0.022+	0.043*	0.008+	0.009+	0.009*	0.000
	(0.013)	(0.013)	(0.013)	(0.020)	(0.004)	(0.004)	(0.004)	(0.008)
mean d.v.	0.55				0.74			
Achieving the expexted level in literacy and numeracy at age 5								
Eligibility	-0.004	-0.005	-0.004	0.037*	0.015**	0.014**	0.013**	0.004
	(0.012)	(0.012)	(0.011)	(0.014)	(0.005)	(0.005)	(0.005)	(0.008)
mean d.v.	0.33				0.53			
Achieving the expexted level in all 6 areas of assessment at age 5								
Eligibility	-0.015	-0.016	-0.015	0.020	0.014**	0.013**	0.013*	-0.002
	(0.011)	(0.012)	(0.011)	(0.015)	(0.005)	(0.005)	(0.005)	(0.009)
mean d.v.	0.27				0.47			
Going beyond the expected level in at least one area of assessment at age 5								
Eligibility	0.012	0.011	0.011	0.025+	0.011**	0.011*	0.010*	0.010
	(0.008)	(0.008)	(0.007)	(0.013)	(0.004)	(0.004)	(0.004)	(0.009)
mean d.v.	0.1				0.24			
Observations	34,458				172,372			

Notes: Dependent variable is 0/1 for achieving the expected level (at least 6) in each of the areas considered or achieving the top score in at least one of the areas (going beyond). Estimation is by OLS. Coefficient shown is the effect of being born 2 weeks before 31st December or 31st March cut-off separately for FSM (left panel) and non-FSM (right panel) pupils. See Table 1 for all other details. Estimation accounts for school-level fixed effects; standard errors are clustered by birth date. Symbols: ** significant at 1% level; * significant at 5% level; + significant at 10% level.

Table 3: Effect of eligibility on educational outcomes at age 5 by FSM status - Girls

	Girls - FSM				Girls - no FSM			
	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Achieving the expexted level in literacy at age 5								
Eligibility	-0.021	-0.022	-0.023*	-0.009	-0.001	0.001	0.001	0.020+
	(0.013)	(0.014)	(0.011)	(0.022)	(0.005)	(0.005)	(0.008)	(0.010)
mean d.v.	0.53				0.72			
Achieving the expexted level in numeracy at age 5								
Eligibility	-0.014	-0.014	-0.014	0.007	-0.007+	-0.006	-0.006	0.003
	(0.013)	(0.013)	(0.011)	(0.019)	(0.004)	(0.004)	(0.005)	(0.007)
mean d.v.	0.63				0.8			
Achieving the expexted level in literacy and numeracy at age 5								
Eligibility	-0.027+	-0.028*	-0.029*	-0.004	-0.003	-0.002	-0.001	0.012
	(0.014)	(0.014)	(0.012)	(0.025)	(0.005)	(0.005)	(0.006)	(0.009)
mean d.v.	0.48				0.69			
Achieving the expexted level in all 6 areas of assessment at age 5								
Eligibility	-0.034*	-0.034*	-0.035*	0.001	-0.004	-0.003	-0.002	0.007
	(0.014)	(0.014)	(0.015)	(0.024)	(0.005)	(0.005)	(0.005)	(0.006)
mean d.v.	0.43				0.64			
Going beyond the expected level in at least one area of assessment at age 5								
Eligibility	-0.002	-0.003	-0.005	-0.006	0.003	0.004	0.003	0.000
	(0.010)	(0.010)	(0.012)	(0.021)	(0.005)	(0.005)	(0.005)	(0.005)
mean d.v.	0.16				0.33			
Observations	33,325				163,694			

Notes: Dependent variable is 0/1 for achieving the expected level (at least 6) in each of the areas considered or achieving the top score in at least one of the areas (going beyond). Estimation is by OLS. Coefficient shown is the effect of being born 2 weeks before 31st December or 31st March cut-off separately for FSM (left panel) and non-FSM (right panel) pupils. See Table 1 for all other details. Estimation accounts for school-level fixed effects; standard errors are clustered by birth date. Symbols: ** significant at 1% level; * significant at 5% level; + significant at 10% level.

Table 4: Effect of eligibility on educational outcomes at age 5 - Placebo

	Boys				Girls			
	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Achieving the expexted level in literacy at age 5								
Eligibility	0.003	0.003	0.003	0.006+	-0.003	-0.003	-0.003	-0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
mean d.v.	0.5				0.68			
Achieving the expexted level in numeracy at age 5								
Eligibility	0.007+	0.007+	0.007+	0.009	0.001	0.001	0.001	-0.001
	(0.004)	(0.004)	(0.004)	(0.005)	(0.002)	(0.002)	(0.002)	(0.003)
mean d.v.	0.7				0.76			
Achieving the expexted level in literacy and numeracy at age 5								
Eligibility	0.002	0.002	0.002	0.004	-0.001	-0.001	-0.001	-0.001
	(0.004)	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)
mean d.v.	0.48				0.34			
Achieving the expexted level in all 6 areas of assessment at age 5								
Eligibility	-0.001	-0.001	-0.001	0.004	-0.001	-0.001	-0.001	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
mean d.v.	0.42				0.59			
Going beyond the expected level in at least one area of assessment at age 5								
Eligibility	-0.004	-0.004	-0.004	-0.003	0.001	0.001	0.001	0.011*
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)
mean d.v.	0.21				0.29			
Observations	217,477				207,350			

Notes: Dependent variable is 0/1 for achieving the expected level (at least 6) in each of the areas considered or achieving the top score in at least one of the areas (going beyond). Estimation is by OLS. Coefficient shown is the effect of being born 2 weeks before 15th January or 15th April cut-off. See Table 1 for all other details. Estimation accounts for school-level fixed effects; standard errors are clustered by birth date. Symbols: ** significant at 1% level; * significant at 5% level; + significant at 10% level.

Table 5: Effect of eligibility on educational outcomes at age 7

	Boys				Girls			
	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Achieving the expexted level in reading test at age 7								
Eligibility	0.003	0.006	0.007+	-0.004	-0.004	-0.004	-0.004	0.007
	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)	(0.006)	(0.010)
mean d.v.	0.7				0.8			
Achieving the expexted level in writing test at age 7								
Eligibility	0.002	0.003	0.004	-0.011	-0.002	-0.001	-0.001	0.010
	(0.005)	(0.005)	(0.005)	(0.008)	(0.005)	(0.005)	(0.004)	(0.008)
mean d.v.	0.55				0.71			
Achieving the expexted level in mathematics test at age 7								
Eligibility	-0.000	0.002	0.002	-0.003	-0.005	-0.005	-0.005	0.004
	(0.004)	(0.004)	(0.002)	(0.005)	(0.004)	(0.004)	(0.005)	(0.008)
mean d.v.	0.75				0.78			
Achieving the expexted level in all 3 subjects tested at age 7								
Eligibility	0.000	0.002	0.002	-0.009	-0.005	-0.005	-0.004	0.003
	(0.005)	(0.005)	(0.005)	(0.007)	(0.005)	(0.005)	(0.004)	(0.007)
mean d.v.	0.53				0.66			
Going beyond the expected level in at least one subject tested at age 7								
Eligibility	0.003	0.003	0.003	-0.007	0.001	0.001	0.001	-0.011+
	(0.005)	(0.005)	(0.006)	(0.010)	(0.005)	(0.005)	(0.005)	(0.007)
mean d.v.	0.31				0.35			
Observations	169,278				162,327			

Notes: Dependent variable is 0/1 for achieving the expected level (at least 2B) in each of the subjects considered or achieving a higher score (3) in at least one of the subjects (going beyond). Estimation is by OLS. Coefficient shown is the effect of being born 2 weeks before 31st December or 31st March cut-off. See Table 1 for all other details. Estimation accounts for school-level fixed effects; standard errors are clustered by birth date. Symbols: ** significant at 1% level; * significant at 5% level; + significant at 10% level.

Table 6: Effect of eligibility on educational outcomes at age 7 by FSM status - Boys

	Boys - FSM				Boys - no FSM			
	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Achieving the expexted level in reading test at age 7								
Eligibility	-0.009	-0.005	-0.002	-0.024	0.006	0.008	0.009+	-0.002
	(0.014)	(0.014)	(0.010)	(0.018)	(0.005)	(0.005)	(0.005)	(0.009)
mean d.v.	0.52				0.74			
Achieving the expexted level in writing test at age 7								
Eligibility	-0.012	-0.009	-0.006	-0.009	0.006	0.008	0.007	-0.008
	(0.013)	(0.013)	(0.007)	(0.014)	(0.006)	(0.006)	(0.007)	(0.012)
mean d.v.	0.35				0.6			
Achieving the expexted level in mathematics test at age 7								
Eligibility	-0.035**	-0.029*	-0.027*	-0.030	0.005	0.006	0.006+	-0.003
	(0.013)	(0.014)	(0.010)	(0.021)	(0.005)	(0.005)	(0.004)	(0.005)
mean d.v.	0.59				0.79			
Achieving the expexted level in all 3 subjects tested at age 7								
Eligibility	-0.025+	-0.021	-0.019**	-0.031*	0.007	0.008	0.007	-0.002
	(0.013)	(0.013)	(0.006)	(0.012)	(0.006)	(0.006)	(0.006)	(0.011)
mean d.v.	0.33				0.57			
Going beyond the expected level in at least one subject tested at age 7								
Eligibility	-0.009	-0.010	-0.011	-0.008	0.007	0.007	0.007	-0.005
	(0.010)	(0.010)	(0.007)	(0.014)	(0.005)	(0.006)	(0.007)	(0.012)
mean d.v.	0.15				0.34			
Observations	30,416				138,193			

Notes: Dependent variable is 0/1 for achieving the expected level (at least 2B) in each of the subjects considered or achieving a higher score (3) in at least one of the subjects (going beyond). Estimation is by OLS. Coefficient shown is the effect of being born 2 weeks before 31st December or 31st March cut-off. See Table 1 for all other details. Estimation accounts for school-level fixed effects; standard errors are clustered by birth date. Symbols: ** significant at 1% level; * significant at 5% level; + significant at 10% level.

Table 7: Effect of eligibility on educational outcomes at age 7 by FSM status - Girls

	Girls - FSM				**Girls - no FSM**			
	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off	linear	quadratic	linear changing at cut-off	quadratic changing at cut-off
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Achieving the expexted level in reading test at age 7								
Eligibility	-0.013	-0.009	-0.006	-0.000	-0.001	-0.002	-0.002	0.012
	(0.015)	(0.015)	(0.012)	(0.019)	(0.004)	(0.004)	(0.006)	(0.008)
mean d.v.		0.65				0.84		
Achieving the expexted level in writing test at age 7								
Eligibility	-0.012	-0.010	-0.008	-0.032*	0.001	0.001	0.001	0.020**
	(0.016)	(0.016)	(0.012)	(0.015)	(0.005)	(0.005)	(0.005)	(0.007)
mean d.v.		0.53				0.76		
Achieving the expexted level in mathematics test at age 7								
Eligibility	-0.027+	-0.027+	-0.026+	-0.021	-0.002	-0.003	-0.003	0.010+
	(0.014)	(0.014)	(0.013)	(0.017)	(0.005)	(0.005)	(0.004)	(0.006)
mean d.v.		0.63				0.81		
Achieving the expexted level in all 3 subjects tested at age 7								
Eligibility	-0.029+	-0.027+	-0.025*	-0.031+	-0.001	-0.001	-0.001	0.012*
	(0.015)	(0.015)	(0.011)	(0.015)	(0.005)	(0.005)	(0.004)	(0.005)
mean d.v.		0.47				0.71		
Going beyond the expected level in at least one subject tested at age 7								
Eligibility	-0.024*	-0.025*	-0.025+	-0.040	0.006	0.006	0.006	-0.001
	(0.011)	(0.011)	(0.015)	(0.027)	(0.006)	(0.006)	(0.005)	(0.008)
mean d.v.		0.18				0.39		
Observations		29,836				131,845		

Notes: Dependent variable is 0/1 for achieving the expected level (at least 2B) in each of the subjects considered or achieving a higher score (3) in at least one of the subjects (going beyond). Estimation is by OLS. Coefficient shown is the effect of being born 2 weeks before 31st December or 31st March cut-off. See Table 1 for all other details. Estimation accounts for school-level fixed effects; standard errors are clustered by birth date. Symbols: ** significant at 1% level; * significant at 5% level; + significant at 10% level.

Appendix Table 1: Summary statistics on outcomes at age 5 and 7

	All	**Boys**	**Girls**	**FSM**	**no FSM**
	mean	mean	mean	mean	mean
Achieving the expexted level in literacy at age 5	0.60	0.52	0.69	0.43	0.63
Achieving the expexted level in numeracy at age 5	0.74	0.71	0.77	0.59	0.77
Achieving the expexted level in literacy and numeracy at age 5	0.57	0.50	0.65	0.40	0.61
Achieving the expexted level in all 6 areas of assessment at age 5	0.52	0.44	0.61	0.35	0.55
Going beyond the expected level in at least one area of assessment at age 5	0.26	0.22	0.30	0.13	0.28
Observations	874,938	448,395	426,543	146,580	728,313
Achieving the expexted level in reading test at age 7	0.75	0.70	0.80	0.58	0.79
Achieving the expexted level in writing test at age 7	0.63	0.55	0.71	0.44	0.67
Achieving the expexted level in mathematics test at age 7	0.76	0.75	0.78	0.61	0.80
Achieving the expexted level in all 3 subjects tested at age 7	0.59	0.52	0.66	0.40	0.64
Going beyond the expected level in at least one subject tested at age 7	0.33	0.31	0.35	0.16	0.37
Observations	706,010	360,778	345,232	128,339	574,543

Notes: Proportion of children achieving the indicated level of attainment (see text for the definition of the achievement measures). Sample of children starting school in academic years 2006/07-2011/12 for age 5 measures and children starting school in academic years 2005/06-2008/09 for age 7 measures. Total number of observations refers to all observations within each group (e.g., all boys), and includes children born 4 weeks before and after the two cut-off points of 31st December and 31st March.

Appendix Table 2: Effect of eligibility on observable variables at age 5

	Boys				Girls			
	linear	quadratic	linear changing	quadratic changing	linear	quadratic	linear changing	quadratic changing
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Free School Meals (FSM)								
Eligibility	0.005	0.005	0.005	0.006	0.002	0.001	0.001	-0.006
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.005)
mean d.v.	0.17				0.17			
Missing FSM status								
Eligibility	0.000	0.000	0.000	-0.000	-0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
mean d.v.	0.00				0.00			
Non white British								
Eligibility	0.005	0.005	0.005	0.006	0.002	0.002	0.003	0.004
	(0.004)	(0.004)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)	(0.007)
mean d.v.	0.28				0.28			
Missing Ethnicity								
Eligibility	-0.002	-0.002	-0.002	0.001	0.001	0.002	0.002	-0.001
	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)
mean d.v.	0.09				0.09			
Special Education Needs								
Eligibility	0.006+	0.007*	0.007*	0.004	0.004+	0.004+	0.004+	-0.000
	(0.003)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)	(0.004)
mean d.v.	0.12				0.06			
Missing Special Education Needs indicator								
Eligibility	0.000	0.000	0.000	-0.000	-0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
mean d.v.	0.00				0.00			
Non English speakers at home								
Eligibility	0.001	0.001	0.002	-0.002	0.002	0.003	0.003	0.002
	(0.003)	(0.003)	(0.004)	(0.005)	(0.003)	(0.003)	(0.003)	(0.005)
mean d.v.	0.18				0.18			
Missing language spoken at home indicator								
Eligibility	-0.002	-0.002	-0.002	0.002	-0.003	-0.002	-0.001	-0.005**
	(0.002)	(0.002)	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)
mean d.v.	0.12				0.11			
Index of deprivation								
Eligibility	0.000	0.001	0.001	-0.001	0.001	0.001	0.001	0.003+
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
mean d.v.	0.23				0.23			
Missing index of deprivation								
Eligibility	-0.002	-0.001	-0.001	-0.005**	0.000	0.001	0.001	0.003+
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
mean d.v.	0.02				0.02			
Observations	206,968				197,148			

Notes: Dependent variable is 0/1 for having or not having a certain attribute, except for deprivation index, which is continuous. Estimation is by OLS. Coefficient shown is the effect of being born 2 weeks before 31st December or 31st March cut-off. Each specification also includes a dummy for the 31st March cut-off and a full set of year dummies. See notes to Table 1 for other details. Estimation accounts for school-level fixed effects; standard errors are clustered by birth date. Symbols: ** significant at 1% level; * significant at 5% level; + significant at 10% level.

Appendix Table 3: Effect of eligibility on educational outcomes (continuous) at age 5

	Boys				Girls			
	linear	quadratic	linear changing	quadratic changing	linear	quadratic	linear changing	quadratic changing
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Total Foundation Stage Profile score at age 5								
Eligibility	0.015+	0.015+	0.015*	0.003	-0.006	-0.003	-0.003	0.007
	(0.008)	(0.008)	(0.007)	(0.013)	(0.008)	(0.008)	(0.009)	(0.012)
Literacy score at age 5								
Eligibility	0.014+	0.015+	0.014+	0.005	-0.003	-0.001	-0.000	0.014
	(0.008)	(0.008)	(0.008)	(0.013)	(0.008)	(0.008)	(0.010)	(0.014)
Numeracy score at age 5								
Eligibility	0.014+	0.014+	0.014*	0.008	-0.009	-0.007	-0.006	0.010
	(0.008)	(0.008)	(0.007)	(0.014)	(0.008)	(0.008)	(0.010)	(0.014)
Social and personal development score at age 5								
Eligibility	0.013	0.014+	0.014+	0.004	-0.005	-0.002	-0.001	-0.011
	(0.008)	(0.008)	(0.008)	(0.012)	(0.008)	(0.008)	(0.008)	(0.008)
Knowledge and understanding of the world score at age 5								
Eligibility	0.013	0.012	0.011	-0.002	-0.007	-0.005	-0.005	0.022
	(0.008)	(0.008)	(0.008)	(0.014)	(0.008)	(0.008)	(0.010)	(0.018)
Creative development score at age 5								
Eligibility	0.018*	0.019*	0.019*	-0.005	-0.003	-0.002	-0.001	-0.004
	(0.008)	(0.008)	(0.008)	(0.013)	(0.008)	(0.008)	(0.006)	(0.008)
Physical development score at age 5								
Eligibility	0.007	0.006	0.005	-0.011	-0.005	-0.003	-0.003	0.004
	(0.008)	(0.008)	(0.006)	(0.010)	(0.008)	(0.008)	(0.007)	(0.012)
Observations		206,845				197,049		

Notes: Dependent variable is the standardised test score (z-score) in the specified area(s). Estimation is by OLS. Coefficient shown is the effect of being born 2 weeks before 31st December or 31st March cut-off (as points of 1 standard deviation). Each specification includes a function of birth date and this is specified as follows: linear trend (1); quadratic trend (2); linear trend with effect allowed to change at cut-off point (3) and quadratic trend with effect allowed to change at cut-off point (4). Each specification also includes a dummy for the 31st March cut-off and full set of dummies for year, free school meal status, ethnicity, language spoken at home, and area of deprivation deciles. Estimation accounts for school-level fixed effects; standard errors are clustered by birth date. Symbols: ** significant at 1% level; * significant at 5% level; + significant at 10% level.