Institute for Fiscal Studies

Identifying the drivers of month of birth differences in educational attainment

IFS Working Paper W13/09

Claire Crawford
Lorraine Dearden
Ellen Greaves

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

Identifying the drivers of month of birth differences in educational attainment[1]

Claire Crawford[2], Lorraine Dearden[3], Ellen Greaves[4]

May 2013

Abstract

Children born at the end of the academic year have lower educational attainment, on average, than those born at the start of the academic year. Previous research shows that the difference is most pronounced early in pupils' school lives, but remains evident and statistically significant in high-stakes exams taken at the end of compulsory schooling. To determine the most appropriate policy response, it is vital to understand which of the four possible factors (age at test, age of starting school, length of schooling and relative age without cohort) lead to these differences in attainment between those born at different points in the academic year. However, research to date has been unable to adequately address this problem, as the four potential drivers are all highly correlated with one another, and three of the four form an exact linear relationship (age at test = age of starting school + length of schooling). This paper is the first to apply the principle of maximum entropy to this problem. Using two complementary sources of data we find that a child's age at the time they take the test is the most important driver of the differences observed, which suggests that age-adjusting national achievement test scores is likely to be the most appropriate policy response to ensure that children born towards the end of the year are not at a disadvantage simply because they are younger when they take their exams.

Key words: Age-period-cohort problem, maximum entropy, month of birth, relative age, educational attainment

JEL classification: I21, J24

---

[1] The authors gratefully acknowledge funding from the Nuffield Foundation (grant number EDU/36559) and from the Economic and Social Research Council (ESRC) via the Centre for the Microeconomic Analysis of Public Policy at IFS (RES-544-28-0001). The help of the Centre for Longitudinal Studies (CLS) in provided access to data is also gratefully acknowledged. The Nuffield Foundation is an endowed charitable trust that aims to improve social well-being in the widest sense. It funds research and innovation in education and social policy and also works to build capacity in education, science and social science research. The Nuffield Foundation has funded this project, but the views expressed are those of the authors and not necessarily those of the Foundation.

[2] Institute for Fiscal Studies.

[3] Institute for Fiscal Studies and Institute of Education, University of London.

[4] Corresponding author: ellen_g@ifs.org.uk, Institute for Fiscal Studies, 7 Ridgmount Street, London, WC1E 7AE, UK.

1. Introduction

Children born at the end of the academic year have lower educational outcomes, on average, than those born at the start of the academic year. This finding, documented in the UK[5] and elsewhere[6], is most pronounced early in pupils' school lives, but remains evident and statistically significant in high-stakes exams taken at the end of compulsory schooling. Crawford et al (2010) show that the impact of month of birth on test scores is approximately linear in large scale administrative data: the disadvantage in educational attainment is greatest for the relatively youngest pupils, but is present for those born across the cohort. Crawford et al (2010) replicate this finding accounting for observable characteristics of the household and neighbourhood using survey data[7]. Research to date has been unable to adequately determine the factors that lead to these differences in outcomes for those born at different points in the academic year, which is necessary to determine the appropriate policy response. This paper is the first to apply the principle of maximum entropy to this problem to determine the main drivers of the differences in a problem analogous to separately identifying age, period and cohort effects in a panel dataset.

There are at least four factors that could potentially lead to lower outcomes, on average, for those that are youngest in the academic year compared to the oldest. These are:

1. they are different ages when they sit the tests;
2. they start school at different ages;
3. the amount of schooling they receive prior to assessment differs;
4. their age relative to others in their class or year group differs.

Identifying which factors are most important is critical for policy-makers. For example, if the age at which pupils sit the test is found to have the largest impact, then tests could be taken at different times of the year (which, incidentally, would also change the length of schooling before the test) or the scores could be appropriately age-adjusted.

There are significant challenges to identifying the factors that affect the difference in attainment between those that are oldest and youngest in the academic year.

The first problem is that all factors are highly correlated with one another, which introduces multi-colinearity. For example if a child is oldest when the sit the test, they are also likely to be relatively old in their class, and to have started school at an older age. There are ways to overcome this problem to some extent, by using all the possible variation in the factors that

---

[5] See, for example, Russell and Startup (1986), Bell and Daniels (1990), Sharp, Hutchison and Whetton (1994), Thomas (1995) and Alton and Massey (1998).

[6] See studies for the US (Datar, 2006; Elder and Lubotsky, 2009; Aliprantis, 2011; Robertson, 2011), Canada (Smith, 2009 and 2010), Germany (Jurges and Schneider, 2007; Puhani and Weber, 2007; Mühlenweg and Puhani, 2010), Sweden (Fredriksson and Ockert, 2005), Norway (Strom, 2004), Chile (McEwan and Shapiro, 2008), Australia (Buddelmeyer and Le, 2011), Italy (Ponzo and Scoppa, 2011), Japan (Kawaguchi, 2011), Hungary (Hamori and Kollo, 2011), Malta (Borg and Falzon, 1995) and Brazil (Sampaio et al., 2011). Studies using cross-country international data sets include Bedard and Dhuey (2006), Borghans and Diris (2010) and Sprietsma (2010).

[7] Crawford et al (2010) also demonstrate the impacts of month of birth on wider outcomes, such as self-esteem, socio-emotional development and engagement in risky behaviours. Results show that the impact of month of birth extends beyond educational attainment.

arise from local authorities' and schools' different admissions policies, and different holiday dates (although the correlation remains high).

Second, and more problematically, there is an exact relationship between three of the four factors listed above; a child's age at test is equal to the age they started school plus the length of schooling they have received (aside from small differences that arise due to differences in holiday dates). This linear dependence is a serious problem, as it is not possible to separately estimate the effect of each factor using standard techniques. For example, using linear regression it is only possible to estimate the impact of two of the three linearly dependent factors, holding the third constant, unless functional form restrictions are imposed. This problem is analogous to attempts to estimate the effect on an outcome of interest (for example wages at time *t*) of age at interview, time period and cohort in a panel dataset. Each factor is thought to have a distinct effect on the outcome observed for each adult; *age* determines stages of life relating to education, work and family; *period* (date) determines influences by aggregate factors such as recessions or booms in the economy; *cohort* determines influences in young life such as access to health and education, institutions and peer-groups (Browning et al, 2012). From now on, we use the term age-period-cohort problem (or APC problem) as short-hand for the problem of identifying the distinct impact of factors that from an exact linear combination. The APC model refers to the method we apply to solve this problem[8], which relies on the assumption that all factors are equally important, unless the data informs us otherwise.

This model is an alternative, and more informative, approach to estimating the most important drivers of the differences in outcomes for those born at the beginning and end of the academic cohort. We are the first to apply the APC model to this context (although there is a rich history in other contexts), and therefore build on previous literature in this area by providing estimates that separate the impact of the three linearly dependent variables.

Datar (2006) relies on a functional form assumption to separate the age start school and age at test effect, by assuming that the age at test effect is linear i.e. that the difference in test scores between two children who are six months apart in age is the same regardless of how old those children are. Under this (strong) assumption, and using the differences in pupils' test scores over time as the dependent variable, the effect of absolute age on test scores is differenced out, leaving only the age of starting school effect. Datar finds that the test scores of older entrants increase by 0.12 standard deviations over and above those of the youngest entrants over a two-year period, implying that it is better for children to start kindergarten when they are older. In Datar's context (using data from the Early Childhood Longitudinal Study in the US) there is no length of schooling effect as all children enter kindergarten at the same time, although it is unclear whether or how the relative age effect features in the analysis.

Using Swedish administrative data on the population born between 1935 and 1984, Fredriksson & Ockert (2005) find that increasing school starting age by one year increases grade point average at age 16 by 0.2 standard deviations, and that relative age accounts for only 6% of the difference in test scores at that age. Black et al. (2008) adopt a similar

[8] Another approach to solve the APC problem would be to impose a parameter restriction of some kind: only one restriction is required, for example restricting the effect of two successive cohorts to be the same, or specifying the form (linear or quadratic, for example) of one of the factors.

approach to identify the impact of school starting age on IQ scores and educational attainment using Norwegian administrative data. They find that starting school younger has a significant positive effect on IQ scores at age 18, but little effect on educational attainment.

Fogelman & Gorbach (1978) use data on a cohort of children born in a particular week in March 1958 in Great Britain (from the National Child Development Study), which effectively enables them to eliminate the absolute and relative age effects. Furthermore, regional variation in school admissions policies creates variation in length of schooling, but also age of starting school. Hence, while they report that the length of schooling effect is positive, their identification strategy would seem to lend itself to estimating a combined effect of the two.

Smith (2010) positions himself as building on Crawford et al. (2007), using variation in school admissions policies in British Columbia to estimate an upper bound on the age at test effect and a lower bound on the age of starting school effect (neither of which can be separated from the length of schooling effect). Using administrative data on grade repetition at grade 3 (age 8-9) and literacy and numeracy scores at grade 10 (age 15-16), he finds relatively large upper bound for the age at test effect and relatively small lower bound for age of starting school effects.

For example, Smith (2010) finds that a one year increase in age at test is associated with a 5.8 percentage point reduction in the likelihood of repeating grade 3, compared with just a 0.6 percentage point reduction associated with a one year increase in age at entry (assuming that the schooling effect is minimal so estimates are close to the bounds). Similarly, he finds a reduction of around 10% of a standard deviation in literacy and numeracy test scores at age 15-16 – very similar to our own estimates of the effect of being born in August relative to September on GCSE attainment at the same age – as a result of a one year reduction in age at test, compared to less than 5% of a standard deviation associated with the same change in age of starting school. Smith summarises the impact of the bounds as: “to the extent that the estimates deviate from the bounds due to positive schooling effects, the true test age effect will be lower than reported, and the true entry age effect will be higher”.

We now describe our preferred methodology (that overcomes the limitations of previous applied approaches), the APC model (Section 2), before a discussion of our two complementary sources of data and educational context (Section 3) and results (Section 4). We conclude that a child’s age when they take a test has the most important factor in the differences observed, on average, for those born at the start and end of the academic year, although a child’s relative age within their cohort is also relevant. Length of schooling generally has a positive effect, while the impact of age start school is close to zero, justifying the move towards single entry points to primary schools in England.

2. The age-period-cohort (APC) model

The APC model has recently been explored by Browning et al (2012), and we refer the reader to this paper for a more formal explanation of the method. In what follows, we present the method more intuitively, with the following notation:

- *y* denotes the dependent variable, in our case test scores
- $\boldsymbol{X}_{LD}$ denotes the vector of linearly dependent variables (age at test, age start school and length of schooling)
- $\boldsymbol{X}_{I}$ denotes the vector of independent variables that include relative age and background characteristics such as family income
- $\boldsymbol{\beta}_{LD}$ denotes the vector of coefficients that represent the degree to which the three linearly dependent variables in $\boldsymbol{X}_{LD}$ affect the dependent variable *y*
- $\boldsymbol{\beta}_{I}$ denotes the vector of coefficients that represent the degree to which the independent variables in $\boldsymbol{X}_{I}$ affect the dependent variable *y*
- For ease of notation below we define $\boldsymbol{\beta} = [\boldsymbol{\beta}_{LD},\ \boldsymbol{\beta}_{I}]$ and $\boldsymbol{X} = [\boldsymbol{X}_{LD},\ \boldsymbol{X}_{I}]$

There are two key insights to the method: first, when the dependent variable has a fixed maximum and minimum, the effect of the independent variables must also be bounded. To see this, consider a dependent variable which can take the values of 0 or 1. If there is only one independent variable, and the independent variable and dependent variable are perfectly positively correlated, then the impact of the dependent variable is one in absolute terms ($\beta = 1$). The bounds for the impact of the independent variable are therefore minus one and one: in mathematical notation, $\beta \in [-1{,}1]$, as a perfect negative correlation implies that ($\beta = -1$).

This intuition can be extended to when the dependent variable takes values other than zero and one; whenever the dependent variable has a maximum and minimum, the impact of the independent variables must also be bounded: If $\mathrm{Y} \in [\mathrm{Y_{min}}, \mathrm{Y_{max}}]$ rather than $\mathrm{Y} \in [0{,}1]$, we can re-arrange it so that $\mathrm{Y} = \mathrm{Y_{min}} + (\mathrm{Y_{max}} - \mathrm{Y_{min}})\overline{\mathrm{Y}}$, where $\overline{\mathrm{Y}} \in [0{,}1]$. This means that if $\overline{\mathrm{Y}}$ equals 0, $\mathrm{Y} = \mathrm{Y_{min}}$ and if $\overline{\mathrm{Y}}$ equals 1, $= \mathrm{Y_{max}}$. Note that the intuition described above applies to $\overline{\mathrm{Y}}$; for $\overline{\mathrm{Y}}$, $\overline{\beta} \in [-1{,}1]$. Applying this result to Y means that $\beta \in [-(\mathrm{Y_{max}} - \mathrm{Y_{min}}), (\mathrm{Y_{max}} - \mathrm{Y_{min}})]$ or $\beta \in [(\mathrm{Y_{min}} - \mathrm{Y_{max}}), (\mathrm{Y_{max}} - \mathrm{Y_{min}})]$, and so β is bounded.

The intuition also extends to the case where more than one independent variable is included in the model: if there are K independent variables in the model, then each β_{k} is bounded. Each β_{k} could only equal the upper or lower bound if it perfectly explained the dependent variable, and no other β_{k} did.

The first insight has given us a way to find a range of estimates that are possible for the impact of the factors that might drive differences between those that are oldest and youngest in the academic year on attainment (i.e. it provides set identification). This first step is helpful, but is not completely satisfactory, as the range of these estimates may be large; if the difference between the maximum and minimum values is large, then so is the possible variation that X_k can explain, as then so is the possible range of β_k.

The second key insight provides a method for selecting the most appropriate parameter vector β within these bounds. The method explored by Browning et al (2012) is based on an idea by Jaynes (1957a, 1957b) whose starting point was that, subject to the known data, the probability distribution that best represents the current state of knowledge is the one with the largest entropy. That is, unless there is a reason for believing otherwise, each possible outcome should be regarded as equally likely (or, in our case, that each independent variable has the same impact on the dependent variable, unless we observe otherwise in the data). This is an extension of LaPlaces' *principle of insufficient reason* (Browning et al, 2012;

Sinn, 1980), and has a natural appeal; why should we consider that any one outcome is more likely than another when we have no evidence to suggest that it is? To illustrate this concept, we repeat an example covered in more depth in Browning et al (2012), known as “Jaynes’ dice problem”, where the objective is to work out the expected value of the next roll of a die.

In Jaynes’ dice problem we have a six sided die that has been rolled a large number of times. We know that the die has six sides and that each side is labelled from one to six, and that the probability distribution meets standard conditions[9], but not whether the sides have equal probability of occurring (that is, whether the die is “fair”). We also know the sample average of the large number of rolls, which gives us some information on which sides occurred more often in the large number of rolls; if the die is fair, we expect the sample average to be 3.5; if the die is weighted towards high numbers, we expect the sample average to be higher than 3.5 and vice versa if it is weighted towards low numbers.

The principle of maximum entropy gives us a starting point for making inferences on the basis of this incomplete information; we should draw them from the probability distribution that has the maximum entropy (or smoothness) permitted by the information we do have (Jaynes, 1982). That is, in Jaynes dice problem, each possible outcome should be regarded as equally likely, unless there is reason for believing otherwise. The problem is solved by constrained optimisation by finding the vector of probabilities for the die landing on each side which are as equal as possible, while at the same time being consistent with the sample average of the number of rolls. The solution is that the probability of the die landing on each side is 1/6 if the sample average is 3.5 (consistent with a fair dice) but the probabilities are higher (lower) for higher values of the dice if the sample average is higher (lower).

This intuition can be applied to our problem at hand: what values of the parameter vector $\boldsymbol{\beta}$ makes outcomes as equally likely as possible, while being consistent with the data we observe? Again, this approach has a natural appeal; why should we consider that any one of the factors is more important than another when we have no evidence to suggest that it is? For example, if the data tells us that children that are older when they sit the test have higher test scores, on average, conditional on other factors, then the proposed solution to the problem would result in a higher element of $\boldsymbol{\beta}$ for the independent variable that represents being older when the test is sat (call this element β_k). If, on the other hand, there was no difference in the test scores of children that sit the test at different ages, conditional on other factors, then β_k will not differ from other β_s, where $s \neq k$.

The link between Jaynes’ dice problem and the *apc* problem (and indeed the age-at-test-age, age-start-school, length-of-schooling problem) is the ability to transform the hunt for the most appropriate parameter vector $\boldsymbol{\beta}$ into a hunt for the most appropriate vector of probabilities (or probability distribution) that are consistent with the data observed (which is exactly the approach used in the dice example above). This transformation, the details of which are not included here but are explained fully in Browning et al (2012), is possible whenever the dependent variable is bounded, so that the parameter vector $\boldsymbol{\beta}$ is also bounded. Within these bounds, it is possible to express each possible parameter vector $\boldsymbol{\beta}$ as a weighted combination of the upper and lower bounds for each independent variable (for example placing more weight on higher or lower values of certain independent variables).

[9] Namely that all probabilities are non-negative and sum to one.

The weighted combination that is most even across all possibilities is preferred, consistent with the data observed (where the preferred "weight" is synonymous with the preferred probabilities). This unique vector of probabilities is then converted to a unique parameter vector $\boldsymbol{\beta}$, our final solution.

Note that in relation to Browning et al (2012) we include additional independent variables in the vector $\boldsymbol{X}$, namely $\boldsymbol{X} = [\boldsymbol{X}_{LD}, \boldsymbol{X}_I]$ rather than $\boldsymbol{X} = \boldsymbol{X}_{LD}$. In practice, this means that the conditional relationships are estimated using a larger number of independent variables, but the same logic applies: unless the (conditional) data give reason to believe that one factor is more important than another, all β_k will be equal.

To identify the impact of the linearly dependent variables, it is important that they have sufficient variation across pupils (for example that length of school prior to the test is not the same for all pupils). In this case, the conditional relationship between length of schooling and pupils' outcomes would be uninformative. An additional concern is the correlation between the three linear dependent variables and the fourth factor of interest, relative age. In the extreme case that relative age is perfectly correlated with one of these factors then the influence of the two can't be separately identified. For example, if age start school and relative age are perfectly correlated, then the model would assign an equal value of β_k to each factor under the principle of maximum entropy. This may create spurious relationships between pupil outcomes and factors (like relative age) that are highly correlated with the linearly dependent variables.

To overcome these problems, we use two complementary sources of data, discussed in more detail below. The first, the National Pupil Database, contains a large number of pupils. For the cohorts of pupils we use there is sufficient variation in the age of starting school across pupils due to differences in admissions policies across different local authorities (administrative bodies) in England. Relative age and age at test are highly correlated, however, as pupils are assessed at the same point in time. The second, the Millennium Cohort Study, contains a sample of pupils much smaller than the National Pupil Database, but has the advantage that pupils are assessed at different points in time, reducing the correlation between relative age and age at test. The relative merits of each source of data are discussed further in the following section.

3. Educational context and sources of data

3.1 School admissions and exit policies

The academic year in England runs from 1 September to 31 August, and is split into three terms (autumn, spring and summer). It is a statutory requirement for children in England to start school by the beginning of the term after they turn five, but within these confines, school admissions policies are set by local (rather than central) authorities, and in most cases children start school considerably earlier than this. The most common admissions policy in

2005 was for all children to start school in the September after they turn four.[10] This means that children born in September are the oldest in the academic year, and children born in August are the youngest.[11]

The fact that admissions policies are determined locally means that there is considerable variation both geographically and over time in the age at which children born at different times of the year start school. For example, the second most common admissions policy in operation in England in 2005 stated that children born between September and February should start school in the September after they turn four, while children born between March and August should start school in the January after they turn four. This means that a child born in August will start school when they are four years and five months old (rather than four years and one month old) if they live in an area which follows this policy rather than the one described above; they would also receive one term less tuition in their first year at school.

This variation is extremely important for our attempts to separately identify the factors that are driving the differences in outcomes that we observe, because it generates variation in the age at which children born on the same day start school and sit the tests, and how much schooling they receive before they do so.

Pupils in England are currently required to remain in school until the last Friday in June of the academic year in which they turn 16.

2.2 National achievement tests

Pupils in state schools in England must participate in a number of national achievement tests during their school career, at ages 5, 7, 11 and 16. Those who choose to stay on beyond compulsory education are additionally tested at ages 17 and 18.

We focus on results in primary school where the relative difference between the oldest and youngest in the academic year is largest. At age 7, pupils are assessed on the basis of reading, writing, speaking and listening, maths and science. At the end of primary school (age 11), they are assessed and tested in English, maths and science. For our cohorts of interest, the school tests are taken at the same point in time for all pupils, and questions are set and marked outside the school. The tests are therefore relatively objective and unbiased.

We compare the results in national tests at age 7 with independent tests administered as part of a nationally representative survey of young children, called the British Ability Scale (BAS). When the children were age 7 this comprised word reading, maths and pattern construction tests. We use the first two of these tests as they are most directly comparable to the national (KS1) tests. For the word reading test the child reads aloud a series of words

---

[10] Only schools under local authority control (e.g. community schools) must follow the local admissions policy, while other schools (such as those that are voluntary aided, voluntary controlled, academies or free schools) can set their own.

[11] It is possible for parents to defer the date on which their child starts school up to and including the statutory date (with the agreement of their local authority), but if they do so, then children tend to be placed into the year above (the correct academic cohort for their age), thus reducing the amount of time that they spend in school overall. This means it is relatively rare for parents to do this: http://media.education.gov.uk/assets/files/pdf/s/school%20admissions%20code%201%20february%202012.pdf.

presented on a card. The maths test is adapted from the NFER Progress in Maths test, where children complete various number based tasks, covering the topics of numbers, shape, space and measures, and data handling[12].

All outcomes are standardised on a nationally representative sample (the whole cohort for national tests and the representative survey for tests from the survey) to have a mean of zero and standard deviation of one. We can therefore directly compare the results for each test, assuming that KS1 and BAS tests are comparable.

3.3 Data

We use two sources of data to estimate the APC (or age at test age, age start school, length of schooling) model. The datasets have complementary strengths and weaknesses, largely determined by the correlation between the four potential driving factors, which arise for three reasons: first, the cohort, as single term entry has become more common over time; second, the timing of the test (either administered at different points in the year or on a single day); third, the child's age, as our measure of relative age becomes more correlated with age at test as children age.

Our first source of data is the National Pupil Database (NPD), which contains information on the performance in national attainment tests for every pupil in a state school in England. These data are large (with information on over half a million pupils per cohort) and accurate, as schools have a statutory duty to report this information to the Department for Education. We focus on three cohorts of pupils: those born between September 1995 and August 1996, September 1996 and August 1997, and September 1997 and August 1998. We choose these cohorts as the geographical variation in admissions policies was relatively high, and for comparability with Crawford et al (2007). For these cohorts, it was relatively common for entry to school to be different for pupils born in different months; only around half of local authorities had a single entry point for the academic year 1999/2000, for example, when the first cohort would have begun school.

The second is the Millennium Cohort Study (MCS), which contains information on one national attainment test and independently assessed cognitive development for a sample of pupils in England born between September 2000 and August 2001. Children born in England that are part of the Millennium Cohort Study (MCS) form a single academic cohort, born between September 2000 and August 2001. The original sample of MCS cohort members born in England was 12,390. Of these, 8,953 remain in the sample when the cohort members are surveyed at wave four (around the age of 7). We adjust for non-random attrition by applying the survey weights provided that correct for the survey design (stratified sampling with over-sampling in neighbourhoods that are relatively deprived and have a high proportion of individuals from ethnic minorities) and non-random attrition on the basis of observable characteristics. Of these, we observe 6,789 with all outcome variables of interest that form our final sample[13]. The majority of these children started school in September 2004, as by this time single form entry was more common across England; around 60% of local authorities had a single entry policy. This reduces the variation in age starting school,

[12] Millennium Cohort Study First, Second, Third and Fourth Surveys: A guide to the datasets, 2010

[13] Two cohort members are dropped as no information on their local area characteristics are available.

but additional variation in length of schooling and age at test arises through the use of comparable attainment tests that were taken at different points in time (rather than at the same point in time as for the national achievement tests).

In each source of data we can infer age at test, age start school, length of schooling and relative age (defined as age relative to others in their year group, as information on class groupings is unfortunately unavailable).

The strength of the NPD is the large and nationally representative sample, and, for the cohorts we select, a high degree of geographical variation in admissions policies. This variation helps to separately identify the impact of the four potential factors as it reduces the correlation between age start school and age at test. The weakness is that the correlation between relative age and age at test is extremely high (between 0.96 and 0.97), as the independent variation in relative age comes solely from the random differences in the distribution of children’s months of birth across schools, and the early variation in relative age under different admissions policies: in single term entry areas relative age will be constant for all terms, but relative age will vary over time in multiple entry areas where younger children enter the cohort at different points through the academic year.

The strength of the MCS is that the high correlation between relative age and age at test is reduced (to between 0.81 and 0.89); the variation in assessment dates means that children with the same relative age could be tested months apart. The disadvantage of the MCS is the smaller sample size, and slightly less geographical variation in admissions policies (as the MCS cohort is slightly younger than our chosen NPD cohorts) which means that the correlation between age start school and relative age is higher (0.83 compared to 0.73). Full details of the correlation between all factors in each source of data are provided in Appendix Table 1.

For each dataset, **age at test** for those taken in school is coded in relation to a single point in May. In practice, the assessments will take place over one week (rather than one day), but abstracting from this minor point should not influence our results and conclusions. For the BAS tests, age at test (in days) was recorded. As noted above, there is more variation in age at test and length of schooling prior to the test for the BAS. To illustrate this, Figure 1 shows the variation in age at test for the national achievement (KS1) and BAS tests, for pupils born in September. Age in days has a range of over 200 days in the BAS test, and by around 30 days (consistent with the variation between the oldest and youngest born in September) for the KS1 tests.

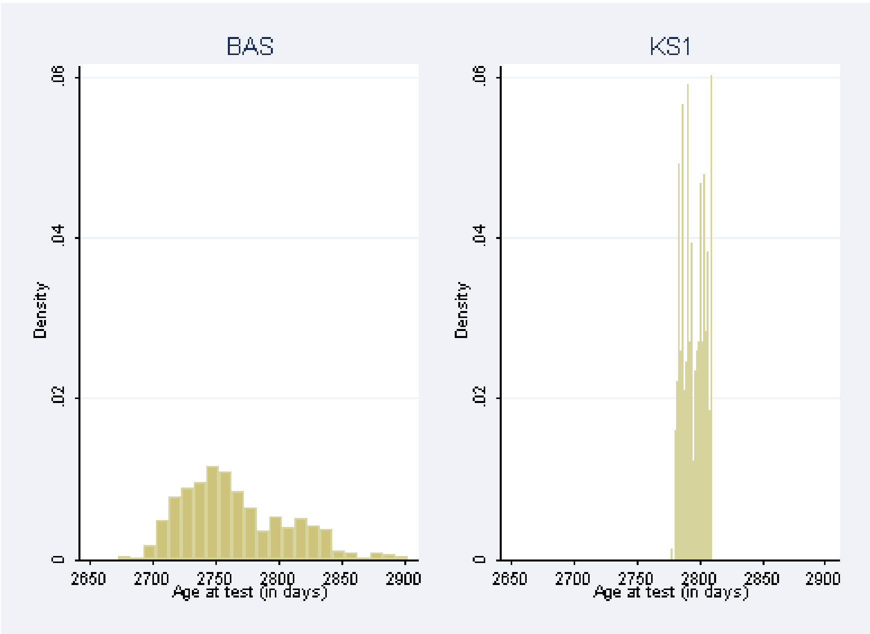
Similarly, the BAS exhibits larger variation in length of schooling prior to the test. Figure 2 shows that length of schooling before the KS1 test is highly concentrated, but, given the variation in date of test, there is a relatively large distribution for the BAS tests.

Date of entry to the school is observed in the NPD for each pupil. The date is incorrect for those that entered a school with an attached nursery, however, as it may instead represent the date that they started nursery. In these problematic cases we impute date of entry to the primary school (in reception year), which is described in Appendix Table 2 for each source of data.

Length of schooling is calculated as the date of test minus the date of entry to the school. We therefore abstract from the small variation in length of schooling that occurs due to variation in holiday dates.

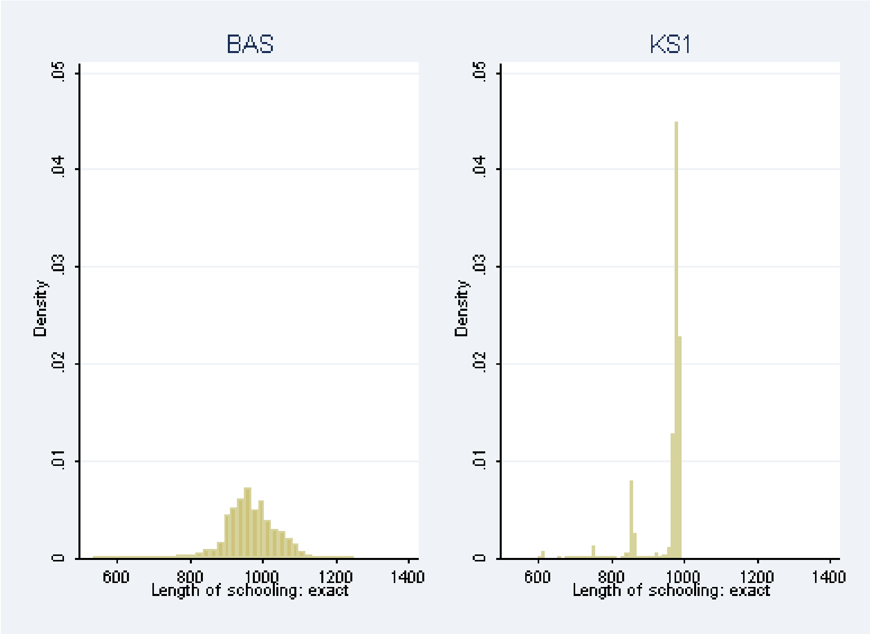
For each pupil, **relative age** is calculated as the proportion of pupils in the school cohort that are younger in each school term prior to the test, weighted by the length of each term. For example, in term one if there are 30 pupils in the school cohort, and a pupil is the 10th oldest, then relative age in this term would be two-thirds ((30-10)/30). For the cohorts of children in the NPD we use, there was significant variation in school entry points for children born in different months: some local authorities operated a single entry point system compared with others that had three entry points (described further in Crawford et al, 2010). This variation means that in some areas relative age will be different across school terms, as younger children join the cohort. There may also be natural variation as some children enter or leave the cohort at later ages. We therefore calculate equivalent figures for all school terms prior to the test. These are all weighted by the length of each term to create the total measure of relative age.

Figure 1: Variation in age at test for those born in September



Note: The sample is MCS cohort members born in September that are present at wave 4 of the survey and for whom we observe KS1 and BAS outcomes.

Figure 2: Variation in length of schooling prior to test for pupils born in September



Note: length of schooling is the exact difference between age at test and age started school. The sample is MCS cohort members that are present at wave 4 of the survey and for whom we observe KS1 and BAS outcomes.

Characteristics of each pupil and their surrounding area are included to account for differences in pupil outcomes that are correlated with our four factors of interest (for example age start school) but only through the characteristics of their home and neighbourhood. For example, more educated parents may be more likely to move to (or campaign for) single form entry areas. We know that children's cognitive development is affected by their parents' level of education. Any correlation between academic outcomes and age start school (and length of schooling prior to the test) may therefore arise because of the admissions policy or because of the parents' characteristics. We can observe a relatively rich set of background characteristics to account for this endogeneity to some extent. In practice, our results are robust to the inclusion or exclusion of background characteristics, suggesting that there is relatively little selection into areas that systematically affects children's test scores.

In the NPD the characteristics we account for are: binary indicators for eligibility for free school meals, English as an additional language, sex, ethnicity, quintiles of neighbourhood deprivation, and continuous variables for the proportion of adults in the local area with a high level of education and the proportion of adults with each level of socio-economic position.

The characteristics we are able to account for in the MSC are the same for the NPD analysis with the addition of: binary indicators for household income quintile, household work status, mother's age at first birth, household marital status, mother and father's level of education, mother and father's occupational status, housing tenure, whether the child was breastfed, birth order (including whether a multiple birth), and the child's birth weight. Full details of these independent variables are included in Appendix Table 3.

The three linearly dependent variables are measured in days, and relative age is a score that ranges from zero to one (where one is the relatively oldest) in each dataset. Age at test, date of entry and length of schooling are converted to be measured in months to increase the precision of our results[14]. We also exclude observations where one of the linearly dependent variables has an uncommon value (where less than 50 observations share that value) and those that started school outside normal dates (where less than two percent of the pupils started school in this month). Our results and conclusions are robust to converting these variables to be measured in fortnights or weeks rather than months and imposing a different rule for restricting to relatively common values only. We present the results without excluding those with non-standard school entry dates in Appendix 1, where the substantive results are unchanged (but the impact of length of schooling becomes non-linear as those with non-standard entry dates have systematically lower levels of attainment).

In both sources of data all outcomes are standardised within cohort to have a mean of zero and standard deviation of one. We can therefore directly compare the results for each cohort (including across the two datasets), and can combine cohorts in the NPD to increase the size of our sample (although our results are robust to using separate cohorts).

4. Results

We present the impact of our four factors of interest estimated using the APC model graphically for ease of presentation, but results are presented in full in the online appendix. In all cases, the excluded category is the first value, so the graphs are plotted relative to the youngest at the test, the youngest when starting school, the fewest months of schooling prior to the test, and the youngest relative to their peers, respectively. The solid line represents the point estimate, while the dotted lines represent the 95 percent confidence intervals.

We are able to document the relative impact of each of the four factors in two samples of pupils (those in the NPD born between September 1995 and August 1998 and those in the MCS with observable KS1 and BAS outcomes) and at two ages (age 6/7 and age 10/11) in the NPD sample. This allows us to explore whether the impact of the four factors is consistent by age, and in different contexts. We begin with a summary of results at age 6/7 (KS1) for the NPD and MCS samples, before discussing results for pupils aged 10/11 and our general conclusions.

NPD: KS1

We observe academic attainment measured by externally set and marked tests at age 7 (KS1) and age 11 (KS2). We focus on total scores, including both maths and English scores, as the patterns for both subjects are consistent (these results are available from the authors on request). The left hand panel of Figure 3 shows the results where only the three linearly

[14] To convert the variables measured in days to variables measured in months we divide each by 30 and round to the nearest month. This improves the precision of the estimates as there are more observations for each value and consequently less noise in the conditional estimate.

dependent variables are included in the model. It is clear that age at test has the largest impact of the three linearly dependent variables: the difference between the oldest and youngest pupil when the test is taken is around 0.4 standard deviations. This is a large effect, equivalent to around 40% of the difference between children born to mothers with a degree and mother's without any formal qualifications at the same age. In contrast, there is a smaller impact of the age of starting school and the length of schooling, around 0.2 standard deviations between the minimum and maximum values.

The right hand panel of Figure 3 shows the results where relative age is additionally included in the model, which allows us to consider the relative impact of all our potential driving factors. However, the results must be taken with some caution given the high correlation between age at test and relative age in this sample. If age at test is perfectly correlated with relative age, then the APC model would allocate equal weight to each factor. In fact, the positive impact of age at test remains large and significant, with around 0.3 standard deviations between the oldest and youngest in the sample. This implies that relative age (which is highly correlated with age at test) accounts for around one quarter of the impact of age at test. The maximum impact of age start school declines by half, to around 0.1 standard deviations from 0.2 standard deviations. As expected, therefore, the difference in outcomes between the relatively oldest and youngest is the remainder: around 0.2 standard deviations. This suggests that the impact of these two factors is at least partly driven by relative age within cohort, while impact of length of schooling is largely unaffected.

As a robustness check, Figure 4 presents the results of the APC model accounting for the observable characteristics of households described in Section 3. Accounting for these additional characteristics of the parents marginally increases the standard errors of the estimates, but the evident patterns are largely unchanged; the impact of the age of starting school is largely zero, while the length of schooling and relative age effects have a maximum impact that is slightly less than the maximum impact of age at test. The robustness of these results to the addition of household characteristics suggests that endogeneity is not a significant problem in this sample; according to observable characteristics of the households at least, the impact of the four factors is not diminished. Appendix 1 presents the results of the model when those with non-standard entry dates are included in the sample. The overall findings are robust, with only the pattern for length of schooling differing, with peaks and troughs that coincide with standard and non-standard entry points, respectively.

There are interesting differences between results for males and females. Figure 5 shows that the impact of age at test is roughly similar for males and females (around 0.3 standard deviations between the oldest and youngest), but relative age is more important for males; the impact of relative age is around double that for females, and significantly different. This suggests that males and females, on average, respond differently to being the youngest in the classroom, with a more detrimental impact for males. The impact of other factors are largely consistent for the two samples, however.

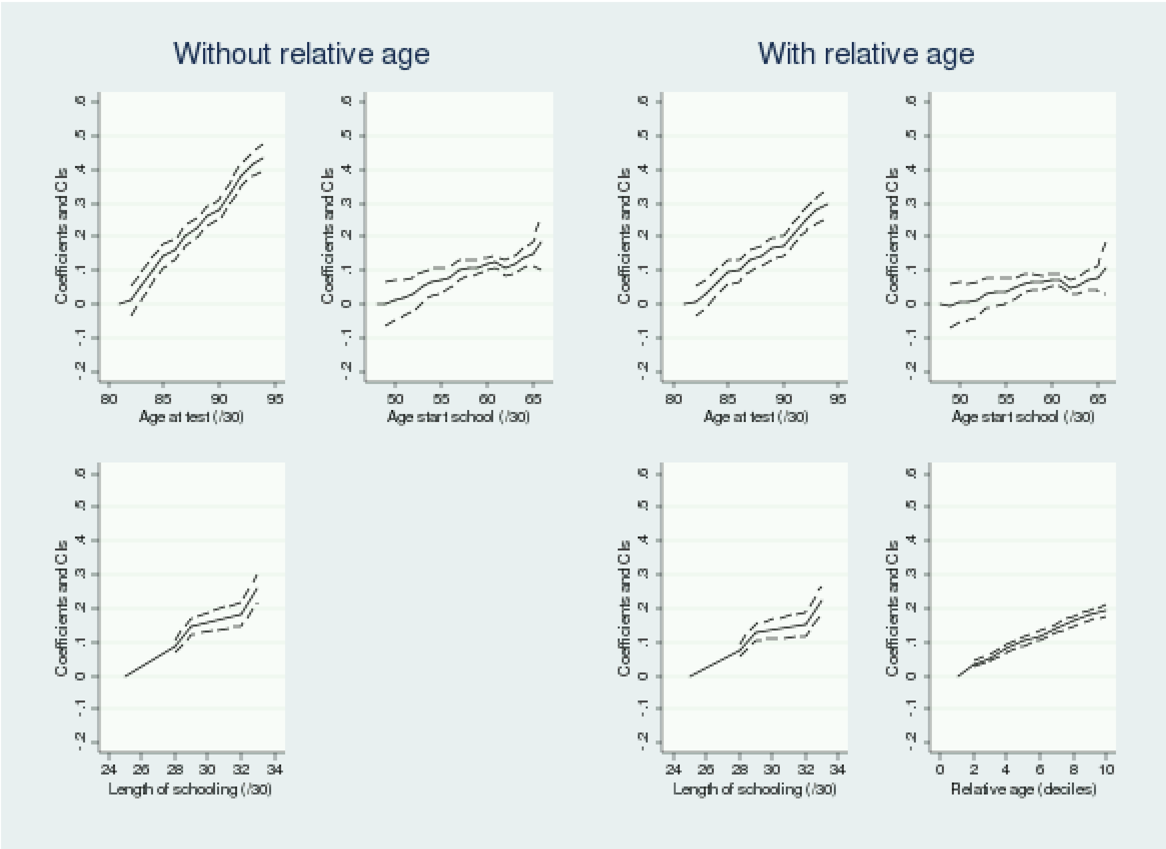
To summarise the results for KS1, the differences in attainment in maths and English for those born in different months of the academic year seem to be driven primarily by a pupil's age at test, while relative age and length of schooling are also important. Table 1 shows the contribution of each factor (age at test, age start school, length of schooling and relative age) to the overall difference for representative pupils born in different months of the academic year[15]. For comparison, the coefficients from linear regression (accounting for the same background characteristics) are presented[16]. As expected the combined impact of the four factors is roughly equal to the coefficient for each month of birth relative to September born pupils (the oldest in the academic cohort). For example, the representative pupil born in August would be 0.6 standard deviations below the representative pupil born in September, according to the estimates from the APC model. This compares with an estimate of -0.583 from the linear regression, which represents the difference, on average, between those born in August and September. Age at test contributes 45% of the total difference between our representative pupils born in August and September implied by the APC model. Relative age is the second most important factor, contributing 32%, and age start school and length of schooling contribute a relatively small amount. There is a similar picture for representative pupils from all months; age at test is the largest contributing factor, followed by relative age. Note that the contribution of length of schooling is zero for some months as the median pupil for this month has the same length of schooling as the median pupil born in September.

Our results show that there are some differences, on average, between males and females; females are largely unaffected by relative age, which in contrast is a relatively important factor for males. Before discussing the policy implications of these results, we present the results for the impact of these factors on attainment using our complementary source of data, the Millennium Cohort Study, which allows us to examine the impact of these factors on an external test taken outside school. As the date of the test varied across pupils, this source of data increases the variation in age at test and length of schooling, and reduces the problem of multi-colinearity, although the sample size is much smaller.

[15] Our representative pupils have the median value of each factor for their specific month of birth. We assume that background characteristics are randomly distributed across pupils born in different months.

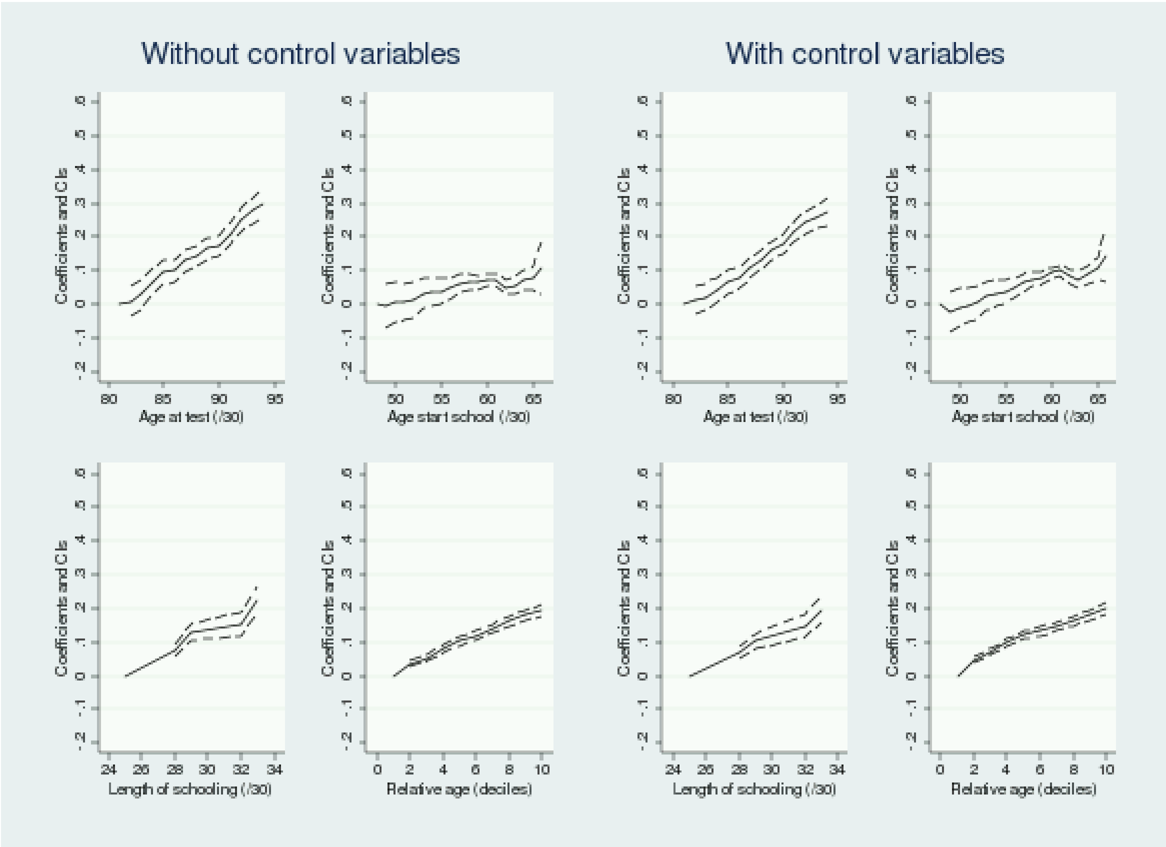
[16] These OLS estimates are based on regression where the outcome of interest is the dependent variable, and the independent variables are a set of dummy variables for month of birth and a set of background characteristics. These OLS coefficients therefore represent the impact of the four factors that vary, on average, between those born in different months of the year.

Figure 3: NPD: Impact of age at test, age start school, length of schooling and relative age on attainment at KS1



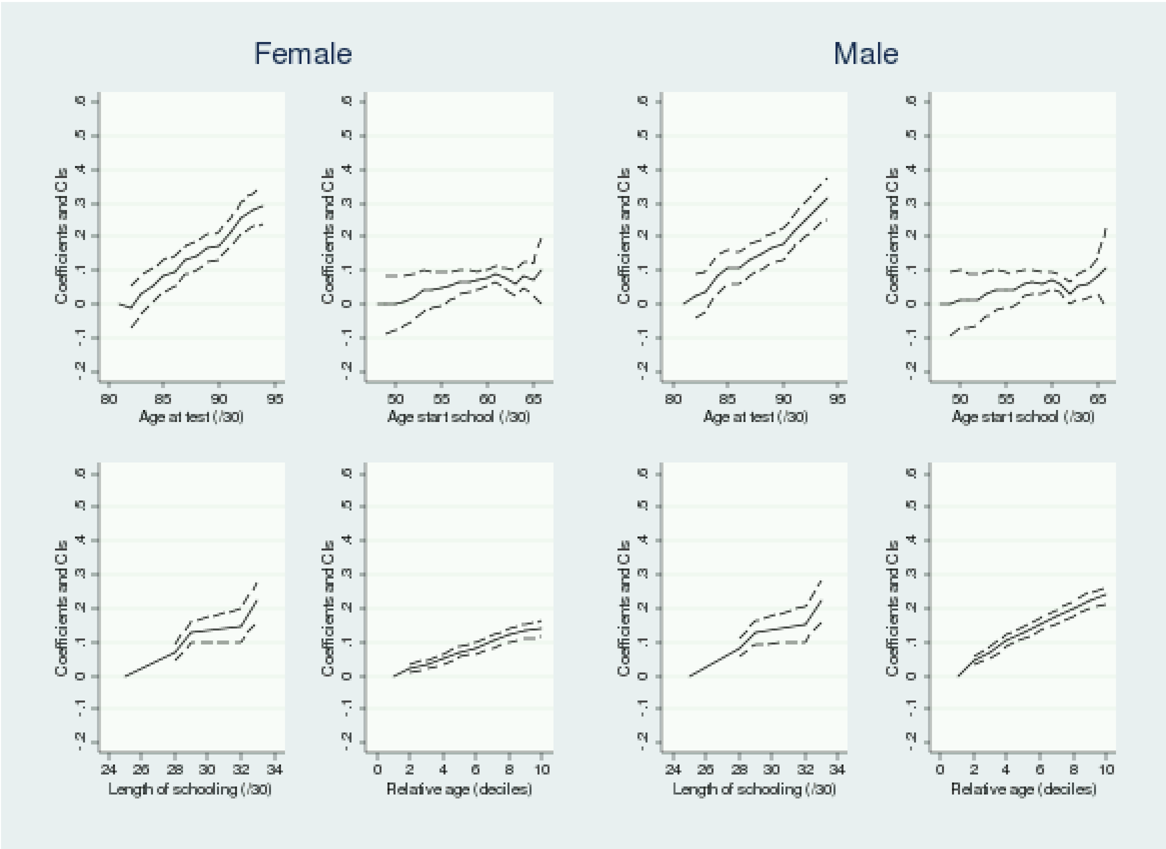
Note: Cohorts of pupils born between September 1996 and August 1999 are included. Cohort dummies are included in the model. Those starting school outside standard dates are excluded from the sample.

Figure 4: NPD: Impact of age at test, age start school, length of schooling, relative age (accounting for independent variables) on attainment at KS1



Note: Cohorts of pupils born between September 1996 and August 1999 are included. Cohort dummies are included in the model. Those starting school outside standard dates are excluded from the sample.

Figure 5: NPD: Impact of age at test, age start school, length of schooling and relative age on for females and males at KS1



Note: Cohorts of pupils born between September 1996 and August 1999 are included. Cohort dummies are included in the model. Those starting school outside standard dates are excluded from the sample.

Table 1: Contribution of each factor to overall difference

Month of birth	OLS (relative to September)	Total implied by APC model (relative to September)	% age at test	% age start school	% length of schooling	% relative age
August	-0.583	-0.600	44.83	11.17	12.17	31.83
July	-0.534	-0.542	45.39	12.36	13.47	28.78
June	-0.477	-0.480	43.75	12.29	15.21	28.75
May	-0.423	-0.409	44.50	10.02	17.85	27.63
April	-0.368	-0.400	44.50	9.00	18.25	28.25
March	-0.313	-0.271	54.61	12.55	0.00	32.84
February	-0.259	-0.241	56.02	14.11	0.00	29.88
January	-0.210	-0.191	59.16	14.14	0.00	26.70
December	-0.149	-0.145	72.41	7.59	0.00	20.00
November	-0.092	-0.103	64.08	7.77	0.00	28.16
October	-0.041	-0.047	57.45	19.15	0.00	23.40
September	Reference	Reference	N/A	N/A	N/A	N/A

Note: based on cohorts of pupils born between September 1996 and August 1999 (who sat Key Stage 1 in 2003-04, 2004-05 or 2005-06). The first column reports the relevant coefficient from a linear regression where the dependent variable is standardised Key Stage 1 scores and independent variables are binary indicators for month of birth and a set of background characteristics. The second column reports the calculated difference based on the median characteristics of those born in each month and the estimates from the APC model. The remaining columns report the calculated contribution of each factor to the median (representative) pupil born in each month. Those starting school outside standard dates are excluded from the sample. Note that the contribution of length of schooling is zero for those born between November and March, as the median pupil in each of these months starts school at the same time as the median pupil born in September.

MCS: KS1

The results of the APC model for the MCS when pooling KS1 and BAS test scores are broadly consistent to those for the NPD: Figure 6 shows that the impact of age at test is the strongest of the three linearly dependent variables, there is a generally positive impact of length of schooling, and the impact of age start school is closest to zero. There is some suggestion that starting school at a late age is detrimental to performance on the KS1 and BAS tests in this sample which was not present in the NPD sample, driven by the select group of pupils that start school in the September after they turn five. The difference between the oldest and youngest when the test is taken appears slightly larger for the MCS sample, although the difference with the NPD sample is not significant.

There is one important distinction in the results, however; relative age is less important for the MCS sample, in fact close to zero and insignificant. To explore whether this is driven by the inclusion of BAS tests or the different sample of pupils, Figure 7 presents the equivalent results using the KS1 tests only and BAS tests only. The standard errors around the estimates increase as a result of the smaller sample sizes (we are now including one observation per child rather than two), so that differences between the minimum and maximum values of all four potential factors are generally insignificant. The positive impact of relative age is more evident for the KS1 tests, however, suggesting that this factor is more important for tests taken in school (or when the correlation between this factor and others is higher). The impact of age at test is more important for the BAS tests, while the impact of this variable for KS1 tests is slightly lower than for the NPD. The slight positive impact of length of schooling and impact close to zero for age start school is more clearly evident for the sub-sample of BAS assessments. This suggests that the driving factors for tests taken in school may be different than tests taken in a different context, in particular that the impact of relative age may be greatest for tests taken in the school environment. Alternatively, the contrasting results between BAS and KS1 assessments may be due to the additional variation and subsequently lower correlation between our four factors of interest in the BAS data, which may make these results more robust. For this reason we focus our discussion on results from the BAS data.

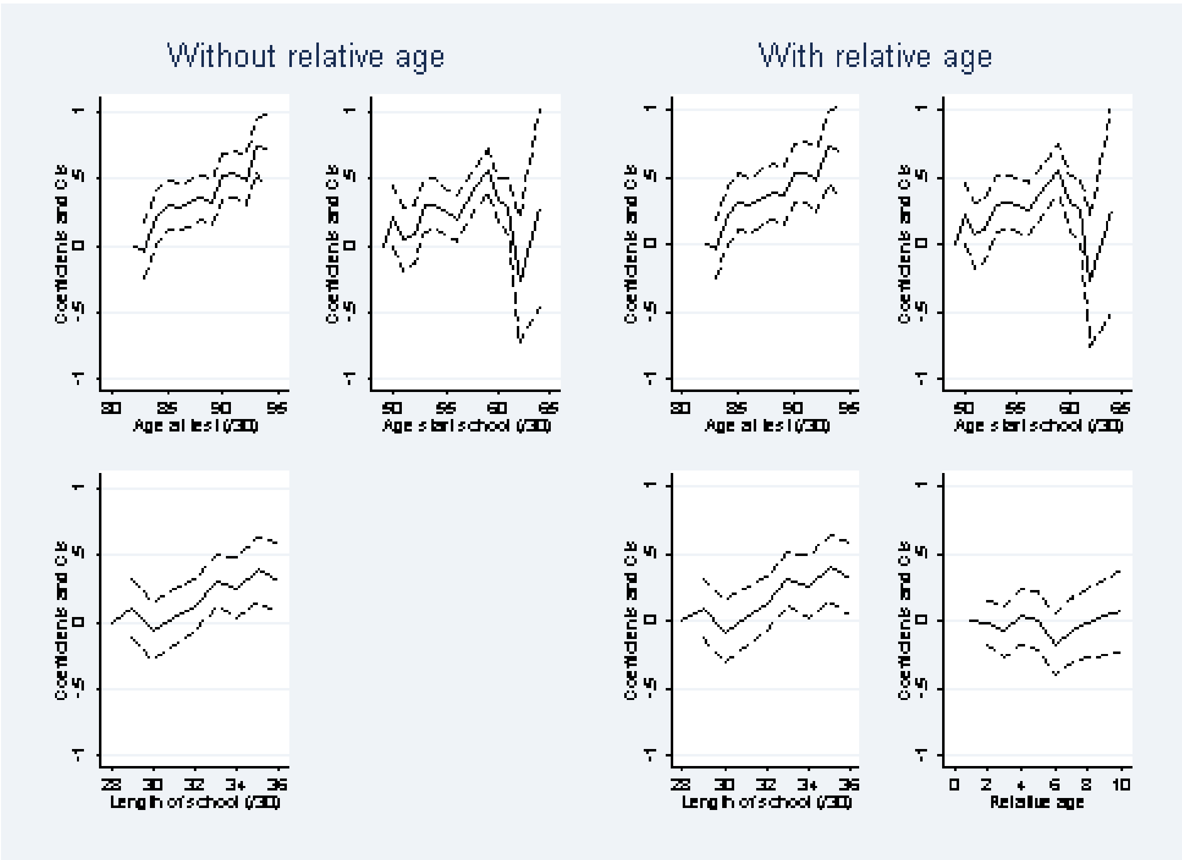
Figure 8 shows that the findings for the BAS assessments are robust to the inclusion of household characteristics observable through administrative data (the same as those used in the analysis of NPD) and more detailed household characteristics available in the survey, although standard errors increase.

Unlike the results in the NPD, males and females are equally affected by their relative age: Figure 9 shows that a child's relative position in their school cohort does not affect either males or females, in contrast to findings in the NPD sample that the impact of relative age for males was particularly strong. Males that start school at an older age (the select group of pupils that begin school in the September after they turn five) do seem to be particularly disadvantaged in terms of KS1 results, while females are largely unaffected.

Results suggest that the driving factor of differences in outcomes at age 7, on average, for those born in different months of birth is the age at test, although length of schooling and relative age also contribute. These results are broadly consistent for different samples of pupils, and for different tests. There is some indication, however, that relative age is a more important factor in tests taken in school (or when the correlation between this and other factors is lower). We therefore interpret the positive impact of relative age with caution, as it could be driven through the correlation with age at test which is especially high in the KS1 sample.

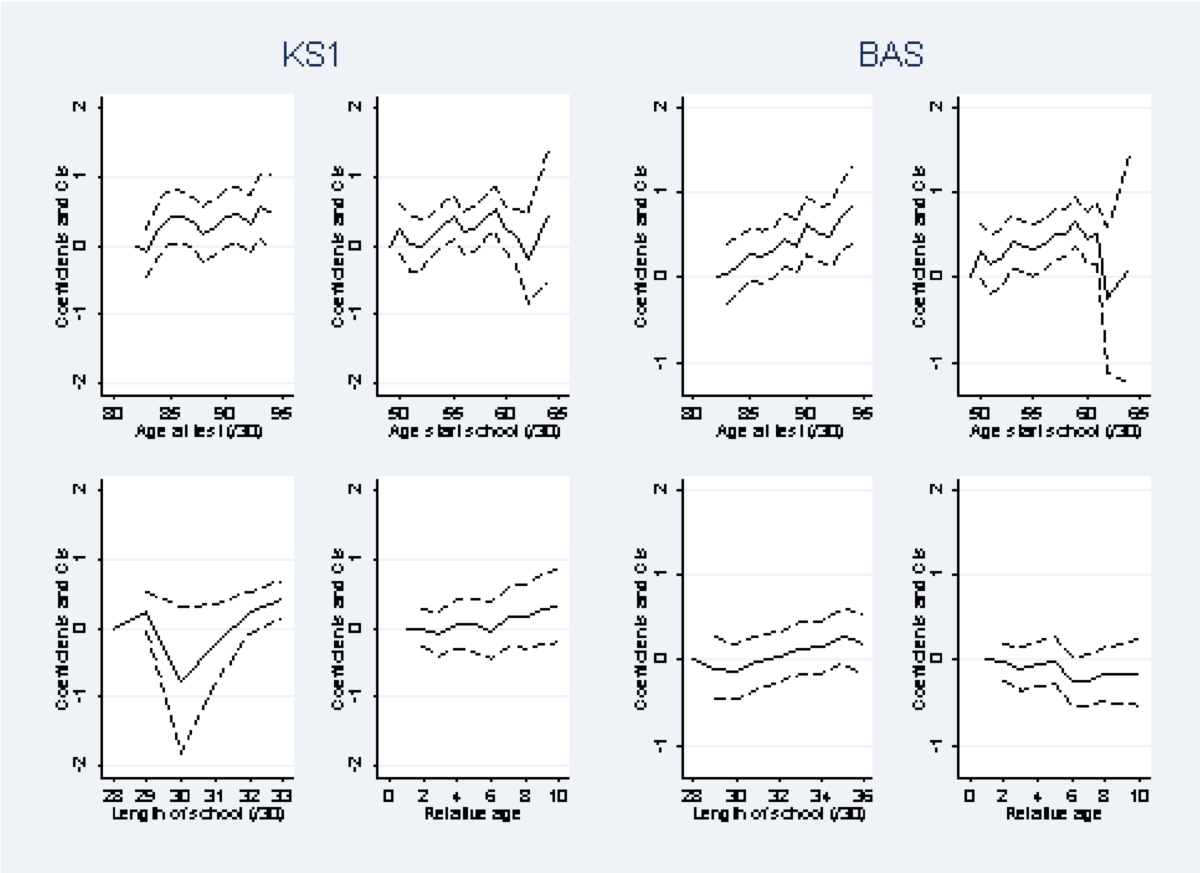
It may be that the some factors become relatively less important as children age, however. To investigate this we now present results for attainment tests taken at age 11, observable in the NPD.

Figure 6: MCS: Impact of age at test, age start school, length of schooling and relative age on attainment at age 7 (KS1 and BAS)



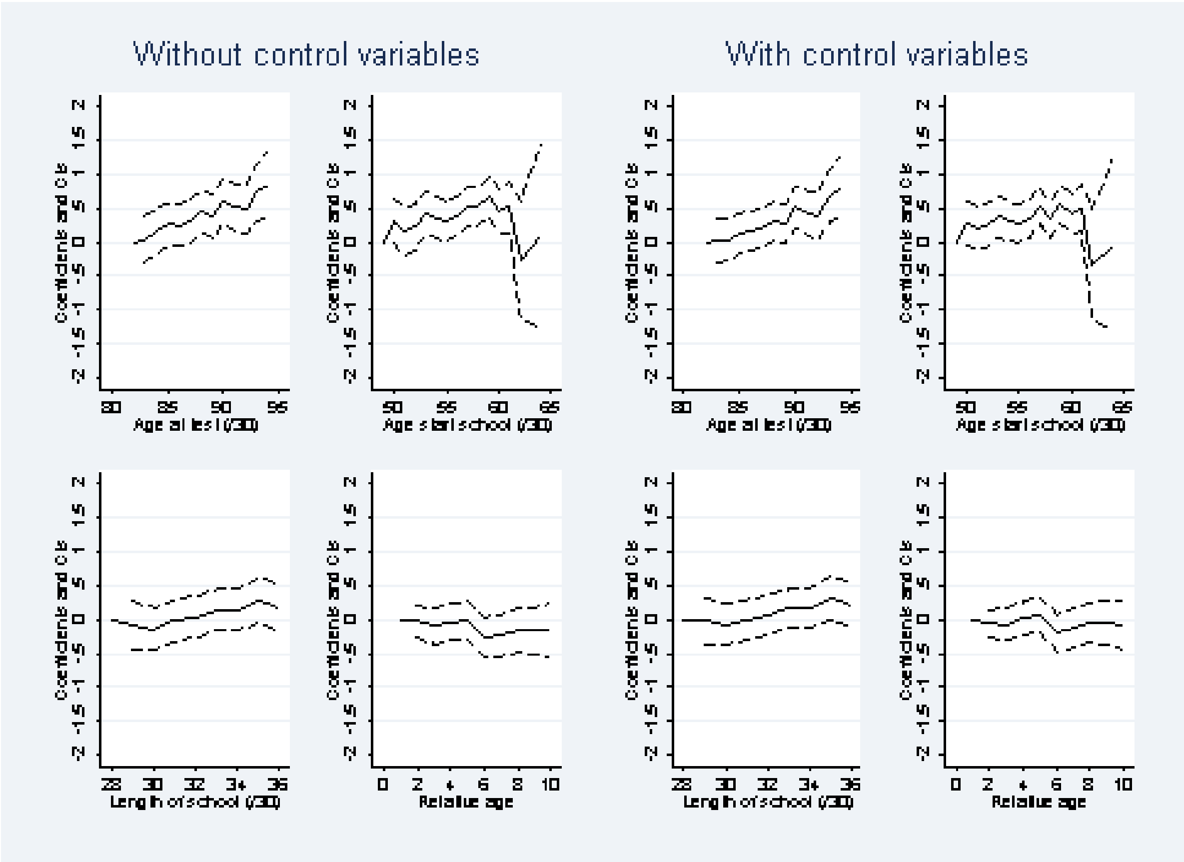
Note: The sample is MCS cohort members that are present at wave 4 of the survey and for whom we observe KS1 and BAS outcomes. Those starting school outside standard dates are excluded from the sample.

Figure 7: MCS: Impact of age at test, age start school, length of schooling and relative age on attainment at age 7 (KS1 and BAS separately)



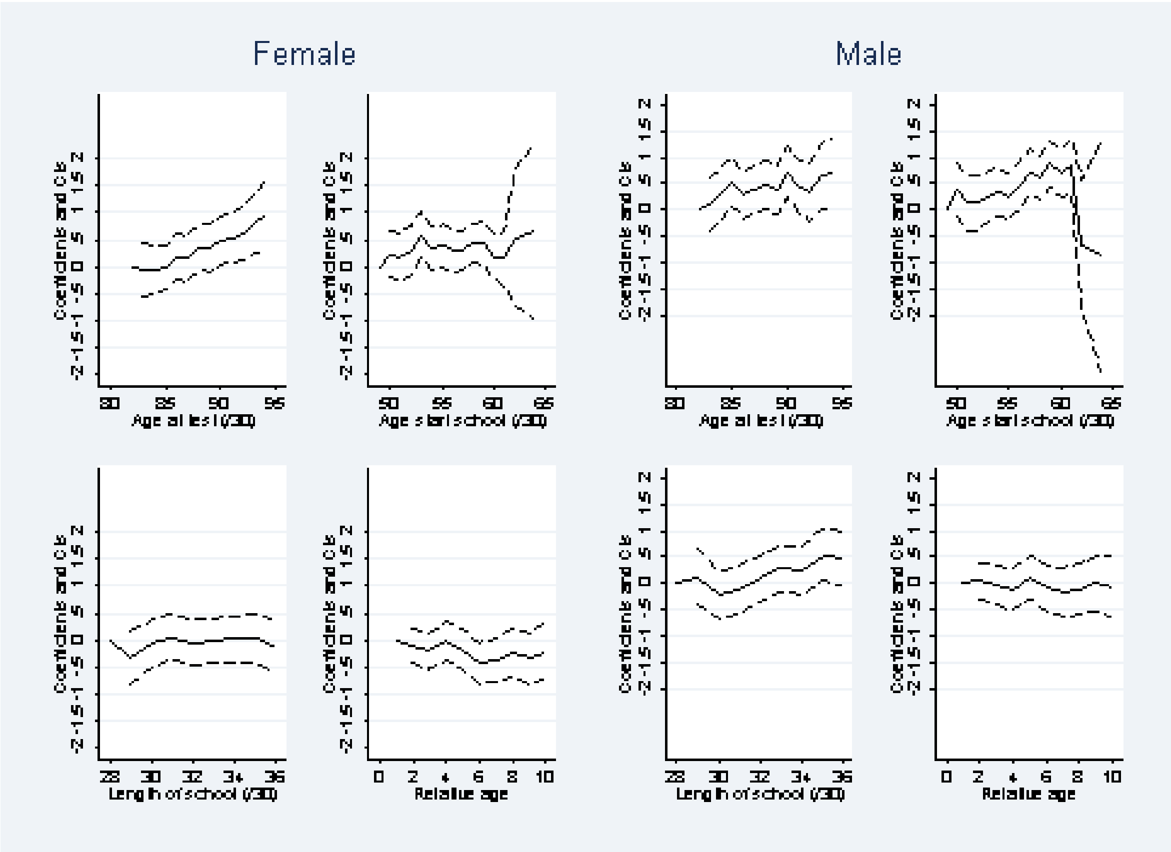
Note: The sample is MCS cohort members that are present at wave 4 of the survey and for whom we observe KS1 and BAS outcomes. Those starting school outside standard dates are excluded from the sample.

Figure 8: MCS: Impact of age at test, age start school, length of schooling and relative age (accounting for independent variables) on attainment at age 7 (BAS)



Note: The sample is MCS cohort members that are present at wave 4 of the survey and for whom we observe KS1 and BAS outcomes. Those starting school outside standard dates are excluded from the sample.

Figure 9: MCS: Impact of age at test, age start school, length of schooling and relative age for females and males on attainment at age 7 (BAS)



Note: The sample is MCS cohort members that are present at wave 4 of the survey and for whom we observe KS1 and BAS outcomes. Those starting school outside standard dates are excluded from the sample.

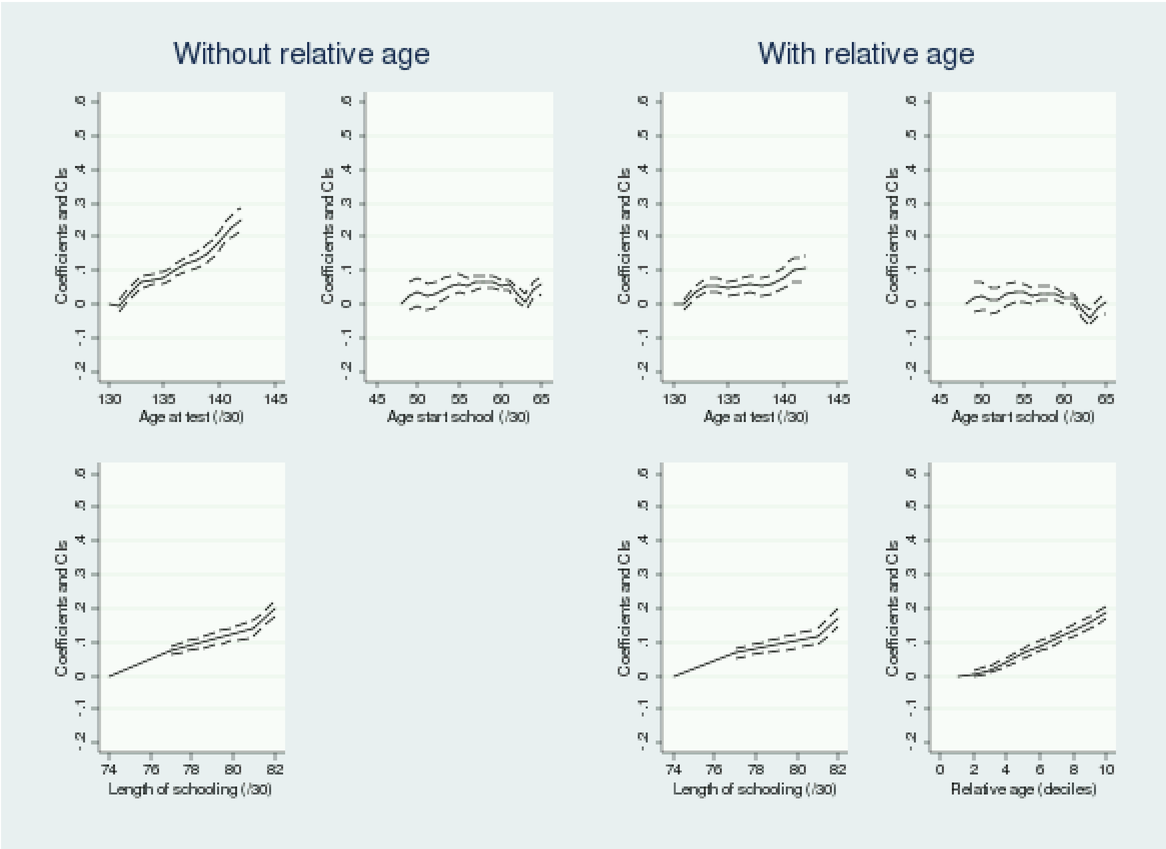
NPD: KS2

The impact of the three linearly dependent variables at KS2 (age 11) are similar to that for KS1 (age 7). Figure 11 presents the explicit comparison between results for KS1 and KS2, while Figure 10 shows that impact of the three linearly dependent factors when excluding and including relative age. The left hand panel of Figure 10 shows the impact of age at test, age start school and length of schooling on the average points score (APS) at KS2, which is similar in pattern and qualitative findings to those for KS1 presented in Figure 3; the impact of age at test is just below 0.3 standard deviations between the oldest and youngest at the test; the impact of length of schooling is around 0.2 standard deviations between the longest and shortest and the impact of age start school is closest to zero. The decline in impact of age start school from KS1 to KS2 is consistent with few cumulative negative effects from starting school at a younger age.

Before discussing further results, we note that age at test and relative age become more slightly more highly correlated as children age (from 0.96 at KS1 to 0.97 at KS2). This is because the pupils have spent relatively more time in cohorts with the full range of ages present, as opposed to some that have a restricted age range in the first year of primary school. The random variation across schools and within schools as pupils move is therefore the only source of variation in later years and this is somewhat limited. Adding this highly co-linear variable to the model is therefore even more problematic than at KS1 (which was also presented with some caution), but we present the results for completeness. As at KS2, relative age has a positive impact on test scores, and reduces the positive impact of age at test by just over one half. This may be because relative age becomes more important as pupils age, or that the two variables become more correlated (and so the model has more difficulty in distinguishing the impact of the two factors). The impact of length of schooling is generally positive, with a similar pattern and magnitude to results for KS1.

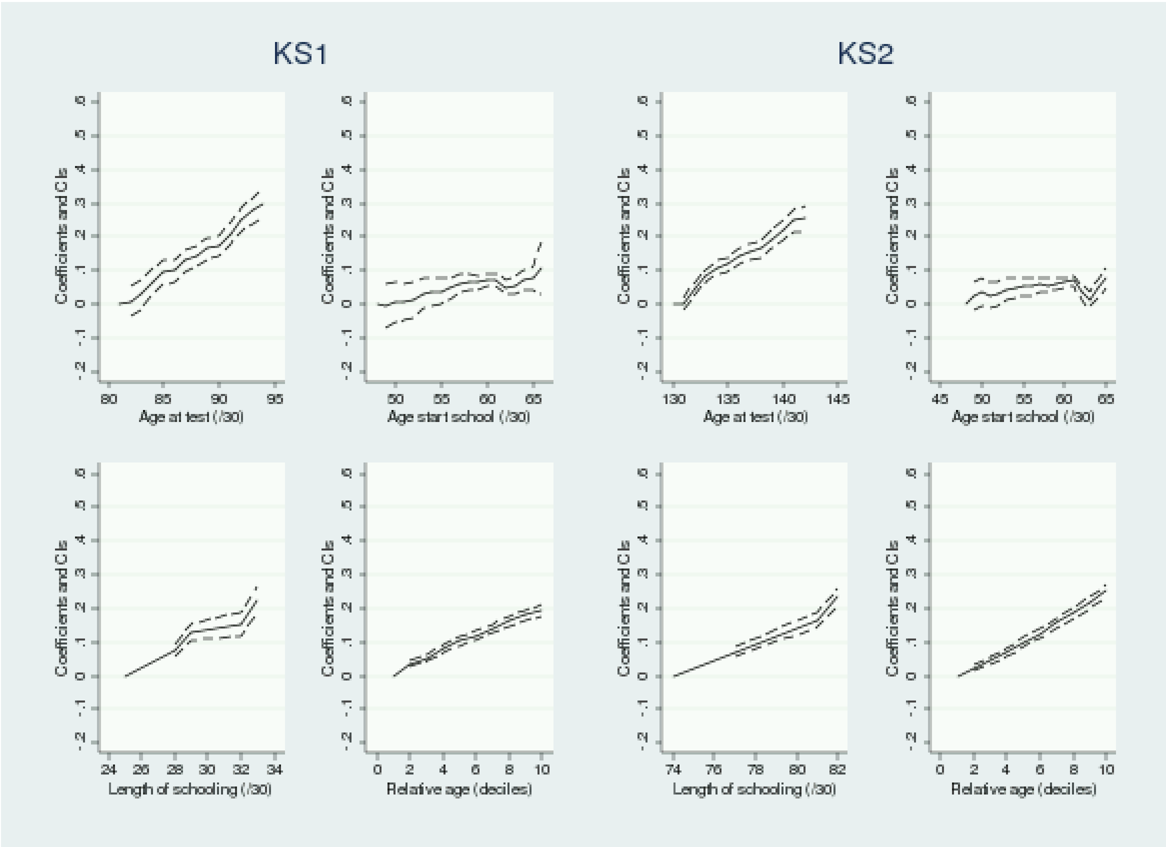
It is striking that the most important factor appears to be relative age, suggesting that the impact of this variable increases as children age (or that the correlation between variables is more problematic). Figure 13 shows that this appears to be driven by males, for who relative age has the largest impact on tests scores (around 0.3 standard deviations between the relatively oldest and youngest) while for females it is less than 0.1 standard deviations. For males, the impact of age at test at KS2 becomes negligible, while for females it remains important, roughly equivalent to the impact of length of schooling.

Figure 10 : NPD: Impact of age at test, age start school, length of schooling and relative age on attainment at KS2



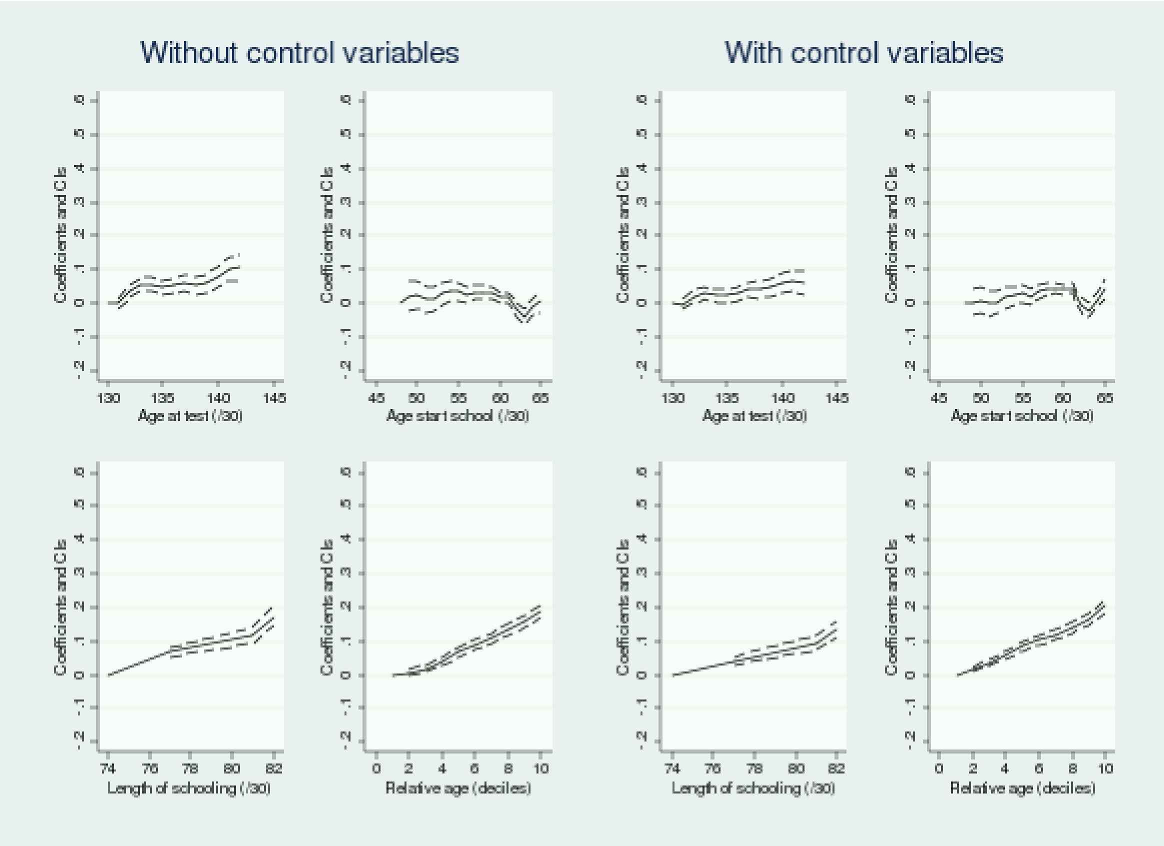
Note: Cohorts of pupils born between September 1996 and August 1999 are included. Cohort dummies are included in the model. Those starting school outside standard dates are excluded from the sample.

Figure 11: NPD: Impact of age at test, age start school, length of schooling, relative age (accounting for independent variables) on attainment at KS1 and KS2



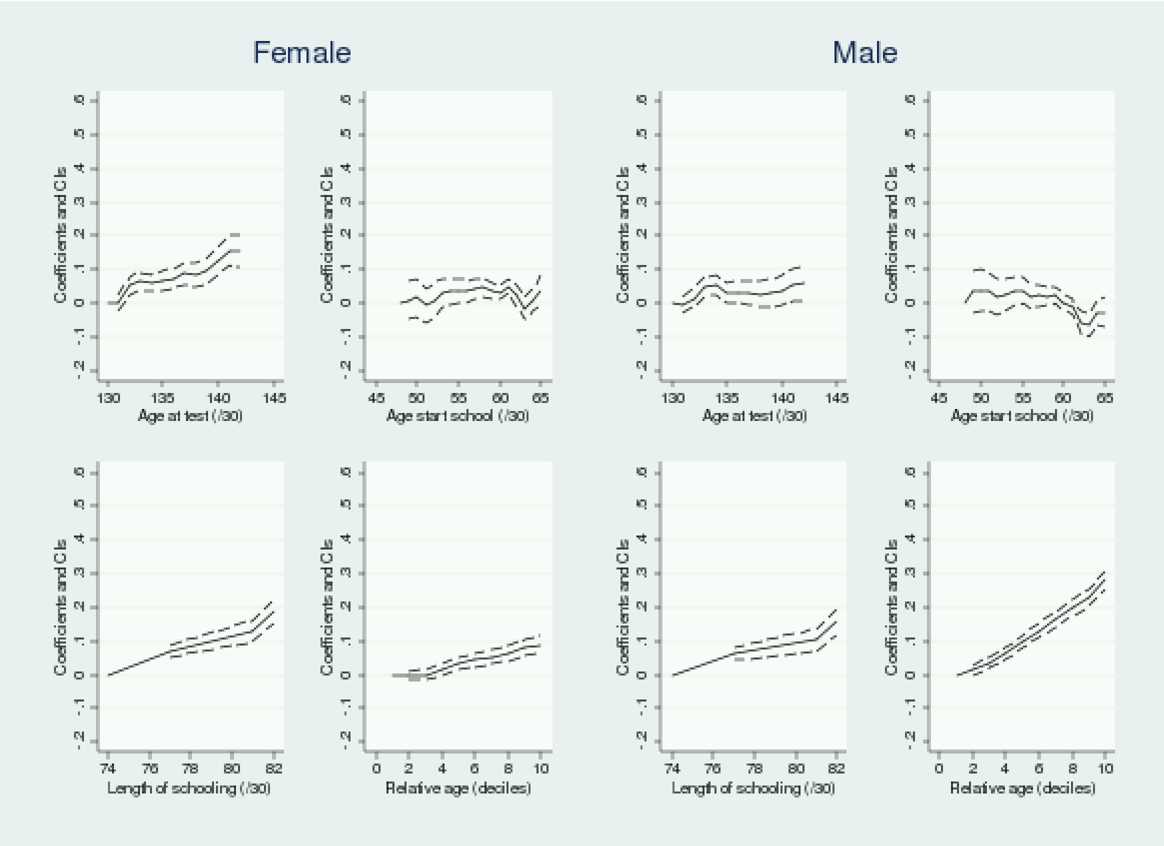
Note: Cohorts of pupils born between September 1996 and August 1999 are included. Cohort dummies are included in the model. Those starting school outside standard dates are excluded from the sample.

Figure 12: NPD: Impact of age at test, age start school, length of schooling, relative age (accounting for independent variables) on attainment at KS2



Note: Cohorts of pupils born between September 1996 and August 1999 are included. Cohort dummies are included in the model. Those starting school outside standard dates are excluded from the sample.

Figure 13: NPD: Impact of age at test, age start school, length of schooling and relative age on for females and males at KS2



Note: Cohorts of pupils born between September 1996 and August 1999 are included. Cohort dummies are included in the model. Those starting school outside standard dates are excluded from the sample.

Summary and policy implications

Results from different samples and different ages are generally consistent. Most definitively, the impact of age of starting school is closest to zero. This suggests that starting school at the age of 5 rather than 4 has a marginal impact on test performance, conditional on all other factors. In contrast, the impact of length of schooling tends to have a positive impact. In combination with results for males that suggest that starting school at an older age may have a negative effect on test performance, this result bolsters support for a single entry point for primary school (and against deferred entry).

Across all samples and ages, the most important of the three linearly dependent factors is age at test. In addition to a policy of single entry rather than staggered entry to primary school (which is now the norm in England), this suggests that an appropriate policy response to the observed month of birth differential is to age-adjust all national test scores. This is a straightforward and cheap policy that is currently used by some selective schools to determine those with highest ability for their age. Considering the three linearly dependent factors only, age-adjustment, which effectively provides each pupil with a score relative to others born in the same month of the year (rather than a score relative to others born in the same academic year), would correct for the difference in test scores, on average, between those born later and earlier in the academic year.

Evidence suggests that the impact of the fourth factor, relative age, also tends to be significant and positive, however, accounting for some of the difference in test scores between children born at the start and end of the academic year. This appears to be especially true as children get older, which may reflect an increasing importance of relative age for test scores, or the increasing correlation between relative age and age at test. In the first case, additional policy responses would be required to address the independent impact of relative age, such as increasing the awareness of relative age effects for teachers and parents. In the second case, the spurious relationship between relative age and test scores, which suggests that no additional policy response would be required. We are cautious about the impact of relative age, as we note that Crawford et al. (2007) found little evidence of a significant role for relative age when adopting a more parametric approach, and it has no significant impact on test scores in a situation when the high level of multi-colinearity is reduced (in the MCS sample), although the context of the test (taken outside of school and away from peers) may also be important.

Is it possible to conclude whether delaying entry to school for summer born children (where these children would join the academic year group below and hence become the relatively oldest in the year) would have a positive impact on test performance? Delaying entry for the youngest pupil in a cohort would be likely to have a positive impact on performance for this pupil: they would be older at each test (which has a positive impact on test scores), their relative age would be higher (which has a positive impact on test scores), their length of schooling prior to the test might be longer (which has a positive impact on test scores), and their age of starting school would be higher (which has a generally positive or at least non-negative impact on test scores).

We are not able to state this definitively, however, as this has not been an option available as part of the school admissions policies commonly used in England, and endogeneity would be a concern if it happened in individual cases (as it is in other countries). It is worth noting that this option could not solve the problem for all pupils; the new youngest pupil in the cohort may suffer more as a result of the policy, as they would now be even younger relative to the oldest in the class and teachers may find it harder to teach classrooms with a larger range of ages. In addition, some parents may not have the option to delay of defer entry (for example for childcare reasons) which may exacerbate inequalities in educational outcomes between the more and less affluent.

5. Conclusions

It is widely observed that a pupil's month of birth relative to the academic year cut-off affects their educational attainment. Although it is clear from most studies that this relates to the structure of the academic year (rather than specific characteristics associated with those born in different months), previous research has been unable to separately identify the precise mechanisms through which a pupil's month of birth affects their educational attainment. Of the four potential mechanisms (a pupil's age at test, length of schooling prior to the test, age of starting school, and relative age within cohort), three are linearly dependent (age at test is equal to age of starting school plus length of schooling), which makes it difficult to observe the independent contribution of each factor. In addition, relative age is highly correlated with these factors, in particular with age at test.

We are the first to apply the principle of maximum entropy to address this problem. As we have seen, this principle chooses the most appropriate probability distribution (and therefore the most appropriate parameter vector) given the information that is observed, where "most appropriate" refers to the one with the highest entropy. This enables us to provide new evidence on the relative contribution of the three linearly dependent variables. The high correlation of the fourth factor (relative age) with the three other factors remains problematic, however, as in cases of perfect correlation the method applies equal weight to each factor. This limits the scope of our conclusions to some extent, although a number of results are clear.

First, the impact of age of starting school is consistently lower than other factors, suggesting that allowing parents greater flexibility over when their children start school is not an appropriate policy response to address the differences in test scores between children born at the start and end of the academic year. In fact, the positive effect of length of schooling consistently outweighs the negative effect of starting school at a younger age, which means that in the current system in England where it is only possible to defer rather than delay entry, it is better for relatively younger pupils to start school earlier in the year that they turn four (in September rather than January or April).

Second, across all samples and ages, the most important of the three linearly dependent factors is age at test. In addition to a policy of single term entry rather than staggered entry to primary school (which is now the norm in England), this suggests that an appropriate policy response to the observed month of birth differential is to age-adjust all national test scores.

Whether additional policy responses are required depends on whether the observed impact of relative age is real or spurious (i.e. driven simply by the high correlation with age at test).

In the latter case, no additional policy response is required. In the former, more comprehensive responses are necessary, such as increasing the awareness of relative age effects for teachers and parents. Given that relative age has almost no impact on test scores where the correlation with age at test is much reduced (observed in the MCS sample), it seems likely that much of the effect of relative age is driven by a spurious rather than a real relationship, although we remain cautious about drawing such a conclusion on the basis of these results alone. We plan to investigate this issue further in future research.

References

Aliprantis, D. (2011), ‘When should children start school?’, Federal Reserve Bank of Cleveland, Working Paper no. 11/26.

Alton, A. and Massey, A. (1998), ‘Date of birth and achievement in GCSE and GCE A-level’, *Educational Research*, vol. 40, pp. 105–9.

Bedard, K. and Dhuey, E. (2006), ‘The persistence of early childhood maturity: international evidence of long-run age effects’, *Quarterly Journal of Economics*, vol. 121, pp. 1437–72.

Bell, J. and Daniels, S. (1990), ‘Are summer-born children disadvantaged? The birthdate effect in education’, *Oxford Review of Education*, vol. 16, pp. 67–80.

Black, S., Devereux, P. and Salvanes, K. (2008), ‘Too young to leave the nest? The effects of school starting age’, National Bureau of Economic Research (NBER), Working Paper no. 13969.

Borg, M. and Falzon, J. (1995), ‘Birth date and sex effects on scholastic attainment of primary school children: a cross-sectional study’, *British Educational Research Journal*, vol. 21, pp. 61–74.

Borghans, L. and Diris, R. (2010), ‘An economic analysis of the optimal school starting age’, University of Maastricht, mimeo.

Browning, M., Crawford, I. and Knoef, M. (2012), ‘The age-period cohort problem: set identification and point identification’, Centre for Microdata Methods and Practice (cemmap), Working Paper no. 02/12.

Buddelmeyer, H. and Le, T. (2011), ‘Effects of age at entry to Year 1 on later schooling outcomes: evidence from Australia’, University of Melbourne, mimeo.

Crawford, C., Dearden, L. and Meghir, C. (2007), *When You Are Born Matters: The Impact of Date of Birth on Child Cognitive Outcomes in England*, Centre for the Economics of Education (CEE) Report to the Department for Children, Schools and Families, http://www.ifs.org.uk/publications/4073.

Crawford, C., Dearden, L. and Meghir, C. (2010), ‘When you are born matters: the impact of date of birth on educational outcomes in England’, Institute for Fiscal Studies (IFS), Working Paper no. 10/06, http://www.ifs.org.uk/publications/4866.

Datar, A. (2006), ‘Does delaying kindergarten entrance age give children a head start?’, *Economics of Education Review*, vol. 25, pp. 43–62.

Elder, T. and Lubotsky, D. (2009), ‘Kindergarten entrance age and children’s achievement: impacts of state policies, family background and peers’, *Journal of Human Resources*, vol. 44, pp. 641–83.

Fogelman, K. and Gorbach, P. (1978), ‘Age of starting school and attainment at 11’, *Educational Research*, vol. 21, pp. 65–7.

Fredriksson, P. and Ockert, B. (2005), ‘Is early learning really more productive? The effect of school starting age on school and labour market performance’, Institute for the Study of Labor (IZA), Discussion Paper no. 1659.

Hamori, S. and Kollo, J. (2011), ‘Whose children gain from starting school later? Evidence from Hungary’, Institute for the Study of Labor (IZA), Discussion Paper no. 5539.

Jaynes, E. T. (1957a), ‘Information theory and statistical mechanics’, *Physical Review*, vol. 106, pp. 620–30.

Jaynes, E. T. (1957b), ‘Information theory and statistical mechanics II’, *Physical Review*, vol. 108, pp. 171–90.

Jaynes, E. T., 1982, `On the Rationale of Maximum-Entropy Methods,' *Proceedings of the IEEE,* 70**,** 939

Jurges, H. and Schneider, K. (2007), ‘What can go wrong will go wrong: birthday effects and early tracking in the German school system’, CESifo, Working Paper no. 2055.

Kawaguchi, D. (2011), ‘Actual age at school entry, educational outcomes, and earnings’, *Journal of the Japanese and International Economies*, vol. 25, pp. 64–80.

McEwan, P. and Shapiro, J. (2008), ‘The benefits of delayed primary school enrollment: discontinuity estimates using exact birth dates’, *Journal of Human Resources*, vol. 43, pp. 1–29.

Mühlenweg, A. and Puhani, P. (2010), ‘The evolution of the school-entry age effect in a school tracking system’, *Journal of Human Resources*, vol. 45, pp. 407–38.

Ponzo, M. and Scoppa, V. (2011), ‘The long-lasting effects of school entry age: evidence from Italian students’, University of Calabria, Working Paper no. 01-2011.

Puhani, P. and Weber, A. (2007), ‘Does the early bird catch the worm? Instrumental variable estimates of early educational effects of age of school entry in Germany’, *Empirical Economics*, vol. 32, pp. 359–86.

Robertson, E. (2011), ‘The effects of quarter of birth on academic outcomes at the elementary school level’, *Economics of Education Review*, vol. 30, pp. 300–11.

Russell, R. and Startup, M. (1986), ‘Month of birth and academic achievement’, *Personality and Individual Differences*, vol. 7, pp. 839–46.

Sampaio, B., da Matta, R., Ribas, R. and Sampaio, G. (2011), ‘The effect of age on college entrance test score and enrollment: a regression discontinuity approach’, University of Illinois at Urbana-Champaign, mimeo.

Sharp, C., Hutchison, D. and Whetton, C. (1994), ‘How do season of birth and length of schooling affect children’s attainment at KS1?’, *Educational Research*, vol. 36, pp. 107–21.

Sinn, H.-W. (1980), ‘A rehabilitation of the Principle of Insufficient Reason’, *Quarterly Journal of Economics*, vol. 94, pp. 493–506.

Smith, J. (2009), ‘Can regression discontinuity help answer an age-old question in education? The effect of age on elementary and secondary school achievement’, *B.E. Journal of Economic Analysis and Policy*, vol. 9, issue 1 (Topics), article 48.

Smith, J. (2010), ‘How valuable is the gift of time? The factors that drive the birth date effect in education’, *Education Finance and Policy*, vol. 5, pp. 247–77.

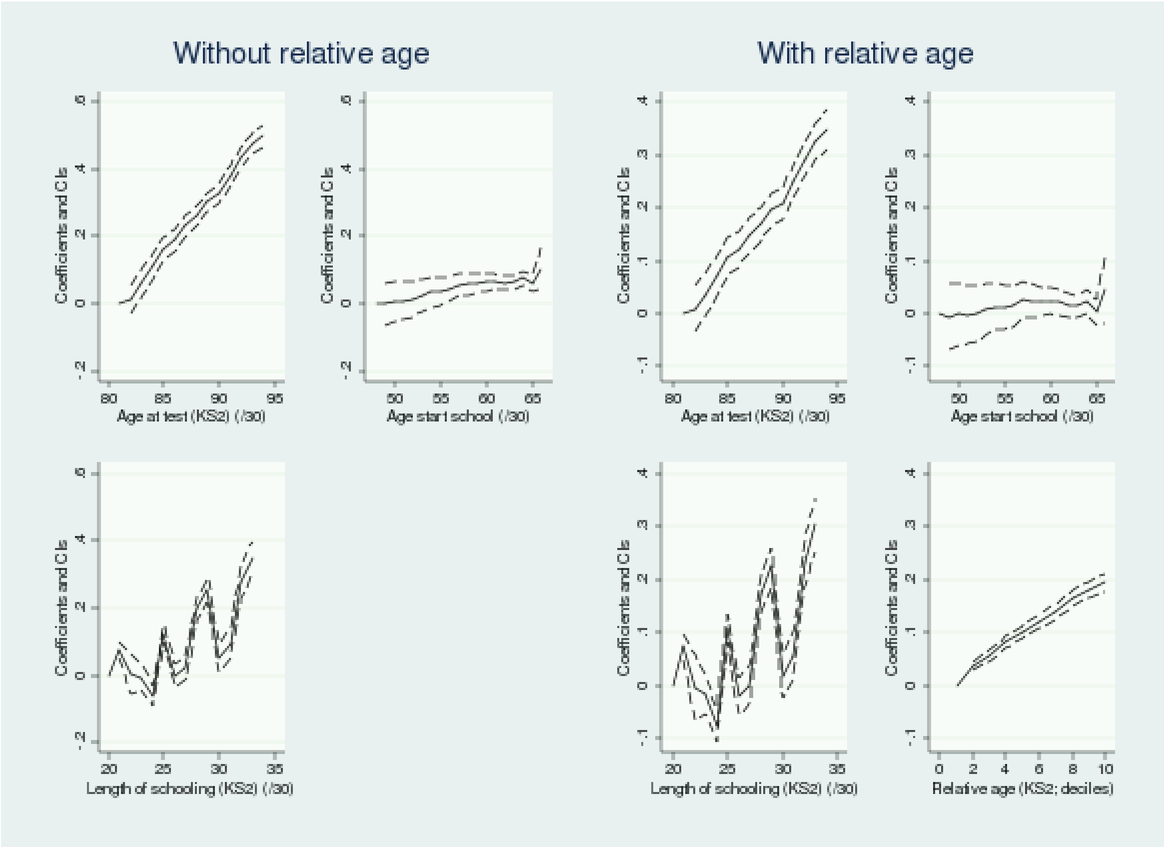
Sprietsma, M. (2010), ‘Effect of relative age in the first grade of primary school on long-term scholastic results: international comparative evidence using PISA 2003’, *Education Economics*, vol. 18, pp. 1–32.

Strom, B. (2004), ‘Student achievement and birthday effects’, Norwegian University of Science and Technology, mimeo.

Thomas, S. (1995), ‘Considering primary school effectiveness: an analysis of 1992 Key Stage 1 results’, *Curriculum Journal*, vol. 6, pp. 279–95.

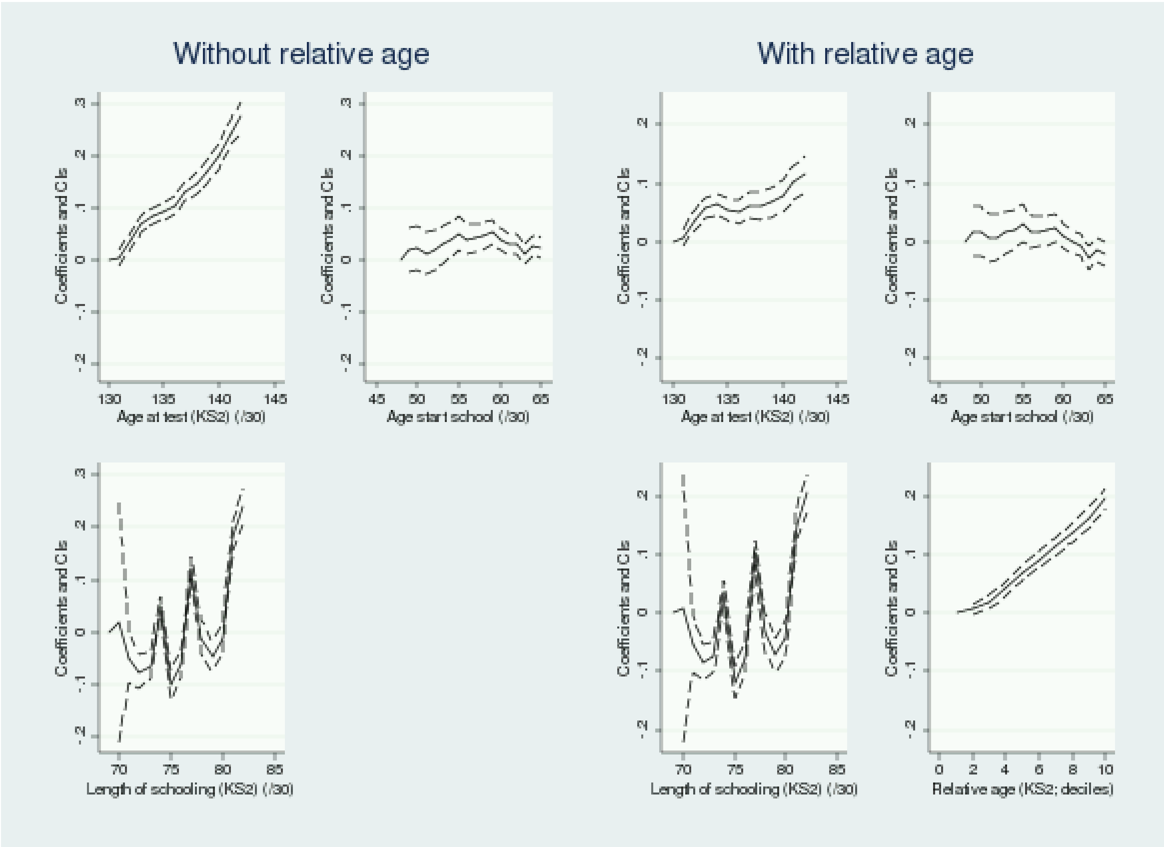
Appendix 1: including those with non-standard entry dates

Figure 1: NPD: Impact of age at test, age start school, length of schooling and relative age on attainment at KS1



Note: Cohorts of pupils born between September 1996 and August 1999 are included. Cohort dummies are included in the model. Those starting school outside standard dates are excluded from the sample.

Figure 2: NPD: Impact of age at test, age start school, length of schooling and relative age on attainment at KS2



Note: Cohorts of pupils born between September 1996 and August 1999 are included. Cohort dummies are included in the model. Those starting school outside standard dates are excluded from the sample.

Appendix 2: Correlation of four potential driving factors in each source of data

Appendix Table 1: Correlation of four potential driving factors in each source of data

	NPD: KS1			
	Age at test	Length of schooling	Age start school	Relative age
Age at test	1			
Length of schooling	0.33	1		
Age start school	0.75	-0.37	1	
Relative age	0.96	0.32	0.72	1
	NPD: KS2			
	Age at test	Length of schooling	Age start school	Relative age
Age at test	1			
Length of schooling	0.33	1		
Age start school	0.75	-0.37	1	
Relative age	0.97	0.33	0.73	1
	MCS: KS1 and BAS tests			
	Age at test	Length of schooling	Age start school	Relative age
Age at test	1			
Length of schooling	0.20	1		
Age start school	0.79	-0.37	1	
Relative age	0.89	0.03	0.83	1
	MCS: KS1 test only			
	Age at test	Length of schooling	Age start school	Relative age
Age at test	1			
Length of schooling	0.25	1		
Age start school	0.85	-0.23	1	
Relative age	0.97	0.24	0.83	1
	MCS: BAS test only			
	Age at test	Length of schooling	Age start school	Relative age
Age at test	1			
Length of schooling	0.16	1		
Age start school	0.72	-0.50	1	
Relative age	0.81	-0.14	0.83	1

Note: all correlations are significant at the 1% level.

Appendix 3: Method for imputing date of entry for those with implausible start dates

Appendix Table 2: Impute date of entry for those with implausible start dates

Step	NPD	MCS
1	Replace all dates of entry to missing for those whose date of entry is unbelievable (for example if start in nursery or if date is outside statutory dates)	
		This is around 37% of the MCS sample. Of the 63% for whom month of starting school is observed, 91% of parent reported start dates are consistent. This is encouraging for the reliability of the data. We do not exclusively use parent reports as only the month and year (rather than date) was reported.
2	Calculate the modal start date for children born in each month for each school without an attached nursery.	
3	Update date of entry to equal the modal start date calculated in step 2 for all pupils with missing dates of entry. All pupils in schools without an attached nursery and with a relevant modal start date now have a date of entry (actual or imputed).	
		All pupils in schools without an attached nursery now have a date of entry (actual or imputed). Only 5% of cohort members are left with missing date of entry after this step.
4	Calculate the modal start date for children born in each month in community schools in each local authority.	
5	Update date of entry to equal the modal start date calculated in step 4 for all pupils with missing dates of entry. All pupils in local authorities with a relevant modal start date now have a date of entry (actual or imputed).	
		All pupils in local authorities with modal start dates now have a date of entry (actual or imputed), except for 29 cohort members without information on school ID or LA.
6	Calculate the modal start date for children born in each month for each area with the same admissions policy	
7	Update date of entry to equal the modal start date calculated in step 6 for all pupils with missing dates of entry. All pupils with non-missing information on admissions policy now have a date of entry (actual or imputed).	

Appendix 4: Independent variables used in the NPD and MCS analysis

Appendix Table 3: Independent variables used in NPD and MCS analysis

Variable	NPD	MCS
Eligibility for free school meals	Binary variable equal to one if the pupil is eligible and zero otherwise.	Binary variable equal to one if the pupil is eligible and zero otherwise.
English is an additional language	Binary variable equal to one if the pupil has English is an additional language and zero otherwise.	Binary variable equal to one if the pupil has English is an additional language and zero otherwise.
Sex	Binary variable equal to one if the pupil is male and zero otherwise.	Binary variable equal to one if the pupil is male and zero otherwise.
Ethnicity	Discrete variable with categories: white British, white other, black African, black Caribbean, black other, Indian, Pakistani, Bangladeshi, Chinese, Asian other, mixed (any), other. Entered as a set of binary variables with White British as the reference category.	Discrete variable with categories: white, black African, black Caribbean, black other, Indian, Pakistani, Bangladeshi, Chinese, Asian other, mixed (any), other Entered as a set of binary variables with white as the reference category
IDACI	Discrete variable with values from one to five representing the quintile of deprivation in the local area (around 750 households). Entered as a set of binary variables with the most deprived (first) quintile as the reference category.	Discrete variable with values from one to five representing the quintile of deprivation in the local area (around 750 households). Entered as a set of binary variables with the most deprived (first) quintile as the reference category.
Proportion of adults with high level of education	Defined as a continuous variable ranging from zero to one, where zero represents no adults in the local area (around 150 households) with a high level of education, and one represents all adults.	Defined as a continuous variable ranging from zero to one, where zero represents no adults in the local area (around 150 households) with a high level of education, and one represents all adults.
Proportion of adults with each level of socio-economic classification	Defined as continuous variables ranging from zero to one, where zero represents no adults in the local area (census area) with each level of socio-economic status, and one represents all adults.	Defined as continuous variables ranging from zero to one, where zero represents no adults in the local area (census area) with each level of socio-economic status, and one represents all adults.
Household income quintile		Discrete variable with values from one to five representing the quintile of household income (equivalised). Entered as a set of binary variables with the most deprived (first) quintile as the reference category.
Household work status		Discrete variable with categories: both in work/on leave, main respondent in work/on leave & partner not, partner respondent in work/on leave & main respondent not, both not in work/on leave. Entered as a set of binary variables with both in work/on leave as the reference category.
Mother's age at		Discrete variable with categories: less

first birth		than 20, between 20 and 25, between 25 and 30, between 30 and 35 and over 35. Entered as a set of binary variables with those over 35 as the reference category.
Household marital status		Discrete variable with categories: lone parent, cohabiting couple, married couple. Entered as a set of binary variables with married as the reference category.
Mother and father's level of education		Discrete variable with categories: none, lower level vocational qualifications, GCSE A*-C (benchmark academic qualification at the end of compulsory schooling), AS/A level (academic qualification taken after compulsory schooling), foundation degree (a vocational qualification in higher education), and degree & higher degree. Entered as a set of binary variables with the most educated (those with at least a degree) as the reference category. Entered as a set of binary variables for mothers and fathers with the most educated (those with at least a degree) as the reference category in each case.
Mother and father's occupational status		Discrete variable with categories: routine, semi-routine, lower supervisory and technical, small employer and self-employed, intermediate, low managerial and professional, high managerial and professional. Entered as a set of binary variables for mothers and fathers with the highest status (high managerial and professional) as the reference category in each case.
Housing tenure		Discrete variable with categories: mortgage/own home, rent privately, rent from local authority, with parents, other. Entered as a set of binary variables with mortgage/own home as the reference category.
Whether child was breastfed		Binary variable equal to one if the pupil was breastfed and zero otherwise.
Child's birthweight		Discrete variable with categories: low, normal, and high. Entered as a set of binary variables with normal as the reference category.
Multiple birth		Binary variable equal to one if the pupil is a twin/triplet and zero otherwise.
Birth order		Discrete variable with categories: first, second and third or more. Entered as a set of binary variables with first as the reference category.