Productivity of school expenditure: Differences across pupils from diverse backgrounds

Cheti Nicoletti

DERS, University of York and ISER, University of Essex

Birgitta Rabe

ISER, University of Essex

This version: February 20, 2013

Abstract

This paper assesses the productivity of school expenditure for pupils from diverse backgrounds. We focus on pupils who are targeted by school policies to receive extra help as well as ethnic minorities and let the productivity of inputs vary by level of lagged cognitive skills. We find that investments in children with high previous attainment are most productive (dynamic complementarity). Pupils with special educational needs benefit immensely from extra help whereas high ability black or disadvantaged pupils tend to get left behind. Spending on education support staff as well as learning resources helps close attainment gaps between pupils.

Keywords: Education production function, test scores, school quality

JEL codes: I22, I24

The corresponding authors are Cheti Nicoletti (cheti.nicoletti@york.ac.uk) and Birgitta Rabe (brabe@essex.ac.uk).

We thank the Department for Education for making available, under Terms and Conditions, data from the National Pupil Database. Financial support from the Nuffield Foundation is 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. More information is available at www.nuffieldfoundation.org.

1 Introduction

Cunha and Heckman (2007 and 2008) and Cunha et al. (2005 and 2010) have emphasized two important features of a child's skill production function. The first is that a child's skills at a specific stage of life depend on skills produced at the previous stage (self-productivity). The second is that higher skills in a stage increase the productivity of inputs in the following stages (dynamic complementarity). In this paper we use rich administrative data on English state schools to evaluate the effect of expenditure per pupil on cognitive skills of children at the end of compulsory schooling when they are about 16 years old. We consider a value added model to allow pupil's cognitive skills at the end of compulsory schooling to depend on his or her lagged cognitive skills, the level of skills at the end of primary school (see for example Hanushek 1986; Hanushek et al. 1996; Todd and Wolpin 2003). We take account of dynamic complementarity by letting the productivity of expenditure per pupil vary across children with different levels of lagged cognitive skills.

Beside the differentials in lagged cognitive skills, another potential cause for a heterogeneous effect of the expenditure per pupil on children's cognitive abilities is the presence of school policies targeted at improving the attainments of specific groups of children. We take account of this by letting the productivity of expenditure per pupil differ for children with special educational needs, belonging to ethnic minorities, with English as additional language and that are identified as being gifted or talented. These are groups of children that are targeted to receive extra support in English state schools and - except for the gifted and talented children - tend to have lower lagged cognitive abilities. We also look at the productivity of school expenditure for children from disadvantaged socio-economic backgrounds although this group is not explicitly targeted by school policies. To net out the part of the productivity gap explained by a different level of lagged cognitive skills, we compare children from these diverse groups but with the same level of lagged cognitive skills.

To control for unobserved family-specific endowment and inputs which do not vary between siblings we estimate a value-added education production model using sibling fixed effect estimation (see Rosenzweig and Wolpin 1994; Altonji and Dunn 1996; Behrmann et al. 1996; Todd and Wolpin 2007). Furthermore, to control for the dependence between the lagged test and the unobserved child-specific endowment, we use a two-step procedure (see

Nicoletti and Rabe 2012; Del Boca et al. 2012). We also control for pupil characteristics and a rich set of variables describing school characteristics and composition, including those school-level variables that are used to determine the allocation of funds to schools. Identification of the expenditure per pupil is possible because during the time-period considered in our sample there is exogenous variation in funding levels received by schools (see section 2 for more details). To allow the expenditure per pupil to differ across different levels of lagged ability and across different groups of pupils, we extend the value added model by considering interactions of the expenditure per pupil with dummy variables indicating different levels of lagged cognitive abilities and different groups of pupils.

To understand better the mechanisms through which school policies work, we also consider different ways in which school funds are used. By assessing the effect of the expenditure per pupil for teaching staff while controlling for the pupil-teacher ratio, we disentangle the effect of quality and quantity of teachers on pupils' test scores (see Rivkin et al. 2005; Gilpin and Kavanagh 2012). Unlike the US, where teacher quality does not appear to be closely related to salaries (Hanushek et al. 2005), in England teachers are paid according to a national pay scale, so that higher expenditure per pupil after controlling for teacher numbers indicates that a school uses more senior and therefore more experienced teachers. We also further disaggregate school expenditure to assess which type of expenditure is more effective for children from disadvantaged backgrounds, dividing the expenditure per pupil in teaching staff, education support staff, supply teachers, learning resources and other expenditure.

This paper provides - to our knowledge - the first comprehensive empirical assessment of (i) dynamic complementarity in school inputs; (ii) heterogeneity of input productivity across pupil types and (iii) the effect of different types of school expenditure on different groups of pupils. We contribute to the literature of skill formation as well as to the literature on the effectiveness of spending in schools. We find that the productivity of the expenditure per pupil is increasing in lagged cognitive skills for all groups of children, indicating strong dynamic complementarity. School policies seem to have a beneficial effect particularly on children with special educational needs, who have a higher productivity of the expenditure per pupil than children without such needs. Pupils at the lower end of the lagged attainment distribution benefit more from putting extra teachers in classes than those at the upper end. We also find that expenditure on education support staff has a higher productivity

for children eligible for free school meals, with English as additional language, gifted and talented and belonging to ethnic minorities, whereas expenditures in teaching staff and learning resources seem to have higher productivity for children with special educational needs.

The paper proceeds as follows. Section 2 gives institutional background on the education and school funding system in England and provides justification of our identification strategy. Section 3 describes the estimation methods. In Section 4, we describe our data sources and variables used, while in Section 5 we present the estimation results for the productivity of per pupil expenditure as well as of different types of expenditure on different groups of pupils. Finally, Section 6 concludes.

2 Institutional background

2.1 Education system in England

School education in England is provided by the state. Approximately 93% of school children in England attend state schools, and about 7% are educated in fee-paying private schools. Full-time education is compulsory for all children aged between 5 and 16. The education during these years is divided into four Key Stages and pupils undergo externally marked National Curriculum Tests at the end of Key Stages 2 and 4. Until recently such tests were also carried out at Key Stages 1 and 3 but today progress at these stages is examined via individual teacher assessment. Key Stage 2 National Curriculum Tests are taken at the end of primary school, usually at age 11, in the core subjects of English, Mathematics and Science. Key Stage 4 tests are taken at age 16 at the end of compulsory schooling, and are either the General Certificate of Secondary Education (GCSE) exams or equivalent vocational or occupational exams. Pupils decide which GCSE courses to take, and because English, Mathematics and Science are compulsory study subjects, virtually all students take GCSE examinations in these topics, plus others of their choice, with a total of ten different subjects normally taken. In addition to GCSE examinations, a pupil's final grade may also incorporate coursework elements. Key Stage 2 and 4 test results receive a lot of attention nationally as they play a prominent role in the computation of so-called school league tables, which are used by policy makers to assess schools and by parents to inform school choice.

There are government policies to provide some groups of pupils with extra help in schools. These include gifted and talented pupils (G&T) that are identified by teachers as having an advanced ability in an academic subject or a talent in areas such as sport and music. Schools are responsible to provide greater challenges in lessons and opportunities to demonstrate and develop abilities for gifted and talented pupils. The other groups routinely identified by schools for extra input are (deemed to be) on the lower end of the ability distribution. These include ethnic minority pupils (for example black Caribbean and African pupils and pupils of Pakistani or Bangladeshi heritage) that have an attainment gap compared to others and pupils with English as an additional language (EAL). Finally, there are particular programs for pupils with special educational needs (SEN), i.e. pupils with learning difficulties. Pupils that have been assessed by local education authorities receive a statement which is usually associated with additional funding received by the school. There are also pupils identified by the schools as having special needs, but without statement. Although pupils from disadvantaged socio-economic backgrounds are not explicitly targeted for extra support, free school meal (FSM) eligibility is a criterion used widely in school reporting, so that this group is an implicit target group for schools.

2.2 School funding in England

In the time-period considered in our empirical analysis (2005-2010) there was considerable exogenous variation in the funding levels received by schools which we exploit to identify the effect of school inputs.[1] First, the substantial real increase in funding per pupil from an average of 4,690 pounds in 2005 to 5,750 pounds in 2010 (23% increase in 2010 prices) means that the same schools will have had differing funding levels over time. Second, the rules used to allocate funding across schools have also changed over time, they vary regionally and they are systematically slow to adapt to varying circumstances.

Most funding for state schools in England comes from central government which hands money to local education authorities, of which there are 154. The central government grant is calculated using a funding formula mostly based on pupil numbers, educational disadvantage and area costs. However, a so-called spend-plus methodology is applied whereby local

[1] In our empirical analysis we consider test scores of four cohorts of pupils, taking exams in 2007, 2008, 2009 and 2010. School inputs are three-year averages of expenditure per pupil, so that for a student taking exams in 2007, inputs will be from the period 2005-2007.

authority grants are determined as flat-rate increases on the grant received the previous year - with a historical starting point in 2005-06 - plus an extra increase based on the formula. "So, current levels of school funding are based on an assessment of needs which is out of date, and on historic decisions about levels of funding which may or may not reflect precisely what schools needed then" (Department for Education 2011, p. 3).

Local authorities then use their own funding formulas to hand out the money received from central government to schools. Apart from pupil numbers, many local authorities historically allow more funding for pupils from deprived backgrounds (eligible for free school meals), with special educational needs and with English as an additional language (Chowdry and Sibieta 2011). There is considerable variation between local authorities in the formula used (West 2009). However, a major constraint that local authorities face when setting their formulas is the Minimum Funding Guarantee, introduced in 2004-05, which guarantees each school a minimum increase per pupil per year. Effectively this largely limits the freedom with which local authorities can choose their funding rules (Levačič 2008). In 2010-11 the Minimum Funding Guarantee accounted for half the increase in the central school grant (Chowdry and Sibieta 2011).

The combination of spend-plus methodology and Minimum Funding Guarantee has weakened the relationship between school funding levels and educational need. The historical anchor of the funding formula leads to a low reactivity to changes in schools such as rising or falling numbers of deprived pupils. "Some areas are now woefully underfunded compared with how they would be if the system reflected need properly, whereas some areas continue to receive funding to which they should no longer be entitled" (Department for Education 2011, p. 4). In 2010-11 7% of secondary schools had a level of funding at least 10% lower than predicted using observable characteristics, and 6% had funding at least 10% higher (Chowdry and Sibieta 2011, p. 12). These non-linearities are perceived as being too complex and essentially unfair by the current UK government, and reforms to the funding system are being introduced. For the purposes of this paper we can conclude that there is exogenous variation in school expenditure after controlling for pupil and school characteristics.

3 Methods

We focus on cognitive development during the stage that goes from the end of primary school to the end of compulsory schooling, i.e. from about 11 to 16 years of age, and adopt the following education production model:

$$Y^{*}_{ih,16} = f(I^{F}_{ih}, I^{S}_{ih}, X_{ih}, Y^{*}_{ih,11}, \mu_{ih}), \tag{1}$$

where $Y^{*}_{ih,16}$ and $Y^{*}_{ih,11}$ are unobserved latent cognitive abilities of child i in family h at ages 16 and 11, I^{F}_{ih} is the family investment in the child cognitive development between ages 11 and 16, I^{S}_{ih} is the corresponding school investment, X_{ih} is a vector of other child, household and school characteristics, which are not direct investments in children's cognitive skill but may affect it (e.g. gender, ethnicity, language spoken at home, free school meal eligibility, number of siblings, school characteristics, and pupil composition), and μ_{ih} is the child time-invariant cognitive endowment.

Our estimation sample consists of all pupils enrolled in state schools in England who took their Key Stage 4 tests in the period 2007-2010. For this sample we are unable to observe family and school investments; but we can observe the school expenditure per pupil, which we use as a measure of school investment, and three measures of cognitive abilities each at ages 11 and 16, which are test scores in Mathematics, English and Science obtained in Key Stage 2 and 4 exams. We assume that the relationship between each of these three test scores observed at age 11 and 16 and the unobserved latent cognitive skill at the corresponding age follows a classical measurement error model[2]

$$Y_{ihs,11} = Y^{*}_{ih,11} + e_{ihs,11} \text{ and } Y_{ihs,16} = Y^{*}_{ih,16} + e_{ihs,16}, \tag{2}$$

where the subscript s indicates the test subject and takes value 1 for Mathematics, 2 for English and 3 for Science, $e_{ihs,16}$ and $e_{ihs,11}$ are subject-specific random components identically and independently distributed across children, households and test subjects with mean zero and variance σ^{2}_{e}, and are independent of the inputs in the production model and of the true

[2]Imposing a classical measurement error model is equivalent to imposing a factor model with a single factor and equal factor loadings. The psychologist Spearman (1904) is the pioneer of the factor analysis and he has been the first to apply it to capture a latent measure of skill which he called general intelligence or g-factor. But single factor models, to take account of measurement errors in observed cognitive skill tests, have also been used more recently by economists (e.g. Cunha and Heckman 2008).

latent skill at age 11 and 16, $Y^*_{ih,11}$ and $Y^*_{ih,16}$. The random components $e_{ihs,16}$ and $e_{ihs,11}$ measures the deviation of the subject specific skill from the general latent skill. We assume that there is no correlation between $e_{ihs,16}$ and $e_{ihs',11}$ if $s \neq s'$, but we allow for persistence in the subject-specific ability across age, i.e. $Cov(e_{ihs,16}, e_{ihs,11}) \neq 0$. Furthermore, we assume that the persistence in $e_{ihs,t}$ be identical to the persistence in $Y^*_{ih,t}$, more precisely we assume that the correlation between $Y^*_{ih,16}$ and $Y^*_{ih,11}$ net of the explanatory variables in the education production model is identical to the correlation in $e_{ihs,16}$ and $e_{ihs,11}$. This implies that the persistence in each of the three test scores is identical to the persistence in the latent cognitive ability.

Specification with homogenous effect of school investment

Under the assumptions defined above and imposing that the production function (1) be additive, separable and linear in its arguments, we can rewrite it as

$$Y_{ihs,16} = \alpha + I^F_{ih}\beta_F + I^S_{ih}\beta_S + X_{ih}\gamma + Y^*_{ih,11}\rho + \mu_{ih} + e_{ihs,16}, \tag{3}$$

where we replaced the unobserved latent cognitive skill at age 16 with the observed test score in subject s and $s = 1, 2, 3$. Model (3) is usually known as the valued added model (see Todd and Wolpin 2003) and it has been extensively used in previous empirical papers to evaluate the contributions of school inputs in a specific stage of the child's school life by controlling for the child's cognitive skill at the beginning of the stage (see Hanushek 1997; Meghir and Rivkin 2011).

Since the lagged cognitive skill $Y^*_{ih,11}$ is unobserved, we replace it with the lagged test in subject s and rewrite equation (3) as

$$Y_{ihs,16} = \alpha + I^F_{ih}\beta_F + I^S_{ih}\beta_S + X_{ih}\gamma + Y_{ihs,11}\rho + \mu_{ih} + u_{ihs,16}, \tag{4}$$

where $u_{ihs,16} = e_{ihs,16} - e_{ihs,11}\rho$. Because $Y_{ihs,11}$ and $u_{ihs,16}$ are correlated we would generally expect the ordinary least square estimation to be biased and inconsistent; but, under the assumption that $Y^*_{ih,t}$ and $e_{ihs,t}$ have equal persistence, the asymptotic bias caused by this correlation cancels out. Indeed we can prove that the asymptotic bias of the ordinary least square estimation of ρ is equal to:

$$plim\ \hat{\rho}_{FFE} = \rho + \left[\frac{Cov(\mu_{ih}, M_W Y_{ihs,11})}{Var(M_W Y_{ihs,11})}\right] + \left[\frac{Cov(e_{ihs,16}, e_{ihs,11})}{Var(M_W Y_{ihs,11})} - \rho\frac{Var(e_{ihs,11})}{Var(M_W Y_{ihs,11})}\right], \tag{5}$$

where W is the vector of explanatory variables in our value added model (4), which excludes the lagged test and the unobserved child specific endowment, and M_W is the projection matrix on the space orthogonal to the one generated by the variables W. The first term between brackets is the asymptotic bias caused by omission of the child endowment μ_{ih}; while the second term between square brackets is the asymptotic bias caused by the correlation between $Y_{ihs,11}$ and $u_{ihs,16}$, which cancels out because the assumption of identical persistence in $Y^*_{ih,t}$ and in $e_{ihs,t}$ implies that

$$Cov(e_{ihs,16}, e_{ihs,11}) = \rho Var(e_{ihs,11}).$$

To take account of the endogeneity of the lagged test caused by the unobserved child-specific endowment, μ_{ih}, we adopt a *two-step estimation*. In the first step we use the three contemporaneous tests and the three corresponding lagged tests for each child to estimate a *child fixed effect model*. This allows us to control for the unobserved child specific endowment that is invariant across subjects and to consistently estimate ρ in the value added model (4). Nevertheless, this estimation is unable to identify the remaining slope coefficients because the corresponding variables do not vary across the three tests.

In the second step we use the estimated coefficient ρ to compute a new dependent variable $(Y_{ihs,16} - Y_{ihs,11}\rho)$ which we regress on the remaining variables,

$$Y_{ihs,16} - Y_{ihs,11}\rho = \alpha + I^F_{ih}\beta_F + I^S_{ih}\beta_S + X_{ih}\gamma + \mu_{ih} + u_{ihs,16}. \quad (6)$$

For this regression we adopt sibling fixed effect estimation to control for potential unobserved variables that do not vary between siblings and in particular to control for the family investment I^F_{ih} which we are unable to observe in our sample. Under the assumption that the inputs in the educational production function may depend on unobserved children's endowment only through past test scores and that the family unobserved endowment does not vary across siblings, the sibling fixed effect estimation is consistent.[3]

Specification with heterogenous effect of school investment

[3]See for two other application examples of the two-step estimation Nicoletti and Rabe (2012) and Del Boca et al. (2012).

We also consider a specification of the education production model which allows the productivity of the school investment to vary by levels of the lagged cognitive ability by letting β_S change across the deciles of the lagged test, i.e.

$$Y_{ihs,16} = \alpha + I_{ih}^F \beta_F + I_{ih}^S \beta_{S,10} + \sum_{j=1}^{9} I_{ih}^S \mathbf{I}(d_{j-1} < Y_{ihs,11} \leq d_j)\beta_{S,j} + X_{ih}\gamma + Y_{ihs,11}\rho + \mu_{ih} + u_{ihs,16}, \quad (7)$$

where $\mathbf{I}(d_{j-1} < Y_{ihs,11} \leq d_j)$ is an indicator function, d_j is the j-th decile of the lagged test (j=1,...,9), d_0 is the minimum value taken by the lagged test, $\beta_{S,10}$ is the effect of the school investment for children whose lagged test is in the top decile, and $\beta_{S,j}$ is the differential effect of the school investment for children whose lagged test is in between the $(j-1)$-th and the j-th decile.

The two-step estimation is performed as before but since the new variables

$$\left[I_{ih}^S \mathbf{I}(d_{j-1} < Y_{ihs,11} \leq d_j)\right] \quad j = 1, ..., 9 \quad (8)$$

can change across the three lagged tests, we estimate both ρ and $\beta_{S,j}$ for $j = 1, ..., 9$ in the first step, i.e. by individual fixed effect estimation of equation (7) using the repeated observations on the three contemporaneous and lagged test scores available for each pupil.

In the second step we use the estimated coefficients ρ and $\beta_{S,j}$ to compute a new dependent variable

$$(Y_{ihs,16} - Y_{ihs,11}\rho) - \sum_{j=1}^{9} I_{ih}^S \mathbf{I}(d_{j-1} < Y_{ihs,11} \leq d_j)\beta_{S,j}$$

which we regress on the remaining explanatory variables,

$$Y_{ihs,16} - Y_{ihs,11}\rho - \sum_{j=1}^{9} I_{ih}^S \mathbf{I}(d_{j-1} < Y_{ihs,11} \leq d_j)\beta_{S,j} = \alpha + I_{ih}^F \beta_F + I_{ih}^S \beta_{S,10} + X_{ih}\gamma + \mu_{ih} + u_{ihs,16}. \quad (9)$$

As before we estimate this model by using sibling fixed effects.

In our empirical application we also consider another set of model specifications which allows the productivity of the school investment to vary by groups of children which have been targeted by policies to improve their educational attainments, and a final set of models in which the productivity is left to change across pupil groups as well as across deciles of their lagged test.

4 Data

We use the National Pupil Database (NPD), which is available from the English Department for Education and has been widely used for education research. The NPD is a longitudinal register dataset for all children in state schools in England, covering roughly 93% of pupils in England. It combines pupil level attainment data - National Curriculum assessments typically taken at ages 7, 11, 14 and 16 - with pupil characteristics as they progress through primary and secondary school. Pupil characteristics are collected in annual school censuses and include, for example, age, gender, ethnicity, the pupil's language group and free school meal eligibility.

Pupil-level variables

Our outcomes of interest are General Certificate of Secondary Education (GCSE) or equivalent vocational test results at the end of compulsory schooling, at age 16 (Key Stage 4). We consider Key Stage 4 results in the core subjects English, Mathematics and Science which are directly comparable to test results at the end of primary school. In Key Stage 4 pupils receive a grade for each GCSE course, where pass grades include A*, A, B, C, D, E, F, G. We transform these grades into a continuous point score which we refer to as the Key Stage 4 score.[4]

We control for lagged cognitive skills using Key Stage 2 National Curriculum tests taken at the end of primary school, usually at age 11, in the three core subjects of English, Mathematics and Science. In the Key Stage 2 exams, pupils can usually attain a maximum of 36 points in each subject, but teachers will provide opportunities for very bright pupils to test to higher levels. All test scores are standardised to have a mean of zero and a standard deviation of one.

Individual and family background variables available in the NPD include the gender of the pupil, ethnicity (white British, black, mixed, Indian, Pakistani and Bangladeshi, Chinese), whether or not the first language spoken at home is English, whether special educational needs have been identified for the child and whether the pupil has been classified as gifted and/or talented. Moreover, we can identify whether or not a pupil is eligible for free school

[4]We use a scoring system developed by the Qualifications and Curriculum Authority which assigns 16 points to pass grade G, and 6 points are added for each unit improvement from grade G.

meals (FSM). FSM eligibility is linked to parents' receipt of means-tested benefits such as income support and income-based job seeker's allowance. We also include the number of months a pupil is older than an August-born (the youngest in a school cohort) to control for difference in age for children in the same school-year. We use as family background variable the number of all siblings in the state school system in 2007. This is an approximation to the true number of siblings as it only includes siblings observable in the NPD. Finally, the NPD contains information on the level of deprivation in the children's residential neighbourhood, assessed by the Income Deprivation Affecting Children Index.

School-level variables

We merge school-level expenditure information from Consistent Financial Reporting data sets for 2004-2010 to the NPD. This data set contains details on different types of income and expenditure for each school, separately for each academic year. The data allow us to derive the expenditure per pupil which excludes capital expenditure such as new construction, but includes expenditure items such as learning resources which may benefit pupils for several years. Therefore we consider the average school expenditure over three years rather than yearly expenditure in our estimates. Expenditure per pupil is expressed in 2010 prices, calculated using the GDP deflator. The data set also allows us to classify how the money was spent in each school. We distinguish teaching staff, education support staff, supply teachers, learning resources (computing and other resources) as well as other expenditure in our analysis.

In addition we add school-level characteristics to the NPD using Schools, Pupils and their Characteristics tables published by the Department for Education (e.g. Department for Education 2010). These tables are derived from the annual school censuses. School-level characteristics include an indicator of whether the school is a community school[5] or not, the number of pupils in the school, single sex schools, and whether the admission to the school is selective. Most selective schools are grammar schools which select pupils by skill at age eleven. We also characterize schools in terms of their pupil composition, using the proportion of pupils that receive free school meals, whose first language is English, that are of white British, black, mixed, Indian, Pakistani/Bangladeshi and Chinese ethnicity and

[5]Community schools are owned, governed and managed by the Local Education Authority, whereas in voluntary aided and voluntary controlled schools as well as in foundation schools some or all of these functions are carried out by other organisations such as the Church of England in faith schools, for example.

that have special educational needs (with and without statements). As for the expenditure we average these variables describing the pupil composition over three years. We also add cohort mean Key Stage 2 test scores in English, Science and Maths as school-level controls for prior attainment within the school.

Sibling definition

We identify siblings by matching pupils living at the same address, using address data released under special conditions. The first year that full address details were collected in the NPD across all pupil cohorts was 2007. Siblings are therefore defined as pupils in state schools aged 4-16 and living together at the same address in January 2007. Siblings that are not school-age, those in independent schools and those living at different addresses in January 2007 are excluded from our sibling definition. Step and half siblings are included if they live at the same address, and we are not able to distinguish them from biological siblings.[6]

Estimation sample

For our analysis we select a sample of the oldest two siblings from each household. Focusing on two siblings from each family avoids having to expand the dataset to include all sibling pair combinations within each household with the risk of over-representing households with a large number of children. In the vast majority of families there are only two siblings observed taking Key Stage 4 exams in the observation period. The sample includes all pupils that took Key Stage 4 exams in 2007 or in one of the three following years (2008, 2009, 2010).

We remove pupils with duplicate data entries or with missing data on any of the background or school-level variables from the dataset. Moreover, we retain only pupils for whom we have non-missing test scores for all outcomes at both Key Stages 2 and 4 which leads to a reduction in sample size of 13%. 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 for children with emotional and/or behavioural difficulties.

[6]The matching of siblings was carried out using 1) postcode and house number/name for addresses with no flat or block number; 2) postcode, house number/name and flat number for addresses without block number; 3) postcode, house number/name, flat and block number; 4) postcode, flat and block number where house number/name was missing. Of the 7.246 million pupil files with address information contained in the 2007 school census, only 4,158 cases had insufficient address information to be used and were therefore excluded, and 1,212 cases were dropped where more than ten siblings were identified at an address, and it is possible that they were falsely identified as siblings (false positives).

Academy schools which have been introduced in 2000 and allow schools more autonomy and flexible governance are also excluded. In 2007 about 1.5% of schools had academy status, rising to about 6.5% in 2010. Finally, we exclude the top 1% of the expenditure per pupil distribution to remove extreme outliers from the data set. The remaining sample contains 425,262 pupils (212,631 sibling pairs).

Tables 1 and 2 describe our sample. Table 1 shows unstandardized test scores for English, Science and Mathematics which are on average highest in Science at the end of primary school and in English at the end of compulsory schooling at age 16. The bottom panel of the Table displays mean school expenditure by expenditure category. The expenditure per pupil, averaged over three years, is around £5,000 in 2010 prices. The bulk of this expenditure is spent on teaching staff (£2,899 per pupil), followed by education support staff (£417 per pupil) and expenditure on learning resources (£327 per pupil). Table 2 displays individual characteristics in the top panel and school characteristics in the bottom one. 91.1% of pupils have English as their first language, 18.5% are classified as gifted and talented and 10.8% have special educational needs. The proportion of White British pupils is 83.0% and 10.8% of pupils in our sample are eligible to receive free school meals. The bottom panel of Table 2 shows that the pupil to teacher ratio in our sample is 16.5 on average. Secondary schools are quite large with more than 1,000 pupils in a school on average. The school-level proportions of pupils with free school meal eligibility, ethnicity and English as their first language are comparable to the individual level means. 4.9% of pupils go to selective schools and the majority of pupils in our sample are educated in community schools.

In Table 3 we show the percentage of pupils belonging to groups that are targeted by school policies by decile groups of the lagged test. The top panel of the Table focuses on pupils that are eligible to receive free school meals (FSM), that have special educational needs (SEN), English as an additional language (EAL) and that are classified as being gifted and talented (G&T). As expected, all groups except G&T are represented more heavily in the bottom end of the attainment distribution than the top. This is most pronounced for SEN pupils of which one in three is in the bottom lagged attainment decile, followed by children from disadvantaged socio-economic backgrounds where 70% can be found in the lower half of the distribution and one in five pupils is in the lowest decile. Among the G&T pupils one in three are in the top decile and two thirds are in the top 30%. When distinguishing pupils by

ethnicity the distribution across lagged test score deciles is more equal for most ethnicities (see bottom panel of Table 3). Notable exceptions are Blacks and pupils of Pakistani and Bangladeshi heritage who are more likely than others to be at the bottom of the distribution and Chinese pupils where the reverse is true.

Another way to look at the sample is to ask how much money is spent on each group of pupils that we distinguish in this paper. Table 4 lists the average expenditure type for each type of pupil. Pupils eligible for free school meals and with English as an additional language tend to be in schools where more money is spent in all expenditure categories - teaching staff, education support staff, supply teachers, learning resources and other expenditure - while this is not true for pupils with special educational needs and with gifts and talents. Higher expenditure on supply teachers is not necessarily a good signal, however, as these are teachers bought in from agencies to cover for absences and unfilled vacancies. In the bottom panel of the Table we list expenditure by ethnicity and find that children that according to Table 3 are from ethnic minority groups that are the lowest performing tend to be in schools that spend more across all expenditure categories. So with the exception of pupils with special educational needs those pupils that are likely to be in the lower end of the lagged test score distribution find themselves in schools where more money is being spent on pupils.

5 Empirical results

5.1 Main results

In this section we report the estimation results for an education production model with child cognitive ability at the end of compulsory schooling as outcome. We adopt the two-step estimation described in Section 3 to control for unobserved school, family and pupil characteristics and to identify the effect of school investments, measured by the school expenditure per pupil. All estimation results reported in this section refer to value added models with the dependent variable given by the standardized test at age 16 in Mathematics, English and Science and with explanatory variables which include the corresponding lagged test score at age 11, the school expenditure per pupil, a set of child's and school characteristics and dummies for the academic year.

The pupil characteristics we control for are: number of school-age siblings; gender; dummies for whether English is an additional language, the pupil has been labeled gifted and talented, is eligible for free school meals, has special educational needs; ethnic group dummy variables (White British, Black, Indian, Pakistani and Bangladeshi, Chinese, Mixed, and Other ethnic minorities); number of months the child is older than August born children and deprivation score of the child's area of residence. The school controls are: number of pupils in the school; proportions of children in the school who are eligible for free school meal, whose first language is English, who have special educational needs (with and without statement), and belonging to different ethnic groups; dummy variables for the type of school (community, selective and single sex school), and school averages of the pupils' test scores in English, Maths and Science at age 11 (see Table 2).

We take account of the fact that the productivity of inputs can be higher for children with higher cognitive ability at the end of primary school (higher lagged test) by letting the expenditure per pupil effect change across deciles of the lagged test computed across all pupils in the sample. We also consider that schools may concentrate their effort to improve the cognitive abilities of certain groups of children and this can lead to a higher effect of expenditure per pupil for these children. For this purpose we identify a set of groups of pupils which are usually targeted by school policies or from disadvantaged socio-economic backgrounds, which are children

1. eligible for free school meal (FSM),
2. with special educational needs (SEN),
3. with English as additional language (EAL),
4. gifted and talented (G&T), and
5. from minority ethnic groups;

and we estimate the productivity of the expenditure per pupil separately for each of these groups.

Table 5 reports the results, which are organized in three panels. The top panel shows the estimation results of four models, labeled model 1 to model 4, each one allowing the effect

of expenditure per pupil to differ between a considered group of pupils (FSM, SEN, EAL and G&T) and the remaining pupils. Under the heading *non-FSM* we report the estimated effect of expenditure per pupil for children who are not eligible for FSM, while under the heading *diff FSM* we report the difference between the estimated effect for FSM and non-FSM children. Next we report the corresponding results for children with SEN, with EAL and G&T. In the middle panel of Table 5, we report the estimation of a model which lets the effect of expenditure per pupil differ across ethnic groups (which is labeled in Table 5 as model 5). More precisely, we report the effect of expenditure per pupil for White British children followed by the differences between each of the minority groups and the White British. Finally, the bottom panel of Table 5 reports the results of a model which allows for dynamic complementarity (labeled model 6), i.e. for a different effect of the expenditure per pupil across deciles of the lagged school test. The first column reports the effect for the children who are at the top decile, while the following columns report the differences between each of the remaining 9 deciles and the top one.

We find that the effect of expenditure per pupil is significantly higher for children with SEN, with EAL, G&T (at 10% significance level) and for all ethnic minorities than for the corresponding comparison groups (non-SEN, non-EAL, non-G&T and White British); whereas children eligible for FSM seem to have a lower effect of expenditure per pupil (see top and middle panels in Table 5). Looking at the bottom panel of Table 5, there is a very sharp decrease in the effect of expenditure per pupil going from the top to the bottom decile of the lagged test distribution. For a 1,000 pounds increase in the expenditure per pupil, children whose lagged test is in the top decile get an improvement in their test scores of 6.2% of a standard deviation. This improvement reduces rapidly moving towards the lower part of the lagged test distribution, it is halved for pupils at the 6th decile and it is basically zero for children who are at the bottom decile.

Two main findings emerge from these estimation results. First, the effect of expenditure per pupil seems to be higher for the groups of pupils analyzed in this paper and this might be explained by school policy interventions targeted at these groups, except perhaps for children eligible for FSM. Second, there is strong evidence of dynamic complementarity, so that if children's lagged ability at the end of primary school is low, school investments during secondary school have a small effect.

Productivity of expenditure across pupil groups and lagged attainment

Although the results considered in Table 5 are very informative, we must interpret them with some caution because of two main restrictions imposed by our models. First, our models do not control for the fact that children belonging to different groups may have different levels of lagged test scores, which implies a potential difference in productivity of the expenditure per pupil. For example, children who are G&T have on average a higher level of lagged test scores (see Table 3), so their estimated higher effect of the expenditure per pupil could be simply a consequence of the fact that higher lagged ability implies a faster learning process. Second, considering one group of pupils at a time is unsatisfactory because we end up with a comparison group which is heterogenous.

For this reason we also estimate a model which allows the productivity of expenditure per pupil to change both across the 10 deciles of the lagged test score and across different groups of pupils. The model considers a triple interaction between the expenditure per pupil, the 10 dummy variables corresponding to the 10 deciles of the lagged test and the 5 dummy variables for the following groups of pupils, FSM, SEN, EAL, G&T and the reference group which is given by all pupils who do not belong to any of the above groups. Results on the effects of the expenditure per pupil are reported in Table A1 in Appendix 1, but it is easier to examine these graphically by looking at Figure 1.

The graphs reveal that the effect of expenditure per pupil increases monotonically and quite sharply across deciles for all groups of pupils. The most striking result is that children with special educational needs seem to be the group of pupils who benefit most from the expenditure per pupil for any level of the lagged test. This is probably the consequence of additional school help received by these children. In contrast, the profile of the effect of expenditure per pupil across deciles for children eligible for FSM, with EAL and G&T is quite close to the corresponding profile for the reference group.

The effect of expenditure per pupil for children with EAL is always higher than for the reference group, except for the two top deciles. This seems to suggest that the gap in test scores observed for children with EAL at the end of primary school with respect to White British children (see Table 3) reduces considerably during secondary school. This may be in part related to school policy interventions and in part related to the fact that children with EAL in our sample have been in the country at least since the end of primary school and

therefore by the end of compulsory school their potential language issues lessen. This helps in explaining not only why ethnic minority pupils improve relative to White British pupils, but also why the relative improvement is smaller for Black Caribbean pupils than others (Dustman et al. 2011).

The expenditure per pupil effect for the G&T pupils is lower than for the reference group except at the two top deciles. This seems in line with the aim of the policy intervention for G&T children, which is to provide exceptionally able children with the chance to develop their talent. Children who have been labeled G&T may have high abilities in subjects different from the ones considered in our analysis (e.g. in sports or arts), in which case they would be pushed to develop their talent in these subjects presumably with little or no effect on the test scores in Maths, English and Science.

Pupils eligible for FSM with a lagged test score over the median have an estimated effect of expenditure per pupil which is lower than the corresponding effects observed for any other group. Children eligible for FSM are more likely to be in poverty and, while this condition does not seem to limit the ability potential of children with test scores below the median at the end of primary school, it does seem to limit the potential of children above the median. It is possible that teachers have lower expectations of FSM pupils and tend not to spot and promote exceptional ability in them.

We also estimate a model with interactions of the expenditure per pupil with the dummy variables indicating the 10 deciles of the lagged test and with dummy variables for different ethnic groups. Again the results are reported graphically in Figure 2, while the full set of estimated interaction effects are reported in Appendix A, Table A2. To make the graph more readable we omit the Mixed and Other Ethnic minority groups which are very similar to the White British group of pupils. The graph displays the profiles of the effect of expenditure per pupil by deciles for White British, Black, Indian, Pakistani and Bangladeshi, and Chinese pupils.

There are substantial differences between White British and some of the other ethnic groups. Specifically, the productivity of the expenditure per pupil is higher for Indians than for White British pupils all along the lagged test distribution, whereas Chinese seem to have an extremely high productivity of the expenditure per pupil from the 5th decile onward and a lower productivity for the first decile. Chinese pupils usually perform better than White

British pupils (see Department for Education and Skills 2005 and Dustmann et al. 2011) and this is probably mainly because of their extremely positive attitudes to education and intense effort in home work (see Francis and Archer 2005). Except for the Chinese pupils all other minorities tend to have lower test results at the end of primary school than White British children (see Table 3), but there is a catch up by the end of compulsory schooling. This catch up is reflected in our estimates by a higher effect of expenditure per pupil especially for ethnic minority children whose test scores at the end of primary school are low. Black, Pakistani and Bangladeshi children, who are the ethnic minority groups performing worse at the end of primary school, seem to have higher effects of the expenditure per pupil than White British, except at the top of the lagged test distribution.

The English education system, as the US one, is characterized by test-based accountability. In England the recommended targets are to achieve 5 or more A*-C grades in exam subjects taken at the end of compulsory schooling. Previous papers have found that accountability of exam results can induce teachers to give additional help to children who are in the middle of the ability distribution, and probably closer to the exam result target, but leave behind children at the bottom and top of the distribution (Lazear 2006; Neal and Whitmore Schanzenbach, 2010). This seems in part true for England, where we find that children from disadvantaged groups such as children eligible for FSM and Black pupils have a higher effect of expenditure per pupil when they are at the bottom of the distribution of the lagged test, but a lower effect at the top of the distribution. This could indicate that children from disadvantaged backgrounds with potentially high levels of ability are left behind because of the test-based school accountability system.[7]

Differences between gender in the effect of expenditure per pupil

In this section we consider again the estimation of the education production model with heterogenous effects of school expenditure but we carry out the analysis separately for boys and girls. This allows us to assess whether there are gender differences in the productivity of inputs during secondary school. For this we select, respectively, all sister and brother pairs in the sample and perform same sex sibling fixed effects estimations. For conciseness, we report only the graphs plotting the effect of school expenditure on the level of the lagged

[7] A look at raw KS2 test scores reveals a discontinuity around the threshold for the level required at the end of primary school. This suggests that teachers were likely teaching to the mean.

test (10 deciles of the lagged test), separately by gender. Figure 3 reports this effect for girls and boys (see top and bottom graphs) for the following 5 different groups of pupils, FSM, FSM, EAL, G&T and the reference groups (i.e. the group of pupils who do not belong to any of the previous groups); whereas Figure 4 reports the equivalent graphs for the main ethnic groups again separately for girls and boys.

The productivity of the expenditure per pupil seems to increase more steeply with the level of lagged test for boys than for girls. For example, for White British boys in the bottom decile of the lagged test distribution, an increase of £1,000 in the expenditure per pupil leads to an increase of 1% of a standard deviation in the test score; whereas this increase rises to 9% for boys at the top decile. The corresponding percentages for White British girls are 1% and 6%. Similar differences between boys and girls can be seen for all other groups.

While there is a literature on gender gaps in productivity, there are no papers that look at explaining differences in productivity that vary along the distribution of the lagged test. We find that boys perform worse than girls and this confirms previous results on gender gap in test scores. Our graphs seem also to suggest that there is a partial catch up of boys with girls, but only for boys who are at the higher level of the lagged cognitive abilities.

There seem also to be some gender differences when comparing boys and girls belonging to the same ethnic group. The productivity of expenditure per pupil appears to be higher for Pakistani and Bangladeshi boys than girls, while the reverse is true when looking at the Indian ethnic group.

5.2 Quality and quantity of teachers

Two main mechanisms that schools may consider to improve the productivity of inputs are reducing the number of pupils per teacher and/or increasing the quality of the teachers. In this section we consider the number of pupils per teacher and the expenditure per pupil on teaching staff, which can be seen as a proxies for the quantity and quality of teaching respectively. The expenditure per pupil on teaching staff depends on the teachers' wages and numbers of teachers. In the English educational system wages depend mainly on teachers' experience (seniority) because teachers are paid according to national wage scales where wages increase with the years of service. This implies that the expenditure per pupil in

teaching is a measure of teachers' experience once we control for the number of pupils per teacher. Teacher experience has been linked to pupil outcomes in some but not all previous papers (e.g. Rockoff 2004) and we will use it to proxy teacher quality. To control for other school resources we also consider all other expenditure per pupil, excluding teaching staff.

In Table 6 we report the two-step estimation results of the effect of these three school resources (pupil-teacher ratio, teaching staff expenditure per pupil and all other expenditure per pupil) on test scores at age 16. The results are organized similarly to Table 5. When looking at the effect of the pupil-teacher ratio, we generally find a negative impact especially for children at the bottom of the lagged test distribution (see bottom panel) and for disadvantaged groups, in particular Pakistani and Bangladeshi pupils, children eligible for FSM and children with EAL (see middle and bottom panels). These pupils do worse if they have to share teachers with other pupils. This is consistent with Lazear (2001) who predicts that disadvantaged pupils benefit from smaller class size through the effects on disruptive behaviour. In contrast, the expenditure on teaching staff has a positive effect on test scores and especially so for the high achievers, i.e. the children at the top of the lagged test distribution, and for non-FSM, non-EAL and non-G&T and White British pupils.

All other expenditure has a larger positive effect for all the four groups of children considered in the top panel (FSM, SEN, EAL and G&T), for all minority groups and for low achievers at the end of primary school than money spent on teachers. This seems to suggest that money spent on staff and items other than teachers includes inputs that are more helpful for these groups of children. In summary, pupils at the lower end of the ability distribution and disadvantaged groups benefit from having more teachers in the classroom, while more able pupils benefit from higher quality teachers as proxied by teacher experience.

To unpack which type of other expenditure produces this beneficial effect for disadvantaged groups, we further disaggregate this into four different categories in next section.

5.3 Effect of different types of school expenditure

In this section we look at the productivity of different types of expenditure for different types of pupils. We disaggregate school expenditure per pupil into expenditure for teaching staff, education support staff, supply teachers, learning resources and other expenditure.

Education support staff are assistants intended to be used in the classroom to release teachers from non-teaching tasks, but are in reality often employed to assist in teaching, particularly of pupils with special educational needs or low ability (Blatchford et al. 2011). Supply teachers are teachers hired from public or private agencies to cover sickness, unfilled vacancies and teacher absence for professional development. Learning resources include books, computing resources and other items and devices used in teaching. Other expenditure includes categories such as premise maintenance, back office and energy use. Note that our variables measuring expenditures per pupil are averages at school level and we do not observe how expenditure was allocated to different grades and to different groups of children.

Tables 7 and 8 follow a similar structure to Table 5 and display the estimation results of the value added model estimated using the two-step procedure, where the effect of school expenditure is now allowed to vary by expenditure category. The top panel of Table 7 shows the results of four models allowing the effect of the expenditure categories to vary between the groups of pupils that we consider (FSM, SEN, EAL and G&T) and the respective reference category. The bottom panel gives the effect of each expenditure category for White British pupils as well as the difference of each of the ethnic minority groups with the White British group.

Focusing first on expenditure on teaching staff, with results reported in the first line of each panel, we find that spending money on teachers is productive for all pupils in school, but not equally so.[8] The effect is larger for the reference groups (non-FSM, non-EAL, non-G&T and White British) than for the pupils on FSM, with EAL, G&T pupils and the ethnic minorities. Looking now at the effect of committing funds to education support staff, the opposite pattern emerges. While the reference groups stand to gain little from expenditure on learning support staff; for pupils on FSM, with EAL, G&T and most ethnic minorities the effect is quite substantial. This seems to suggest that within the classroom education support staff is employed to provide extra input and support for the pupils that are singled out to receive extra help, whereas teachers dedicate their attention to the rest of the pupils. The exception are pupils with SEN who benefit more from spending on teachers than pupils without SEN but not from spending on education support staff.

[8]Chinese pupils are an exception.

Temporary absences of teachers, even if filled by supply teachers, can have negative effects because of the disruption to teaching plans and pupil-teacher interaction and the likely lower quality of the supply teachers.[9] Furthermore, any increased use of supply teachers in schools can indicate that there is absenteeism or delays in filling vacant positions, so that expenditure on supply teachers may be a proxy for other underlying problems in the school. It is therefore not surprising that we find a negative productivity effect of spending money on supply teachers. The size of the effect does not differ hugely between different groups of pupils. Exceptions are gifted and talented children and some ethnic minorities (Indians, Pakistani/Bangladeshi and other ethnicities) who seem unaffected by being taught by supply instead of permanent teachers.

The effect of spending money on learning resources is generally positive. However, some groups of children benefit less from these resources than others. These include children eligible to receive FSM, with EAL, G&T and some ethnicities (Mixed and Pakistani/Bangladeshi).

These results must be interpreted with caution, however, because we are comparing groups of pupils without controlling for the different levels of lagged test scores which imply a different productivity in learning by group. For example we find that teaching staff expenditure has been particularly productive for pupils not targeted by school policies, but we would have expected this as these pupils are likely to be from the higher end of the attainment distribution. The same is true for spending on learning resources. What we can say, however, is that expenditure on education support staff has been equally or more productive for those pupils that are targeted by school policies, and was therefore successful in attenuating the cognitive achievement gaps for the targeted children. Furthermore we can see that pupils with SEN have benefited hugely from inputs by teachers and learning resources.

This is supported by Table 8 which reports the results obtained when interacting type of expenditure with lagged attainment deciles. Expenditure on teaching staff has the highest productivity for the pupils in the top lagged attainment decile, and this productivity

[9]Issues with the quality of agencies that hire out supply teachers have led to the introduction of a quality standard in 2002.

decreases almost monotonically across deciles, indicating dynamic complementarity. In contrast, spending on education support staff has the same productivity across the whole attainment distribution. Given that we would expect productivity to be decreasing in lagged attainment because of dynamic complementarity this seems to support our earlier inference that education support staff is mostly employed to help the weakest pupils in class, whereas teachers concentrate on the whole attainment distribution. The use of supply teachers is generally detrimental, and most harmful to pupils in the 8th and 9th attainment deciles. The benefit of learning resources, on the other hand, seems to accrue equally to all pupils regardless of their lagged attainment, so their use in the classroom seems to have an equalizing effect.

One could argue that decisions of schools on how to spend their funds are potentially endogenous because they could be correlated with other unobserved school characteristics that affect the productivity of skill formation. The sibling fixed effects estimation only omits potential unobserved time-invariant school characteristics which differ between siblings. Given that the majority of siblings in our sample go to the same school (84%), the time-invariant school effect is likely to be very similar for siblings and the sibling difference estimates therefore widely unaffected by school characteristics. Nonetheless we expect to have variation in the expenditure per pupil between siblings because of the strong increasing trend in school expenditure and because siblings attend secondary school in different periods.

6 Conclusions

While most previous papers have taken into account the dynamics of the educational production function by allowing the lagged cognitive ability to affect current ability, they have generally neglected the dynamic complementarity of school inputs. To our knowledge this paper is the first to take account of three important factors that may affect the productivity of school expenditure per pupil, which are (1) the level of the child's lagged cognitive ability (dynamic complementarity), (2) the extra help schools usually give to subgroups of pupils, especially low achievers, and (3) the way schools spend their money.

We let the productivity of school inputs vary by levels of lagged achievement, by pupil groups (those pupils targeted by school policies for extra help and ethnic minorities in par-

ticular) and by expenditure category. To net out the part of the productivity gap explained by a different level of lagged cognitive skills, we compare children from diverse groups but with the same level of lagged cognitive skills.

We find that the productivity of the expenditure per pupil is increasing in lagged cognitive skills for all groups of children, indicating strong dynamic complementarity. The relationship between skills at the end of primary school and the productivity of school inputs during secondary school is stronger for boys than for girls. School policies seem to have a beneficial effect particularly on children with special educational needs, who have a higher productivity of the expenditure per pupil than children without such needs at the same levels of lagged attainment. Most ethnic minority pupils have lower test results at the end of primary school than White British pupils, but there is a catch-up by the end of compulsory school which is reflected in higher productivity of expenditure per pupil for these pupils. Corresponding to this result we find that children with English as an additional language make up in the gap that initially separates them from native English speakers. Pupils from disadvantaged backgrounds (as measured by eligibility for free school meals) and black pupils who were high achievers at the end of primary school have an estimated effect of expenditure per pupil which is lower than the corresponding effects observed for any other group. This suggests that the potential of these pupils is not reached and they are being left behind.

We take a closer look at the mechanisms through which school policies work by considering the effect of different uses of school funds. We find that pupils at the lower end of the ability distribution and disadvantaged groups benefit from having more teachers in the classroom, while more able pupils benefit from higher quality teachers as proxied by teacher experience. We also find that expenditure on education support staff has a higher productivity for children eligible for free school meals, with English as additional language, gifted and talented and belonging to ethnic minorities than other pupils not targeted by school policies, and is therefore successful in attenuating the cognitive achievement gaps for the targeted children. The same is true for spending funds on learning resources - their use in the classroom seems to have an equalizing effect as they benefit pupils equally, regardless of lagged attainment. We also find that that pupils with special educational needs have benefited hugely from inputs by teachers and learning resources.

References

Altonji J.G. and T.A. Dunn (1996),"Using Siblings to Estimate the Effect of School Quality on Wages" *The Review of Economics and Statistics*, 78(4): 665-671.

Behrman J.R., M.R. Rosenzweig and P. Taubman, (1996), "College Choice and Wages: Estimates Using Data on Female Twins" *The Review of Economics and Statistics*, 78(4): 672-685.

Blatchford, P., Bassett, P., Brown, P., Martin, C., Russell, A. and P. Webster (2011) "The impact of support staff on pupils' 'positive approaches to learning' and their academic progress", *British Educational Research Journal* 37, (3): 443-464

Chowdry, H. and L. Sibieta (2011), "School Funding Reform: An Empirical Analysis of Options for a National Funding Formula" IFS Briefing Note BN123.

Cunha, F. and J.J. Heckman (2007), "The Technology of Skill Formation" *American Economic Review*, 92(2): 31-47.

Cunha F. and Heckman J. (2008), Formulating, "Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation" *Journal of Human Resources*, 43(4), pp. 738-782.

Cunha F., Heckman J., Lochner L. and Masterov D. (2005), Interpreting the Evidence on Life Cycle Skill Formation. in Hanushek E. and Welch F. (eds.), *Handbook of the Economics of Education*, North Holland.

Cunha F., J.J. Heckman and S. Schennach (2010), "Estimating the Technology of Cognitive and Noncognitive Skill Formation" *Econometrica*, 78 (3), 883-931.

Del Boca, D., C. Monfardini and C. Nicoletti (2012), "Children's and Parent's Time-Use Choices and Cognitive Development during Adolescence" Human Capital and Economic Opportunity Working Group working paper 2012-006, Chicago: University of Chicago.

Department for Education (2011): A Consultation on School Funding Reform: Rationale and Principles. http://www.education.gov.uk/consultations/downloadableDocs/School%20Funding %20Reform%20consultation%20final.pdf, accessed 18.1.2012

Department for Education and Skills (2005): Ethnicity and Education: The Evidence on Minority Ethnic Pupils. DfES Topic Paper RTP01-05.

Dustmann, C., Machin, S. and U. Schönberg (2011), "Ethnicity and Educational Achievement in Compulsory Schooling" *Economic Journal*, 120(August): F272-297.

Gibbons, S., S. McNally and M. Viarengo (2011), "Does Additional Spending Help Urban Schools? An Evaluation Using Boundary Discontinuities" SERC Discussion Paper 90, London School of Economics.

Francis, B. and Archer, L. (2005), "British-Chinese Pupils' and Parents' Constructions of the Value of Education" *British Educational Research Journal*, 31 (1), 89-108.

Gilpin, G. and M. Kaganovich (2012), "The quantity and quality of teachers: Dynamics of the trade-off", *Journal of Public Economics*, 96, 417-429.

Hanushek E.A. (1986), "The Economics of Schooling: Production and Efficiency in Public Schools" *Journal of Economic Literature*, 24, 1141-1177.

Hanushek E.A., S.G. Rivkin and L.L. Taylor (1996), "Aggregation and the Estimated Effects of School Resources" *The Review of Economics and Statistics* 78: 4, 611-627.

Hanushek, E., Kain, J., O'Brien, D., Rivkin, S. (2005), "The market for teacher quality" NBER Working Paper 11154.

Hobbs G., and A. Vignoles (2007), "Is Free School Meal Status a Valid Proxy for Socio-economic Status (in Schools Research)?" CEEDP, 84. Centre for the Economics of Education, London School of Economics and Political Science, London, UK.

Lazear, E.P. (2001), "Educational Production" *Quartlerly Journal of Economics*, 116 (August), 777-803.

Lazear, E.P. (2006), "Speeding, terrorism, and teaching to the test" *Quarterly Journal of Economics*, 121, 1029-61.

Levačič, R. (2008), "Financing Schools. Evolving Patterns of Autonomy and Control" *Educational Management Administration & Leadership*, 36(2): 221-234.

Neal, D.A. and D. Whitmore Schanzenbach (2010), "Left behind by design: proficiency counts and test-based accountability" *Review of Economics and Statistics*, 92(2), 263-83.

Nicoletti, C. and B. Rabe (2012), "The effect of school resources on test scores in England" ISER working paper 2012-13, Colchester: University of Essex.

Rockoff, J.E. (2004), "AssociationThe Impact of Individual Teachers on Student Achievement: Evidence from Panel Data" *American Economic Review*, 94, 247-252.

Rosenzweig, M. and K.I. Wolpin (1994), "Are there Increasing Returns to the Intergenerational Production of Human Capital? Maternal Schooling and Child Intellectual Achievement" *The Journal of Human Resources*, 29(2): 670-693.

Todd, P.E. and K.I. Wolpin (2007), "The Production of Cognitive Achievement in Children: Home, School and Racial Test Score Gaps" *Journal of Human Capital*, 1(1): 91-136.

West, A. (2009), "Redistribution and Financing Schools in England under Labour. Are Resources Going Where Needs Are Greatest?" *Educational Management Administration & Leadership* 37(2): 158-179.

Tables

Table 1: Descriptive statistics: Unstandardized test scores and school expenditure

	mean	std. deviation
Unstandardized test scores		
Key stage 2 English score	27.1	4.1
Key stage 2 Science score	29.1	3.7
Key stage 2 Maths score	27.7	4.6
Key stage 4 English score	40.8	9.1
Key stage 4 Science score	40.6	10.2
Key stage 4 Maths score	40.1	10.6
School expenditure		
Expenditure per pupil(£)	4,932	678
Expenditure teaching staff (£)	2,899	308
Expenditure education support staff (£)	417	165
Expenditure Supply teachers (£)	105	67
Expenditure learning resources (£)	327	120
Other expenditure (£)	1,511	370
Number of observations	425,262	

Notes: National Pupil Database, 2007-2010; Consistent Financial Reporting Data 2005-2010; Schools, Pupils and their Characteristics Data 2005-2010. Pupil expenditure in 2010 prices, calculated using GDP deflator.

Table 2: Descriptive statistics: Individual and school-level controls

	mean	std. deviation
Individual characteristics		
Male	0.503	
N school-age siblings in state schools	2.589	0.854
First language English	0.911	
Gifted and talented	0.185	
White British	0.830	
Black	0.024	
Mixed	0.023	
Indian	0.024	
Pakistani/Bangladeshi	0.047	
Chinese	0.003	
Other ethnicity	0.049	
Free school meal eligible	0.108	
Special Educational Needs	0.165	
Deprivation score of residence	0.192	0.166
n months older than August-born	5.503	3.479
School characteristics (3 year averages)		
Pupil to teacher ratio	16.5	1.5
Number of pupils (full time eq.)	1,153	349
Prop. free school meal eligible	0.124	0.107
Prop. first language English	0.900	0.176
Prop. Special Educational Need, with statement	0.021	0.013
Prop. Special Educational Need, no statement	0.161	0.086
Prop. white British	0.808	0.230
Prop. black	0.030	0.069
Prop. mixed	0.027	0.024
Prop. Indian	0.025	0.070
Prop. Pakistani/Bangladeshi	0.038	0.113
Prop. Chinese	0.004	0.006
Community school	0.572	
Selective school	0.049	
Single sex school	0.113	
KS2 English scores, by cohort	27.0	1.5
KS2 Maths scores, by cohort	27.5	1.7
KS2 Science scores, by cohort	29.0	1.3
Number of observations	425,262	

Notes: National Pupil Database, 2007-2010; Schools, Pupils and their Characteristics Data 2005-2010.

Table 3: Percentages of pupils in different decile groups of the lagged test score

	FSM	SEN	EAL	G&T			
bottom decile	21.4	32.3	17.7	1.0			
2nd decile	15.4	17.2	13.3	1.8			
3rd decile	12.6	11.8	11.7	2.8			
4th decile	10.9	9.1	10.5	3.9			
5th decile	9.4	7.5	9.4	5.2			
6th decile	9.1	6.7	9.6	8.3			
7th decile	6.5	4.6	7.3	9.3			
8th decile	6.0	4.4	7.6	14.8			
9th decile	5.0	3.5	6.9	20.8			
top decile	3.7	2.9	6.2	32.2			
N obs.	137,241	210,807	114,126	236,142			
	White British	Black	Mixed	Indian	Pakistani/Bangladeshi	Chinese	Other ethnicity
bottom decile	8.8	16.6	9.2	10.5	20.4	5.5	11.2
2nd decile	9.2	13.1	9.3	10.6	15.1	7.5	9.6
3rd decile	9.3	11.7	9.2	10.2	12.5	8.0	9.5
4th decile	9.6	11.1	9.6	10.2	10.9	7.6	9.2
5th decile	9.7	9.9	9.4	10.1	9.3	8.5	9.3
6th decile	11.1	9.9	10.7	10.9	9.1	10.9	10.6
7th decile	9.4	7.7	8.9	8.7	6.7	8.3	8.7
8th decile	10.6	7.9	10.7	9.9	6.4	11.9	10.2
9th decile	10.9	6.8	11.0	9.8	5.4	12.5	10.5
top decile	11.4	5.2	12.1	9.2	4.4	19.4	11.2
N obs.	1,058,850	30,084	29,538	30,777	60,240	3,936	62,085

Notes: National Pupil Database, 2007-2010; Schools, Pupils and their Characteristics Data 2005-2010. FSM, SEN, EAL and G&T indicate children eligible for free school meal, with special educational needs, with English as additional language and labeled gifted and talented.

Table 4: Mean expenditure per pupil (in £) by type of expenditure

	FSM	SEN	EAL	G&T			
Teaching Staff	3,041	2,955	3,162	2,898			
Education Support Staff	481	465	481	407			
Supply Teachers	136	116	149	101			
Learning Resources	351	335	368	329			
Other expenditure	1,691	1,597	1,770	1,505			
N	45,747	70,269	38,042	78,714			
	White British	Black	Mixed	Indian	Pakist/Bangl	Chinese	Other ethnicity
Teaching Staff	2,859	3,312	3,003	3,071	3,111	3,009	3,032
Education Support Staff	409	476	436	443	500	413	420
Supply Teachers	100	160	116	122	152	111	115
Learning Resources	321	373	339	359	373	341	338
Other expenditure	1,475	1,775	1,575	1,662	1,808	1,569	1,595
N	352,950	10,028	9,846	10,259	20,080	1,312	20,695

Notes: National Pupil Database, 2007-2010; Consistent Financial Reporting Data 2005-2010; Schools, Pupils and their Characteristics Data 2005-2010. Pupil expenditure in 2010 prices, calculated using GDP deflator. FSM, SEN, EAL and G&T indicate children eligible for free school meal, with special educational needs, with English as additional language and labeled gifted and talented.

Table 5: The effect of expenditure per pupil on test scores at age 16

Target group	Model 1		Model 2		Model 3		Model 4		
	non-FSM	diff FSM	non-SEN	diff SEN	non-EAL	diff EAL	non-G&T	diff G&T	
Exp. p. pupil	0.029**	-0.003**	0.023**	0.023**	0.028**	0.004**	0.028**	0.006+	
	(0.003)	(0.001)	(0.003)	(0.003)	(0.003)	(0.001)	(0.003)	(0.003)	
Ethnicity	Model 5								
	White Brit.	Difference to White British for							
		Black	Mixed	Indian	Pak/Bangl	Chinese	Other		
Exp. p. pupil	0.027**	0.006*	0.003*	0.010**	0.011**	0.020**	0.002*		
	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.008)	(0.001)		
Lagged test	Model 6								
	Top decile	Difference to top decile for the							
		9th	8th	7th	6th	5th	4th	3rd	2nd	1st
Exp. p. pupil	0.062**	-0.011**	-0.020**	-0.023**	-0.031**	-0.035**	-0.041**	-0.047**	-0.053**	-0.062**
	(0.003)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Number of observations	1,275,786									

Notes: Pupil expenditure in 2010 prices, calculated using GDP deflator. FSM: free school meal eligible; SEN: special educational needs; EAL: English as additional language; G&T: gifted and talented. + p < .10, * p < .05, ** p < .01. Tests are standardized. Robust standard errors in parenthesis. Control variables include all variables listed in Table 2 plus the standardized lagged test and dummies for academic year.
Data source: National Pupil Database, 2007-2010; Consistent Financial Reporting Data 2005-2010; Schools, Pupils and their Characteristics Data 2005-2010.

[9]blah blah blah

Figure 1: Effect of expenditure per pupil by deciles of the lagged test for different groups of pupils.

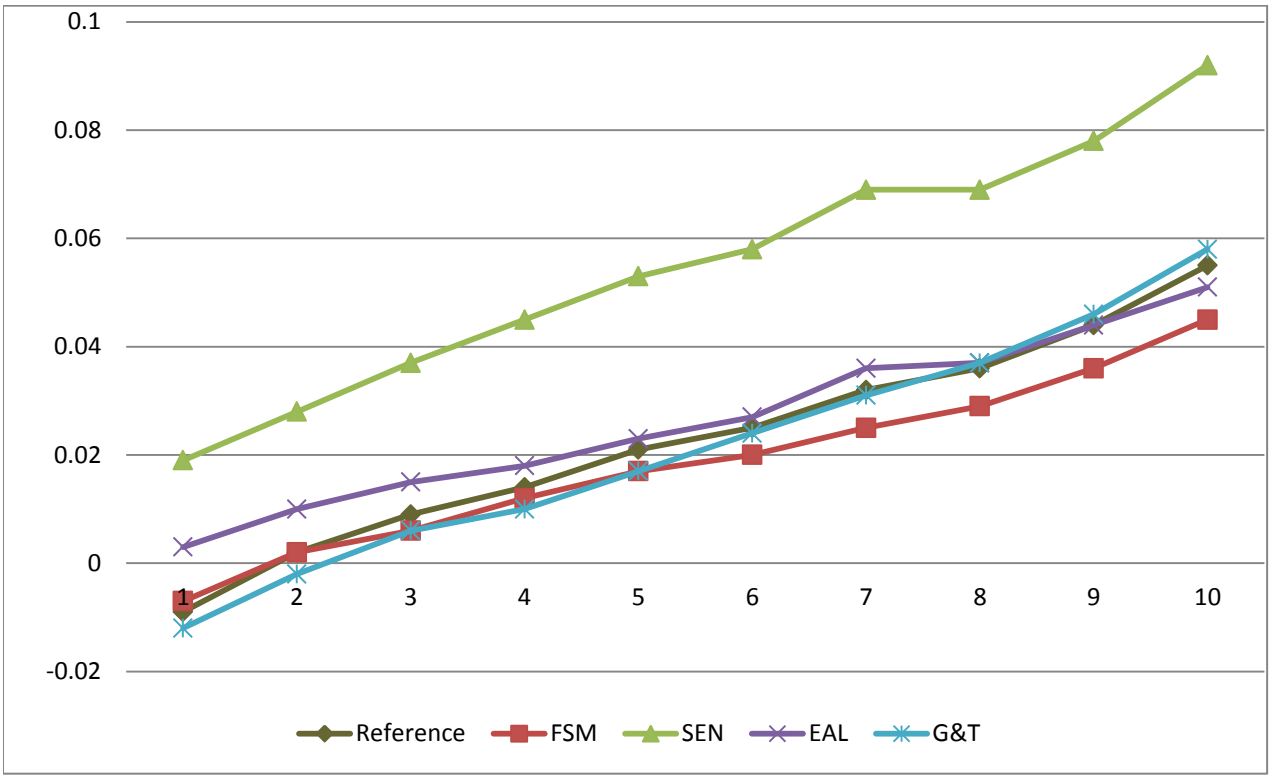


Figure 2: Effect of expenditure per pupil by deciles of the lagged test for different ethnic groups.

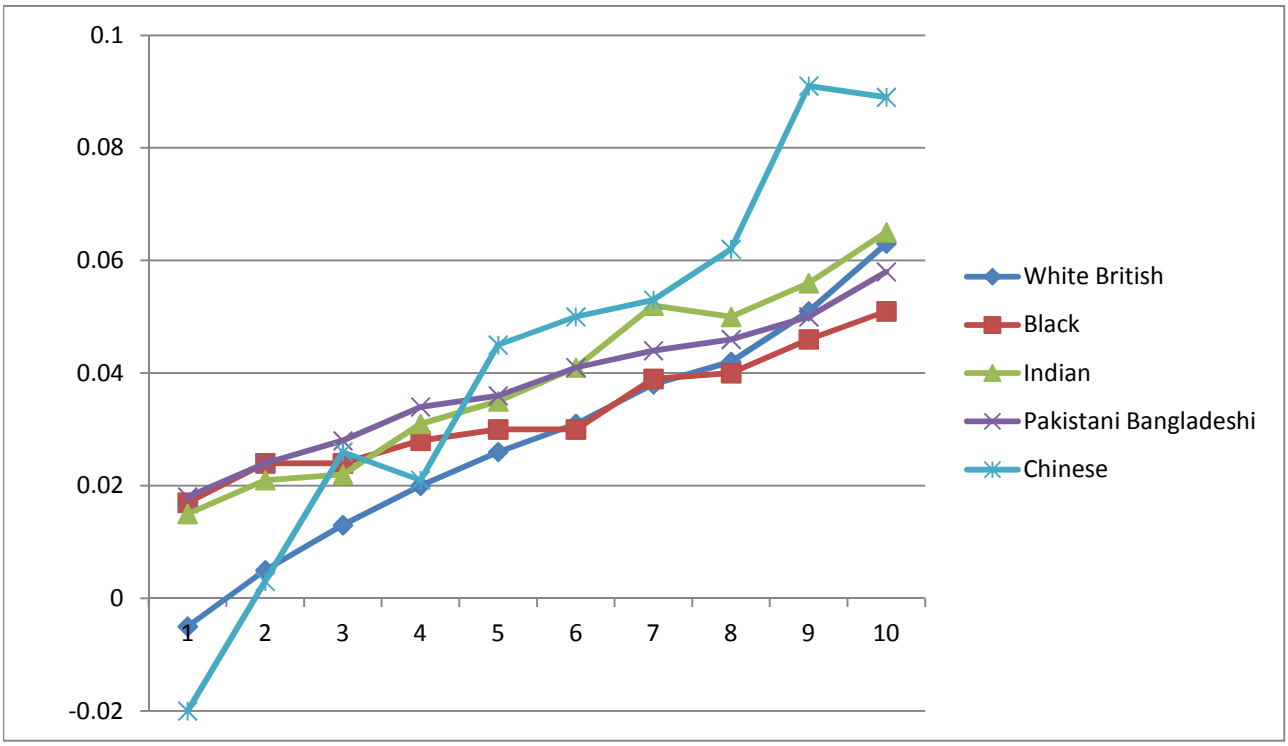


Figure 3: Effect of expenditure per pupil by deciles of the lagged test for different groups of pupils: girls and boys

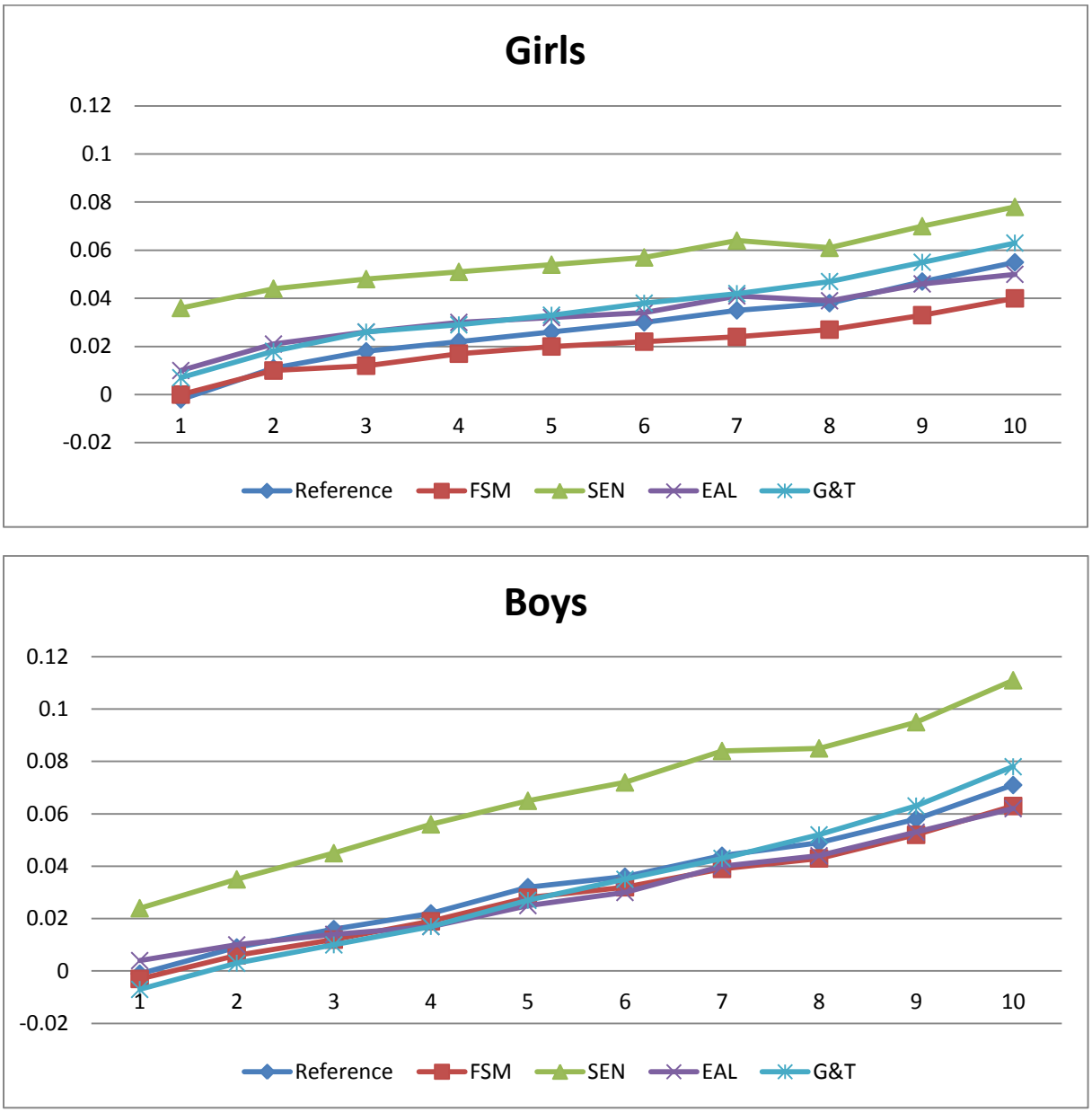
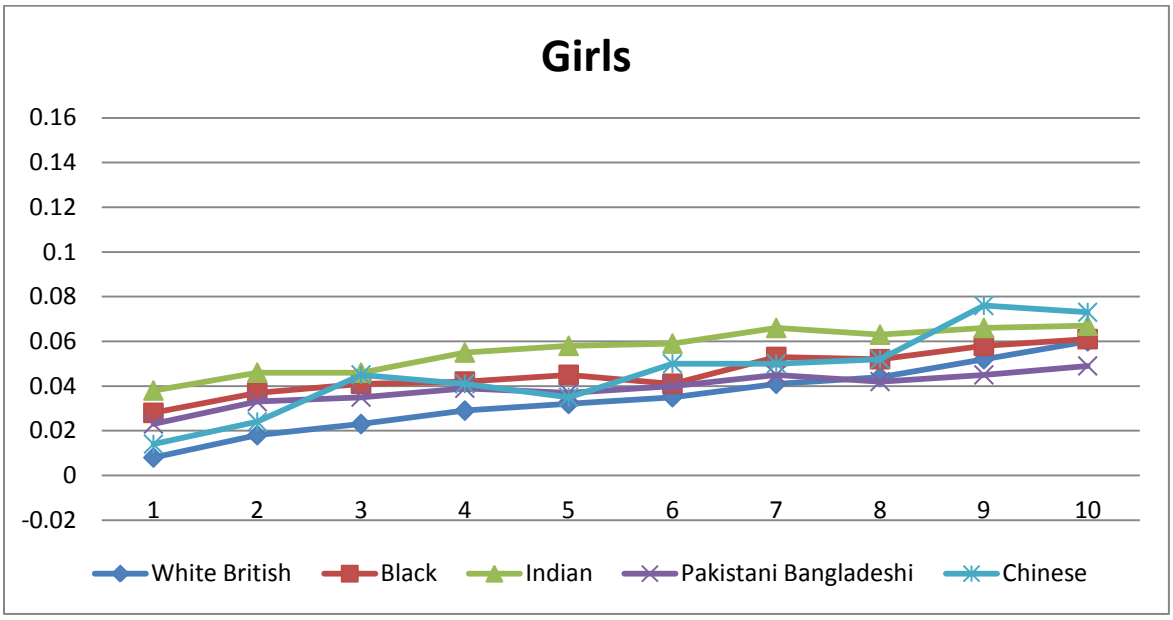


Figure 4: Effect of expenditure per pupil by deciles of the lagged test for different ethnic groups: girls and boys



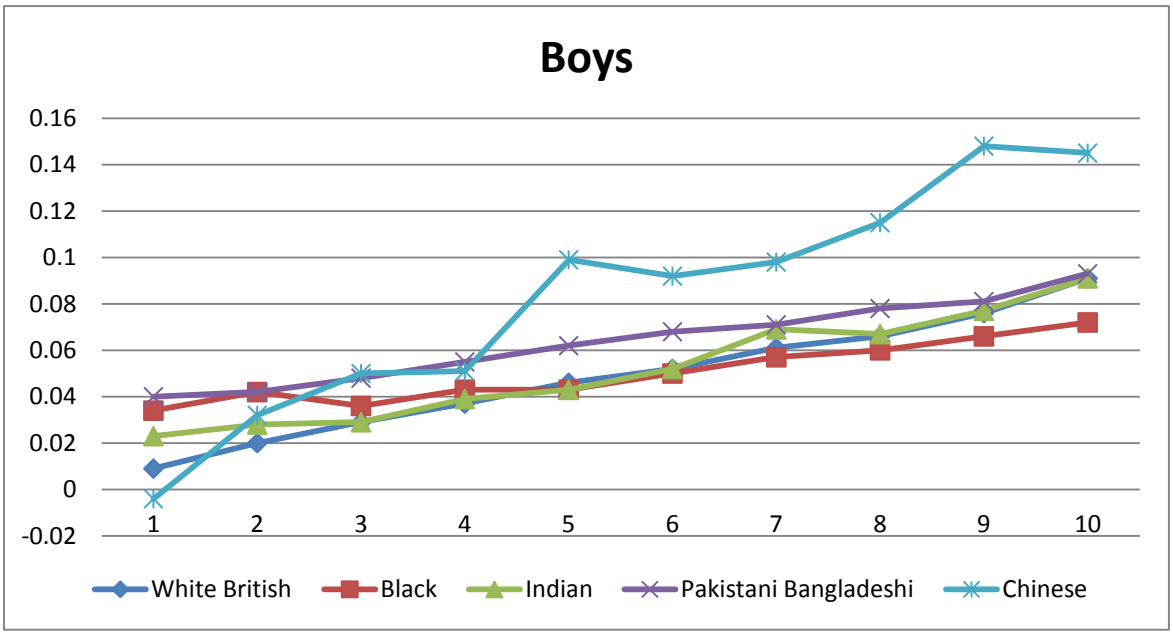


Table 6: Effects of the quality and quantity of teachers on test scores at age 16

Target group	Model 1		Model 2		Model 3		Model 4			
	non-FSM	diff FSM	non-SEN	diff SEN	non-EAL	diff EAL	non-G&T	diff G&T		
Pupil-teacher ratio	-0.008**	-0.006**	-0.008**	-0.002	-0.008**	-0.005**	-0.009**	-0.001		
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)		
Teaching Staff exp.	0.020**	0.004	0.016*	0.021*	0.025**	-0.017*	0.026**	-0.027**		
	(0.007)	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)		
All other exp.	0.011**	0.030**	0.011**	0.021**	0.008+	0.061**	0.012**	0.022**		
	(0.004)	(0.007)	(0.004)	(0.004)	(0.004)	(0.008)	(0.004)	(0.005)		
Ethnicity	Model 5									
	White Brit.	Difference to white British for								
		Black	Mixed	Indian	Pak/Bangl	Chinese	Other			
Pupil-teacher ratio	-0.009**	-0.004+	-0.003	0.007**	-0.006**	-0.005	-0.001			
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.007)	(0.001)			
Teaching Staff exp.	0.027**	-0.003	-0.013	-0.081**	-0.014	-0.056	-0.020*			
	(0.008)	(0.015)	(0.015)	(0.016)	(0.014)	(0.048)	(0.010)			
All other exp.	0.002	0.044**	0.044**	0.078**	0.074**	0.156**	0.037**			
	(0.004)	(0.014)	(0.014)	(0.015)	(0.011)	(0.048)	(0.009)			
Lagged test	Model 6									
	Top decile	Difference to top decile for the								
		9th	8th	7th	6th	5th	4th	3rd	2nd	1st
Pupil-teacher ratio	0.000	-0.003**	-0.006**	-0.007**	-0.008**	-0.009**	-0.010**	-0.013**	-0.015**	-0.017**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Teaching Staff exp.	0.045**	-0.004	-0.005	-0.015*	-0.026**	-0.028**	-0.035**	-0.040**	-0.043**	-0.052**
	(0.007)	(0.005)	(0.006)	(0.006)	(0.007)	(0.007)	(0.008)	(0.008)	(0.009)	(0.011)
All other exp.	-0.003	-0.001	0.003	0.015*	0.021**	0.019**	0.020**	0.032**	0.029**	0.034**
	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.009)	(0.010)
Number of observations	1,275,786									

Notes: Pupil expenditure in 2010 prices, calculated using GDP deflator. FSM: free school meal eligible; SEN: special educational needs; EAL: English as additional language; G&T: gifted and talented. + p < .10, * p < .05, ** p < .01. Tests are standardized. Robust standard errors in parenthesis. Control variables include all variables listed in Table 2 plus the standardized lagged test and dummies for academic year. Data source: National Pupil Database, 2007-2010; Consistent Financial Reporting Data 2005-2010; Schools, Pupils and their Characteristics Data 2005-2010.

Table 7: The effect of different types of expenditure on test scores at age 16

Target group	Model 1		Model 2		Model 3		Model 4	
	non-FSM	diff FSM	non-SEN	diff SEN	non-EAL	diff EAL	non-G&T	diff G&T
Teaching staff	0.047**	-0.027**	0.038**	0.028**	0.050**	-0.035**	0.049**	-0.025**
	(0.006)	(0.005)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)
Educ. support staff	0.026*	0.070**	0.034**	0.001	0.023+	0.124**	0.018	0.110**
	(0.012)	(0.020)	(0.012)	(0.013)	(0.012)	(0.024)	(0.012)	(0.014)
Supply Teachers	-0.227**	-0.054	-0.229**	-0.026	-0.251**	0.076	-0.280**	0.275**
	(0.023)	(0.044)	(0.022)	(0.030)	(0.023)	(0.047)	(0.022)	(0.032)
Learning Resources	0.072**	-0.085**	0.051**	0.062**	0.082**	-0.180**	0.070**	-0.052**
	(0.013)	(0.031)	(0.013)	(0.020)	(0.013)	(0.035)	(0.013)	(0.020)
Other exp.	0.006	0.040**	0.008	0.020**	0.003	0.073**	0.017**	-0.020**
	(0.006)	(0.011)	(0.006)	(0.007)	(0.006)	(0.012)	(0.006)	(0.007)
Ethnicity	Model 5							
	White Brit.	Difference to white British for						
		Black	Mixed	Indian	Pak/Bangl	Chinese	Other	
Teaching staff	0.052**	-0.025*	-0.022*	-0.047**	-0.036**	-0.074*	-0.023**	
	(0.006)	(0.011)	(0.010)	(0.012)	(0.010)	(0.036)	(0.007)	
Educ. support staff	0.008	0.102*	0.106*	0.106*	0.195**	0.191	0.082**	
	(0.013)	(0.044)	(0.042)	(0.048)	(0.035)	(0.147)	(0.028)	
Supply Teachers	-0.291**	0.124+	0.112	0.346**	0.232**	-0.213	0.136*	
	(0.025)	(0.074)	(0.090)	(0.108)	(0.071)	(0.296)	(0.063)	
Learning Resources	0.091**	-0.089	-0.153*	-0.078	-0.299**	0.042	-0.035	
	(0.014)	(0.064)	(0.060)	(0.068)	(0.049)	(0.206)	(0.040)	
Other exp.	-0.001	0.046*	0.046*	0.075**	0.081**	0.156+	0.024+	
	(0.006)	(0.023)	(0.023)	(0.023)	(0.017)	(0.088)	(0.015)	
Number of observations	1,275,786							

Notes: Pupil expenditure in 2010 prices, calculated using GDP deflator. FSM: free school meal eligible; SEN: special educational needs; EAL: English as additional language; G&T: gifted and talented. + p < .10, * p < .05, ** p < .01. Tests are standardized. Robust standard errors in parenthesis. Control variables include all variables listed in Table 2 plus the standardized lagged test and dummies for academic year. Data source: National Pupil Database, 2007-2010; Consistent Financial Reporting Data 2005-2010; Schools, Pupils and their Characteristics Data 2005-2010.

Table 8: The effect of different types of expenditure on test scores at age 16 by deciles of the lagged test scores

	Top decile	Difference to top decile for the								
		9th	8th	7th	6th	5th	4th	3rd	2nd	1st
Teaching staff	0.109**	-0.017**	-0.034**	-0.048**	-0.066**	-0.071**	-0.083**	-0.101**	-0.112**	-0.131**
	(0.006)	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)	(0.008)
Educ. support staff	0.043**	0.008	-0.004	-0.018	-0.010	-0.016	-0.004	0.010	-0.014	-0.013
	(0.012)	(0.015)	(0.017)	(0.018)	(0.019)	(0.021)	(0.022)	(0.023)	(0.026)	(0.034)
Supply Teachers	-0.230**	-0.112**	-0.090*	-0.003	-0.027	-0.005	-0.002	0.010	0.037	0.106
	(0.022)	(0.037)	(0.041)	(0.044)	(0.046)	(0.049)	(0.051)	(0.056)	(0.062)	(0.075)
Learning Resources	0.081**	-0.007	-0.008	-0.032	-0.019	-0.037	0.001	-0.036	-0.003	-0.055
	(0.013)	(0.022)	(0.023)	(0.026)	(0.027)	(0.028)	(0.030)	(0.031)	(0.035)	(0.043)
Other exp.	-0.009	0.003	0.008	0.025*	0.030**	0.029**	0.020+	0.035**	0.031*	0.039*
	(0.006)	(0.008)	(0.009)	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)	(0.014)	(0.016)
Number of observations	1,275,786									

Notes: Pupil expenditure in 2010 prices, calculated using GDP deflator. + p < .10, * p < .05, ** p < .01. Tests are standardized. Robust standard errors in parenthesis. Control variables include all variables listed in Table 2 plus the standardized lagged test and dummies for academic year. Data source: National Pupil Database, 2007-2010; Consistent Financial Reporting Data 2005-2010; Schools, Pupils and their Characteristics Data 2005-2010.

Appendix A: Triple interaction results

Table A1: The effect of expenditure per pupil by deciles of the lagged test, by groups of pupils

	Top decile	Difference to top decile for the								
		9th	8th	7th	6th	5th	4th	3rd	2nd	1st
row j, column i	*i=1*	*i=2*	*i=3*	*i=4*	*i=5*	*i=6*	*i=7*	*i=8*	*i=9*	*i=10*
Reference, *j=1*	0.055**	-0.011**	-0.019**	-0.023**	-0.030**	-0.034**	-0.041**	-0.046**	-0.053**	-0.064**
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Diff. to Reference										
FSM, *j=2*	-0.010**	0.002	0.003	0.003	0.005*	0.006**	0.008**	0.007**	0.010**	0.012**
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
SEN, *j=3*	0.037**	-0.003	-0.004	0.000	-0.004+	-0.005*	-0.006*	-0.009**	-0.011**	-0.009**
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
EAL, *j=4*	-0.004**	0.004*	0.005**	0.008**	0.006**	0.006**	0.008**	0.010**	0.012**	0.016**
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
G&T, *j=5*	0.003	0.004*	0.005**	0.008**	0.006**	0.006**	0.008**	0.010**	0.012**	0.016**
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)

Notes: For the reference group (row *j=1*) the effect is given by adding cell *(1,1)* to the cell with the decile of interest, *(1,i)*. For pupils who belong to a non-reference group (*j>1*) and who are in the top decile, the effect is given by adding cell *(1,1)* to cell *(j,1)*. For pupils who belong to a non-reference group (*j>1*) and who are in a lower decile (*i>1*) the effect is given by adding cells *(1,1)*, *(1,i)*, *(j,1)* and *(j,i)*. Pupil expenditure in 2010 prices, calculated using GDP deflator. + p < .10, * p < .05, ** p < .01. Tests are standardized. Robust standard errors in parenthesis. Control variables include all variables listed in Table 2 plus the standardized lagged test and dummies for academic year. Data source: National Pupil Database, 2007-2010; Consistent Financial Reporting Data 2005-2010; Schools, Pupils and their Characteristics Data 2005-2010.

Table A2: The effect of expenditure per pupil by deciles of the lagged test, by ethnicity

	Top decile	Difference to top decile for the								
		9th	8th	7th	6th	5th	4th	3rd	2nd	1st
row j, column i	*i=1*	*i=2*	*i=3*	*i=4*	*i=5*	*i=6*	*i=7*	*i=8*	*i=9*	*i=10*
White British, *j=1*	0.063**	-0.012**	-0.021**	-0.025**	-0.032**	-0.037**	-0.043**	-0.050**	-0.058**	-0.068**
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Diff. to White British for										
Black, *j=2*	-0.012**	0.007*	0.010**	0.013**	0.011**	0.016**	0.020**	0.023**	0.031**	0.034**
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Mixed, *j=3*	-0.001	0.001	0.003	0.002	0.002	0.004	0.001	0.005	0.010*	0.011*
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)
Indian, *j=4*	0.002	0.003	0.006*	0.012**	0.008*	0.007*	0.009**	0.007*	0.014**	0.018**
	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Pak./ Bangl., *j=5*	-0.005+	0.004	0.009**	0.011**	0.015**	0.015**	0.019**	0.020**	0.024**	0.028**
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Chinese, *j=6*	0.026**	0.014**	-0.006	-0.011	-0.007	-0.007	-0.025**	-0.013	-0.028**	-0.041**
	(0.008)	(0.005)	(0.006)	(0.007)	(0.007)	(0.008)	(0.008)	(0.010)	(0.010)	(0.013)
Other, *j=7*	0.001	-0.001	-0.002	0.003	-0.000	0.001	-0.001	0.002	0.004	0.007+
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)

Notes: For the reference group (row *j=1*) the effect is given by adding cell (*1,1*) to the cell with the decile of interest, (*1,i*). For pupils who belong to a non-reference group (*j>1*) and who are in the top decile, the effect is given by adding cell (*1,1*) to cell (*j,1*). For pupils who belong to a non-reference group (*j>1*) and who are in a lower decile (*i>1*) the effect is given by adding cells (*1,1*), (*1,i*), (*j,1*) and (*j,i*). Pupil expenditure in 2010 prices, calculated using GDP deflator. + p < .10, * p < .05, ** p < .01. Tests are standardized. Robust standard errors in parenthesis. Control variables include all variables listed in Table 2 plus the standardized lagged test and dummies for academic year. Data source: National Pupil Database, 2007-2010; Consistent Financial Reporting Data 2005-2010; Schools, Pupils and their Characteristics Data 2005-2010.