Sibling spillover effects in school achievement

Cheti Nicoletti
DERS, University of York and ISER, University of Essex

Birgitta Rabe
ISER, University of Essex

This version: May 12, 2016

Abstract

This paper provides empirical evidence on direct sibling spillover effects in school achievement using administrative data on 220 thousand siblings in England. We extend previous strategies to identify peer effects by exploiting the variation in school test scores across three subjects observed at ages 11 and 16 as well as variation in peer quality between siblings. We find a statistically significant positive spillover effect from the older to the younger sibling but not vice versa. Spillover effects from high achieving older siblings are larger than from low achieving ones, but this relationship is weaker for students from disadvantaged backgrounds.

Keywords: Family effects, peer effects, social interaction, education

JEL codes: I22, I24

Contact: Cheti Nicoletti: Department of Economics, University of York, York YO10 5DD, cheti.nicoletti@york.ac.uk

We thank the Department for Education for making available, under terms and conditions, data from the National Pupil Database. This research was supported by core funding of the Research Centre on Micro-Social Change at the Institute for Social and Economic Research by the University of Essex and the UK Economic and Social Research Council (awards RES-518-28-001 and ES L009/53/1).

1 Introduction

In this paper we study the extent to which school achievements of an older sibling directly improve the school outcomes of their younger sibling. Assessing the magnitude of sibling spillover effects is important to understand whether interactions between siblings are a relevant mechanism through which intergenerational transmission of disadvantage operates. It also helps us to understand whether sibling interactions are a mechanism through which the effect of investments in children may be amplified by the so called social multiplier effect (see Manski 1993 and 2000; Glaeser et al. 2003). A large positive spillover effect would suggest that there are externalities of parental and public investments into children through their positive effects on siblings.

While the economic literature recognizes the important role of parent-child interactions for child development,[1] the role of sibling interactions is yet to be clearly established. Previous economic papers concentrating on siblings have mainly focused on the intrafamily allocation of resources (Becker 1981), where parental investments into children's human capital depend on parental preferences regarding inequality between children,[2] on birth order and the number and gender composition of siblings.[3] It is only recently that economists have begun to look at the effect of interactions between siblings on educational outcomes. Qureshi (2015a) and Joensen and Nielsen (2015) provide evidence on the causal sibling spillover effect on years of schooling and subject choices. We add to this literature by providing empirical evidence on the extent to which *cognitive ability* of a child is transmitted to his/her younger sibling. More precisely, we estimate the sibling spillover effect of a child's school test scores at age 16 on her younger sibling's test scores at the same age. By focusing on spillover effects in compulsory study subjects (English, Science and Maths), we are able to capture sibling influence on skills, effort and motivation rather than on subject choice.

There is an extensive literature providing evidence on the existence of school peer effects in cognitive ability (see Sacerdote et al. 2011 and Epple and Romano 2011 for reviews).

[1] See, for example, Akabyashi (2006), Hotz et al. (2008), Heckman and Mosso (2014).

[2] See Behrman et al. (1982), Behrman et al. (1994), Rammohan and Robertson (2011), Del Bono et al. (2012), Almond and Mazumder (2013).

[3] See e.g. Becker and Thomes (1976), Behrman and Taubman (1986), Butcher and Case (1994), Kessler (1991), Black et al. (2005).

However, interactions between siblings are arguably the most frequent and relevant interactions a child may have with other children, and they could therefore result in substantially larger spillover effects. From a policy point of view this would suggest that family-based interventions have a larger social multiplier effect than school-based ones. This paper sheds light on this question by providing, to our knowledge for the first time, estimates of sibling spillover effects in school achievement.

A direct causal link between the cognitive ability of the older and younger sibling may exist for several reasons. For example, there may be productivity spillovers from the older to the younger sibling through teaching, help with homework and joint formative and leisure activities. Conversely, a badly behaved older sibling might cause disruption at home, causing negative productivity spillovers. Another mechanism may be imitation, which happens because a sibling gains utility from behaving similarly to their sibling (see Barr and Hayne 2003, Joensen and Nielsen 2015). Here a sibling may be a role model for academic behavior, educational aspirations and values. Finally, a further mechanism is information transmission. A sibling may share information about the costs and benefits of exerting effort, as well as knowledge pertaining to a school or specific teachers.

In this paper we are interested in estimating the role of interactions between siblings in the transmission of cognitive abilities from older to younger siblings. For this reason we aim at producing an estimate of the sibling spillover effect that is cleaned as far as possible of indirect effects caused by behavioral responses or other confounding factors. We have the advantage of access to census-level information on the whole population of children in state schools in England, with outcomes that are measured in externally marked national tests for several subjects and at several points in time. Our estimation strategy allows us to improve on the previous literature in the field by identifying the spillover effect of school attainment between siblings while minimizing biases caused by omitted variables, in particular those related to parental investments. This strategy can be viewed as a combination of two different methods. The first method exploits the variation of school test scores across subjects to eliminate individual fixed effects, while the second instruments the older sibling's school test score with the predetermined school performance of the older sibling's peers.

Simply regressing a child test score on the older sibling's corresponding test score would not produce a consistent estimate of the sibling spillover effect because the estimated sibling association would be in part explained by similarities in inherited abilities, in school and family investments and characteristics, and in the environment siblings are exposed to. To clean the sibling association in test scores of these confounding factors, we make use of test scores at the end of compulsory schooling (at about age 16) in Mathematics, English and Science. We regress a child's test score on her older sibling's test score using within-pupil between-subject estimation, i.e. estimating child fixed effects.[4] The two main gains of this fixed effect estimation are that it allows us to (i) control for the younger child's unobservable average ability and other characteristics that are invariant across the three subjects and may confound the spillover effect because they are similar between siblings, (ii) clean the sibling spillover effect of the impact of investments by schools and parents between siblings that do not vary across subjects. Further, to account for subject-specific school characteristics we rely on school-by-cohort-by-subject fixed effects.

To further take account of *subject-specific* skills that are acquired from parents through family investments and/or inheritance and shared by siblings, we instrument the older sibling's test scores at age 16 using the average test scores at age 11 of her school mates. Our instrumental variable strategy exploits idiosyncratic differences in peer group quality across schools and/or across cohorts, and we will present descriptive statistics of the variation in our instrument conditional on individual and school-by-cohort-by-subject fixed effects. To avoid reverse causality running from the older sibling to her school mates, we consider only the performance of new peers that the older sibling first encountered in secondary school and we use the new peers' prior test scores obtained in primary school to measure attainment (see Lavy et al. 2012; Gibbons and Telhaj, 2015).

Our peer identification strategy is similar to that adopted by Kelejian and Prucha (1998), Lee (2003), Bramoullé et al. (2009), De Giorgi et al. (2010), De Giorgi et al. (2015) and Nicoletti et al. (2016) and is based on the presence of some intransitivity in the network of peers. Intransitivity occurs if a person interacts with her peers but not with all of the peers of her peers. In our application we have intransitivity because we assume that the older sibling's school mates do not interact directly with the younger sibling. This implies that, while the older sibling's test scores can be affected directly by her school mates' results,

[4] This estimation is similar in spirit to the within-pupil between-subject estimation used by Dee (2005) and (2007), Clotfelter et al. (2010) and Slater et al. (2010) and it has been used by Lavy et al. (2012) to estimate school peer effect on test scores.

there is no effect from the older sibling's school mates on the younger sibling (other than indirectly through the older sibling). We scrutinize this identifying assumption by performing sensitivity checks on the data, for example by excluding from the estimation sample school mates of the older sibling who live in the same area and might therefore interact directly with the younger sibling.

Empirical researchers estimating a causal effect between individuals' outcomes are usually concerned with the reflection problem (see Manski 1993; Brock and Durlauf 2001; Moffitt 2001), i.e. simultaneity of the individuals' behavior and potential reverse causality. We cannot exclude the existence of spillover effects going from the younger to the older sibling,[5] but in our application the younger sibling's age 16 exam is in the future with respect to the corresponding older sibling's exam at age 16. Therefore reverse causality seems unlikely. In any case, our instrumental variable estimation is able to control for the potential reflection issue because the prior average test scores of the older sibling's new school peers cannot be affected by the younger sibling.

Based on our instrumental variable estimation we find that an increase of a standard deviation in a child's test score at age 16 leads to a statistically significant increase in the corresponding test score observed for his/her younger sibling of about 10% of a standard deviation. This means that for each exam grade improvement of the older sibling - for example from a B to an A - the younger sibling's exam marks increase by about 10% of a grade, which is equivalent to the impact of increasing yearly per pupil school expenditure in the younger sibling's school by around £1,000 (see Nicoletti and Rabe 2012).

We investigate potential mechanisms explaining the sibling spillover effect from the older to the younger sibling by performing sub-group analysis (see Dahl et al. 2014). We find suggestive evidence of strong productivity spillovers for children whose older siblings are at the top of the achievement distribution. This indicates that older siblings who perform well in school are more effective teachers for their younger siblings. We also find that spillover effects of older siblings who perform badly in school are stronger for disadvantaged children, suggesting that there are externalities from investing in learning of disadvantaged children which have so far been overlooked.

[5] Ewin Smith (1993) suggests that the cognitive abilities of older children might improve thanks to teaching younger siblings.

2 Related literature

Over many years, social scientists have used sibling correlations in socio-economic and educational outcomes to measure the importance of family background, where any sibling resemblance indicates that family background matters. Since Solon et al. (2000) introduced the variance decomposition approach to put bounds on the possible magnitude of family and neighborhood effects, using correlations between siblings and between unrelated neighbors, a large number of empirical papers have analyzed sibling correlations in different outcomes (see Raaum et al. 2006; Mazumder 2008; Björklund et al. 2002; Björklund et al. 2009; Lindahl 2011; Björklund and Salvanes 2011; Nicoletti and Rabe 2013). Furthermore, sibling correlations have been decomposed into the part that is related to intergenerational transmission and a residual part, see Corcoran et al. 1990; Solon 1999; Bingley and Cappellari 2013). Nevertheless, this decomposition does not provide estimate of the causal sibling spillover effect directly attributable to sibling interactions.

In Table 1 we summarize the results of previous papers on sibling spillover effects that have identified a causal spillover effect. In all papers, the reflection problem is dealt with by using instrumental variables that explain the outcome of one sibling but not the other.[6] These papers study a wide range of outcomes, including high school graduation (Oettinger 2000), teenage motherhood (Monstad et al. 2011), paternity leave take-up (Dahl et al. 2014), years of schooling (Adermon 2013; Qureshi 2015a), school subject choices (Joensen and Nielsen 2015), and household formation (Aparicio-Fenoll and Oppedisano 2016).

The majority of these papers use the introduction of policy reforms that changed the conditional probability (cost) of a given outcome for a random portion of siblings to identify sibling spillover effects. For example, Joensen and Nielsen (2015) look at a pilot school reform implemented in Denmark which reduced the cost for students of choosing advanced Mathematics and Science courses because of the introduction of a more flexible choice set for subject combinations. The probability of choosing advanced Mathematics and Science increases by about 33.4 percentage points for children whose older sibling chose these advanced subjects and this spillover effect is statistically significant at the 10% level. Both

[6] Another two papers looking at spillover effects of siblings which do not use instrumental variables for their estimation are Kuziemko (2006) and Altonji et al. (2016). Both papers use panel data and estimate dynamic models to identify sibling spillover effects on fertility and teenage substance use, respectively.

Adermon (2013) and Monstad et al. (2011) exploit the increase in the minimum school leaving age, in Sweden in the 1940s and 1950s and Norway in the 1960s. An increase of two years in the school leaving age was introduced at different times in different municipalities creating variation over time and space in years of schooling of the older sibling (and therefore of fertility in the case of Monstad et al. 2011). Adermon (2013) finds no effect of an older sibling's years of schooling on the younger sibling while Monstad et al. (2011) find positive effects on teenage motherhood.

One possible concern with using policy reforms to identify the direct effect of sibling interactions is that some of these reforms are implemented over long time-periods, leaving parents time to adjust their investments between siblings. For example, in the case of a rise in the school leaving age parents might motivate the older sibling not affected by the reform to stay in school for longer and/or discourage the younger sibling from staying on after compulsory schooling ends in a bid to equalize between siblings, leading to estimates of the sibling spillover effect that are biased downwards.[7] This paper is affected by this problem to a much lesser extent because our identification is based on an instrumental variable that exploits group membership and idiosyncratic differences in the average subject-specific peer quality across cohorts that we will argue are not readily observable by parents and therefore less susceptible to behavioral responses.

Another advantage of our approach is that our identification strategy does not require a policy reform to generate variation in the outcome for the older sibling and is therefore applicable regardless of specific policies or contexts. Assuming that there is idiosyncratic variation in peer quality over time and across schools and our identifying assumption holds, the identification strategy we propose can be applied at any time in any country where appropriate data are available.

There are some recent papers that have estimated the total effect on a child's educational outcomes of conditions or policy reforms affecting his/her sibling (see Fletcher et al. 2012; Breining 2014; Breining et al. 2015; Qureshi 2015b). The estimation of such total spillover

[7] Joensen and Nielsen (2015) who look at the effect on siblings' subject choices of improved access to advanced Mathematics and Science courses avoid the issue of parental responses by focusing on the first year of implementation only.

effects, which include both the direct causal effect and the indirect effect through e.g. behavioral responses by parents, can be of policy interest; but it is difficult if not impossible to generalize the indirect effect to other contexts.

Other papers we are aware of that focus on the direct, but likely not causal, effect of sibling interactions on child cognitive development belong to the psychology literature and usually focus on early child development. Cicirelli (1972) and Dunn (1983) provide evidence that young children are effective teachers for their younger siblings. Gregory and Williams (2001) emphasize the importance of older siblings in transmitting school values to their younger siblings, especially in immigrant households where parents have difficulty speaking the language used at school. Azmitia and Hesser (1993) compare sibling and peer influence on child cognitive development and find that older siblings are more effective in teaching their younger siblings than unrelated children of the same age.

A strand of the economic literature closely related to our paper is the literature on peer effects in education. Sacerdote (2011) and Epple and Romano (2011) provide comprehensive surveys of this literature. We are interested in comparing our results on sibling spillovers to estimates of school peer effects to assess whether sibling interactions are more relevant than school peer interactions. For this comparison, we focus on papers that use a similar methodology to us, i.e. papers based on fixed effects estimation that exploits idiosyncratic variation in school peer quality across school cohorts and/or variation in the child's ability across subjects or grades, and papers which use data on nominated school friends and exploit intransitivity in friendship networks.

Hoxby (2000) is the first to exploit differences in peer quality within schools and across cohorts, and finds a school peer effect varying between 0.10 and 0.55 percentage points for a point increase in the reading and mathematics score averaged across school peers. Hanushek et al. (2003), who use the same data (Texas Schools Project) but control for both student and school-by-grade fixed effects, find that an increase of 1 standard deviation in the average achievement of peers leads to an increase in the child's achievement of about 20% of a standard deviation. Lin (2010), who considers the network of nominated friends in the National Longitudinal Survey of Adolescent Health (AddHealth), find an effect of school friends of about 27% of a standard deviation on the child's grade point average. The magnitude of the estimated school peer effects reduce substantially in more recent papers that have used similar approaches but have controlled more extensively for issues of reverse causality (e.g. Lavy et al. 2012) or for the endogeneity of the friendship network (Del Bello et al. 2015). Using the same school administrative data as in our application, Lavy et al. (2012) find a zero school peer effect when considering the average ability across all school peers in a cohort, and Gibbons and Telhaj (2015) find a small effect of 2% of a standard deviation for a standard deviation increase in peers' prior scores. Similarly, using AddHealth, Del Bello et al. (2015) find a zero effect of friends living in the same neighbourhood on the child's grade point average. In contrast, the effect of friends in the same school estimated by Patacchini et al. (2011) and Del Bello et al. (2015) are somewhat higher at 10% of a standard deviation in completed years of schooling and in average grade points, respectively.

The literature on educational production and child development (see Todd and Wolpin 2003; Cunha and Heckman 2007; Cunha and Heckman 2008; Hanushek and Woessmann 2011) is also closely related to our paper. This has provided a theoretical framework to model the production of children's cognitive abilities taking account of family and school inputs, but has not focused on the potential effect of interactions between siblings. In this paper, we extend the recent work on education production models by evaluating, to our knowledge for the first time, the effect of sibling interactions on child development during adolescence.

3 Identification strategy

To identify the sibling spillover effect on test scores at the end of compulsory schooling (at about age 16) we consider the following value added model:[8]

$$Y_{1,isqt,16} = \alpha + Y_{1,isqt,11}\rho + Y_{2,is'qt',16}\gamma + I^F_{1,it}\beta_{1,F} + I^S_{1,ist}\beta_{1,S} + X_{1,isqt}\beta_{1,X} + \mu_{sqt} + \mu_{1,i} + e_{1,isqt,16}, \quad (1)$$

where:

- $Y_{1,isqt,16}$ is the age 16 test score of the younger child of the sibling-pair i, in school s and subject q, who belongs to the cohort t;[9]
- $Y_{1,isqt,11}$ is the corresponding test score at age 11;
- $Y_{2,is'qt',16}$ is the test score at age 16 of the older sibling, who might have attended a different school s' and belongs to a different cohort t';
- $I^F_{1,it}$ is the family investment in the younger child of the sibling-pair i between age 11 and 16;
- $I^S_{1,ist}$ is the corresponding school investment that is not subject-specific;
- $X_{1,isqt}$ is a row vector of other child, household and school characteristics, which are not direct investments in a child's cognitive skills but may affect them;
- μ_{sqt} are unobserved investments that vary by school, subject and cohort;
- $\mu_{1,i}$ is the younger child's unobservable ability;
- $e_{1,isqt,16}$ is an error term which is assumed to be identically and independently distributed with mean zero and homoscedastic.

In this model ρ measures the persistence in test scores between age 11 and 16; γ is our main parameter of interest which measures the spillover effect from the older to the younger sibling; $\beta_{1,F}$ and $\beta_{1,S}$ are the productivity of family and school investments; and $\beta_{1,X}$ is a

[8] See Todd and Wolpin (2003) for a definition.

[9] Two students belong to the same school cohort if they began school in the same year. We do not consider twins or siblings whose age gap is such that they begin school in the same year.

column vector with the effects of the remaining explanatory variables $X_{1,i}$, and α is the intercept. We observe for each sibling-pair their test scores in Mathematics, English and Science so that q takes value 1 for Mathematics, 2 for English and 3 for Science.

Identifying the causal spillover effect in test scores from the older to the younger sibling, γ, is challenging because of three main issues: (i) unobserved correlated effects, i.e unobserved common characteristics of two siblings that may explain their similar test scores; (ii) the reflection problem, i.e. reverse causality; (iii) the endogeneity of the network, i.e. non-random sorting of individuals into groups.

To address (i), we control for unobserved child-specific endowments and characteristics that do not vary across subjects but that could be similar between siblings by transforming model (1) in deviations from the mean across subjects, i.e. we transform the dependent variable in $DevY_{1,isqt,16} = Y_{1,isqt,16} - \sum_{j=1}^{3} Y_{1,isjt,16}/3$ and we apply an analogous transformation to all right hand side variables, leading to

$$DevY_{1,isqt,16} = DevY_{1,isqt,11}\rho + DevY_{2is'qt',16}\gamma + Dev\mu_{sqt} + Deve_{1,isqt,16}. \qquad (2)$$

This transformation eliminates from the equation all inputs that do not vary across subjects as well as the unobserved child endowment, $\mu_{1,i}$. This comprises cognitive and non-cognitive abilities and health which could be similar between siblings, therefore confounding the sibling spillover effect.[10] Furthermore, the transformation also eliminates the effect of changes in parental investments between siblings, as long as they do not vary across subjects.

However, the deviation from the mean across subjects is unable to eliminate unobserved characteristics that do vary by subject. In particular we might be concerned about subject-specific *school* investments that are shared by siblings because they attend schools with similar (or indeed same) characteristics that are unobserved by us. We partial out shared subject-specific school background by using school-by-cohort-by-subject fixed effects that control for μ_{sqt}, i.e. for unobserved subject-specific school investments and characteristics for the cohort t.

The issue of unobserved subject-specific *family* investments and skills inheritance is more challenging. By controlling for the lagged test score, i.e. the test score in subject q at age

[10] If we chose to include controls pertaining to the older sibling in the model $(I^F_{2,it}, I^S_{2,ist}, X_{2,isqt})$, these would also be eliminated with this transformation.

11, we estimate a spillover effect that is purged of the influence of such family characteristics up to the age of 11.[11] To also control for the effect of these unobserved subject-specific characteristics between ages 11 and 16, we adopt instrumental variable estimation where we instrument the subject-specific test scores of the older sibling at age 16 using the average attainment of the school-by-cohort peers of the older sibling.

We need to make sure that our instrumental variable estimation is not affected by reverse causality (the reflection problem, issue (ii)). While we can be quite confident that there is no reverse causality from the younger to the older sibling because the younger sibling's test scores are in the future; reverse causality between the older sibling and her peers could affect the validity of our instrument. Therefore we adopt the strategy used by Lavy et al. (2012) and Gibbons and Telhaj (2015) who measure peers' ability using prior achievements in end-of-primary-school national tests at age 11 but only considering *new* peers that a pupil (in our case the older sibling) encounters for the first time in secondary school (in the compulsory transition from primary to secondary school a major reshuffling of pupils takes place so that on average students meet more than 80% new peers.) This instrument is immune to reflection problems because the older sibling's test score at age 16 cannot affect her new school peers' test scores at age 11.

We therefore instrument the subject-specific test scores of the older sibling at age 16 using the average of $DevY_{js'qt',11}$ (measured in attainment percentiles) over the *new* school-by-cohort peers of the older sibling, which we call $NewMDevY_{2,s'qt',11}$. Because in equation (1) we control for both child fixed effects and school-by-cohort-by-subject fixed effects, the instrument captures whether the older sibling's new school-cohort mates were relatively better in a specific subject than the younger sibling's school-cohort mates, after partialling out the effect of any older sibling's mates who have a younger sibling in the same school-cohort as our reference sibling pair.[12] The variation in the instrument is caused by idiosyncratic differences in average subject-specific peer quality between groups of school peers within and across schools and cohorts. These differences can occur, for example, because of changes in the quality of teaching in a specific subject (e.g. because of teacher turnover) or in the

[11] In our empirical model we control both for same-subject and cross-subject past test scores, in other words we let the age 16 English score depend on age 11 scores in English, Science and Mathematics, and the same for the other subjects.

[12] Partialling out the latter effect avoids any influence of the older sibling's mates through their younger siblings who are in the same school-cohort as our reference child.

composition of the new school-cohort mates in terms of subject-specific abilities. We use the same type of instrumental variable estimation to compute the spillover effect from the younger to the older sibling. The model specification is identical to model (1) with the subscripts 1 and 2 exchanged to swap the role of the younger sibling with the one of the older sibling.

Our identifying assumption is that a student can be affected by the test scores of the new school peers of her sibling only through her sibling. This assumption could be invalid if there is direct interaction between the older sibling's new school mates and the younger sibling, for example. We discuss this and other possible threats to identification in section 5.2 and present a number of robustness checks. For example, we exclude the older sibling's school peers who live in the same neighborhood from the computation of $NewMDevY_{2,s'qt',11}$ to assess whether possible interaction within a neighborhood may affect results. We conclude from these checks that our estimated sibling spillover effect holds across a number of specifications.

To be a valid instrumental variable $NewMDevY_{2,s'qt',11}$ must be uncorrelated with any unobserved variables that affect the younger sibling's test results. The younger sibling's test result can be affected by the prior average test results of her school peers at the end of primary school, which we call $MDevY_{1,sqt,11}$, and more in particular by the prior average test results of her new school peers, $NewMDevY_{1,sqt,11}$, which is the variable corresponding to $NewMDevY_{2,sqt,11}$ of the older sibling. Because siblings tend to attend similar schools with similar subject-specific characteristics, our instrument is likely to be correlated with both $MDevY_{1,sqt,11}$ and $NewMDevY_{1,sqt,11}$. Therefore we explicitly control for $NewMDevY_{1,sqt,11}$ in our equation (2). The subject-specific average test results of peers, $MDevY_{1,sqt,11}$, get wiped out by estimating school-by-cohort-by-subject fixed effects for the younger sibling.

In order to isolate the causal effect of the older sibling's attainment on the younger we also need to assume that there are no behavioral responses by parents. Our identification strategy is more immune to this problem than previous papers that rely on policy reforms that affect one sibling and not the other, for example an increase in the school leaving age. In the case of policy reforms we expect parents to adjust the allocation of their investments between siblings so that the instrumental variable is not independent of these investments. Of course parents could also react to our instrument, which captures whether the older sibling's new school-cohort mates were relatively better in a specific subject than the younger sibling's school-cohort mates, after partialling out the effect of older sibling mates who have a younger sibling in the same school-cohort as our reference sibling pair. It does not seem plausible that parents are able to perceive this type of variation however.

The final concern when estimating sibling spillover effects is the endogeneity of the network (issue (iii)). Our network of peers is endogenous if the likelihood of peer interaction depends on unobserved characteristics which also affect school test scores. In our case the network of peers for each child is defined by her school-cohort peers and her sibling, therefore the likelihood to form interactions depends on selection into the family and into the school.

In our sample a high percentage of siblings, 83.5%, attend the same secondary school, but even if two siblings attend two different schools they might sort into schools with similar characteristics, e.g. similar quality of teachers in a particular subject or peers with similar subject-specific abilities. As we have noted, controlling for school-by-cohort-by-subject fixed effects allows us to clean the sibling spillover effect from the confounding effect of such similarities in school characteristics and school peers between siblings. The selection into families could bias our results if older siblings have school-by-cohort peers with unobserved genetic traits and family characteristics similar to their younger siblings. Nevertheless, this potential bias is eliminated because our estimation takes account of these unobservables by considering individual and school-by-cohort-by-subject fixed effects.

4 Data

The empirical analysis is based on 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 English students. It combines student level attainment data with student characteristics as they progress through primary and secondary school.

Educational system in England

Full-time education is compulsory for all children aged between 5 and 16, with most children attending primary school from age 5 to 11 and secondary school from age 11 to 16. The education during these years is divided into four Key Stages. Students take externally marked National Curriculum Tests at the end of Key Stages 2 and 4. Until recently such national 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. Pupils take tests in the three core subjects of English, Mathematics and Science. Key Stage 4 tests are taken at age 16 at the end of compulsory schooling. Pupils enter General Certificate of Secondary Education (GCSE) or equivalent vocational or occupational exams at this stage. They 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.

Outcome and observed background

We focus on GCSEs (Key Stage 4) because they mark the first major branching point in a young person's educational career, and lower levels of GCSE attainment are likely to have a longer term impact on experiences in the labor market. We consider results in the core subjects English, Mathematics and Science which are directly comparable to test results at the end of primary school. Students receive a grade for each GCSE course, where pass grades include A*, A, B, C, D, E, F, G. We use a scoring system developed by the Qualifications and Curriculum Authority to transform these grades into a continuous point score which we refer to as the Key Stage 4 score.[13]

We control for lagged cognitive achievement using Key Stage 2 National Curriculum tests taken at the end of primary school, usually at age 11, in 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. We control for past test scores in the same as well as the other subjects (same- and cross-subject

[13] A pass grade G receives 16 points, and 6 points are added for each unit improvement from grade G.

effects). All test scores are standardized to have a mean of zero and a standard deviation of one.

The NPD annual school census provides a number of individual and family background variables. These include month and year of birth and gender of the student, ethnicity, whether or not the first language spoken at home is English, any special educational needs identified for the child, eligibility for free school meals (FSM)[14], area of residence and the number of siblings in the family. As we control for child fixed effects in all our models we do not use these variables as explanatory variables, apart from an interaction term between pupil gender and subject-specific effects to control for gender differences in attainment. We do use several of the background characteristics in the estimation of heterogeneous spillover effects by subgroups and in our robustness checks.

Sibling definition

The NPD includes address data, released under special conditions, which allows us to match siblings in the data set in the year 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 (see Nicoletti and Rabe 2013 for details).

Peer Ability

For each older sibling in our data set we construct a measure of peer ability based on the peers' end-of-primary school test scores that are unaffected by the older sibling. By using information on the primary schools attended by all pupils we restrict this measure to the new peers encountered by the older sibling for the first time in secondary school. Each student in our sample has 187 cohort peers on average, of which 160 (86%) are new peers. As class identifiers are not available in the data we use grade-level ability to proxy the quality of peers experienced by the older sibling. Measurement error in peer quality may bias estimates downward but should not affect the validity of our instrument, and unlike Lavy et al. (2012)

[14]FSM eligibility is linked to parents' receipt of means-tested benefits such as income support and income-based job seeker's allowance and has been used in many studies as a low-income marker (see Hobbs and Vignoles 2010 for some shortcomings).

we do not need to restrict our sample to small secondary schools because our aim is not to identify an endogenous school mate effect in the first stage. We do follow Lavy et al. (2012) in expressing peer ability in terms of percentiles by subject. Moreover, for robustness analysis we adopt their definition of the very worst performing students as the fraction of new peers that were in the bottom 5th percentile of the subject ability distribution at Key Stage 2.

Sample restrictions

The main sample for our analysis includes all sibling pairs taking their Key Stage 4 exams in 2007, 2008, 2009 or 2010. We remove from the data all twins and siblings attending the same academic year. When we have multiple pairs of siblings from one family in the observation window we consider the two oldest students to avoid any multiplier spillover effects (what Dahl et al. 2014 call the snowball effect).[15] We also remove pupils with duplicate data entries or with missing data on background variables from the dataset (4% of the sample). 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 10.5%. Missing cases are concentrated among low attaining students that are more likely to be absent at the exams or, at Key Stage 4, choose not to take exams in one or more of the core subjects. Comparing the original with the retained sample the average test score is increased by about 1%. 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 behavioral difficulties. Further we adopt some of the sampling restrictions used in Lavy et al. (2012), namely we exclude secondary schools with fewer than 15 pupils and schools where the fraction of pupils below the 5th or above the 95th percentile exceeds 20%. The final sample contains 414,360 siblings (207,180 sibling pairs) who go to 2,948 secondary schools in England. We use data that is pooled across 3 subjects, so that we have 621,540 sibling pair observations in total.

Descriptives

Table 2 reports the means and standard deviations of the unstandardized test scores at age 11 and 16 (Key Stages 2 and 4) respectively; but in all our estimated models we consider the

[15] The percentage of siblings who have more than one older siblings is about 1.4% of the sample which already excludes twins and siblings attending the same academic year.

standardized test scores by subject. The bottom panel of the Table also provides mean and standard deviation of other characteristics used for the estimation of heterogeneous spillover effects and in our robustness analysis.

Table 3 gives an overview of the identifying variation in our dependent variable, i.e. the younger sibling's standardized school test score at age 16, and in our instrumental variable, which is the average of the subject-specific Key Stage 2 test score percentile across the older sibling's new school peers. The top panel of Table 3 shows the mean and total variation measured by the standard deviation of the younger sibling's test score, the variation net of the child fixed effect and finally the variation net of both the child and of the school-by-cohort-by-subject fixed effects (see first, second and third rows). The variation net of the child fixed effect is the within individual (child) variation measured by the standard deviation of the residuals of a child fixed effect regression. Similarly the variation net of the child and of the school-by-cohort-by-subject fixed effects is measured by the standard deviation of the residuals of the regression that controls for both child fixed effect and school-by-cohort-by-subject fixed effects. The within child variation is quite substantial, about 40% of the overall variation. Further applying a child and school-by-cohort-by-subject fixed effects estimation does not reduce the variation in the data by much, there is still over a third of the original variation left.

The bottom panel of Table 3 shows the variation in our instrument. First we show the total variation in the mean test score percentiles of the older siblings' new peers, excluding the older sibling. On average the older sibling has 160 new peers in the same school-cohort. The total variation in the average peer test score percentile (49.07) is 9.35. By considering the variation net of the child fixed effect we capture the extent to which the older sibling's peers are relatively better in one subject than the others, for example because they have a good teacher in a particular subject. The standard deviation net of the child fixed effect is 2.34 percentiles. The last row of the Table shows the variation in the data net of both the child and school-by-cohort-by-subject fixed effect. This is the instrument we use to identify our model. As we can see, about 20% of the original variation in the data is remaining after applying the various fixed effects.

5 Empirical Results

5.1 Main empirical results

We begin by reporting in Table 4 the correlations in sibling's test scores which are a general measure of the importance of background shared between siblings on educational outcomes. Since the test scores at ages 11 and 16 are standardized by subject to have zero mean and unit variance, we can estimate the raw correlation in test scores by a simple regression of the test scores at age 16 on the sibling's test score at age 16.[16] This produces the so called sibling intraclass correlation that does not generally capture a causal peer effect (see Angrist 2014). The raw correlation in test scores is shown in column (1) of Table 4 and estimated to be 0.478 which is in line with results of previous papers (e.g. Nicoletti and Rabe 2013; Björklund et al. 2010).

In column (2) we display the sibling correlation in test scores net of the effect of past test scores obtained by the younger sibling at the end of primary school, which we estimate by using a value added model, i.e. by regressing the test scores at 16 on the sibling's test scores at 16 and controlling for same and cross-subject test scores at age 11 (as well as subject-gender interactions). This sibling correlation captures the effect of shared family and environment characteristics which operate between ages 11 and 16. We can see that the net sibling correlation is 0.285. In column (3) we show the correlation estimated using the same value added model as in column (2) and controlling for the younger child fixed effects (0.136). This eliminates the influence of all environment, family and child characteristics that are invariant across subjects, including the intra-household allocation of resources between siblings.

Finally, in column (4) we show the sibling correlation estimated using both child fixed effects and school-by-cohort-by-subject fixed effects. The latter net out subject-specific school characteristics. This correlation (0.102) therefore comes closest to capturing a causal relationship, but it can still be overestimated because of unobserved subject-specific skills transmitted in the family that are similar between siblings. Families are likely to have

[16]For all regression models we take account that the error terms are clustered at school-cohort-subject level and report robust standard errors. We do not find any difference in the standard errors if we allow for cluster correlation in the error terms within school or within school-cohort level, rather than within school-cohort-subject.

subject-specific traits - being a 'maths' or 'music' family, for example - which can affect both subject-specific inherited child endowments and subject-specific family investments.

In Table 5 we present our main estimates of the sibling spillover effect in school test scores from the older to the younger sibling at age 16 (end of compulsory schooling) when controlling for individual fixed effects as well as for school-by-cohort-by-subject fixed effects and using instrumental variable estimation to eliminate the bias caused by omitted subject-specific family investments and characteristics (see column 1). Furthermore in column (2) we report the corresponding instrumental variable estimation for the sibling spillover effect going from the younger to the older sibling, using the average ability of the younger sibling's new school peers as instrument. For both estimations we consider the value added model (1), where the control variables include past (same and cross-subject) test scores obtained at the end of primary school, the younger (older in column 2) sibling's average peer performance and gender-subject interactions. We are not concerned about the endogeneity of the lagged test caused by the fact that child unobserved endowments influence both the test scores at ages 11 and 16 because all our estimations control for child fixed effects and therefore eliminate child unobserved endowments.[17]

Our instrumental variable estimation is a two-stage least square (2SLS) estimation with fixed effects. The first stage consists in the regression of the older sibling's test score at 16 on all explanatory variables plus an instrument given by the average subject-specific ability at age 11 of the older sibling's new school peers encountered for the first time in secondary school; whereas the second stage is the regression of the younger sibling's test score on all explanatory variables and with the older sibling's test score replaced by its prediction from the first stage regression. Both first and second stage regression control for the individual fixed effect and the younger siblings' school-by-cohort-by-subject fixed effect. Furthermore, we control for the subject-specific average test score of the new peers of the younger sibling, i.e. her secondary school cohort peers who attended a different primary school. Therefore, our instrument captures whether the older sibling's school-cohort new mates were relatively better in a specific subject than the younger sibling's new school-cohort mates, for example because of changes in the quality and quantity of school inputs in a specific subject or in the subject-specific abilities of the school-cohort mates.

[17] This method to correct for the endogeneity of the lagged test has already been applied for example in Nicoletti and Rabe (2012), Slater et al. (2012) and Del Boca et al. (2012).

The first stage regression is very similar to the model adopted in Lavy et al. (2012) to estimate school peer effects using the same school administrative data that we use, but looking at test scores at age 14 (Key Stage 3) rather than at age 16 (Key Stage 4) as outcomes. As in Lavy et al. (2012) we are concerned with the reverse causality that goes from older siblings to their school peers and deal with this by using predetermined peer ability measures.[18] But, contrary to them, we are not interested in interpreting the coefficient of the ability of the older sibling's school peers as an endogenous school peer effect. What we are concerned with is the validity of our instrument, which holds even if the effect of the older sibling's school peers is explained by unobserved contextual factors such as the subject-specific teaching ability of the older sibling's teachers. Indeed our first stage regression controls only for unobserved contextual factors that may affect the older sibling's new school peers who have a younger sibling in the same school and cohort as our reference younger sibling. Therefore we do not expect to exactly replicate in our first stage estimation the results of Lavy et al. (2012) who find no effect of average peer ability but a negative effect of being surrounded by very badly performing peers.[19]

The first row of Table 5 shows the estimated sibling peer effect. Looking first at the sibling spillover effect from the older to the younger sibling (column 1), we find that an increase of one standard deviation in the test score of the older sibling leads to an increase of 9.8% of a standard deviation in the corresponding test score of the younger sibling, and this effect is strongly statistically significant. This spillover effect seems fairly small, indeed it is smaller than most sibling peer effects estimated in previous papers on other outcomes, listed in Table 1. In contrast, there is no statistically significant spillover effect in test scores going from the younger to the older sibling (see column 2). This is in line with expectations, as we would not expect the age 16 test scores of the older sibling to be affected by their younger sibling's tests that take place in the future.

[18]Recall that to avoid reverse causality in the first stage equation our instrument is computed averaging the subject-specific ability at the end of primary school (age 11) only of new school peers, i.e. excluding all old peers that went to the same primary school as the older sibling.

[19]There are also other differences between our first stage estimation and the main equation estimated in Lavy et al. (2012), including differences in (i) the estimation method (we control for younger sibling's school-by-cohort-by subject fixed effects additionally to the child fixed effect), (ii) the point in time when school test scores are measured (at 16 rather than 14), the selection of the sample (e.g. we do not focus on small schools).

The F-statistics for the significance of the instrumental variable in the first stage is large and does not leave any doubt on the validity of the instrument. In contrast to Lavy et al. (2012) we find a positive and statistically significant effect of average peer ability on the older sibling's test scores (in line with Gibbons and Telhaj, 2015). We report the first stage results when considering as instrument the older sibling's new school peers average test score measured in attainment percentiles, as well as when considering as instrumental variables the percentage of school peers in the top and bottom 5% of the test score distribution (the latter is the main peer effect identified by Lavy et al., 2012) respectively in panels A, B and C in Table A1 in the Appendix. In our first stage equation it is the average ability of school peers rather than the percentage of badly performing peers that matters.

The endogeneity test reported in Table 5 suggests that after controlling for child fixed effects and school-by-cohort-by-subject fixed effects there is no residual endogeneity of the older sibling's test score and we cannot reject the equality of the estimation with fixed effects and the 2SLS estimation with fixed effects. The difference between the two estimates is very small and equivalent to just 0.004 of a standard deviation in the younger sibling's test score for an increase of a standard deviation in the older sibling's test score (compare column 4 in Table 4 and column 1 in Table 5). Although we prefer the 2SLS estimation, the estimation with fixed effects is more precise, and we will therefore use it to produce estimates that allow for a heterogenous sibling spillover effect (see Section 5.3).

How do the sibling effects we estimate compare to those obtained for school peers and school friends in previous papers? Our review of the relevant literature in Section 2 showed that school peer effects estimated in recent paper are zero or very small, while effects based on nominated school friends are higher at around 10% of a standard deviation. This seems to suggest that sibling interactions are comparable to interactions between school friends and more relevant than interactions between school peers.

5.2 Threats to identification: Robustness checks

In this section we discuss threats to the validity of our identification strategy and probe the stability of our benchmark estimates to alternative specifications.

Direct influence of older sibling's school mates on the younger sibling

Our identifying assumption is that the older sibling's peers have no direct influence on the younger sibling's test score. We investigate here the possibility that older sibling's school mates could directly interact with the younger sibling in the neighborhood and therefore violate the exogeneity assumption. Although we are excluding peers from the older sibling's primary school including 'forever friends' who the younger sibling may know and have interacted with as a child, it may be that some new secondary school peers live in the same neighborhood and interact with each other even if they do not belong to the same cohort. Evidence for England shows that there are no neighborhood peer effects in school achievement (Gibbons et al. 2013), but we still want to test this possibility. In our data, we can define neighborhoods based on Lower Level Super Output Areas which are statistical geographies created to reflect proximity and social homogeneity and have an average of roughly 1,500 residents and 650 households. In our sample, an average of 9 peers from the same school and cohort live in a neighborhood defined in this way (a school cohort comprises 188 pupils on average). Among these, 5 students are old and 4 are new peers. Secondary students may interact within a wider geographical area, so we also look at Middle Layer Output Areas (with a minimum size of 5,000 residents and 3,000 households with an average population size of 7,500). An average of 34 peers from the same school and cohort live in an area thus defined, of which 22 are new peers. We take this as the maximum proportion of the older sibling's school mates a (very sociable) younger sibling could be exposed to within the residential area. Note that MSOAs are quite large geographical areas, with an average size of 1,958 hectares across England (1 hectare=10,000m^2) which cannot easily be covered by a child on a regular basis, in particular in rural areas.

To test the possibility of neighborhood interaction, we exclude the older sibling's new school peers living in the same neighborhood from the computation of the instrumental variable to remove the potential direct effects that go from children living in the same neighborhood to the younger sibling. We also perform the same test by excluding older sibling's new school peers living in the same area, defined at the Middle Layer Super Output Area (MSOA) level. Table 6 displays the results of this exercise. Excluding older sibling's new school mates living in the same neighborhood from the calculation of the instrument changes the estimated sibling spillover effect by very little. Excluding older sibling's school mates living in the same area again produces a result that is comparable to the benchmark estimate. This suggests that direct interaction within neighborhoods and wider areas does not threaten our identifying assumption.

Another possibility is that younger siblings directly interact with their older sibling's school mates at school. However, unlike the cases of Bramoullé et al. (2009) and Calvó-Armengolo et al. (2009), where the unrelated peers of peers can be taught in the same school class, in our case the older sibling's peers are in different classes and cohorts than the younger sibling, sometimes several years apart. In English schools cohorts are taught strictly separately, and because of the large cohort size of secondary schools even school assemblies and trips usually take place separately by cohort. This means that interactions that are relevant for learning are unlikely to take place across cohorts in school.

Our instrumental variable could also fail because of the way our sample is constructed. We have data for four cohorts of students taking age-16 exams, and it is possible that an older sibling has school mates whose younger siblings are in the same cohort and same school as her younger sibling. In this case there could be a direct effect of the older sibling's school mates on the younger sibling through their younger siblings. However, because we control for the younger sibling's school-by-cohort-by-subject fixed effects, any link to the older sibling's school mates through the younger siblings' school mates is broken.

Exploring additional instruments

Next we check the validity of our instrument further by using additional instruments, which allows us to test the over-identifying restrictions. We consider as first additional instrument the proportion of the older sibling's school mates that had a particular subject as their best subject. This may reflect the selection of similarly talented students into the same school or the presence of better teachers in a specific subject within a school. As we can see in the first row of the bottom panel of Table 6, the F-test of the excluded instruments is very high, indicating that the instruments are relevant, and the estimated sibling spillover effect remains the same as before. The Hansen's J test shows that the null that the instruments are exogenous cannot be rejected.

We consider as our second additional instrumental variable the fraction of the older sibling's new peers that were in the bottom 5-th percentile of the subject ability distribution at the end of the primary school. This variable is identical to the one used in Lavy et al.

(2012) to estimate school peer effects. We show the results of our IV estimates using both our original instrument and the fraction of bad peers of the older sibling as instruments in the second row of the bottom panel of Table 6. As we can see, this does not change the results and the Hansen's J test suggests that our instruments are valid.

5.3 Heterogeneous spillover effect and possible mechanisms

In this section we perform sub-group analysis to explore the heterogeneity of the results and to assess what we can learn about possible mechanisms that may drive the sibling spillover effects (imitation, productivity spillovers and information transmission).

We begin by estimating spillover effects from the older to the younger sibling by sex composition and age gap between the siblings (measured in academic years) using fixed effect estimation.[20] Results are shown in the second and third panels of Table 7 whereas in the first panel we report for comparison the homogenous spillover effects obtained using our preferred estimates, i.e. the fixed effect estimations without and with instrumental variable. We might expect siblings who are of the same sex or closer in age to interact more and feel closer to each other and therefore to be more likely to engage in imitation, direct help/teaching or information sharing. Indeed we find that the sibling spillover effect is substantially higher for siblings of the same gender (brother and sister pairs) than for mixed gender siblings, and somewhat larger for siblings who are more closely spaced. However, this analysis does not allow us the discriminate between the different candidate mechanisms.

To assess the possible role of productivity spillovers we split the sample by the older sibling's attainment. Productivity spillovers are produced through learning of the younger sibling from their older sibling, for example by spending time together in doing formative activities, by being taught or by receiving help with their homework. This type of spillover should arguably be larger when the older sibling performs well at school as this will affect the quality of the interaction. In the fourth panel of Table 7 we report the sibling peer effect separately by the position of the older sibling in the distribution of school test scores at age 16 (Key Stage 4). More precisely, we report the sibling spillover effect for older siblings with average school test score across the three subjects (English, Science and Maths) in the

[20] The heterogeneous sibling spillover effects are estimated by interacting the older sibling's subject-specific test score with dummy variables for different subgroups.

bottom 5th percentile, between the 5th and 95th percentiles, and in the top 5th percentile of the distribution.[21] The Table shows that the sibling peer effect for older siblings who are in the top 5th percentile of the distribution is almost three times larger than the peer effect observed for older siblings in the bottom 5th percentile. This seems to suggests that at least part of the sibling spillover we estimate in this paper is caused by productivity spillovers, particularly through teaching and help with homework provided by older siblings.

Another possible mechanism explaining the sibling spillover effect is information transmission. Some of the information transmitted from the older to the younger sibling may include information on subject-specific exam preparation, homework requirements, teachers and learning techniques that is particular to the school attended. Therefore we might expect spillover effects to be larger for siblings going to the same school than for siblings at different schools if information transmission is an important channel for sibling spillovers. In the fifth panel of Table 7 we report the sibling peer effect estimated separately for siblings going and not going to the same school. We find that siblings attending the same school have larger spillover effects which could suggest that part of the spillover effect is caused by information transmission, which is likely to be more effective for children going to the same school whatever the older sibling's school achievements are. Clearly the subsample of siblings attending different secondary schools is not a random sample, therefore we might be concerned that the lower spillover effect observed for this subample could be in part caused by endogenous selection into schools. Nevertheless, because our model controls for the subject specific test scores obtained at the end of primary school, any unobserved subject-specific and cohort-specific school characteristics and any subject-invariant child characteristics, we think that the larger peer effect for siblings going to the same school is unlikely to be exclusively caused by endogenous sample selection and at least in part related to information transmission.

To further assess the relevance of imitation, productivity and information spillovers, we next examine effects by family background and the older sibling's position in the age 16 exam score distribution. We expect that children from disadvantaged families have parents who are less likely to help them in their learning and to transmit subject-specific information and therefore sibling interactions to play a larger role. This may vary by older sibling's school

[21] We choose these bands following Lavy et al., 2012. Notice that we do not have an issue of regression to the mean because the sibling spillover effect is estimated by regressing the relative advantage of the younger sibling in a subject-specific test score on the corresponding older sibling's relative advantage, i.e. considering the difference between test scores in two different pairs of subjects.

performance. We measure family disadvantage in three different ways, by deprivation of neighborhood of residence,[22] eligibility for free school meals and by whether the language spoken at home is English. Neighborhood deprivation captures income deprivation of the area while free school meal eligibility indicates low income in the student's household. Families who do not speak English at home are not necessarily income deprived, but they likely lack knowledge of the English education system as the parents in such families will in most cases not have been raised and educated in England.

Table 8 shows results for all children, and for children by neighborhood deprivation, free school meal status and language spoken at home in separate panels. Within each panel the first row gives results for all children. We see that the sibling spillover effect is larger for children who are not speaking English at home than it is for those who do, but it is lower for children who live in deprived areas or are eligible for free school meals than for children from more affluent backgrounds. When focusing on older siblings who are in the bottom 5th percentile of the attainment distribution, all three disadvantaged groups of siblings have a stronger sibling spillover effect than the corresponding advantaged groups. Conversely, the spillover from a high achieving older sibling (in the top 5th percentile) is lower for children on free school meals and living in deprived neighborhoods. This seems to suggest that imitation of bad behavior and performance is more prevalent in disadvantaged families, whereas teaching and help by siblings is less common (except in families that do not speak English at home). Given that disadvantaged children are more likely to have an older sibling who is not performing well in school and to be more exposed to other examples of bad performing peers in school and in the neighborhood than affluent children, this asymmetry seems to exacerbate performance gaps between students by background. Students with English as additional language seem to be badly affected by poorly performing older siblings also, but results for the middle and to of the older sibling's attainment distribution indicate that productivity spillover are at work for them.

Taken together, the results on heterogenous spillovers by groups seem to suggest that all three mechanisms of imitation, productivity spillovers and information transmission play a role in driving the sibling spillover effect with perhaps a more dominant role of productivity

[22] Deprivation is measured using the Income Deprivation Affecting Children Index at the Lower Level Super Output Area, which is a sub-domain of the English Indices of Deprivation. We divide children's neighborhoods into the most, middle and least deprived tertiles.

spillovers for children whose older siblings are at the top of the achievement distribution and for this reason probably more effective teachers for their younger siblings. A positive and encouraging result is that the sibling peer effect is higher for children whose older siblings are performing well rather than badly in school. Nevertheless, this positive result is less pronounced among children who are economically disadvantaged. We observe a larger role played by positive productivity spillovers and information transmission among children who are economically more advantaged and a larger role of imitation of the older sibling's bad school performance among disadvantaged children. This seems to suggest the need for more positive role models for disadvantaged children.

6 Conclusions

In this paper we provide empirical evidence of sibling spillover effects in school achievement based on administrative data of 220 thousand siblings taking their end-of-compulsory schooling (age 16) exams in a four year time-window. We measure school achievement using test scores obtained in national exams in England in the compulsory subjects English, Mathematics and Science. We find strong evidence of direct sibling spillover effects in school achievement. An increase in the test scores of an older sibling of one standard deviation leads to an increase in the corresponding test score of the younger sibling of about 10% of a standard deviation. In terms of one grade improvement (e.g. from a grade B to an A) this effect is equivalent to the effect of increasing school expenditure per pupil by about £1,000 (see Nicoletti and Rabe 2012). As expected, we find no spillover effect going from younger to older siblings.

We extend the literature on the effect of social networks by focusing on the role of interactions between siblings rather than between school peers in producing cognitive ability. Compared to results found for school peer effects, the sibling spillover we estimate is larger than school peer effects and similar to school friend effects. Our paper also adds to the economic literature on child development by highlighting the direct role of sibling interactions, whereas the previous literature has mainly focused on parent-child interactions.

Our main methodological contribution is to propose a new strategy to identify sibling spillover effects in education which can be universally applied because it does not rely on context-specific instruments. Moreover it does not rely on the introduction of policy reforms where parents may reallocate resources between siblings in reaction to the policy, confounding the direct spillover effect. Our main concern when estimating the spillover effect is unobserved heterogeneity, in particular potential unobserved family and school investments that are shared by siblings and that can cause a spurious association between siblings. We use within-pupil between-subject estimation to control for child, school and family characteristics that are subject-invariant. Furthermore, we control for subject-specific school characteristics by applying school-by-cohort-by-subject fixed effects estimation.

We account for subject-specific skills acquired from parents through inheritance or investments and shared by siblings by using an instrumental variable strategy. We instrument the older sibling's test scores using the average prior test scores of her new school mates encountered for the first time in secondary school, exploiting idiosyncratic changes in average peer quality across schools and/or cohorts. We make use of the fact that the older sibling's test scores can be affected directly by her school mates' results, whereas we assume there is no direct effect of the older sibling's school mates on the younger sibling. We present checks testing this assumption, and the results lend credibility to the causal interpretation of our results. Our instrument is very relevant; nevertheless the results suggest that after applying our multiple fixed effect strategy there is no residual bias caused by subject-specific family investments.

The large sample size available in our data allows us to perform subgroup analysis and to explore potential mechanisms behind the sibling spillover effect. We find evidence of a larger sibling spillover for closely spaced and same gender siblings who we might expect to have closer interactions than different gender and widely spaced siblings. Siblings going to the same school have larger spillover effects than siblings going to different schools, which may mean that transmission of school-specific information is one channel through which spillovers are created. We find substantially larger sibling spillovers when focusing on older siblings who are high achievers while the effect of a badly performing older sibling is smaller on average. This seems to suggest that older sibling are effective teachers for their younger siblings especially when they perform well in school. Nevertheless, the positive effect of high achieving older siblings is reduced for children from economically disadvantaged backgrounds whereas the effect of badly performing older siblings is amplified in these families.

Taken together, our paper has important implications for policy that seeks to narrow the attainment gaps between children from different socio-economic backgrounds. Our results indicate that we need to be more concerned about older siblings being negative role models for disadvantaged children who are likely less exposed to positive role models in their school and in their neighborhood. This suggests that investments into students from deprived families can have considerable externalities through their benefits on younger siblings.

References

Adermon A. (2013). “Sibling Spillover in Education: Causal Estimates from a Natural Experiment.” PhD Dissertation, *Uppsala University*.

Akabyashi, H. (2006). “An equilibrium model of child maltreatment.” *Journal of Economic Dynamics and Control*, 30, 6, 993-1025.

Almond D. , Mazumder B. (2013). “Fetal origins and parental responses”, *Annual Review of Economics*, 5, 37-56.

Altonji, J.G., Cattan, S. and I. Ware (2016). “Identifying Sibling Influence on Teenage Substance Use”, *Journal of Human Resources*, forthcoming.

Angrist J. (2014), “The Perils of Peer Effects”, *Labour Economics* 30: 98-108.

Aparicio-Fenoll, A., and V. Oppedisano (2016). ”Should I stay or should I go? Sibling effects in household Formation.” *Review of Economics of the Household*, forthcoming.

Azmitia M. and J. Hesser (1993). “Why siblings are important agents of cognitive development: A comparison of siblings and peers.” *Child Development*, 64, 2, 430-444.

Barr R. and H. Hayne (2003). “It’s not what you know, it’s who you know: older siblings facilitate imitation during infancy.” *International Journal of Early Years Education*, 11, 7-21.

Becker, G.S. (1981), A Treatise on the Family. Cambridge, MA: Harvard University Press.

Becker, G.S. and N. Thomes (1976). “Child endowments and the quantity and quality of children.” *Journal of Political Economy*, 84 (2), S143-S162.

Behrman, J., R. Pollak, and P. Taubman (1982). “Parental Preferences and Provision for Progeny.” *Journal of Political Economy*, 90(1), 52-73.

Behrman, J.R., and P. Taubman (1986). “Birth order, schooling, and earnings”. *Journal of Labor Economics*, 4 (3), S121-S150.

Behrman, J.R., Rosenzweig, M.R. and P. Taubman (1994).“Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment.” *The Journal of Political Economy*, 102, 6, 1131-1174.

Bingley, P., and L. Cappellari (2013). “Correlations of Brothers Earnings and Intergenerational Transmission.” Working Paper Series of the Università Cattolica del Sacro Cuore, Dipartimento di Economia e Finanza, Working Paper 6, November 2013.

Björklund, A., Raaum, O. and E. Österbacka (2002). “Brother correlations in earnings in Denmark, Finland, Norway and Sweden compared to the United States.” *Journal of Population Economics*, 15 (4), 757-772.

Björklund, A., Jäntti, M. and M. Lindquist (2009). “Family background and income during the rise of the welfare state: trends in brother correlations for Swedish men born 1932-1968.” *Journal of Public Economics*, 93 (5-6), 671-680.

Björklund, A., Eriksson, K.H. and M. Jäntti (2010). "IQ and family background: Are associations strong or weak?" *The B.E. Journal of Economic Analysis and Policy*, (Contributions), 10(1), Article 2.

Björklund, A. and K.G. Salvanes (2011). "Education and Family Background: Mechanisms and Policies." in E.A. Hanushek, S. Machin, and L. Woessmann (Eds.) *Handbook of the Economics of Education*, , Amsterdam: North Holland, vol. 3, chapter 7, 201-247.

Black, S., P. Devereux and K.G. Salvanes (2005). "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education.", *Quarterly Journal of Economics*, 120, 2, 669-700.

Bramoullé, Y., H. Djebbari and B. Fortin (2009)."Identification of peer effects through social networks", *Journal of Econometrics*, 150, 41-55.

Breining, S. (2014). "The presense of ADHD: Spillovers between siblings, *Economics Letters*, 124, 469-473.

Breining, S., Daysal, N.M., Simonsen, M., Trandafir, M. (2015). "Spillover Effects of Early-Life Medical Interventions" IZA Discussion Paper No. 9086.

Brock W., and S.N. Durlauf (2001). Interactions-Based Models, in James J Heckman and Edward Learner (Eds.), *Handbook of Econometrics*, Volume 5, 3297-3380, Amsterdam Elsevier Science, North-Holland.

Butcher, K.F., and A. Case (1994). "The effect of sibling sex composition on women's education and earnings." *Quarterly Journal of Economics*, 109, 3, 531-563.

Calvó-Armengolo, A., Patacchini, E., Zenou, Y. (2009). "Peer Effects and Social Networks in Education." *Review of Economic Studies*, 76, 1239-1267.

Cicirelli V. G. (1972). "The effect of sibling relationships on concept learning of young children taught by child teachers." *Child Development*, 43, 1, 282-287.

Clotfelter, C.T., H.F. Ladd and J.L. Vigdor (2010), "Teacher Credentials and Student Achievement in High School. A Cross-Subject Analysis with Student Fixed Effects." *Journal of Human Resources*, 45(3): 655-681.

Corcoran, M., Gordon, R., Laren, D., and G. Solon (1990), "Effects of Family and Community Background on Economic Status." *The American Economic Review*, 80, 2.

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

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

Dahl, B.G., Løken K.V., and M. Mogstad (2014)."Peer Effects in Progam Participation". *American Economic Review*, 104(7): 2049-2074..

Dee, T.S. (2005), "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?", *American Economic Review Papers and Proceedings* 95(2), 158-165.

Dee, T.S. (2007), “Teachers and the Gender Gaps in Student Achievement.” *Journal of Human Resources* 42 (3), 528-554.

De Giorgi, G., M. Pelizzari and S. Redaelli (2010), “Identification of social interactions through partially overlapping peer groups”, *American Economic Journal: Applied Economics*, 2, 241-275.

De Giorgi G. , Fredriksen A., and L. Pistaferri (2015), ”Consumption Network Effects”, unpublished manuscript, Stanford University.

Del Bello, C.L., Patacchini, E. and Zenou, Y., 2015. “Neighborhood Effects in Education.” Unpublished manuscript.

Del Bono, E., Ermisch, J. and M. Francesconi (2012). “Intrafamily Resource Allocations: A Dynamic Structural Model of Birth Weight.” *Journal of Labor Economics*, 30, 3, 657-706.

Del Boca, D., Monfardini C. and Nicoletti C. (2012) Childrens and parents time-use choices and cognitive development during adolescence, IZA Discussion Paper No. 6868.

Dunn J. (1983). “Sibling relationships in early childhood.” *Child Development*, 54, 4, 787-811.

Epple, D. and Romano, R., (2011). “Peer effects in education: A survey of the theory and evidence.” Handbook of social economics, 1(11), pp.1053-1163.

Ewin Smith T. (1993). “Growth in academic achievement and teaching younger siblings.” *Social Psychology Quarterly*, 56, 1, 77-85.

Fletcher, J., Hair, N.L., Wolfe, B.L. (2012). “Am I My Brother’s Keeper? Sibling Spillover Effects: The Case of Developmental Disabilities and Externalizing Behavior.” NBER Working Paper 18279.

Gibbons, S., and S. Telhaj (2015). “Peer Effects: Evidence from Secondary School Transition in England”, Oxford Bulletin of Economics and Statistics. DOI: 10.1111/obes.12095.

Gibbons, S., O. Silva and F. Weinhardt (2013). “Everybody needs good neighbours? Evidence from students’ outcomes in England”, *Economic Journal*, 123(571): 831-874.

Glaeser, J., A. Scheinkman and B.I. Sacerdote (2003). “The Social Multiplier”, *Journal of the European Economic Association*, 1, 2/3, Papers and Proceedings of the Seventeenth Annual Congress of the European Economic Association, 345-353.

Gregory E. and A. Williams (2001). “Siblings bridging literacies in multilingual contexts.” *Journal of Research in Reading*, 24, 3, 248-265.

Hanushek E.A. and L. Woessmann (2011). “The Economics of International Differences in Educational Achievement.” In Eric A. Hanushek, Stephen Machin and Ludger Woessmann (Eds.), *Handbook of the Economics of Education*, Amsterdam: North Holland, vol. 3, chapter 2, 89-200.

Hanushek, E., Kain, J., Markman, J. and S. Rivkin. (2003). “Does peer ability affect student achievement?”, *Journal of Applied Econometrics*, 18, 527544.

Heckman J.J. and S. Mosso (2014). "The Economics of Human Development and Social Mobility", NBER working paper No. w19925.

Hobbs G., and A. Vignoles (2010), "Is free school meal 'eligibility' a good proxy for family income?" *British Educational Research bJournal*, 36(4): 673-690.

Hotz, J., Hao, L. and G. Z. Jin (2008). "Games parents and adolescents play: Risky behaviors, parental reputation, and strategic transfers." *Economic Journal*, 118, 528, 515-555.

Hoxby, C. (2000). "Peer effects in the classroom: Learning from gender and race variation." National Bureau of Economic Research Working Paper No. w7867z.

Joensen, J.S. and H.S. Nielsen (2015) "Peer Effects in Math and Science". Available at SSRN: http://ssrn.com/abstract=2548702 or http://dx.doi.org/10.2139/ssrn.2548702.

Kelejian, H.H., and I.R., Prucha (1998). "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances." *Journal of Real Estate Finance and Economics*, 17, 99 121.

Kessler, D. (1991). "Birth order, family size, and achievement: family structure, and wage determination." *Journal of Labor Economics*, 9, 4, 413-426

Kuziemko, I. (2006). "Is Having Babies Contagious? Fertility Peer Effects Between Adult Siblings.", Harvard University.

Lavy, V., Silva, O. and F. Weinhardt (2012). "The good, the bad, and the average: evidence on ability peer effects in schools." *Journal of Labor Economics*, 30 (2), 367-414.

Lee L.F. (2003). "Best Spatial Two-Stage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances." *Econometric Reviews*, 22, 307-335.

Lin, X. (2010)."Identifying Peer Effects in Student Academic Achievement by Spatial Autoregressive Models with Group Unobservables." *Journal of Labour Economics*, 28,4, 825-860.

Lindahl, L. (2011)."A comparison of family and neighbourhood effects on grades, test scores, educational attainment and income - evidence from Sweden." *Journal of Economic Inequality*, 9 (2), 207-226.

Mazumder, B. (2008). "Sibling similarities and economic inequality in the U.S." *Journal of Population Economics*, 21 (3), 685-701.

Manski, C.F. (1993). "Identification of endogenous social effects: The reflection problem." *Review of Economic Studies*, 60:531-42.

Manski C.F. (2000). "Economic Analysis of Social Interactions." *Journal of Economic Perspectives*, 14, 115-136.

Moffitt, R. (2001). "Policy Interventions, Low-level Equilibria, and Social Interactions." in Steven H. Durlauf and H. Peyton Young (Eds.) *Social Dynamics*, 45-82. Cambridge: MIT Press.

Monstad, K., Propper, C., and K.G. Salvanes (2011), "Is Teenage Motherhood Contagious? Evidence from a Natural Experiment." CMPO, Working Paper 11/262.

Nicoletti, C. and B. Rabe (2012), The effect of school resources on test scores in England, Discussion Paper in Economics, no. 12-19, DERS, University of York and ISER working paper series, 2012-13, University of Essex.

Nicoletti, C. and B. Rabe (2013), "Inequality in Pupils' Test Scores: How Much do Family, Sibling Type and Neighbourhood Matter?" *Economica*, 80(318), 197-218.

Nicoletti, C., Salvanes, K.G and E. Tominey (2016), "The Family Peer Effect on Mothers' Labour Supply." Discussion Papers in Economics No 16/04, University of York.

Oettinger, G. (2000). "Sibling Similarity in High School Graduation Outcomes: Causal Interdependency or Unobserved Heterogeneity?", *Southern Economic Journal*, 66 (3), 631-648.

Patacchini, E., Rainone, E. and Y. Zenou (2011), "Dynamic Aspects of Teenage Friendships and Educational Attainment." CEPR Discussion Paper No. DP8223. Available at SSRN: http://ssrn.com/abstract=1758432.

Qureshi, J.A. (2015a). "Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital", unpublished masucript, $https://sites.google.com/site/javaeriaqureshi/qureshi_{s}is$
0 (last accessed 14.2.2016).

Qureshi, J.A. (2015b). "The Influence of Siblings on Academic Achievement", unpublished manuscript, $http://javaeria.people.uic.edu/Qureshi_{s}iblingeffects.pdf$ (last accessed 14.2.2016).

Raaum, O., Salvanes, K.G. and E. Sørensen (2006). "The neighborhood is not what it used to be." *Economic Journal*, 116 (508), 200-222.

Rammohan, A. and E.P. Robertson (2011). "Human capital, kinship, and gender inequality." *Oxford Economic Papers*, 64, 3, 417-438.

Sacerdote B. (2011). "Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" in Eric A. Hanushek, Stephen Machin and Ludger Woessmann (Eds.), *Handbook of the Economics of Education*, vol. 3, chapter 4, 249-277.

Slater H., Davies N.M., and S. Burgess (2010), "Do Teachers Matter? Measuring the Variation in Teacher Effectiveness in England", *Oxford Bulletin of Economics and Statistics*, 74 (5), 629-645.

Solon, G. (1999)."Intergenerational Mobility in the Labor Market." in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3A, 1761-1800. Amsterdam: North Holland.

Solon, G., Page, M.E. and G.J. Duncan (2000). "Correlations between neighboring children in their subsequent educational attainment." *Review of Economics and Statistics*, 82, 3, 383-392.

Todd P. and K.I. Wolpin (2003), "On the Specification and Estimation of the Production Function for Cognitive Achievement", *Economic Journal*, 113, 485, F3-F33.

Table 1: Previous papers on sibling spillover effects

Authors	Outcome	Method	Data	Effect from older sibling	Effect from younger sibling
Oettinger (2000)	High school graduation	2SLS, using child-specific variables as IV	NLSY79	0.188-0.240+	0.043-0.082
Monstad, Propper, Salvanes (2011)	Teenage motherhood	2SLS, using a schooling reform as IV	Register data from Norway	0.242**	
Qureshi (2015a)	Years of schooling	2SLS, using distance from closest school	Learning and Educational Achievement in Punjab Schools	0.420**	0.163
Adermon (2013)	Years of schooling	2SLS, using a school reform as IV	Register data from Sweden	not sig	not sig
Aparicio-Fennoll and Oppedisano (2016)	Leaving parental home	2SLS, using rental subsidy eligibility as IV	Spanish Survey on Income and Living Conditions	-0.493+ negative	
Dahl, Loken Mogstad (2014)	Paternity leave take-up	2SLS, using a father's leave reform as IV	Register data from Norway	0.153* older or younger	
Joensen and Nielsen (2015)	Advanced Maths and Science choice	2SLS, using a high school reform as IV	Register data from Denmark	2-3* percentage points	

Notes: + $p < .10$, * $p < .05$, ** $p < .01$. 2SLS=two-stage least squares, IV=instrumental variable

Table 2: Descriptive statistics

Unstandardized test scores	Older		Younger	
	mean	std dev	mean	std dev
Key Stage 2 English score (age 11)	27.1	4.5	26.4	5.0
Key Stage 2 Science score (age 11)	29.0	4.1	28.6	4.8
Key Stage 2 Maths score (age 11)	27.5	4.9	27.3	5.3
Key Stage 4 English score (age 16)	40.6	9.1	40.2	8.9
Key Stage 4 Science score (age 16)	40.1	10.4	40.4	9.8
Key Stage 4 Maths score (age 16)	39.7	10.6	39.5	10.4
Sibling characteristics	%			
Same school	84.9			
Brothers	25.8			
Older brother, younger sister	24.9			
Older sister, younger brother	24.2			
Sisters	25.1			
Age gap 1 year	29.4			
Age gap 2 years	49.4			
Age gap 3 years	21.2			
2 children in family	59.4			
3+ children in family	40.6			
Urban neighbourhood	78.3			
Free School Meal eligible	10.5			
English additional language	8.1			
No. of observations pooled across subjects	621,540			
No. of sibling pairs	207,180			
No. of schools	2,948			

Notes: National Pupil Database, 2007-2010.

Table 3: Identifying variation in test scores and instrumental variable

	mean	std. dev.
Younger sibling's test scores at 16		
Total variation	0.090	0.899
Variation net of child fixed effect	0.000	0.357
Variation net of child and school-cohort-subject fixed effects	0.000	0.330
Instrumental variable: KS2 percentiles		
Total variation	49.07	9.352
Variation net of child fixed effect	0.000	2.340
Variation net of child and school-cohort-subject fixed effects	0.000	1.709
No. of observations	621,540	

Notes: National Pupil Database, 2007-2010. *The instrumental variable is the average of the subject-specific Key Stage 2 test score percentiles across the older sibling's new school peers, excluding the older sibling.

Table 4: Sibling correlations in test scores

	(1) Raw correlation	(2) Correlation value added	(3) Correlation value added Child FE	(4) Correlation value added Child-School-Coh-Subj FE
Corr.	0.478**	0.285**	0.136**	0.102**
	(0.001)	(0.002)	(0.001)	(0.001)
Observations	621,540			

Notes: + p < .10, * p < .05, ** p < .01. National Pupil Database, 2007-2010. Column (3) includes child fixed effects, column (4) Child-by-school-by-cohort-by-subject fixed effects. Standard errors clustered at school-cohort-subject level in parentheses. Pooled sample, pooling the observations for Mathematics, English and Science. Value added model in column (2) and subsequent models in columns (3) and (4) control for younger siblings' same-subject and cross-subject age 11 test scores and subject-by-gender dummies.

Table 5: Sibling spillover effect: Main results

	(1) From older to younger Child-School-Coh-Subj FE with IV	(2) From younger to older Child-School-Coh-Subj FE with IV
γ	0.098**	-0.123
	(0.033)	(0.145)
F-test first stage	485.6	30.73
Endogeneity test	0.017	2.548
p-value	(0.896)	(0.110)
Observations	621,255	629,925

Notes: + p < .10, * p < .05, ** p < .01. National Pupil Database, 2007-2010. IV regression with child-by-school-by-cohort-by-subject fixed effects. Dependent variables are standardised Key Stage 4 scores in English, Science and Maths. Value added models control for same-subject and cross-subject age 11 test scores and subject-by-gender dummies. Standard errors clustered at school-cohort-subject level in parentheses. Pooled sample, pooling the observations for Mathematics, English and Science. The instrument is the average Key Stage 2 attainment percentile of the older sibling's new peers in secondary school. The F-test is the Angrist-Pischke multivariate F-test of excluded instruments in the first stage. The endogeneity test is the robust Durbin-Wu-Hausman test.

Table 6: Robustness checks

(1) Sibling spillover effect	(2) F-test	(3) Endogeneity test	(4) Hansen's J test	(5) Number of observations
Benchmark estimation				
0.098**	485.6	0.0172		621,255
(0.033)		(0.896)		
Excluding older sibling's school mates living in the same neighbourhood				
0.095**	480.0	0.0468		621,255
(0.033)		(0.829)		
Excluding older sibling's school mates living in the same area				
0.100**	478.1	0.00702		621,255
(0.034)		(0.933)		
Using additional instruments				
1. New peers' KS2 percentiles and best KS2 subject				
0.089**	259.1	0.231	3.635	621,255
(0.032)		(0.631)	(0.057)	
2. New peers' KS2 percentiles and percentage of KS2 bottom 5% pupils				
0.088**	241.5	0.174	1.399	621,255
(0.033)		(0.677)	(0.237)	

Notes: + p < .10, * p < .05, ** p < .01. National Pupil Database, 2007-2010. Sibling spillover effects from the older to the younger sibling using instrumental variable estimation child and school-by-cohort-by-subject fixed effects. Standard errors clustered at school-cohort-subject level and p-values in parentheses. Pooled sample, pooling the observations for Mathematics, English and Science. The instrument in the benchmark estimation is the average Key Stage 2 attainment percentile of the older sibling's new peers in secondary school. The F-test is the Angrist-Pischke multivariate F-test of excluded instruments in the first stage. The endogeneity test is the robust Durbin-Wu-Hausman test. Neighborhood refers to the Lower Level Super Output Area, Area to the Middle Layer Output Area of residence. Additional instruments are the proportion of new peers that had English, Science or Maths as their best subject at the end of primary school and that was in the bottom 5% of pupils in Key Stage 2 respectively.

Table 7: Heterogenous sibling spillover effects

	(1)	(2)	(3)	(4)
Benchmark	IV	FE		
Sibling spillover	0.098**	0.102**		
	(0.033)	(0.001)		
F-Test	485.6			
Sex combination:	brother→brother	brother→sister	sister→brother	sister→sister
Sibling spillover	0.122**	0.083**	0.093**	0.111**
	(0.002)	(0.002)	(0.003)	(0.003)
Age gap:	1 year	2 years	3 years	
Sibling spillover	0.108**	0.102**	0.096**	
	(0.002)	(0.002)	(0.003)	
Older sib KS4 results:	bot 5th percentile	5-95th percentile	top 5th percentile	
Sibling spillover	0.049**	0.110**	0.142**	
	(0.003)	(0.002)	(0.003)	
Sibling's school:	same	different		
Sibling spillover	0.107**	0.081**		
	(0.001)	(0.003)		
Observations	621,540			

Notes: + p < .10, * p < .05, ** p < .01. National Pupil Database, 2007-2010. Benchmark estimates reported in column (1) are IV estimations using child-by-school-by-cohort-by-subject fixed effects. All other results are from child-by-school-by-cohort-by-subject fixed effect estimation. Standard errors clustered at school-cohort-subject level in parentheses. Pooled sample, pooling the observations for Mathematics, English and Science. Each panel represents one estimation with interaction terms used to derive coefficients by sub-group. Age gap is measured in academic years. Older siblings attainment at Key Stage 4 grouped by whether the pupil was in the top (bottom) 5% of the national attainment distribution for his/her cohort or in the midle 90%.

Table 8: Sibling spillovers in disadvantaged families

	(1)	(2)	(3)
Benchmark	IV	FE	
Sibling spillover	0.098**	0.102**	
	(0.033)	(0.001)	
F-Test	485.6		
Neighborhood deprivation:	most deprived tertile	middle tertile	least deprived tertile
All	0.093**	0.106**	0.112**
	(0.002)	(0.002)	(0.002)
Older sibling bottom 5%	0.058**	0.038**	0.039**
	(0.004)	(0.005)	(0.007)
Older sibling middle 90%	0.104**	0.116**	0.110**
	(0.003)	(0.003)	(0.003)
Older sibling top 5%	0.125**	0.145**	0.149**
	(0.006)	(0.005)	(0.004)
Free School Meal status:	FSM eligible	not FSM eligible	
All	0.087**	0.105**	
	(0.004)	(0.001)	
Older sibling bottom 5%	0.061**	0.045**	
	(0.007)	(0.003)	
Older sibling middle 90%	0.097**	0.112**	
	(0.005)	(0.002)	
Older sibling top 5%	0.133**	0.143**	
	(0.014)	(0.003)	
Language at home:	not English	English	
All	0.130**	0.100**	
	(0.004)	(0.001)	
Older sibling bottom 5%	0.080**	0.047**	
	(0.012)	(0.003)	
Older sibling middle 90%	0.139**	0.107**	
	(0.005)	(0.002)	
Older sibling top 5%	0.144**	0.142**	
	(0.011)	(0.003)	
Observations	621,540		

Notes: + p < .10, * p < .05, ** p < .01. National Pupil Database, 2007-2010. Benchmark estimates reported in column (1) are IV estimations using child-by-school-by-cohort-by-subject fixed effects. All other results are from child-by-school-by-cohort-by-subject fixed effect estimation. Standard errors clustered at school-cohort-subject level in parentheses. Pooled sample, pooling the observations for Mathematics, English and Science. Each panel includes results from two regressions, one with interaction terms capturing family disadvantage, the other using disadvantage x older sibling attainment interaction terms. Older siblings attainment at Key Stage 4 grouped by whether the pupil was in the top (bottom) 5% of the national attainment distribution for his/her cohort or in the midle 90%.

Appendix Table A1: First stage results

	(1) Effect on older sibling's attainment	(2) Standard error	(3) F-test
Panel A			
New peers' average of the KS2 percentile	0.0074	(0.000)	485.65
Panel B			
Percentage new peers in top 5%	0.0057	(0.848)	
Percentage new peers in bottom 5%	-0.3435	(0.000)	26.71
Panel C			
Percentage new peers in top 5%	-0.4618	(0.000)	
New peers' average of the KS2 percentile	0.0100	(0.000)	
Percentage new peers in bottom 5%	0.1374	(0.007)	214.33
Number of observations	621,255		

Notes: + p < .10, * p < .05, ** p < .01. National Pupil Database, 2007-2010. Coefficients are first stage results from IV estimations using child-by-school-by-cohort-by-subject fixed effects. Standard errors clustered at school-cohort-subject level in parentheses. Pooled sample, pooling the observations for Mathematics, English and Science. Panels A, B and C report the estimates of three different first stage equations. Panel A shows the effect on the older sibling's attainment of average new peers' KS2 percentiles. This is the first stage results of our benchmark model reported in Table 5. Panel B shows first stage results when using the percentage of new peers that were in the top (bottom) 5% of the national attainment distribution for his/her cohort at the end of primary school; Panel B shows the first stage results using the top and bottom 5% as well as the average KS2 percentiles of new peers.