Sibling spillover effects in school test scores

Cheti Nicoletti

DERS, University of York and ISER, University of Essex

Birgitta Rabe

ISER, University of Essex

Preliminary Draft: March 19, 2014

Abstract

We provide the first empirical evidence on direct sibling spillover effects in cognitive abilities by using school administrative data in England with information on test scores at age 16 in Mathematics, English and Science. We use child fixed effects, school-by-cohort-by-subject fixed effects as well as instrumental variable estimation to take account of unobserved confounding factors and of the reflection problem. We exploit the fact that students interact both with their siblings and with their school mates and that school mates of a child's older sibling do not usually interact directly with the younger sibling. This allows us to use the average test scores of the older sibling's school mates to derive an instrumental variable. Our empirical results confirm that there is a small statistically significant sibling spillover effect going from older siblings to younger siblings but not vice versa.

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

JEL codes: I22, I24

Contact: Cheti Nicoletti: cheti.nicoletti@york.ac.uk; Birgitta Rabe: brabe@essex.ac.uk

We thank the Department for Education for making available, under terms and conditions, data from the National Pupil Database. 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 (award RES-518-28-001).

1 Introduction

While the recent economic literature recognizes the important role of interactions between parents and children for the child's development,[1] the role of sibling interactions is yet to be clearly established. Previous economic papers concentrating on siblings have focused on either the differential parental investments into siblings to reinforce or compensate for sibling differences in abilities,[2] or on the effects of birth order, number and gender composition of siblings on parental investments and child's later life outcomes.[3] It is only more recently that researchers have begun to look at the effect of interactions between siblings and there are some recent papers on spillover effects in education (see Oettinger 2000; Qureshi 2011; Joensen and Nielsen 2013; Adermon 2013). Nevertheless, none of these papers have focused on the effect of sibling interactions on educational achievements measured by school test scores.

In this paper we provide the first empirical evidence on the extent to which cognitive abilities of a child can be transmitted to his/her younger sibling through interactions between the two. More precisely we estimate the direct sibling spillover effect of a child's school test scores at age 16 on her younger sibling's test scores at the same age. Contrary to previous papers our identification strategy is not based on instrumental variables or policy reforms that are context specific, but it exploits the fact that pupils interact with their school mates as well as with their siblings and that children do not usually interact with their sibling's school mates. This implies that characteristics of the older sibling's school mates cannot affect directly the younger sibling and therefore can be used as instruments to explain the older siblings' test score once controlled for unobserved school and child characteristics.

We are interested in the spillover effect that is caused by the direct causal effect of a child on his/her younger sibling rather than the indirect effect that can be mediated by common background characteristics or by intra-household allocation of investments between siblings (see Behrman et al. 1982, Solon et al. 2000, Björklund and Salvanes 2011). Being able to identify a sibling spillover effect in test scores that is causal is important for at least

[1] See, for example, Akabyashi (2006), Hotz et al. (2008) and Coscosnati (2012).

[2] See Behrman et al. (1982), Ermisch and Francesconi (2000), Bernal (2008), and Rammohan and Robertson (2011)

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

two different reasons: (i) to assess whether the interactions between siblings is a mechanism through which intergenerational transmission of cognitive abilities may amplify(see Solon et al. 2000, Björklund and Salvanes 2011, Adermon 2013), (ii) to evaluate whether interventions targeted at reducing inequalities in school test results may have a multiplier effect through sibling interactions (see Manski 1993 and 2000; Glaeser et al. 2003).

Possible causal mechanisms of the direct spillover effect include (1) children's imitation of their older sibling's behaviour (role model) (Widmer 1997; Altonji et al. 2013; Adermon 2013); (2) children learning from their older siblings through playing and doing other activities together (Joensen and Nielsen 2013) or by receiving help with their home work (Qureshi 2011; Adermon 2013; Joensen and Nielsen 2013); (3) transmission of information, beliefs and attitudes from the older sibling to the younger, e.g. by providing advice to the younger sibling on educational choices (Joensen and Nielsen 2013); (4) differentiation in behavior between siblings to avoid direct competition which may lead them to specialize in different areas and can cause a negative spillover effect; (5) signalling effects from the older to the younger sibling on the returns to effort.

Simply regressing a child test score on the older sibling's corresponding test score would not produce a consistent estimation 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 they are exposed to. To clean the sibling association in test scores of these confounding factors and estimate the causal sibling spillover effect, we use school register data in England which provides information on tests scores at the end of compulsory schooling, at about age 16, in the subjects of Mathematics, English and Science for the full population of students in state schools.

By using the three test scores available for each pair of siblings at age 16, we are able to regress a child's test score on her older sibling's test score, controlling for individual fixed effects. This estimation is equivalent to the within-pupil between-subject estimation used among others by Dee (2005) and (2007), Clotfelter et al. (2010) and Slater et al. (2010). The two main gains of this individual fixed effect estimation are that it allows us to (i) control for the effect of any unobserved investments or characteristics of the younger child that are invariant across subjects and might confound the spillover effect because they are similar to the corresponding characteristics of his/her sibling, (ii) clean the sibling spillover effect of the impact of differential allocation of resources between siblings that are related to factors such as sibling differences in abilities, number of siblings in the family or sibling sex composition. The effect of intra-household allocation of resources are especially important in the context of child cognitive abilities because it has been found that parents invest differentially in two siblings in an attempt to either compensate or reinforce for differences in their abilities (see Behrman et al. 1982 and 1994).

We also want to consider unobserved school characteristics that are subject specific. If siblings receive similar subject-specific investments, for example from good Mathematics teachers, then the sibling spillover effect could be overestimated. This is likely to happen because siblings tend to go to the same school or to sort into similar schools. To take account of this and of other subject-specific school characteristics we consider school-by-cohort-by-subject fixed effects.[4] Notice that even after controlling for individual and school-by-cohort-by-subject fixed effects, we are still unable to control for potential subject-specific skills acquired from parents through subject-specific family investments and/or inheritance and shared by siblings. To take account of subject-specific skills that are transmitted from parents to their children in the period that goes from birth to age 11, we condition on the subject-specific test score of the child and her sibling at age 11. This is still not enough to identify a causal sibling spillover effect because there could be an intergenerational transmission of subject-specific skills between age 11 and 16, which is probably similar between siblings.

To tackle this last issue, we instrument the older sibling's test scores at age 16 using the average test scores of her school mates.[5] This identification strategy is in the spirit of the peer effect identification and estimation framework based on social networks (see Kelejan and Prucha 1998; Lee 2003; Bramoulle et al. 2009; De Giorgi et al. 2010). We make use of the fact that older sibling's test scores can be affected directly by her school mates' results, whereas we assume that there is no effect from the older sibling's school mates to the younger sibling if not indirectly through the older sibling. We test this assumption by performing a

[4]We say that two children belong to the same cohort if they take their age 16 exams in the same academic year.

[5]More precisely, in our model we control for school-by-cohort-by-subject fixed effects, and our instrument captures whether the older sibling's school-cohort mates were relatively better in a specific subject than the younger sibling's school-cohort mates. The variation in the instrument is caused by idiosyncratic changes in the average subject-specific test score across schools and/or across cohorts.

number of sensitivity checks on the data, for example by excluding school mates of the older sibling who live in the same neighbourhood and might therefore interact directly with the younger sibling. Note that we are not interested in estimating a causal school peer effect for the older sibling, as the validity of the instrument does not rely on this. In fact, the effect of the older sibling's school mates on the older sibling can be the consequence of a causal school peer effect but also of factors such as unobserved contextual variables, e.g. the quality of the teachers in the school-cohort of the older sibling, or of school composition effects.

Except for the use of instrumental variable estimation our strategy is similar to the one used by Lavy et al. (2012) to identify the effect of school peers on children's abilities. The main difference is that, because we are estimating sibling rather than school peer effects, our identification is more challenging and requires to take into consideration the confounding effect caused not only by similar schools characteristics of the two siblings but also caused by intergenerational transmission, differential parental investments to compensate or reinforce for differences in siblings' abilities, and subject-specific parental investments; but, contrary to Lavy et al. (2012), we are less concerned about the reflection problem.

Empirical researchers estimating a causal effect between individuals' outcomes are usually concerned with the reflection problem (see Manski 1993), i.e. the simultaneity of the individuals behaviours and the potential reverse causality. We cannot exclude the existence of spillover effects going from the younger to the older sibling,[6] 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. Nevertheless, if this was not enough to correct for the potential reflection issue, our instrumental variable estimation would be able to control for it because the average test scores of the the older sibling's school peers cannot be affected by the younger sibling and therefore there cannot be any reverse causality.

Using school test scores at age 16 we find that an increase of a standard deviation in a child's test score leads to a small increase in the corresponding test score observed for his/her younger sibling of about 2% of a standard deviation and this sibling spillover effect is strongly statistically significant. Interestingly, if the two siblings attended the same school,

[6] Ewin Smith (1990) and (1993) suggests that the cognitive abilities of older children might improve thanks to teaching younger siblings.

the spillover effect almost doubles (about 4%). This seems to suggest that there is a more effective transmission of abilities between siblings attending the same school, possibly because the older sibling has direct information relevant for succeeding in the shared school.

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, there have been a plethora of empirical papers analyzing sibling correlations in different outcomes (see Raaum et al. 2006; Mazumder 2008; Björklund et al. 2002; Björklund et al. 2009; Björklund et al. 2010; Lindahl 2011; Björklund and Salvanes 2011; and Nicoletti and Rabe 2013). However, only more recently have economists tried to identify the part of the sibling correlation or association which is explained by sibling interactions, i.e. the causal sibling spillover effect.

In Table 1 we summarize the results of previous papers on sibling spillover effects that have identified a causal spillover effect. The sibling outcomes analyzed by these papers are: high school graduation (Oettinger; 2000), years of schooling (Qureshi 2011; Adermon 2013), school subject choices (Joensen and Nielsen 2013), teenage motherhood (Monstad et al 2011), paternity leave take-up (Dahl et al. 2013) and maternal labour participation after childbirth (Nicoletti et al. 2014). In all papers, the reflection problem is dealt with by using instrumental variables that explain the outcome of one child but not of her sibling. Furthermore, by assuming that their instrumental variables are uncorrelated with potential omitted characteristics, at least after controlling for the observed explanatory variables, they can claim that their instrumental variable estimation corrects also for the bias that might be caused by correlated unobservables.

Oettinger (2000) uses the US National Longitudinal Survey of Youth 1979 (NLSY79) to estimate the sibling spillover effect in the probability of a child graduating from high school by age 19 and uses as instrumental variables background characteristics that are sibling specific, such as whether the family was intact during childhood and the unemployment rate at age 18. He does not find any statistically significant spillover effect from the younger to the older sibling but finds some significant effects from the older to the younger sibling (See Table 1). Unfortunately, it is difficult to believe that the instrumental variables used by Oettinger (2000) are uncorrelated with unobserved family and neighbourhood characteristics that can explain both siblings outcomes.

Qureshi (2011) focuses on years of schooling and estimates the spillover effect going from the oldest sister to younger brothers in rural Pakistan using the Learning and Educational Achievement in Punjab Schools (LEAPS) data. Exploiting the fact that there is a strong gender segregation of schools in Pakistan, she uses school distance to the closest girls' school for the oldest sister as instrument for her years of schooling. Even if the school distance is usually not random for children living in developed countries, this instrumental variable seems plausible in the context of rural Pakistan. Qureshi (2011) finds that a one year increase in the schooling of the oldest sister leads to almost half a year increase of schooling for her younger brothers and to statistically significant large effects on her younger sibling's literacy and school enrollment. These large effects are in part explained by the fact that, in developing countries, child care is not exclusively a parental responsibility and older sisters often have caring responsibilities for their younger siblings.

Monstad et al. (2011), Adermon (2013), Dahl et al. (2013), and Joensen and Nielsen (2013) all use administrative data and methods that are similar to the 'partial-population' identification approach (see Moffitt 2001) to estimate the sibling spillover effect on decisions about becoming a teenage mother, years of schooling, taking paternity leave, and choosing advanced Mathematics and Science as school subjects. All these papers use the introduction of policy reforms that changed the probability (cost) of adopting the given decision for a random (at least after conditioning on the explanatory variables) portion of siblings.

Monstad et al. (2011) consider a reform that raised the minimum school leaving age in Norway by two years, therefore decreasing the probability of girls to become a teenage mother. This Norwegian school reform was implemented by different municipalities in different years and the timing of implementation by different municipalities does not seem related to fertility, education level, income level or the size of the municipalities (see Monstad et al. 2008; Black et al. 2005 and 2008; and Aakvik et al. 2010). Monstad et al. (2011)

use the school reform dummy for the older sister, i.e. a dummy taking value 1 if the older sister was affected by the reform and zero otherwise, as instrumental variable to estimate the sibling spillover effect of teenage motherhood that goes from older to younger sisters. They assume that the school reform dummy for the older sister explains the older sister probability of becoming a teenage mother but does not explain directly the probability of the younger sister of becoming a teenage mother, at least after controlling for municipalities dummies, the younger sister school reform dummy and a set of family and individual characteristics. Having an older sister who became a teenage mother leads to a statistically significant increase in the probability of becoming a teenage mother of about 24.2 percentage points.

Adermon (2013) exploits the increase in the minimum school leaving age in Sweden. Similarly to Norway, an increase of two years in the school leaving age was introduced at different times in different municipalities. The timing of the implementation by different municipalities is not completely random; but, once controlling for birth cohort and municipality fixed effects and municipality-specific trends, Holmlund (2008) suggests that it is exogenous. Using the school reform dummy for the older sibling as instrument and controlling for municipality fixed effects and trends, Adermon (2013) does not find any significant sibling spillover effect on years of schooling.

Joensen and Nielsen (2013) 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. This reform was adopted only by some schools and there do not seem to be any systematic difference between schools that introduced this reform and schools that did not (for details on potential differences see Joensen and Nielesen 2013). Furthermore, to avoid any potential bias caused by an endogenous selection of students in schools that implemented the reform, Joensen and Nielesen (2013) consider only the first year of the implementation i.e. they consider only children who could not anticipate the reform at the time of school enrollment. 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.

Dahl et al. (2013) consider the introduction of paid paternity leave, which only fathers of children born after the first of April 1993 can take, and adopt for the first time a regression discontinuity design to estimate sibling spillover effects. Their estimation method is equivalent to a two-stage least square estimation with the discontinuity in eligibility for paternity leave around the first of April 1993 used as instrumental variable for the sibling decision to take paternity leave. Having a brother who took paternity leave seems to increase the probability to take paternity leave of about 15.3 percentage points and it is statistically significant at the 5% level.

Nicoletti et al. (2014) use Norwegian administrative data to evaluate the spillover effect of sisters in maternal labour participation one year after child birth using unrelated neighbours' labor market participation to instrument a sister's participation. Because mothers who have a sister living in a different neighborhood can be affected by the behavior and characteristics of other mothers living in this neighborhood only through the behaviour of her sister, Nicoletti et al. (2014) use the average characteristics of mothers living in the sister's neighborhoods as instrumental variables. Furthermore, to avoid the potential bias caused by unobserved neighborhood characteristics (especially if mothers and their sisters tend to sort into similar neighborhoods), they control for the mothers' neighborhood fixed effects. They estimate the sister spillover effect on the probability to work one year after child birth to be about 30 percentage points and significant at the 1% level.

Another two papers looking at spillover effect of siblings, which are worth consideration even if they do not use instrumental variables for their estimation are Kuziemko (2006) and Altonji et al. (2013). Both papers use panel data and estimate dynamic models to identify the sibling spillover effect. Kuziemko (2006) evaluates the effects of having a sibling who had a child in the last 12 months, 13-14 months and 25-36 months on the probability of having a child in the current month. She uses repeated observations of the dependent variable across time to control for unobserved individual and family fixed effects therefore exploiting the variation in fertility of siblings across time to identify the spillover effects. Using the Panel Study of Income Dynamics (PSID) and the NLSY79 in the US, she finds that there is a statistically significant effect on child birth of having a sibling who had a child either in the last 12 months or in between 13 and 25 months ago of about 15% and 17% respectively. Altonji et al. (2013) estimate the effect of having an older sibling smoking (drinking, and making use of and selling drugs) on his/her younger sibling using the National Longitudinal Survey of Youth 1997 (NLSY97). They consider a joint model for the two siblings and assume that the smoking behavior by the older sibling in year $(t-1)$ has an effect on the younger sibling's smoking in year t but not vice versa, and they find some significant sibling spillover effects on smoking and other substance use. The major concern with the approaches adopted by Kuziemko (2006) and Altonji et al. (2013) is that they may lead to a biased estimation if there are spillovers effects going from the younger to the older sibling or if there are some omitted variables.

While the above papers on sibling spillover effect aims at identifying the "direct" effect of interactions mechanisms between siblings such as the mechanisms of imitation, teaching and learning; previous economic papers have almost exclusively focused on the"indirect" spillover effects of siblings and in particular on the effect caused by intra-household resources allocation among children with differential abilities (e.g. Behrman et al. 1982; Rosenzweig and Schultz 1982; Rosenzweig and Wolpin 1988; Behrman et al . 1994; Ermisch and Francesconi 2000; Behrman and Rosenzweig 2004; Del Bono et al. 2012) and by the potential decrease in parental investment per child when the number of siblings increase (Becker and Thomes 1976; Kessler 1991; Black et a. 2005 and 2010; Angrist et al. 2010; Cáceres-Delpiano 2006; Lee 2008).

Some of the sibling spillover effects estimated in the above papers might reflect the combination of both "direct" effect of sibling interactions and "indirect" effect mediated by parental allocation of resources between siblings. This can happen if there are unobserved parental investments and the instrumental variables considered are not independent of these investments. For example, in the case of an instrumental variable given by a school reform the sibling spillover effect might capture the effect caused by the parental reallocation of resources between siblings in reaction to the introduction of a school reform. Reforms raising the school leaving age used by Monstad et al. (2013) and Adermon (2013) were implemented over long time-periods, and 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 equalise between siblings. In this case, the sibling spillover effects computed by Monstad et al. (2011) and by Adermon (2013) would be underestimated. This paper is free of this problem because our identification is based on an instrumental variable that exploits group membership rather than a reform that may allow parents to reallocate resources between siblings to attenuate the potential sibling differences caused by the reform.

Other strands of the economic literature closely related to our paper on sibling spillover effects in test scores are the literature on educational production and child development (see Todd and Wolpin 2003; Cunha and Heckman 2007; Cunha and Heckman 2008; Hanusheck and Woessmann 2011) and on school peer effects (see for a review Sacerdote 2011). These research strands have provided a theoretical framework to model the production of children's cognitive abilities taking account of family and school inputs and of the possible school peer effects, but they have so far ignored the potential effect of interactions between siblings. In this paper, we extend the recent work on education production models and school peer effects by Nicoletti and Rabe (2012) and Lavy et al. (2012), who both use school register data for England as in our application, and we provide detail on how to identify the causal sibling spillover effect in school test scores at age 16, i.e. the effect of sibling interactions on child development during adolescence.

The only other papers we are aware of that focus on the direct effect of siblings interactions on child 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 to talk the language spoken at school. Azmitia and Hesser (1993) compare sibling and peer influence on children cognitive development and find that older siblings are more effective in teaching their younger siblings than unrelated children of the same age.

3 Identification strategy

To identify the effect of sibling spillover on test score results at the end of compulsory school (at about age 16) we consider the following value added model[7]:

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

where $Y_{1,isqt,16}$ is the test score in subject q, measured at age 16 for the younger child of the sibling-pair i who attended school s and belong to the cohort t;[8] $Y_{1,isqt,11}$ is the corresponding

[7] See Todd and Wolpin 2003 for a definition.

[8] Two students belong to the same school cohort if they began school in the same year.

test score at age 11; $Y_{2,is'qt',16}$ is the test score at age 16 for the older sibling, who might have attended a different school s' and belongs to a different cohort t';[9] $Y_{2,is'qt',11}$ is the corresponding test score of the older sibling at age 11; $I^F_{1,it}$ 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; $X_{1,i}$ is a row vector of other child, household and school characteristics, which are not direct investments in children's cognitive skill but may affect it; μ_{sqt} is the unobserved school-by-cohort-by-subject investment effect; $\mu_{1,i}$ is the younger child subject-invariant unobserved endowment; α is the intercept; ρ measures the persistence in test scores between age 11 and 16; γ is our main parameter of interest which measures the spillover effect of older siblings; $\beta_{1,F}$ and $\beta_{1,S}$ are the productivity of family and school investments; and $\beta_{1,X}$ is a column vector with the effects of the remaining explanatory variables $X_{1,i}$. 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 two main issues: (i) unobserved correlated effects, i.e unobserved common characteristics of two siblings that may explain their similar test scores and (ii) the reflection problem.

We control for unobserved child specific endowments and characteristics that are invariant 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 $DevY_{1,isq,16} = Y_{1,isq,16} - \sum_{j=1}^{3} Y_{1,isj,16}/3$ and we apply an analogous transformation to all right hand side variables, leading to

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

This transformation eliminates from the equation all inputs that are invariant across subjects as well as $\mu_{1,i}$, i.e. the unobserved child specific and subject-invariant endowment, which captures unobserved cognitive abilities and other unobserved child capabilities such as non-cognitive abilities and health, which could be similar between siblings and therefore confound the sibling spillover effect. Notice that this transformation also eliminates possible spillover effects between siblings that are not causal but are the consequence of changes in the intra-household allocation of resources between siblings. Parents may re-allocate resources between

[9] We do not consider twins or siblings whose age gap is such that they begin school in the same year.

siblings because of differences in their abilities, for example investments in one child might decrease if her sibling develops disability or health issues, or because they have more children and therefore decrease the parental investment per child.

Nevertheless, the deviation from the mean across subjects is unable to eliminate unobserved characteristics which are subject-specific. We are in particular concerned about unobserved subject-specific abilities shared by the siblings because of similar family and school investments, which might favor one subject over another, or because of family inheritance of subject specific skills.

By using school-by-cohort-by-subject fixed effects we are able to control for μ_{sqt}, i.e. for unobserved subject-specific school investments and characteristics for the cohort t. In our sample a high percentage of siblings, 83.5%, attend the same secondary school, and even if two siblings attend two different schools they might still have similarities in subject-specific characteristics of their school, e.g. similar quality of teachers in Mathematics or similar composition of school peers in terms of subject-specific abilities. Controlling for school-by-cohort-by-subject fixed effects allows us to compute a sibling spillover effect which is net of the effect of such school similarities between siblings.

The issue of unobserved subject-specific family investments and subject-specific skill inheritance is more challenging. By controlling for the lagged test score, i.e. the test score in subject q at age 11, we at least control for potential effect of subject-specific family investments and subject-specific skill inheritance up to the age of 11. To control also for the effect of these unobserved subject-specific characteristics between age 11 and 16, we adopt an instrumental variable estimation. More precisely we instrument the subject-specific test score of the older sibling at age 16 using the average of $DevY_{js'qt',16}$ over the school-by-cohort peers of the older sibling, excluding the test score of the older sibling, which we call $MDevY_{2,s'qt',16}$. We assume that a child can be affected by the test scores of the school peers of her siblings only through her sibling. This is because a child does not usually interact with the school peers of her sibling. This assumption could be invalid in a situation where the sibling's school peers live in the same neighborhood so that some interactions between the child and his/her sibling's school peers are possible. To investigate this we perform robustness checks where we exclude from the computation of $MDevY_{2,s'qt',16}$ the older sibling's school peers who live in the same neighborhood, and we do this by adopting two different definitions of neighborhood. The assumption could also be invalid if there is interaction between a younger sibling and older siblings' peers at school. To investigate this we perform a robustness check where we estimate sibling spillover effects separately for siblings going to schools that offer post-16 schooling and for schools that do not. In schools that do not offer post-16 schooling the younger siblings in our sample will have had no school-based contact with older siblings' peers for 1-3 years before taking her exams, depending on the sibling age gap.

Notice that if siblings tend to sort into similar schools with similar subject-specific investments and characteristics, then $MDevY_{1,sqt,16}$ and $MDevY_{2,s'qt',16}$ could be correlated and confound the spillover effect, hence the importance to control for $MDevY_{1,sqt,16}$ in our equation. We actually control for $MDevY_{1,sqt,16}$ by using school-by-cohort-by-subject fixed effects. This implies that the sibling spillover effect is identified by idiosyncratic changes in $MDevY_{1,sqt,16}$ across cohorts and across schools.

The use of the instrumental variable estimation and the fact that the older sibling test scores at 16 are observed earlier in time than the younger sibling test scores allows us to address also any potential reflection issue, i.e. allows us to cancel any potential causal relationship that goes from the younger to the older sibling rather than vice versa.

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.

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 pupils in England. It combines pupil level attainment data with pupil 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. Pupils undergo 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

Our outcomes of interest are General Certificate of Secondary Education (GCSE) or equivalent vocational test results at the end of compulsory schooling, usually taken at age 16 (Key Stage 4). We focus on GCSEs 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 adult labour market. We consider Key Stage 4 results in the core subjects English, Mathematics and Science which are directly comparable to test results at the end of primary school. In Key Stage 4 pupils receive a grade for each GCSE course, where pass grades include A*, A, B, C, D, E, F, G. We use a scoring system developed by the Qualifications and Curriculum Authority to transform these grades into a continuous point score[10] which we refer to as the Key Stage 4 score.

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 the three core subjects of English,

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

Mathematics and Science. In the Key Stage 2 exams, pupils can usually attain a maximum of 36 points in each subject, but teachers will provide opportunities for very bright pupils to test to higher levels. All test scores are standardized to have a mean of zero and a standard deviation of one.

The NPD annual school census allows identification of a number of individual and family background variables. These include gender of the pupil, a binary variable coding ethnicity (White British, Black, Mixed, Indian, Pakistani/Bangladeshi, Chinese), whether or not the first language spoken at home is English, whether special educational needs have been identified for the child[11], whether or not a pupil is eligible for free school meals (FSM)[12] and the number of all siblings in the state school system in 2007. As we control for child fixed effects in all our models we do not use these variables as explanatory variables, but we use some of them for heterogeneity and sensitivity analysis.

Sibling definition

The NPD includes address data, released under special conditions, which allows us to match siblings in the data set. The first year that full address details were collected in the NPD across all pupil cohorts was 2007. Siblings are therefore defined as pupils in state schools aged 4-16 and living together at the same address in January 2007. Siblings that are not school-age, those in independent schools and those living at different addresses in January 2007 are excluded from our sibling definition. Step and half siblings are included if they live at the same address, and we are not able to distinguish them from biological siblings (see Nicoletti and Rabe 2013 for details).

Estimation sample

The main sample for our analysis includes all siblings pairs which we identify as children living together at the same address in 2007 and taking their Key Stage 4 exams in 2007, 2008, 2009 or 2010. When we have multiple pairs of siblings in this observation window we consider the two oldest ones to avoid any further multiplier spillover effects (what Dahl et al.

[11]These are pupils with learning difficulties, including behavioral and health conditions. Those that have been assessed by local education authorities receive a statement which is usually associated additional funding received by the school. There are also pupils identified by the schools as having special needs, but without statement.

[12]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 2007 for some shortcomings).

2013 call the snowball effect) with additional siblings interactions.[13] We remove pupils with duplicate data entries or with missing data on any of the background or school-level variables from the dataset. Moreover, we retain only pupils for whom we have non-missing test scores for all outcomes at both Key Stages 2 and 4 which leads to a reduction in sample size of 13.7%. 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 behavioural difficulties. The final sample contains 435,896 siblings (217,948 sibling pairs).

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 standardized test scores by subject.

5 Empirical Results

5.1 Main empirical results

We begin by reporting in Table 3 the correlations in siblings test scores which are a general measure of the importance of background shared between siblings on educational outcomes. In column (1) of Table 3 we show the raw correlation in tests scores (0.50) which is in line with previous papers (e.g. Nicoletti and Rabe 2013, Bjoerklund 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 two siblings 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 both siblings test scores at 11.[14] This sibling correlation is capturing the effect of shared family and environment characteristics which operates between ages 11 and

[13]After removing twins and siblings attending the same academic year, only 1.4% of siblings are higher-order siblings that we remove from the sample

[14]Since the test scores at ages 11 and 16 are standardized by subject to have mean 0 and variance 1, we can estimate the raw correlation in test scores by a simple regression of the test scores at 16 on the sibling's test score at 16 and the net correlation by estimating the value added model.

16. We can see that the net sibling correlation is 0.31. Finally, in column (3) we show the correlation estimated using the value added model with lagged test scores and controlling for the younger child fixed effects. This nets out the influence of all environment, family and child characteristics that are invariant across subjects, including the intra-household allocation of resources between siblings. This correlation (0.12) therefore comes closer to capturing a sibling causal interdependence, but there are likely other factors that are similar between siblings but cannot be attributed to a direct spillover effect. These factors include subject-specific school characteristics and subject-specific skills transmitted in the family.

In Table 4 we present our main estimates of the sibling spillover effect in school test scores from the older to the younger sibling at 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. For both estimations we consider the value added model (1) that control for past test scores obtained at the end of primary school. 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.[15] Our instrumental variable estimation is a two-stage least square (2SLS) estimation with fixed effects and the instrument we use is the subject-specific average test score for the school-cohort peers of the older sibling (younger sibling in column 2). Because in the equation (1) we control for both child fixed effects and school-by-cohort-by-subject fixed effects, the instrument captures whether the older (younger in column 2) sibling's school-cohort mates were relatively better in a specific subject than the younger sibling's school-cohort mates. The variation in the instrument is caused by idiosyncratic changes in the average subject-specific test score across across schools or within the same school but across different cohorts. These changes can occur because of changes in the quality of teaching in a specific subject (e.g. because of teacher turnover) or in the composition of the school-cohort mates in terms of subject-specific abilities across schools or within schools across cohorts.

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

The top panel of Table 4 shows the first stage results. We find that there is a strong relationship between our instrument and the older sibling's test scores. The coefficient is statistically significant at the 1% level and the F-statistics for the significance of the instrumental variable in the first stage is huge and does not leave any doubt on the validity of the instrument. Second stage results are displayed in the bottom panel of the Table. Looking first at the sibling spillover effect from the older to the younger sibling (column 1), we find that an increase of 1 standard deviation in the test score of the older sibling leads to an increase of 1.9% of a standard deviation in the corresponding test score of the younger sibling. This spillover effect seems small, but it is strongly statistically significant. 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 siblings to be affected by their younger sibling's tests taking place in the future. The endogeneity test reported in Table 4 indicates that we can strongly reject the equality of the estimation with fixed effects and of the 2SLS estimation with fixed effects, therefore rejecting the exogeneity of the sibling test score. Comparing the sibling spillover effect reported in column (1) of Table 4 with the sibling correlation in test scores net of a child fixed effect (Table 3, column 3) we can see that subject-specific family and school investments that are similar between siblings and that explain the sibling association in test scores are quite important. After controlling for these subject-specific family and school investments the correlation is much reduced.

5.2 Threats to identification: Robustness checks

Our identification strategy relies on the assumption that our instrumental variable is valid, i.e. relevant and exogenous. Our instrumental variable is based on the average test scores observed for the older sibling's school mates and we use it to instrument the older sibling's test score. While there are no doubts on the relevance of our instrument in explaining the older sibling's test score (see F-test in Table 4), there are potential threats to its exogeneity caused by (i) the potential correlation of our instrument with child specific unobserved endowments that are potentially correlated between siblings and also between school mates; (ii) the sorting of siblings into schools with similar investments and characteristics that are subject specific; (iii) the potential direct influence of the older sibling's school mates on the younger sibling. The first two threats are eliminated by our estimation procedure that controls for child fixed effects and for school-by-cohort-by-subject fixed effects, whereas the last threat is not addressed by our estimation. We investigate this threat further and present robustness checks.

We can think of three situations where the assumption that older sibling's school mates have no direct influence on younger sibling's test scores may be violated. This could be through direct interaction between the older sibling's school mates and the younger sibling in the neighborhood, the school or the home. We will look at each of these possibilities in turn.

It may be the case that children living in the same neighborhood interact and play with each other outdoor even if they do not belong to the same cohort, for example by meeting up in parks or hanging out near local shops.[16] 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. This is only 5% of a school cohort which comprises 181 pupils on average. So the interaction within a neighborhood is limited to a small fraction of the peers the older sibling is exposed to at school while learning. Taking into account that students may interact within a wider geographical area, 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 33 peers from the same school and cohort live in an area thus defined, which is 18% of an average school cohort. We take this as the maximum proportion of the older sibling's school mates a (very sociable) younger sibling could interact with in the residential area.

To test the possibility of neighborhood interaction more formally, we exclude the older sibling's school peers living in the same neighborhood in the computation of the instrumental variable with the aim to net out the potential direct effects that go from the children living in the same neighborhood to the younger sibling. We also perform the same test by excluding older sibling's school peers living in the same area, defined at the Middle Layer Output

[16]Students from the same neighborhood could also share the journey to school. The average distance from home to school is 3.5 miles for secondary school students in England, and 34% of pupils go to school on buses, 24% by car, 36% on foot (Department for Transport 2010).

Area level. Table 5 displays the results of this exercise. Excluding older sibling's 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.

The next possibility we want to consider is interaction of the younger sibling with her older sibling's school peers at school, for those siblings that attend the same school. Teaching in English secondary schools is separate by cohort, so any interaction would have to take place during the lunch break which takes place between the morning and the afternoon session of the school day, and is about 30 minutes long. While the scope for interaction in lunch break is limited through the available time, there is also the possibility that students meet in after school clubs organised by the school. To satisfy ourselves that interaction at school is no issue, we estimate the sibling spillover separately for siblings going and not going to schools that offer post-16 schooling (in so-called sixth forms working towards A-levels, the University entry exams). Younger siblings in schools without post-16 schooling will not be exposed to older siblings' school mates at school in the last years leading up to their age 16 exam, as these will have left the school.[17] Depending on the age gap between older and younger sibling, there will be no school-based contact for 1-3 years. Table 5 shows the results of these estimates. The estimated sibling spillover effect is even larger at schools that do not offer post-16 schooling than at those that do, indicating that school-based interaction is not a problem for identification.

Finally, it is possible that interaction between younger siblings and older siblings' peers takes place at home, for example when friends visit. Survey evidence for England shows that about a third of children aged 11-15 had no friends round their house in a reference week, another third had friends round 1-2 times, and the remaining third three or more times (Jamieson and McKendrick 2005). This seems a low frequency of contact within the home, and again would comprise only a fraction of the relevant peer group of the older sibling (181 students on average). Research based on AddHealth data (The National Longitudinal Study of Adolescent Health in the USA) shows that among siblings from grade 7 to 12 about half

[17] Even in schools that offer post-16 schooling a considerable proportion of pupils will leave school at the end of compulsory schooling, at age 16.

have no or few mutual friends, 30% have some mutual friends and 20% have mostly mutual friends (Rende et al. 2005). An older study based on the Arizona sibling study of 10-16 year olds finds similar proportions of mutual friends as in the AddHealth data. When asked how often children interact with the mutual friends shared with siblings, however, 50-70% of children say they never or rarely interact (Rowe et al. 1993). This evidence indicates that interaction in the home is unlikely to threaten our identification strategy.

We are also interested in checking the validity of our instrument further by using additional instruments and computing a Hansen's J test on the over-identifying restrictions. There are different potential alternative instruments (e.g. other average characteristics of the older sibling's school mates) that we can use to explain the older sibling's test score, and we experimented with several different options. Eventually we chose our additional instrument among potential candidates to ensure high statistical and substantive significance in explaining the older sibling's test score. Specifically, we consider 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 school in a specific subject. We use this instrument together with our original instrument to allow us to carry out an over-identification test of the validity of the instrument. The Hansen's J test reported at the bottom of Table 5 shows that the null that the instruments are exogenous cannot be rejected.

5.3 Heterogeneity analysis

In this section we look at sibling spillover effects for sub-groups of the population. In particular we are interested in sibling spillover effects separately by

- gender composition of sibling pairs (female-female, male-female, female-male and male-male). Although we have no prior as to how effects may vary by gender composition, we do know that raw correlations are higher for sisters than for brothers or mixed gender siblings (Nicoletti and Rabe 2013).
- sibship size. We would expect to see larger sibling spillover effects for 2-child than for larger families if the interaction within each sibling pair decreases in the number

of pairs living in a family. On the other hand, it is possible that the older sibling compensates for lack of parental time in larger families.

- free school meal status. It is possible that in low income families older siblings transmit positive school attitudes and learning to compensate for a lack of transmission from parents to children. On the other hand there could be multiplier effects in higher income families whereby investments into children are high, and these get reinforced through sibling spillover effects.

- language spoken at home. Similarly to low income families, older siblings can be expected to either compensate or reinforce advantages/disadvantages associated with speaking English or another language at home.

- whether siblings go to the same school. We expect interactions between two siblings going to the same school to result in a larger spillover effect because of school specific information that can be transmitted from the older to the younger child.

Table 6 shows the results of the heterogeneity anaysis. All results are estimated using 2SLS estimation of the value added model (1) with individual and school-by-cohort-by-subject fixed effects. We perform the sub-group analysis by interacting both the older sibling's test score and the instrument with indicator variables indicating group membership. These interactions allow us to estimate the sibling spillover effect simultaneously for various groups. The benchmark estimates are reported in the first row of Table 6, followed by results for the different sub-samples. Looking first at estimates by sibling gender combination, we see that there are positive and statistically significant effects of an older brother to his younger brother and from an older sister to a younger brother, but no statistically significant sibling spillover effect beween same-sex sibling pairs. Given the many mechanisms through which sibling spillover effects may operate, including the older sibling serving as role model, teaching, playing/interacting, giving information and signalling to the younger sibling and competition between siblings, it is hard to give an interpretation to these differential effects by gender combination. They are, however, interesting and worth further investigation.

Regarding sibship size, Table 6 displays separate effects for siblings from 2-child and from larger families. The Table shows that the results are very similar between the groups, and equality of the coefficients is not rejected. This result does not allow us to distinguish whether the sibling spillover effect is driven by interaction intensity - which presumably is larger in smaller families - or by older siblings trying to compensate for lack of parental time - which is presumably more often the case in larger families.

The next two estimates reported in Table 6 look at differential sibling spillover effects by free school meal status and language spoken at home. Interestingly, the sibling spillover effect is larger in low income families and in families that do not speak English at home, and this difference is statistically significant. This may suggest that in more deprived families older siblings compensate for the lack of parental inputs, and sibling spillover effects may be a mechanism that make children from these families more resilient. In other words, sibling spillover effects are concentrated less deprived families where they produce multiplier effects and reduce the intergenerational transmission of inequality.

Finally we look at the differences in sibling spillover effects for siblings going to the same and different schools. With this we want to assess whether a more direct knowledge by the older sibling of the specific school rules and teachers lead to a higher level of sibling transmission of test scores. This seems the be the case, we find that the spillover effect is sizeably larger than the benchmark estimation (see Table 5). An increase of the test score of the older sibling attending the same school by one standard deviation increases the test score of the younger sibling by 3.2% of a standard deviation, compared to 1.9% when considering the whole sample. This seems to suggest that the effect of interactions between siblings is stronger when the older sibling is able to help the younger sibling by transmitting school-specific information e.g. on school rules and teachers. For siblings going to different school (a minority of 15% in our sample) we find no statistically significant spillover effect.

6 Conclusions

In this paper we propose a new strategy to identify the sibling spillover effects in education and we apply it to provide the first empirical evidence on the sibling spillover effect in school test scores.

Our main concern when estimating this spillover effect is unobserved heterogeneity, in particular the potential unobserved family and school investments that are shared by siblings and that can cause a spurious association between siblings. The sibling correlation reduces from roughly 50% to about 12% when we control for child, school and family characteristics that are subject-invariant. Furthermore, when we also control for family and school characteristics that are subject-specific, the sibling correlation reduce to 1.9% and we can confidently attribute this net correlation to a causal sibling spillover effect going from older siblings to younger siblings. While we find a statistically significant spillover effect of 1.9% from older to younger siblings, there is no significant spillover effect going from younger to older siblings.

To summarize, our analysis provides evidence on the existence of a sibling spillover effect in test scores but of modest magnitude with respect to the importance of the effect of family and school characteristics that are shared by the siblings. Nevertheless, there are specific situations where the sibling spillover effects can get more important, e.g. we find that the spillover effect is considerably larger for siblings who attend the same school and for siblings from families with low-incomes and a language other than English spoken at home.

References

Aakvik, A., K.G. Salvanes and K. Vaage (2010). “Measuring heterogeneity in the returns to education in Norway using educational reforms.” *European Economic Review*, 54(4), 483-500.

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.

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.

Altonji, J.G., Cattan, S. and I. Ware (2013). “Identifying Sibling Influence on Teenage Substance Use”, *IFS Working Paper* W13/04.

Angrist J., V. Lavy and A. Schlosser (2010), “Multiple Experiments on the Causal Link between the Quantity and Quality of Children.” *Journal of Labor Economics*, 28, 4, 773-824.

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.

Behrman, J.R. and M.R. Rosenzweig (2004). “Returns to Birth Weight.” *The Review of Economics and Statistics*, 86(2), 586-601.

Bernal, R. (2008). “The Effect of Maternal Employment and Child Care On Children’s Cognitive Development.” *International Economic Review*, 49(4): 1173-1209.

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

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.

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.

Black, S., P. Devereux and K.G. Salvanes (2008). "Staying in the classroom and out of the maternity ward? The effect of compulsory schooling laws on teenage births." *The Economic Journal*, 118 (530), 1025-1054.

Black S., P. Devereux and K. Salvanes (2010). "Small Family, Smart Family? Family Size and the IQ Scores of Young Men." *Journal of Human Resources*, 45, 1, 33-58.

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

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.

Cáceres-Delpiano J. (2006). "The Impacts of Family Size on Investment in Child Quality." *Journal of Human Resources*, 41, 4, 738-754.

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.

Cosconati, M. (2012). "Skill Formation and Parenting Styles: Evidence From a Structural Supermodular Game with Multiple Equilibria." forthcoming???

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 (2013)."Peer Effects in Progam Participation". *American Economic Review*, forthcoming.

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, 241275.

Del Boca D., C. Flinn C. and M. Wiswall (2014), "Household Choices and Child Development" *Review of Economic Studies*, forthcoming.

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

Ermisch, J. and M. Francesconi (2000). "Educational Choice, Families, and Young People's Earnings." *The Journal of Human Resources*, 35, 1, 143-176.

Ewin Smith T. (1990). "Academic achievement and teaching younger siblings." *Social Psychology Quarterly*, 53, 4, 352-362.

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

Glaeser, Jose A. Scheinkman and Bruce 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.

Jamieson, L. and J.H. McKendrick (2005). "Teenagers' relationships with peers and parents", in J.F. Ermisch and R.E. Wright (eds.) *Changing Scotland. Evidence form the British Household Panel Survey.* The Policy Press, Bristol, p. 17-32.

Joensen, J.S. and H.S. Nielsen (2013). "Peer Effects in Math and Science", *Fourth International Workshop on Applied Economics of Education*, Catanzaro.

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.

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

Holmlund, H. (2008)."A Researcher's Guide to the Swedish Compulsory School Reform." *CEE Discussion Paper*, 87, Centre for the Economics of Education, London School of Economics.

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.

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.

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 J. (2008). "Sibling Size and Investment in Children's Education: An Asian Instrument." *Journal of Population Economics*, 21, 4, 855-875.

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

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.

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.

Fletcher, J., Hair, N.L. and B.L. Wolfe (2012). “Am I my Brother’s Keeper? Sibling Spillover Effects: The Case of Developmental Disabilities and Externalizing Behavior”, *National Bureau of Economic Research*, Working Papers Series 18279.

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

Manski, C. (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 (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 (2014), “Mother’s labor supply: the spillover effects of sisters and female cousins.”

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

Qureshi, J.A. (2011). “Additional Returns to Investing in Girls’ Education: Impact on Younger Sibling Human Capital”, PhD thesis, Harris School of Public Policy Studies. University of Chicago

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, 383-392.

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 (2012). “Human capital, kinship, and gender inequality.” *Oxford Economic Papers*, 64, 3, 417-438.

Rosenzweig, M.R. and T.P. Schultz (1982). "Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India." *American Economic Review*, 72(4), 803-815.

Rosenzweig, M.R. and K.I. Wolpin (1988). "Heterogeneity, Intrafamily Distribution, and Child Health." *The Journal of Human Resources*, 23(4), 437-461.

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., 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.

Widmer, E. (1997). "Influence of Older Siblings on Initiation of Sexual Intercourse." *Journal of Marriage and the Family*, 59(4): 928-938.

Tables

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	Registers data from Norway	0.242**	
Qureshi (2011)	Years of schooling	2SLS, using distance from closest school	Learning and Educational Achievement in Punjab Schools	0.420**	0.163
Joensen and Nielsen (2013)	Advanced Math and Science choice	2SLS, using a high school reform as IV	Registers data from Denmark	0.334+	
Adermon (2013)	Years of schooling	2SLS, introduction school reform	Registers data from Sweden	not sig	not sig
Dahl, Loken Mogstad (2013)	paternity leave take-up	2SLS, introduction of paid father's leave	Register data from Norway	0.153* older or younger	
Nicoletti, Salvanes Tominey (2014)	Maternal labor participation	2SLS, peers of peers outcome	Register data from Norway	0.270**-0.299** older or younger	

Notes: + $p < .10$, * $p < .05$, ** $p < .01$.

Table 2: Descriptive statistics: Unstandardized test scores

	Older		Younger	
	mean	std dev	mean	std dev
Key Stage 2 English score	27.4	4.1	26.9	4.1
Key Stage 2 Science score	27.8	4.6	27.7	4.6
Key Stage 2 Maths score	29.2	3.6	29.0	3.8
Key Stage 4 English score	40.9	9.2	40.6	9.0
Key Stage 4 Science score	40.4	10.5	40.9	9.9
Key Stage 4 Maths score	40.1	10.7	40.0	10.5
N	217,948		217,948	

Notes: National Pupil Database, 2007-2010.

Table 3: Sibling correlations in test scores

	(1) Raw correlation in sibling test scores	(2) Correlation in sibling test scores value added	(3) Correlation in sibling test scores value added child FE
Test score older sibling	0.495**	0.305**	0.124**
	(0.001)	(0.001)	(0.001)
N	653,835		

Notes: + p < .10, * p < .05, ** p < .01. National Pupil Database, 2007-2010. We report standard errors between parenthesis. Pooled sample, pooling the observations for Mathematics, English and Science. Value added model controls for test scores at age 11 for both siblings.

Table 4: 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
First Stage		
$MDevY_{2,s'qt',16}$	0.84**	0.83**
	(0.005)	(0.006)
F-test	26,781	18,825
p-value	(0.000)	(0.000)
Second Stage		
γ	0.019**	0.005
	(0.006)	(0.008)
Endogeneity test	117.8	119.8
p-value	(0.000)	(0.000)
N	653,613	653,178

Notes: + p < .10, * p < .05, ** p < .01. National Pupil Database, 2007-2010. We report standard errors between parenthesis. Pooled sample, pooling the observations for Mathematics, English and Science. The F-test is to test the significance of the excluded instrument in the first stage. The endogeneity test is the robust Durbin-Wu-Hausman test.

Table 5: Sibling spillover effects: Robustness checks

	(1) Spillover effect FE IV	(2) F-test	(3) Endog. test	(4) Hansen's J test	(5) No. of obs.
Benchmark estimation					
	0.019**	26,781	117.8		653,613
	(0.006)	(0.000)	(0.000)		
Excluding older sibling's school mates living in same neighborhood					
	0.017**	26,422	123.6		653,598
	(0.006)	(0.000)	(0.000)		
Excluding older sibling's school mates living in same area					
	0.016**	25,397	125.6		653,577
	(0.006)	(0.000)	(0.000)		
Siblings at same school with and without post-16 education					
with	0.018*	5,558	117.4		578,268
	(0.009)	(0.000)	(0.000)		
without	0.036**				
	(0.012)				
Using two instruments to test overidentification					
	0.020**	13,689	117.6	0.291	653,613
	(0.006)	(0.000)	(0.000)	(0.590)	

Notes: + p < .10, * p < .05, ** p < .01. We report standard errors and p-values between parenthesis. Pooled sample, pooling the observations for Mathematics, English and Science. The F-test is to test the significance of the excluded instrument 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. Estimate comparing siblings at same school with and without post-16 education also includes siblings at diferent schools, coefficient not shown.

Table 6: Sibling spillover effect: Estimation for sub-groups

	(1)	(2)	(3)	(4)
Benchmark	all			
coeff	0.019**			
	(0.006)			
F-Test	26,781			
Sex combination:	brother→brother	brother→sister	sister→brother	sister→sister
coeff	0.057**	0.001	0.028**	-0.005
	(0.008)	(0.008)	(0.009)	(0.009)
F-Test	6,680			
Ho: (1)=(3), χ^2	8.08 (0.005)			
Ho: (2)=(4), χ^2	0.35 (0.841)			
Agegap:	1 year	2 years	3 years	
coeff	0.042**	0.004	0.019+	
	(0.009)	(0.007)	(0.010)	
F-Test	8,971			
Ho: (1)=(3), χ^2	25.66 (0.000)			
Ho: (2)=(3), χ^2	3.86 (0.146)			
Family size:	2 siblings	3+ siblings		
coeff	0.019**	0.020**		
	(0.006)	(0.007)		
F-Test	13,421			
Ho: (1)=(2), χ^2	0.02 (0.878)			
Free School Meal:	FSM eligible	not FSM		
coeff	0.064**	0.013*		
	(0.012)	(0.006)		
F-Test	13,483			
Ho: (1)=(2), χ^2	20.22 (0.000)			
Language at home:	not English	English		
coeff	0.031*	0.018**		
	(0.014)	(0.006)		
F-Test	13,322			
Ho: (1)=(2), χ^2	12.51 (0.002)			
Siblings school:	same school	diff. school		
coeff	0.032**	-0.003		
	(0.007)	(0.009)		
F-Test	11,180			
Ho: (1)=(2), χ^2	10.50 (0.001)			

Notes: + p < .10, * p < .05, ** p < .01. We report standard errors and p-values between parenthesis. Pooled sample, pooling the observations for Mathematics, English and Science. The F-test is to test the significance of the excluded instrument in the first stage. Hypotheses test gives χ^2 and p-value.