Parents, siblings and schoolmates

The effects of family and school interactions on educational achievement and long-term labor market outcomes*

Marco Bertoni
University of Padova

Giorgio Brunello
University of Padova, IZA, CESifo and ROA

Lorenzo Cappellari
Catholic University of Milan and IZA

August 2017

PRELIMINARY DRAFT

---

*Address for correspondence: Marco Bertoni, Department of Economics and Management, University of Padova, via del Santo 33, 35123 Padova IT, phone: +39 049 827 4002, email: marco.bertoni@unipd.it. Giorgio Brunello: giorgio.brunello@unipd.it. Lorenzo Cappellari: lorenzo.cappellari@unicatt.it. We are very grateful to Paul Bingley for constant help with the data. We also thank Olmo Silva and seminar participants at Maastricht, Berlin (ESPE 2016) and Rome (*Visitinps)* for comments and suggestions. Lotte Kofoed Jorgensen provided excellent research assistance. Marco Bertoni and Giorgio Brunello acknowledge the financial support by Fondazione Cariparo "Starting Grants". The usual disclaimer applies.

Abstract

We investigate whether school peer effects on educational achievement and labor market outcomes vary with family characteristics, including the gender composition of siblings and parental background. Using Danish register data, we find that the benefits from interacting with “privileged” students accrue only to “disadvantaged” males. We also find that girls benefit from exposure to a higher share of girls at school in terms of employment at age 31 to 40 only if they have sisters. Boys with sisters face instead poorer long term employment prospects. The size of the estimated effects is small.

Keywords: education peer effects, gender, parental background, human capital production, long term outcomes
JEL codes: I21, J16, J24

1. **Introduction**

Recent empirical research has shown that social interactions at school can affect academic achievement and ultimately labor market outcomes. Two measures of school peer characteristics that have attracted attention are the share of girls and average parental background in the class/grade/school attended by an individual.[1] There is evidence that a higher share of girls positively affects the learning outcomes of both girls and boys (see Lavy and Schlosser, 2011) and later outcomes of girls (see Black, Devereux and Salvanes, 2013). Positive average parental background effects on individual learning have been detected by Ammermueller and Pischke, 2009, using PIRLS data for Europe, and by Black, Devereux and Salvanes, 2013, who have shown that the average earnings of peers' fathers matter for the education and labor market outcomes of boys in Norway.

A potentially relevant question that has attracted little attention so far is whether the benefits and costs of peer interactions at school are affected by earlier interactions occurring within the family and involving both parents and siblings. For example, it is important to know whether the benefits from interacting with schoolmates having well educated parents accrue mainly to those coming from a similar "privileged" background or to the "disadvantaged", as this may affect policies targeted at desegregating schools. Since peer interactions among differently gendered individuals start in the family, shaping the goals and expectations of girls and boys, we would also like to know whether having interacted with at least one sister in the family affects the benefits and costs of having many female schoolmates at school.

In this paper we provide evidence on the effects of school-family interactions on educational attainment and longer term labor market outcomes. We use Danish register data to investigate whether and how the effects of peer

[1] Another peer characteristic that has been shown to generate relevant spillover effects is ability (see Lavy et al., 2012a, and Lavy et al., 2012b, Booij et al., 2017). However, we do not have measures of ability in our data.

characteristics at school vary with the type of family interactions involving parents and siblings. Identification of peer effects exploits plausibly random variation in peer composition within schools and between cohorts, controlling also for school-specific trends. We partial out unobserved heterogeneity between siblings by including family fixed effects.

There are several economic and social mechanisms explaining why peer characteristics affect individual performance. A higher share of girls in the class or school can improve the learning environment by reducing disruption (Lazear, 2000). Individual behavior may also change. For instance, pupils with more female schoolmates – or with more sisters at home - may change their attitude toward risk (see Booth and Nolen, 2012) and competitiveness (see Gneezy, Niederle and Rustichini, 2003). On the other hand, peers with a better parental background may act as positive role models in education and facilitate access to economic and social networks.

Understanding the effects of school peer characteristics is important for the design of school admission and class formation policies. On the one hand, these effects help informing the ongoing debate about single-sex schools. A report by the UK Department of Education has suggested that boys should be taught separately (see Doris et al, 2013). According to the European Association of single-sex education, co-education should be challenged, especially for girls, although there is evidence that single-sex schooling may reduce the gender math gap (see Fryer and Levitt, 2010). On the other hand, they shed light on the benefits of de-segregating schools and – assuming that parental education is a relevant measure of family background – they also inform about the intergenerational social returns to education, about which relatively little is known to date (see Oreopoulos an Salvanes, 2011).

We measure school peer characteristics in the 9th grade (age 15) with the share of female schoolmates and the average parental background of schoolmates, and family interactions with parental education and the presence of at least one sister in the family. We draw our data from administrative registers of the Danish population. The key advantage of these data is that, for each individual, we have information on parents, siblings and schoolmates.

When we do not distinguish males and females according to their interactions in the family, we find that the share of female schoolmates affects only the educational attainment of females and has no statistically significant effect on labor market outcomes, contrary to what found by Black, Devereux and Salvanes, 2013. We also find that average parental background in the school increases both educational attainment and lifetime earnings for males but not for females, similarly to what found by Black, Devereux and Salvanes, 2013, using average log father's earnings to measure background. As in Black, Devereux and Salvanes, 2013, the statistically significant effects that we find are always very small.

However, our data indicate that the effects of school peer characteristics vary with individual parental education. In particular, we find that, for males, a higher share of "privileged" schoolmates increases years of education and lifetime earnings for the "disadvantaged", but have no statistically significant effect on the "privileged".[2] For females, we find instead that having a sister in the family significantly affects the impact of the share of female schoolmates on years of completed education and employment at age 31 to 40.

In particular, we find that females (males) with at least one sister are more (less) likely to be employed after experiencing a higher share of female schoolmates.[3] We believe that, for females, the combination of many female schoolmates and the presence of sisters in the family may attenuate feminine stereotypes and positively affect competitiveness and risk taking, with positive effects on long term employment prospects. On the other hand, males with

[2] As discussed below, we define as "privileged" an individual with at least one parent having higher education and as "disadvantaged" an individual whose parents have less than higher education.

[3] Since these estimated effects are statistically significant at the 10 percent level of confidence, we need to interpret these results with some caution.

sisters may be not only less disruptive but also acquire more feminine traits, for instance in terms of lower competitiveness and higher risk aversion. Interacting with a higher share of girls at school could foster these traits, with negative labor market consequences.

With some extrapolation, our results point out that the benefits and costs of single schools are likely to be small and to vary not only with individual gender but also with the gender composition of siblings. They also indicate that attracting students from “privileged” background can produce small benefits for (male) “disadvantaged” students in terms of education and lifetime earnings, without having statistically significant consequences on the “privileged”. Thus, they provide an argument in favor of mixing students from different backgrounds.

The paper is organized as follows: we review the literature in Section 1 and describe our data and the empirical setup in Sections 2 and 3. Results are presented in Section 4. Conclusions follow.

2. The Literature

Three strands of literature are particularly relevant for our paper. The first strand looks at how social interactions at school affect academic achievement, career choice and labor market outcomes. Prominent contributions in this area include Ammermueller and Pischke, 2009, Lavy and Schlosser, 2011, and Black, Devereux and Salvanes, 2013. The second strand looks at how the gender composition of siblings affects educational choice and later outcomes – see for instance Butcher and Case, 1994. Finally, the third strand considers intergenerational mobility and the role played by family background in the determination of long run economic outcomes (see Black and Devereux, 2011, for a review).

Ammermueller and Pischke, 2009, estimate background peer effects using international data for European fourth graders from the Progress in International Reading Literacy Study (PIRLS). In their study, they exploit the variation in the composition of peers among classes within schools – which they argue to be random – and find positive effects of peers' family background, measured in terms of the number of books at home, on individual learning.

Lavy and Schlosser, 2011, study the effects of the percentage of female schoolmates in elementary, middle and high schools of school performance in Israel. The selected outcomes include test scores, matriculation status and number of credits earned. They find that the proportion of girls in a class has a positive and significant effect on the academic achievement of girls and boys. Benefits are larger for students with low parental education and for new immigrants. They argue that these gains are mediated through lower level of classroom disruption and violence, improved relationships within the class and lessened teacher's fatigue.

Black, Devereux and Salvanes, 2013, substantially extend this literature by focusing on the role of lower secondary school peers, aged between 14 and 16, on schoolmates in the same grade, who as teenagers are expected to be particularly affected by peer influences, and by considering post-school outcomes – including teenage childbearing, educational attainment and average earnings in a three years window - rather than test scores. Using Norwegian data, they find relatively small peer effects and that, while females benefit from having a higher proportion of female peers, males are negatively affected. Therefore, moving to single sex schools would benefit both girls and boys. They also find that, while maternal education has no detectable impact on outcomes, the father's income matters for boys.

In a recent contribution,[4] Schone et al, 2017, use Norwegian data to investigate the impact of female peers of the choice of scientific and technological

[4] Additional contributions in this area include Hoxby, 2000; Whitmore, 2005; Anelli and Peri, 2017; Oosterbeek at al, 2014; Ciccone et al, 2015; Park, 2015; Eisenkopf et al, 2015 and Feld and Zoelitz, 2017.

(STEM) fields in upper secondary schools. They find that more female peers in lower secondary schools increase the probability that female students choose STEM courses, and explain this result by arguing that a higher share of females may weaken gender stereotypes that keep females away from STEM fields.

If the gendered aspects of individual behavior are brought into play by the gender of others with whom they interact, we expect the gender composition of siblings in the family to play an important role in the development of personality and cognitive traits. Socialization at home may shape the goals and expectations of girls and boys (see Rosenberg and Sutton-Smith, 1968; Stoneman et al, 1986), and alter the effect of peer characteristics at school on individual outcomes.

On the one hand, being less exposed to stereotyped behaviors, females growing up in a whole-sisters family may be less inclined to acquire traditional gender roles. On the other hand, the interaction between sisters and brothers could produce relevant externalities, with females learning to male-behave from brothers and brothers learning female attitudes from their sisters. In line with this argument, psychological studies show that girls with older brothers develop more masculine traits, while boys with older sisters are characterized by more feminine traits (Koch 1955; Brim, 1958).

Empirical research has investigated the effects of siblings on educational outcomes. While the negative relationship between the number of siblings, birth order and educational outcomes is well documented (see Steelman et al., 2002 for a review), results are less conclusive for the role played by the sex composition of siblings. Using US data, Butcher and Case, 1994, find that a daughter with a brother receives half a year more education than if she had a sister. Conley, 2000, finds instead that girls' educational attainment is lowered by the presence of brothers. These conflicting results could be due to selection based on family size, or to the way in which joint investments are made in a family. Focusing just on the arguably random gender of the first born child, Dahl and Moretti, 2008, find that siblings with a first born sister have lower education on average compared to siblings with a first born brother (see Bharadwaj, Dahl and Sheth, 2014).

Women and men may differ in their propensity to choose a risky outcome because of innate preferences or because pressure to conform to gender-stereotypes encourages girls and boys to modify their innate preferences. Single-sex environments may modify the risk-taking preferences of students in economically important ways. Booth and Nolen, 2012, use a controlled experiment in which subjects were given an opportunity to choose a risky outcome in a real-stakes gamble with a higher expected monetary value than the alternative outcome with a certain payoff and in which the sensitivity of observed risk choices to environmental factors could be explored.

They show that gender differences in preferences for risk-taking are sensitive to whether the girl attends a single-sex or coed school. Girls from single-sex schools are as likely to choose the real-stakes gamble as boys from either coed or single sex schools, and more likely than coed girls. They also find that gender differences in preferences for risk-taking are sensitive to the gender mix of the experimental group, with girls being more likely to choose risky outcomes when assigned to all-girl groups.

Peter, Lundborg and Webbink, 2015, investigate how the gender of a sibling affects individual education, earnings and family formation. They find that the gender of the co-twin influence males and females in different ways. On the one hand, females with sisters obtain lower education and give birth earlier. On the other hand, males with brothers earn more and are more likely to get married and have children. They argue that males are likely to be less risk averse, more competitive, less socially minded, less agreeable and less neurotic than females. If these traits spill-over to other siblings, this may explain why those with brothers have higher income.

Starting from the theoretical contribution of Becker and Tomes, 1979, a large strand of literature in labor economics has studied how family background affects inequality in individual outcomes. In a model where parents care about their children's lifetime earnings, transmit their own productive endowments and can invest in children's earnings capacity, lifetime earnings are transmitted inter-generationally. Solon, 1999, Björklund and Jännti, 2009, and Black and Devereux. 2011, review the relevant empirical literature.

Whether and how the own and peers' background interact in the production of individual long run outcomes is less well known. Malamud et al, 2016, find that, although access to abortion and access to better schools in Romania each have positive impacts on human capital investments, there is no evidence of significant interactions between these shocks shocks. Rossin-Slater and Wust, 2015, find that the long-run effects of a high quality preschool childcare in Denmark is stronger for pupils not exposed to a nurse home visiting program enhancing parenting skills.

3. The Data

We use data from administrative registers of the Danish population. Since 1968, the civil registration system attributes a unique personal identifier to all residents, which we use to reconstruct families and track individuals across various registers. We merge these data with individual tax declarations, which include information on individual earnings, and with school registers to associate individuals to their schoolmates. These registers were introduced in the country in 1973 to monitor compliance with compulsory school reforms.

Our data consist of 18 cohorts of individuals born between 1958 and 1975, for whom we observe labor incomes between 1989 and 2015. There are 1,009,924 individuals in our sample, and 860,879 non-missing observations for real earnings. For each individual, we observe her completed education at age 31, well after the completion of highest statutory education in Denmark.

The school registers allow us to link each individual to her schoolmates on October 31 of the calendar year when she turned 15, typically corresponding to enrolment in the 9th grade of compulsory education. There are 1,459 schools in our sample. For each individual in our sample, we compute the share of female schoolmates *SG* by considering all the pupils born in the same year and enrolled in the same school rather than those belonging to the same grade, as done for instance by Black, Devereux and Salvanes, 2013.

While our measure of the share of female schoolmates is not exposed to the endogeneity threats discussed by Black, Devereux and Salvanes, 2013, who argue that parents could choose strategically the school starting age in order to place their children in the desired group, it is very close to the proportion of females in the grade, because the vast majority of children in Denmark start school at the prescribed age.[5]

Moreover, there are very few grade retentions, and most Danish students complete primary and secondary education in the same school. Therefore, mobility between school is less common. Figure 1 shows the distribution of *SG* by school and cohort. The average share is 0.489, and the standard deviation is 0.09. Using the normal distribution as a first order approximation, 95 percent of the distribution of *SG* lies within the interval 0.310-0.670.

We combine school registers with household information to retrieve for each schoolmate information on parental education, that we use to compute the share of schoolmates *SH* who have at least one parent with higher education.[6] Figure 2 shows the distribution of *SH*, with mean 0.238 and standard deviation 0.15. Compared to *SG*, peer characteristic *SH* is skewed to the left and with a long and thin right tail.

The two measures of peer characteristics, *SG* and *SH*, are similar to those adopted by Black, Devereux and Salvanes, 2013, who use the share of females

[5] According to Bingley, Cappellari and Tatsiramos, 2016, in the 1980s 95 percent of pupils aged 15 were enrolled in grade 9th.
[6] We define higher education as at least 13 years of schooling.

in the grade, average age in the class, average mother's education in the grade and average log father's earnings in the grade. Compared to Black, Devereux and Salvanes, 2013, we also investigate the relationship between school and family interactions.

We use information on siblings and parents to compute two indicators of family interactions: a dummy *FG* equal to 1 if the individual has at least one sister and 0 otherwise, and a dummy *FH* equal to 1 if at least one parent has higher education and 0 otherwise. In the parlance of this paper, we define individuals with *FH=1* as "privileged" and individuals with *FH=0* as "disadvantaged". In our sample, 54.6 percent of individuals have at least one sister, and 23.9 percent have at least one parent with higher education.

In families with two siblings, the average age spacing between siblings is 3.7 years. In families with four siblings, age spacing to the closest sibling is 2.6 years. On average, age spacing with the closest sibling is 3.4 year, and with the closest sister is 4 years. The average distance between the first and the last born is 5.2 years.

Following Black, Devereux and Salvanes, 2013, we select as individual outcomes educational attainment and labor market variables.[7] Schools in Denmark are compulsory until age 16. Post – secondary education can be general (gymnasium and higher preparatory), technical (commercial, technological and scientific programs) or vocational. These programs typically last three years and are followed by college. We classify the fields of study at college in the following groups: scientific and technological (STEM); humanities; health related; law and social sciences (the residual sectors are agriculture, environmental protection, other minor fields).

We construct the following measures of educational attainment: a) completed years of education; b) a dummy indicating the completion of tertiary education; c) an indicator of whether the highest degree attained by the

[7] Lavy and Schlosser, 2011, use instead test scores and high school achievements.

individual is vocational or academic;[8] d) the field of study in college. Turning to labor market outcomes, the most comprehensive measure of individual earnings over the life cycle is lifetime earnings. Black, Devereux and Salvanes, 2013, use instead average earnings over a three-year window. It is questionable whether the selected window provides a good proxy of lifetime earnings. Following the literature on the life cycle bias in earnings (see Haider and Solon, 2006, Bhuller et al, 2017, and Nybom and Stuhler, 2016, among others), we use instead average earnings from age 31 to age 40. By choosing this broader window, we minimize the gap between the selected measure and lifetime earnings.[9]

We use tax records to obtain pre-tax annual labor earnings – or total income from labor - at 2012 prices. We consider an observation to be valid when measured annual earnings is above 35,000 Danish Crowns (about 4,700 euro). Consistently, we define long term employment as a dummy equal to 1 if the individual has had at least five valid earnings observations between age 31 and 40.

The summary statistics of the variables used in the empirical analysis are presented in Table 1. Independently of the selected indicator, educational attainment is always higher among females, who have on average 13.11 years of education, compared to 12.94 years for males. In spite of the fact that 38.3 percent of females have completed college, only 3 percent have completed a scientific field and more than 20 percent has completed instead a degree in health and related fields. The percent of males who have completed a college degree is 30.3 percent. Of those, close to 9 percent have chosen either a STEM

[8] At the secondary level, academically oriented education is considered academic, while education in vocational schools, often in combination with apprenticeship, is regarded as vocational. At the tertiary level, academy professional degrees, professional bachelor degrees, top-up degrees, and business academy bachelor's degrees - which all include internship periods - are counted as vocational. University bachelor's degrees and postgraduate degrees, as well as artistic bachelor's degrees and master's degrees, are considered academic.
[9] We compute this indicator if at least five valid observations in the age interval are available. See Solon, 1992, and Mazumder, 2005.

field or law and social sciences. Close to 65 percent of males and 62 percent of women have vocational education as their highest degree. The probability of employment for five years in the age range 31 to 40 is close to 0.85 for both genders. Finally, our proxy of average real pre-tax lifetime earnings is 2.5 percent higher for males than for females.

4. Empirical Methodology

Following Lavy and Schlosser, 2011, and Black, Devereux and Salvanes, 2013, our research design exploits plausibly random variation in peer composition between cohorts within schools, that is likely to arise because of demographic factors. We start from the following baseline empirical specification

$$Y_{ics} = \alpha_c + \beta_s + \lambda X_{ics} + \pi_1 SG_{cs} + \pi_2 SH_{cs} + \theta_1 SG_{cs} * F_{ics} + \theta_2 SH_{cs} * F_{ics} + \varepsilon_{ics} \quad (1)$$

where Y is the outcome; X is a vector of controls, that includes the gender dummy F, paternal and maternal education, the number of siblings, school enrolment and the interactions of these variables with the gender dummy; α_c is a vector of cohort dummies, that we also interact with gender; β_s is a vector of school dummies and of school specific linear trends; SG is the share of female schoolmates; SH is the average parental background of schoolmates; ε is the error term and the indices *i*, *c* and *s* are for the individual, the cohort and the school respectively. We estimate Eq. (1) by ordinary least squares and cluster standard errors at the school level.[10]

As in Lavy and Schlosser, 2011, by including school fixed effects we control for time invariant determinants of endogenous sorting of students across schools. Since in Denmark most students attend primary and secondary education in the same school, school choice decisions are made early on.

[10] Average enrolment by school and cohort is close to 47 students. To increase precision, we treat the effects of school dummies, school specific trends and of the variables in vector X in Eq. (1) as invariant across genders. However, we also present as robustness the results of models estimated separately by gender.

Additionally, school admission policies are based on the place of residence. These institutional features lend support to the school fixed effects model, because it is unlikely that parents react to transitory school quality shocks correlated with *SG* and *SH* by changing their place of residence.

We also include in our specification linear school specific cohort trends, which control for school specific and time varying unobserved factors that could correlate with *SG* and *SH* and affect individual outcomes. In our model, identification hinges upon the presence of cohort specific "jumps" in *SG* and *SH* from each school specific long run trend, most likely induced by random differences across cohorts in the demographic composition of the population residing in the catchment area of each school.

If this identification strategy is valid, peer characteristics *SG* and *SH* are "as good as random", and should therefore be uncorrelated with predetermined characteristics such as gender, parental education and the number of siblings. We report in Table A1 the results of regressing these characteristics on *SG*, *SH*, school enrolment, cohort and school dummies and school specific trends. We find that *SG* affects negatively the gender dummy *F* (female) and that *SH* affects negatively parental education.

With random assignment, however, these negative correlations are only to be expected. To illustrate, assume that all schools have 100 pupils and that each school has 50 girls and 50 boys, independently of the predetermined characteristics of boys and girls. While for each boy the share of girl schoolmates in the school is 50/99, for each girl this share is 49/99. A similar argument holds for the parental background of schoolmates.

A potential drawback of our strategy is that, once school and year fixed effects, school specific trends and individual covariates are controlled for, there is little remaining variation in the peer variables. We address this concern by regressing peer characteristics on the covariates and by computing the associated R squared, that turns out to be equal to 20 percent for *SG* and to

83 percent for *SH*.[11] We conclude that there is still enough remaining variation in the peer variables to guarantee that peer effects are identified after conditioning for school and year effects, school specific trends and individual covariates.

The model of equation (1) illustrates the sources of variation that we use for identifying the effects of school peers characteristics. However, our main interest in this paper is to investigate whether interactions at home – involving individuals, their parents and siblings - may affect the direction and size of the effects of interactions at school. We therefore extend the model of equation (1) to include family characteristics and estimate

$$\begin{aligned} Y_{ics} &= \alpha_c + \beta_s + \lambda X_{ics} + \pi_1 SG_{cs} + \pi_2 SH_{cs} + \gamma_1 SG_{cs} * GF_{ics} + \gamma_2 SH_{cs} * FH_{ics} + \\ &+ \theta_1 SG_{cs} * F_{ics} + \theta_2 SH_{cs} * F_{ics} + \theta_3 SG_{cs} * GF_{ics} * F_{ics} + \theta_4 SH_{cs} * FH_{ics} * F_{ics} + \\ &+ \varepsilon_{f(i)} + \eta_{ics} \end{aligned} \tag{2}$$

that adds to (1) the interactions of *SG* with a dummy indicating the presence of a sister at home (*FG*) and of *SH* with a dummy indicating whether at least one parent has higher education (*FH*). These additional variables are further interacted with the gender dummy.

Since parental background *FH* and the share of sisters in the family are not randomly allocated but are likely to be correlated with observable and unobservable family characteristics, we estimate (2) by including not only school but also family fixed effects $\varepsilon_{f(i)}$ (where *f(i)* denotes the family of person *i*), which capture all these characteristics. We estimate Eq. (2) by ordinary least squares and use two-way clustered standard errors, by school and by family (see Cameron, Gelbach and Miller, 2011). There are close

[11] Given that school allocation mechanisms are residence-based, it is not surprising that school fixed effects explain a larger share of the variation in parental background than in gender composition.

315,000 families with more than one child in our data, and the average number of children among them is 2.6.

5. Results

We organize the presentation of results in three sub-sections. In the first sub-section, we discuss the estimated effects of peer characteristics at school on the selected outcomes. In the second sub-section, we consider whether and how family interactions affect the impact of school interactions on these outcomes. In the final brief sub-section, we discuss a sensitivity to our baseline estimates.

5.1 Average gender and background peer effects

Table 2 reports our estimates of Eq. (1), separately for males and females. For each gender, we show the estimated effects on the selected outcomes of the share of female schoolmates *SG* and of the average school parental background *SH* in the 9^{th} grade of compulsory education.

Consider first the effects of *SG*. For males, we only detect a statistically significant, small and positive effect on the probability of having vocational education as the highest degree. For females, we find that adding ten percent to the share SG increases completed years of education by 0.0075, a tiny effect that is only statistically significant at the 10 percent level of confidence.[12]

Turning to the effects of *SH*, we find that a higher share of "privileged" schoolmates – with at least one parent having higher education - affects males by increasing educational attainment and lifetime earnings and by reducing the percent having vocational education as their highest degree. The size of the estimated effects is small. For example, we estimate that a ten percentage points increase in *SH* (equal to two-thirds of a standard deviation) increases years of completed education and the probability of completing tertiary

[12] There is also a small negative effect on the conditional probability of choosing law and social science at college.

education by 0.015 years and by 0.36 percentage points, and real pre-tax earnings by 0.22 percent.

For females, we find no statistically significant effect of *SG* and *SH* on educational attainment and labor market outcomes and small effects on the choice of college field.[13] In particular, we estimate that adding ten percentage points to *SH* increases the probability of choosing humanities by slightly less than one percentage point and reduces the probability of choosing law and social science by about half a percentage point.

Our findings that average parental background in the school increases educational attainment and earnings for males but not for females are similar to the results by Black, Devereux and Salvanes, 2013, on the effects of average log father's earnings. Furthermore, the size of our estimated effects is of the same order of magnitude as theirs. Conversely, we do not confirm for our data either their finding that a higher share of female schoolmates reduces the educational attainment of males or that it increases female earnings. Overall, our evidence suggests that, independently of gender, this share has little impact on education and labor market outcomes, especially for males.

In the estimates of Table 2, school fixed effects and school specific trends are assumed to be invariant across genders, mostly to increase precision. As a robustness, we relax this assumption by estimating separate regressions by gender. Overall, the results reported in Table A2 in the Appendix are

[13] We wish to estimate the effects of *SG* and *SH* on the probability of choosing a field of study at college, *conditional* on college attendance. Focusing only on the selected sub-sample of college students is problematic if there are effects of SG or SH on college attendance (the "conditional on positive" problem defined by Angrist and Pischke, 2008). We proceed as follows. Let F be the field in question, and C be college attendance. We are interested in $\partial \mathrm{P}(\mathrm{F} \mid \mathrm{C}) / \partial \mathrm{SC}$, C=G, H. Using the definition of conditional probability, P(F|C)= P(F,C)/P(C), we obtain $\partial \mathrm{P}(\mathrm{F} \mid \mathrm{C}) / \partial \mathrm{SC} = [\partial \mathrm{P}(\mathrm{F,C}) / \partial \mathrm{SC}] * \mathrm{P}(\mathrm{C}) - [\partial \mathrm{P}(\mathrm{C}) / \partial \mathrm{SC}] * \mathrm{P}(\mathrm{F,C}) / \mathrm{P}(\mathrm{C})^2$. The right hand side is identified because we observe in the data both P(F,C) and P(C), and can estimate the effects of SC on P(F,C) and P(C) using simple regressions. We do inference using the delta method, assuming zero covariance between the estimated effects of SC on P(F,C) and P(C).

comparable with those in Table 2, but generally have higher standard errors. We therefore retain the pooled specification in analysis that follows.[14]

5.2 *Do school peer effects vary with family characteristics?*

The estimation of Eq. (2) adds family to cohort and school fixed effects and therefore excludes from the sample all individuals without any sibling, for whom there is no within family variation. For comparison, we replicate in Table A3 for the smaller sample the estimates of Table 2 and show that, although precision declines, qualitatively results are similar.

We present the estimates of Eq. (2) in Table 3 for males and in Table 4 for females. Each table is organized in six columns. Columns (1) and (2) show the estimated effects of the share of female schoolmates on the selected outcomes for individuals without and with at least one sister in the household. Column (3) indicates the differences between the coefficients reported in the previous two columns. Columns (4) and (5) show the effects of the share of "privileged" schoolmates, and Column (6) indicates the differences between coefficients.

In the case of males (Table 3), we find that the impact of the share of female schoolmates never varies in a statistically significant way with the presence of a sister at home. In addition, the significant effect of SG on vocational education as the highest degree does not survive in either of the two sub-samples. On the other hand, the impact of the share of "privileged" schoolmates varies in a statistically significant way (at the 5 or 10 percent level of confidence) with the presence of a higher educated parent in the family for completed years of education, the probability of having a vocational degree and long-term real earnings.

[14] Despite the higher standard errors, we notice that in the separate regressions the effect of SG on years of education is for females statistically significant at the 5 percent level of confidence.

In particular, we find that the effect of a higher *SH* on years of education and lifetime earnings is positive and statistically significant (at the 5 or 10 percent level of confidence) for the “disadvantaged”, and not statistically different from zero for the “privileged”. On the other hand, interacting with more “privileged” schoolmates reduces in a statistically significant way the probability that “privileged” males have vocational education as their highest degree. For the “disadvantaged”, the effect is also negative but imprecisely estimated.

In the case of females (Table 4), we find that the presence or absence of sisters in the family induces statistically significant differences in the impact of female schoolmates on years of completed education, the choice of field of study at college and employment at age 31 to 40. While the differences in coefficients are statistically significant, the coefficients themselves are mostly imprecise, perhaps with the exception of the one associated to employment in the age range 31 to 40 for females with at least one sister, which turns out to be positive and statistically significant at the 10 percent level of confidence.

Interestingly, this effect is negative for males with at least one sister. To explain these findings, we argue that the combination of many female schoolmates and the presence of sisters in the family may attenuate feminine stereotypes and positively affect competitiveness and risk taking, with positive effects on long term employment prospects. On the other hand, males with sisters may be not only less disruptive but also acquire more feminine traits, for instance in terms of lower competitiveness and higher risk aversion. Interacting with a higher share of girls at school could foster these traits, with negative labor market consequences.

For females, we also find that the effects of the share of “privileged” schoolmates vary with individual parental background only for the choice of the college field. In particular, a higher share reduces the probability of choosing scientific fields (STEM) and increases the probability of choosing humanities and arts for those with good parental background. For the "disadvantaged", it decreases instead the probability of attending a health, law or social sciences degree, and increases their odds of attending STEM, humanities and arts.

5.3 *A Sensitivity*

The idea that interacting with a sister or a brother in the family may affect interactions at school requires that age spacing within the family is not too broad. We have computed the dummy FG in Tables 3 and 4 using information on completed fertility within a mother-father couple and disregarding spacing. However, younger siblings may not have yet been born when older ones attended grade 9 – when we measure the variable *SG* – and it is questionable whether sisters who have large spacing with their siblings can generate meaningful within-family interactions.

To take these considerations into account, we have re-defined the dummy FG using the gender composition of siblings when the individuals was aged 15 and excluding sisters who were at the time more than 10 years younger or older than the individual. However, it turns out that this correction is empirically negligible and does not affect our empirical results.[15]

6. Conclusions

We have estimated school peer effects on educational achievement and labor market outcomes using Danish register data, asking whether these effects vary with family characteristics, including the composition of siblings and parental background.

We have found that better average parental background in the school increases educational attainment and earnings for males but not for females, and that a higher share of female schoolmates increases the educational attainment of girls but have no statistically significant effect on the attainment of boys. The

[15] Results available from the authors upon request.

positive effect of a higher share of “privileged” schoolmates on years of education and lifetime earnings does not apply to all, but only to “disadvantaged” students. For females, we find instead that having a sister in the family significant affects the impact of the share of female schoolmates on years of completed education and employment at age 31 to 40.

Our data suggest that females with at least one sister are more likely to be employed if they have experienced a higher share of female schoolmates in the 9^{th} grade. The opposite is true for males with at least one sister. Boys and girls without any sister are virtually unaffected by the share of female schoolmates. We explain these findings by arguing that the combination of many female schoolmates and the presence of sisters in the family may attenuate feminine stereotypes and positively affect competitiveness and risk taking, with positive effects on long term employment prospects. On the other hand, males with sisters may be not only less disruptive but also acquire more feminine traits, for instance in terms of lower competitiveness and higher risk aversion. Interacting with a higher share of girls at school could foster these traits, with negative labor market consequences.

Our results suggest that the benefits and costs of single-sex schools are likely to be small and to vary not only with individual gender but also with the gender composition of siblings. They also indicate that attracting students from privileged background produces small benefits for (male) “disadvantaged” students in terms of education and earnings, without having statistically significant consequences on the “privileged”. Peers’ parental background appears to exert a remedial effect for the lack of own parental human capital, which highlights the existence of non-negligible intergenerational social returns of education and supports social mixing in schools.

References

Ammermueller, A., & Pischke, J. S. (2009). Peer effects in European primary schools: Evidence from the progress in international reading literacy study. *Journal of Labor Economics*, *27*(3), 315-348.

Anelli, M., & Peri, G. (2017). The long run effects of high-school class gender composition. *Economic Journal*, forthcoming.

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Becker, G. S., & Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of political Economy*, *87*(6), 1153-1189.

Bharadwaj, P., Dahl, G. B., & Sheth, K. (2014). Gender discrimination in the family. *The economics of the family: How the household affects markets and economic growth*, *2*, 237.

Bhuller, M., Mogstad, M., & Salvanes, K. G. (2011). Life cycle earnings, education premiums, and internal rates of return. *Journal of Labor Economics*, forthcoming.

Bingley, P., Cappellari, L., & Tatsiramos, K. (2014). *Family, Community and Long-Term Earnings Inequality*. IZA DP n. 10089.

Björklund, A., & Jäntti, M. (2009). Intergenerational income mobility and the role of family background. *Oxford Handbook of Economic Inequality*, Oxford University Press, Oxford, 491-521.

Black, S. E., & Devereux, P.J. (2011) Recent Developments in Intergenerational Mobility, in *Handbook of Labor Economics*, Orley Ashenfelter and David Card, eds, North Holland Press.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2013). Under pressure? The effect of peers on outcomes of young adults. *Journal of Labor Economics*, *31*(1), 119-153.

Booij, A. S., Leuven, E., & Oosterbeek, H. (2017). Ability peer effects in university: Evidence from a randomized experiment. *The Review of Economic Studies*, *84*(2), 547-578.

Booth, A. L., & Nolen, P. (2012). Gender differences in risk behaviour: Does nurture matter? *The Economic Journal*, *122*(558).

Brim, O. G. (1958). Family structure and sex role learning by children: A further analysis of Helen Koch's data. *Sociometry*, *21*(1), 1-16.

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

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, *29*(2), 238-249.

Ciccone, A., & Garcia-Fontes, W. (2015). Peer Effects when Girls and Boys Learn Together, a Birth Cohort Approach. *Mimeo*, University of Barcelona GSE.

Conley, D. (2000). Sibship sex composition: Effects on educational attainment. *Social Science Research*, *29*(3), 441-457.

Dahl, G. B., & Moretti, E. (2008). The demand for sons. *The Review of Economic Studies*, *75*(4), 1085-1120.

Doris, A., O'Neill, D., & Sweetman, O. (2013). Gender, single-sex schooling and maths achievement. *Economics of Education Review*, *35*, 104-119.

Eisenkopf, G., Hessami, Z., Fischbacher, U., & Ursprung, H. W. (2015). Academic performance and single-sex schooling: Evidence from a natural experiment in Switzerland. *Journal of Economic Behavior & Organization*, *115*, 123-143.

Feld, J., & Zoelitz, U. (2017). The Effect of Peer Gender on Major Choice and Occupational Segregation. *Mimeo*, IZA.

Fryer, R. G., & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, *2*(2), 210-240.

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, *118*(3), 1049-1074.

Haider, S.,& Solon, G. (2006). Life-Cycle Variation in the Association between Current and Lifetime Earnings. American Economic Review, 96(4): 1308-1320.

Hoxby, C. (2000). *Peer effects in the classroom: Learning from gender and race variation*. National Bureau of Economic Research WP n.7867.

Koch, H. L. (1955). Some personality correlates of sex, sibling position, and sex of sibling among five-and six-year-old children. *Genetic Psychology Monographs*.

Lavy, V., Paserman, M. D., & Schlosser, A. (2012a). Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *The Economic Journal*, *122*(559), 208-237.

Lavy, V., & Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, *3*(2), 1-33.

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

Lazear, E. P. (2001). Educational production. *The Quarterly Journal of Economics*, 116(3), 777-803.

Malamud, O., Pop-Eleches, C., & Urquiola, M. (2016). *Interactions Between Family and School Environments: Evidence on Dynamic Complementarities?* National Bureau of Economic Research WP n. 22112

Mazumder, B. (2005). Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data. *The Review of Economics and Statistics*, *87*(2), 235-255.

Nybom, M., & Stuhler, J. (2016). Heterogeneous income profiles and life-cycle bias in intergenerational mobility estimation. *Journal of Human Resources*, 51 (1): 239-68.

Oosterbeek, H., & Van Ewijk, R. (2014). Gender peer effects in university: Evidence from a randomized experiment. *Economics of Education Review*, *38*, 51-63.

Oreopoulos, P., & Salvanes, K. G. (2011). Priceless: The nonpecuniary benefits of schooling. *The journal of economic perspectives*, *25*(1), 159-184.

Park, S., 2015. Gender Peer Effects: Evidence from a Quasi-Random Classroom Assignment Policy. *Mimeo*, Northwestern University.

Peter, N., Lundborg, P., & Webbink, D. H. (2015). *The Effect of Sibling's Gender on Earnings, Education and Family Formation.* IZA Discussion Paper n. 9128.

Rosenberg, B. G., & Sutton-Smith, B. (1968). Family interaction effects on masculinity-femininity. *Journal of Personality and Social Psychology*, *8*(2p1), 117.

Rossin-Slater, M., & Wüst, M. (2015). Are different early investments complements or substitutes? long-run and intergenerational evidence from Denmark. *Mimeo,* Research Institute of Industrial Economics.

Schøne, P., von Simson, K., & Strøm, M. (2016). *Girls Helping Girls-The Impact of Female Peers on Grades and Educational Choices*. IZA Discussion Paper n. 10586.

Solon, G. (1992). Intergenerational income mobility in the United States. *The American Economic Review*, 393-408.

Solon, G. (1999). Intergenerational mobility in the labor market. *Handbook of labor economics*, *3*, 1761-1800.

Steelman, L. C., Powell, B., Werum, R., & Carter, S. (2002). Reconsidering the effects of sibling configuration: Recent advances and challenges. *Annual Review of Sociology*, *28*(1), 243-269.

Stoneman, Z., Brody, G. H., & MacKinnon, C. E. (1986). Same-sex and cross-sex siblings: Activity choices, roles, behavior, and gender stereotypes. *Sex Roles*, 15(9), 495-511.

Whitmore, D. (2005). Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment. *American Economic Review*, 199-203.

Tables and figures

Table 1. Descriptive statistics

Variables	Males - means	Males - standard deviations	Females - means	Females - standard deviations
Years of education	12.944	(2.28)	13.113	(2.20)
Completed tertiary education	0.303	(0.45)	0.383	(0.48)
Highest degree is vocational	0.649	(0.48)	0.621	(0.49)
Tertiary STEM (unconditional)	0.088	(0.28)	0.030	(0.17)
Tertiary Humanities (unconditional)	0.049	(0.21)	0.054	(0.22)
Tertiary Health and Related Fields (unconditional)	0.051	(0.22)	0.221	(0.42)
Tertiary Law and Social Sciences (unconditional)	0.088	(0.28)	0.072	(0.26)
Average log income age 31 to 40	12.781	(0.46)	12.462	(0.38)
Employed at least 5 years in age range 31 to 40	0.848	(0.36)	0.856	(0.35)
At least one sister in the family	0.546	(0.49)	0.527	(0.50)
At least one parent with higher education	0.239	(0.42)	0.231	(0.42)
Years of education of mother	10.321	(3.15)		
Years of education of father	11.166	(3.29)		
% of girls in school and cohort (SG)	0.489	(0.09)		
% with at least one higher educated parent in school and cohort (SH)	0.238	(0.15)		
Enrollment in school and cohort	47.044	(21.28)		

Note: the total number of observations is 513,485 for males and 496,439 for females. The number of observations for average log income age 31 to 40 is 435,717 for males and 425,162 for females.

Table 2. Estimated effects of SG and SH on school and labor market outcomes. By gender. With school and cohort dummies, no family fixed effects.

	(1)	(2)	(3)	(4)
Gender	Male	Female	Male	Female
Peer variable	SG	SG	SH	SH
Dependent Variable				
Years of education	0.038	0.075*	0.147***	0.041
	(0.042)	(0.041)	(0.044)	(0.043)
Completed tertiary education	0.001	0.011	0.036***	0.008
	(0.009)	(0.009)	(0.010)	(0.009)
Highest degree is vocational	0.018**	0.002	-0.030***	-0.007
	(0.009)	(0.009)	(0.010)	(0.009)
STEM at tertiary (conditional)	0.013	0.004	-0.015	-0.027
	(0.014)	(0.014)	(0.014)	(0.018)
Humanities at tertiary (conditional)	-0.012	-0.001	-0.012	0.086***
	(0.011)	(0.015)	(0.013)	(0.016)
Health at tertiary (conditional)	0.000	0.035	-0.027	-0.023
	(0.018)	(0.027)	(0.023)	(0.026)
Law and Social at tertiary (conditional)	-0.001	-0.037*	0.065***	-0.052***
	(0.015)	(0.019)	(0.015)	(0.020)
Average log income, age 31-40	0.002	0.002	0.022**	0.008
	(0.009)	(0.008)	(0.009)	(0.009)
Employed at least 5 years in age range 31 to 40	-0.004	0.005	0.003	-0.002
	(0.007)	(0.006)	(0.007)	(0.007)

Note: each regression includes school dummies, school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, the education of the father and the mother and the number of siblings. Standard errors clustered by school. Total number of observations: 1,009,924. Observations for average log income 31-40: 860,879. ***: p<.01; **: p<.05; *:p<.10

Table 3. Estimated effects of SG and SH on school and labor market outcomes for males. With school and cohort dummies, family fixed effects and interactions of SG with FG and of SH with FH.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	SG*(1-FG)	SG*FG	Difference (2)-(1)	SH*(1-FH)	SH*FH	Difference (4)-(3)
Years of education	-0.046	-0.093	-0.046	0.185***	0.016	-0.168*
	(0.088)	(0.064)	(0.106)	(0.063)	(0.088)	(0.090)
Completed tertiary education	-0.024	-0.020	0.004	0.052***	0.030	-0.022
	(0.019)	(0.013)	(0.023)	(0.013)	(0.020)	(0.019)
Highest degree is vocational	0.032	-0.002	-0.034	-0.020	-0.057***	-0.037*
	(0.021)	(0.016)	(0.025)	(0.015)	(0.021)	(0.021)
STEM at tertiary (conditional)	-0.011	0.021	0.032	-0.017	-0.025	-0.008
	(0.033)	(0.023)	(0.039)	(0.025)	(0.021)	(0.027)
Humanities at tertiary (conditional)	-0.000	-0.006	-0.006	-0.016	-0.015	0.001
	(0.027)	(0.020)	(0.032)	(0.022)	(0.022)	(0.025)
Health at tertiary (conditional)	0.052	0.012	-0.040	-0.019	-0.010	0.009
	(0.039)	(0.035)	(0.051)	(0.042)	(0.031)	(0.042)
Law and Social at tertiary (conditional)	-0.018	-0.028	-0.010	0.072***	0.055**	-0.017
	(0.035)	(0.024)	(0.041)	(0.027)	(0.024)	(0.031)
Average log income, age 31-40	0.004	-0.011	-0.015	0.025*	-0.022	-0.047**
	(0.020)	(0.014)	(0.024)	(0.014)	(0.021)	(0.021)
Employed at least 5 years in age range 31 to 40	-0.001	-0.024*	-0.023	0.002	0.000	-0.002
	(0.016)	(0.013)	(0.020)	(0.011)	(0.016)	(0.015)

Note: each regression includes family and school dummies, school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, the education of the father and the mother and the number of siblings. Estimation is carried out jointly for males and females. The table reports the estimated marginal effects for females. Standard errors clustered by school and family. Total number of observations: 725,722 of both genders (369,471 males). Observations for average log income 31-40: 565,284 for both genders (287,084 males)***: p<.01; **: p<.05; *:p<.10

Table 4. Estimated effects of SG and SH on school and labor market outcomes for females. With school and cohort dummies, family fixed effects and interactions of SG with FG and of SH with FH.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	SG*(1-FG)	SG*FG	Difference (2)-(1)	SH*(1-FH)	SH*FH	Difference (4)-(3)
Years of education	-0.007	0.052	0.059**	-0.112*	-0.069	0.043
	(0.057)	(0.053)	(0.030)	(0.061)	(0.089)	(0.088)
Completed tertiary education	0.000	0.012	0.012	0.003	-0.019	-0.022
	(0.013)	(0.012)	(0.006)	(0.014)	(0.020)	(0.019)
Highest degree is vocational	-0.006	-0.008	-0.002	-0.031**	-0.026	0.005
	(0.014)	(0.013)	(0.007)	(0.015)	(0,021)	(0.021)
STEM at tertiary (conditional)	0.028	0.001	-0.027***	0.110	-0.110***	-0.220***
	(0.024)	(0.025)	(0.013)	(0.037)	(0.025)	(0.037)
Humanities at tertiary (conditional)	0.005	-0.008	-0.013	0.084***	0.072***	-0.012
	(0.021)	(0.022)	(0.011)	(0.031)	(0.026)	(0.033)
Health at tertiary (conditional)	-0.047	0.020	0.067***	-0.104**	0.045	0.149***
	(0.036)	(0.038)	(0.018)	(0.049)	(0.033)	(0.048
Law and Social at tertiary (conditional)	0.004	-0.004	-0.008	-0.116***	-0.023	0.093**
	(0.027)	(0.029)	(0.014)	(0.040)	(0.029)	(0.041)
Average log income, age 31-40	0.007	0.010	0.003	0.013	-0.010	-0.023
	(0.012)	(0.012)	(0.006)	(0.013)	(0.020)	(0.021)
Employed at least 5 years in age range 31 to 40	-0.005	0.017*	0.022***	0.002	0.014	0.012
	(0.011)	(0.010)	(0.005)	(0.011)	(0.016)	(0.017)

Note: each regression includes family and school dummies, school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, the education of the father and the mother and the number of siblings. Estimation is carried out jointly for males and females. The table reports the estimated marginal effects for females. Standard errors clustered by school and family. Total number of observations: 725,722 of both genders (356,251 females). Observations for average log income 31-40: 565,284 of both genders (287,200 females). ***: p<.01; **: p<.05; *:p<.10

Figure 1. Kernel density estimate of the share of female schoolmates in the school and grade (SG)

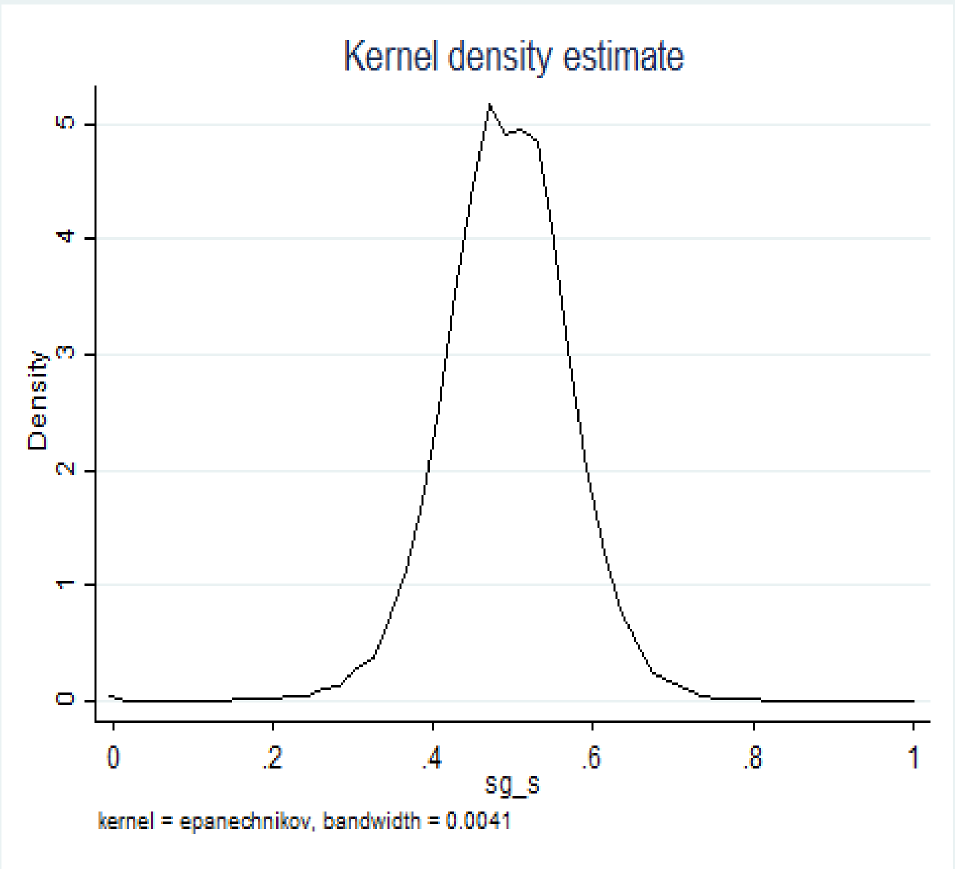
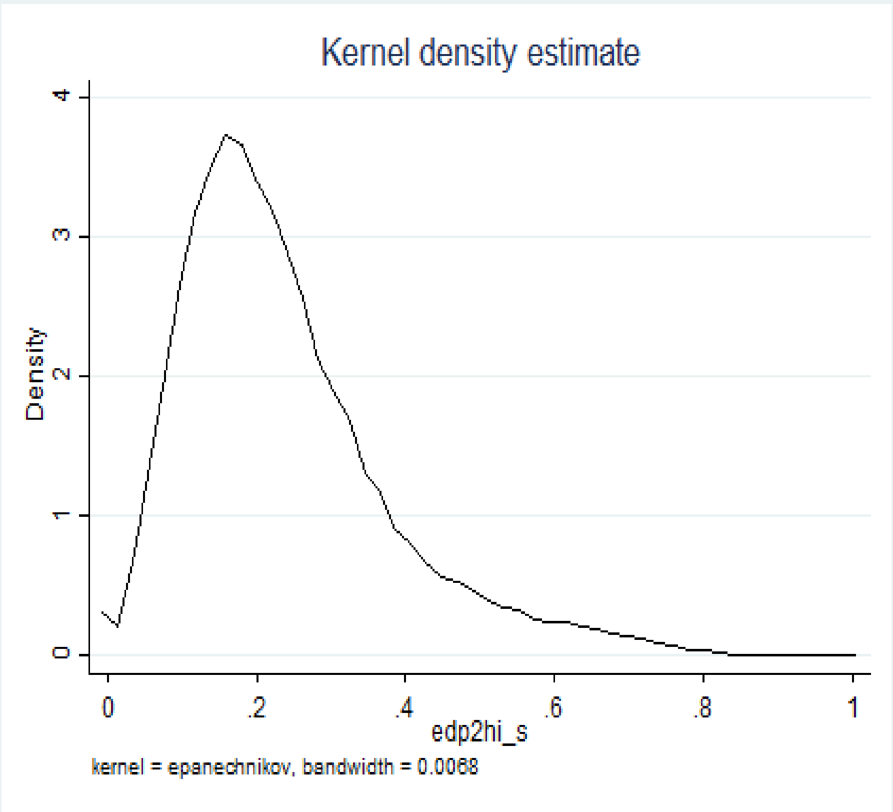


Figure 2. Kernel density estimate of the share of students in the school and grade with at least one higher educated parent (SH)



Appendix

Table A1. Balancing tests. Reverse regressions of individual covariates on the peer variables SG and SH.

	(1)	(2)
Peer variable	SG	SH
Dependent variable		
Female	-0.100***	0.009
	(0.011)	(0.010)
Mother's years of schooling	0.080*	-0.202***
	(0.042)	(0.065)
Father's years of schooling	0.049	-0.231***
	(0.044)	(0.069)
Number of siblings	0.014	-0.028*
	(0.013)	(0.016)

Note: each regression includes school enrolment, cohort and school dummies and school-specific trends. Standard errors are clustered by school. Total number of observations: 1,009,924. ***: p<.01; **: p<.05; *:p<.10

Table A2. Estimated effects of SG and SH on school and labor market outcomes. Models estimated after splitting the sample by gender. With school and cohort dummies, no family fixed effects.

	(1)	(2)	(3)	(4)
Gender	Male	Female	Male	Female
Explanatory variable	SG	SG	SH	SH
Dependent Variable				
Years of education	0.037	0.086**	0.111**	0.079
	(0.043)	(0.043)	(0.056)	(0.053)
Completed tertiary education	0.000	0.013	0.023**	0.022*
	(0.009)	(0.010)	(0.011)	(0.012)
Highest degree is vocational	0.017*	0.004	-0.018	-0.020*
	(0.09)	(0.009)	(0.012)	(0.012)
STEM at tertiary (conditional)	0.014	0.011	-0.052*	-0.007
	(0.020)	(0.009)	(0.027)	(0.012)
Humanities at tertiary (conditional)	-0.013	-0.002	0.019	0.040**
	(0.015)	(0.012)	(0.019)	(0.017)
Health at tertiary (conditional)	0.009	0.007	0.011	-0.039
	(0.015)	(0.026)	(0.020)	(0.033)
Law and Social at tertiary (conditional)	-0.005	-0.022	0.030	-0.000
	(0.021)	(0.015)	(0.025)	(0.018)
Average log income, age 31-40	0.003	0.002	0.025**	0.006
	(0.009)	(0.008)	(0.01)	(0.010)
Employed at least 5 years in age range 31 to 40	0.001	0.000	0.000	0.000
	(0.007)	(0.007)	(0.009)	(0.009)

Note: split sample estimation by gender. Each regression includes cohort and school dummies, school specific trends, school enrolment, the education of the father and the mother and the number of siblings. Standard errors clustered by school. Total number of observations: 513,485 for males and 496,439 for females. Observations for average log income 31-40: 435,717 for males and 425,162 for females. ***: p<.01; **: p<.05; *:p<.10

Table A3. Estimated effects of SG and SH on school and labor market outcomes. By gender. With school and cohort dummies, and family fixed effects.

	(1)	(2)	(3)	(4)
Gender	Male	Female	Male	Female
Explanatory variable	SG	SG	SH	SH
Dependent Variable				
Years of education	-0.077	0.031	0.123**	-0.083
	(0.053)	(0.052)	(0.058)	(0.058)
Completed tertiary education	-0.022*	0.007	0.047***	-0.003
	(0.011)	(0.012)	(0.013)	(0.013)
Highest degree is vocational	0.009	-0.008	-0.033**	-.026*
	(0.013)	(0.013)	(0.014)	(0.014)
Tertiary STEM (conditional)	0.010	0.012	-0.005	-0.012
	(0.020)	(0.024)	(0.020)	(0.027)
Tertiary Humanities (conditional)	-0.004	-0.003	0.022	0.086***
	(0.017)	(0.021)	(0.018)	(0.024)
Tertiary Health (conditional)	0.026	-0.005	-0.026	-0.019
	(0.027)	(0.036)	(0.032)	(0.036)
Tertiary Law and Social (conditional)	-0.024	-0.01	0.067***	-0.076**
	(0.021)	(0.027)	(0.022)	(0.030)
Average log income 31-40	-0.006	0.009	0.011	0.008
	(0.012)	(0.011)	(0.013)	(0.012)
Employed at least 5 years in age range 31 to 40	-0.017	0.009	0.000	0.007
	(0.010)	(0.010)	(0.007)	(0.010)

Note: each regression includes family and school dummies, school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, the education of the father and the mother and the number of siblings. Standard errors clustered by school and family. Total number of observations: 725,722. Observations for average log income 31-40: 565,284. ***: p<.01; **: p<.05; *:p<.10