I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11200

Parents, Siblings and Schoolmates: The Effects of Family-School Interactions on Educational Achievement and Long-Term Labor Market Outcomes

Marco Bertoni
Giorgio Brunello
Lorenzo Cappellari

DECEMBER 2017

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11200

Parents, Siblings and Schoolmates: The Effects of Family-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, IZA and CESifo

DECEMBER 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Parents, Siblings and Schoolmates: The Effects of Family-School Interactions on Educational Achievement and Long-Term Labor Market Outcomes*

We use Danish register data to investigate whether the effects of schoolmates' gender and average parental education on individual educational achievement, employment and earnings vary with individual family characteristics such as the gender of siblings and own parental education. We find that boys with sisters have worse employment prospects than boys with no sisters when exposed to a higher share of girls at school. The opposite is true for girls who have sisters. We also show that the benefits from exposure to "privileged" peers accrue mainly to "disadvantaged" students. These benefits decline when the dispersion of parental education increases. Overall, the size of the estimated effects is small.

JEL Classification: I21, J16, J24

Keywords: education peer effects, gender, parental background, human capital production, long term outcomes

Corresponding author:
Marco Bertoni
Department of Economics and Management
University of Padova
via del Santo 33
35123 Padova
Italy
E-mail: marco.bertoni@unipd.it

* We are very grateful to Paul Bingley for constant help with the data. We also thank Victor Lavy, Michele Pellizzari, Olmo Silva and seminar participants at Catania (SIEP 2017), Copenhagen (SFI), Cosenza (AIEL 2017), Maastricht, Mannheim (ZEW), 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". Lorenzo Cappellari acknowledges financial support by the Danish Council for Independent Research (grant DFF-6109-00226). The usual disclaimer applies.

1. Introduction

Recent empirical research has shown that social interactions at school can affect individual 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 in Israel (Lavy and Schlosser, 2011) and that the average earnings of peers' fathers matter for the education and labor market outcomes of boys in Norway (Black, Devereux and Salvanes, 2013; see also Ammermueller and Pischke, 2009).

Understanding the effects of school peers 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. On the other hand, the presence and nature of school peer effects shed light on the consequences of de-segregating schools. Assuming that parental education is a relevant measure of family background, they also inform about the intergenerational social returns to education, an area where relatively little is known to date (see Oreopoulos and Salvanes, 2011).

A potentially relevant but less studied issue in this area of research is whether the effects of peer interactions at school on individual outcomes are influenced by earlier interactions occurring within the family and involving both parents and siblings. Do the benefits from interacting with schoolmates having well educated parents accrue mainly to those with a similar "privileged" background or to the "disadvantaged"?[2] Answering this question has implications for 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, does the interaction with at least one sister in the family affect the benefits and costs of having many female schoolmates

[1] Another peer characteristic that has been shown to generate relevant spillover effects is ability (see Lyle, 2009, Lavy et al., 2012a, and Lavy et al., 2012b, Booij et al., 2017). We do not have measures of ability in our data.

[2] In the parlance of this paper, we consider individuals as "privileged" or "disadvantaged" on the basis of parental average education.

at school? Adapting the framework of Cunha and Heckman, 2007, exposure to girls or to a privileged background at home may foster / hamper the effects of later exposure to girls or privileged peers in schools.

In this paper, we study the effects of school-family interactions on educational attainment and long term labor market outcomes. Using Danish register data that contain information on parents, siblings and schoolmates, we investigate whether and how the effects of peer characteristics at school vary with family characteristics involving parents and siblings. We measure the former when individuals are aged 15 (normally attending the 9th grade) with the share of female schoolmates and the average parental background of schoolmates, and the latter with parental education and the presence of at least one sister in the family. The identification of peer effects exploits plausibly random variation in peer composition within schools and between cohorts, controlling both for school - specific trends and for 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.

When we ignore the impact of family characteristics on school peer effects, we find that, in our data, the share of female schoolmates has no statistically significant effect on education or labor market outcomes, contrary to what found by Lavy and Schlosser, 2011, and Black, Devereux and Salvanes, 2013, and independently of whether we control or not for family fixed effects.

We also find that, when family fixed effects are not included in our regressions, “better” average parental background in the school increases both individual educational attainment and lifetime earnings, similarly to the results by Black, Devereux and Salvanes, 2013, who measure background with log father’s earnings. As in Black, Devereux and Salvanes, 2013, the statistically significant effects that we estimate are very small.

Accounting for the interactions of family characteristics and school peer effects significantly affects our estimates. Our empirical evidence suggests that males (females) with at least one sister are less (more) likely to be employed after experiencing a higher share of female schoolmates than individuals with only brothers. On the one 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. On the other hand, females growing up with sisters and interacting with female schoolmates may be less exposed to stereotyped behaviors and less inclined to acquire traditional gender roles, with positive employment effects.

We also show that the effects of schoolmates' average parental education vary with individual parental education. Schoolmates with higher average parental education improve both attained years of education and earnings for "disadvantaged" males, and earnings for "disadvantaged" females, but have effects of opposite sign for "privileged" males and females. Although the estimated effects are small, these results indicate that assigning "disadvantaged" students to schools with high average parental background is likely to contribute to reducing the inequality of outcomes.

In addition, our evidence indicates that individuals are affected not only by the average parental education of their peers, but also by its standard deviation. In particular, we find negative effects of the dispersion of peers' parental education on disadvantaged students. Although our data do not allow us to discriminate among alternative channels, we speculate that this result could be due to loosened social ties or to reduced teacher effectiveness in more heterogeneous groups.

Overall, our results suggest that the benefits and costs of single sex schools are likely to be small, that average peers' parental background can partially compensate for the lack of own parental human capital, and that the effectiveness of de-segregation policies could be lower than expected if they increase within-school inequality.

The remainder of the paper is organized as follows: we review the relevant literature in Section 2 and describe our data and the empirical setup in Sections 3 and 4. Results are presented in Section 5. Conclusions follow.

2. The Literature

Three strands of literature are particularly relevant for our paper. The first 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, Black, Devereux and Salvanes, 2013, and Booij, Leuven and Oosterbeek, 2017.[3] The second strand focuses on 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 the elementary, middle and high schools of Israel on test scores, matriculation status and number of credits earned. They find that a higher proportion of girls in a cohort increases 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 by lower level of disruption and violence at school, improved relationships among students and lessened teacher's fatigue.

In a substantial extension of this literature, Black, Devereux and Salvanes, 2013, investigate the effects of lower secondary school peers, aged between 14 and 16, on

[3] Additional contributions in this area include Hoxby, 2000; Whitmore, 2005; Lyle, 2009; Oosterbeek et al, 2014; Ciccone et al, 2015; Park, 2015; Eisenkopf et al, 2015; Anelli and Peri, 2017; Feld and Zoelitz, 2017 and Schone et al, 2017.

schoolmates in the same grade, who as teenagers are expected to be particularly exposed to peer influences. Rather than test scores, they consider post-school outcomes - including teenage childbearing, educational attainment and average earnings in a three years window. 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, Booij, Leuven and Oosterbeek, 2017, estimate peer effects originating from the experimentally manipulated ability composition of tutorial groups for undergraduate students in economics. Going beyond the standard linear-in-means model of peer effects, they find that the impact of group composition on achievement is captured by the mean and standard deviation of peers' prior ability, their interaction, and interactions with students' own prior ability. Their results suggest that ability tracking is beneficial for low and medium ability students.

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 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 joint investments are made in a family. Focusing on the arguably random gender of the first born child, Dahl and Moretti, 2008, find that siblings with a first born sister have on average lower education than 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 influences 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 and can invest in their 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 school and family environments 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 that they significantly interact. Rossin-Slater and Wust, 2015, on the other hand, 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 which we observe labor incomes between 1989 and 2015. We start with those born in 1958 because it is only from this cohort that linkages to parents (and therefore to siblings) are complete. Also, the birth cohort 1958 is the first being matched to the identifier of the school attended at the end of compulsory education. Our last cohort is 1975 because our earnings data end in 2015 (see below) and we wish to observe earnings until age 40. Overall, 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. We define school peers as individuals aged 15 who are born in the same year and enrolled in the same school, and the share of female schoolmates *SG* as the percentage of females among these peers.

Our definition of *SG* differs from the one used by Black, Devereux and Salvanes, 2013, who consider pupils attending the same grade, and is not exposed to the endogeneity threats induced by parents' strategic choice of school starting age. Nonetheless, our measure 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 and there are very few grade retentions. Since most Danish students complete primary and secondary education in the same school, our measure is a good proxy of peer composition throughout compulsory school.

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 obtain data on parental education, that we use to compute peers' average parental education, *E(PE)*, or the average number of years of education completed by parents. Figure 2 shows the distribution of *E(PE)*, with mean 10.743 and standard deviation 1.23.

We use household information on siblings and parents to compute two indicators of family characteristics: a dummy *FG* equal to 1 if the individual has at least one sister and 0 otherwise, and *PE*, parental education measured as the average number of years of education completed by parents. In our sample, 45.6 percent of individuals have at least one sister, and average parental education is equal to 10.743 years.[4]

Following Black, Devereux and Salvanes, 2013, we select as individual outcomes educational attainment, earnings and employment status. As additional outcomes, we also consider an indicator of whether the highest degree attained by the individual is

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

vocational or academic and the field of study in college.[5] We use tax records to obtain pre-tax real annual labor earnings - or total income from labor - at 2012 prices. The most comprehensive measure of individual earnings over the life cycle is 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), showing that this bias is minimized when considering earnings in the thirties, we proxy lifetime earnings with average earnings between age 31 to age 40.[6] We define "long term" employment status as a dummy equal to 1 if the individual has had at least five valid earnings observations between age 31 and 40, and to 0 otherwise.

The summary statistics of the variables used in the empirical analysis are presented in Table 1. Years of education are on average 12.94 for males and 13.11 for females. Average log earnings between age 31 and 40 are equal to 12.78 for males and to 12.42 for females. Finally, the probability of being employed at least five years between age 31 and 40 is 84.8 percent for males and 85.6 percent for females.[7]

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

[5] 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). 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.

[6] We compute this indicator if at least five valid observations in the age interval are available. See Solon, 1992, and Mazumder, 2005. We only retain measured annual earnings above 35,000 Danish Crowns (about 4,700 euro). Black, Devereux and Salvanes, 2013, use average earnings over a three-year window

[7] 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 field or law and social sciences. Close to 65 percent of males and 62 percent of women have vocational education as their highest degree.

$$Y_i = \pi_1 SG_i + \pi_2 E(PE)_i + X_i'\lambda + \alpha_{c(i)} + \beta_{s(i)} + \gamma_{f(i)} + \varepsilon_i \qquad (1)$$

where Y is the outcome; X is a vector of controls, that includes the gender dummy F, paternal and maternal education, the age of the mother at birth, 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 school specific linear trends; γ_f is a vector of family fixed effects; *SG* is the share of female schoolmates; *E(PE)* is schoolmates' average parental education, that we standardize to have zero mean and unit standard deviation; ε is the error term and the indices i, c, f and s are for the individual, the cohort, the family and the school respectively.

As in Lavy and Schlosser, 2011, we control for the endogenous sorting of students across schools by using school fixed effects. Since in Denmark most students attend primary and secondary education in the same school, school choice decisions are made early on. 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 *E(PE)* 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 *E(PE)* and affect individual outcomes. In our model, identification hinges upon the presence of cohort specific "jumps" in *SG* and *E(PE)* 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. Finally, we control for the correlation between peer and family characteristics as well as for any unobservable determinant of school choice that is family-specific and constant across siblings by including in our models also family fixed effects. There are close 315,000 families with more than one child in our data, and the average number of children among them is 2.6.

Identification in two-way fixed effects models is driven by movers, defined as families with children who do not attend the same school. In our sample, 28 percent of families with more than one child are movers. Furthermore, we have checked that there are movers in each school of the sample. Table A2 compares movers with non-movers and shows that the differences in observed characteristics such as the number of children, the age of the mother at birth, the year of birth of the first child, the average education of parents, albeit statistically significant, are small. We control for these characteristics in our regressions either directly or absorbing them with family fixed effects.

If our identification strategy is valid, peer characteristics *SG* and *E(PE)* are "as good as random", and will therefore be uncorrelated with predetermined characteristics such as gender, parental education, the number of siblings and birth order. We report in Table A1 the results of regressing these characteristics on *SG*, *E(PE)*, school enrolment, cohort and school dummies and school specific trends, with and without family fixed effects. We find that *SG* and *E(PE)* correlate negatively with the gender dummy *F* (female) and with 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, family and year fixed effects, school specific trends and individual covariates are controlled for, there is little remaining variation in the peer variables. This is not the case, however, as the R squared of the regressions of peer characteristics on the covariates is equal to 50 and 90 percent for *SG* and *E(PE)*, suggesting that enough residual variation remains.[8]

Since the main purpose of 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 extend the model of equation (1) to include family characteristics and estimate

[8] These results are in line with Black et al, 2013. 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.

$$Y_i = \pi_1 SG_i + \pi_2 E(PE)_i + \pi_3 SG_i * FG_i + \pi_4 E(PE)_i * PE_i * FG_i + \\ + X_i'\lambda + \alpha_{c(i)} + \beta_{s(i)} + \gamma_{f(i)} + \varepsilon_i \quad (2)$$

that adds to (1) the interactions of *SG* with a dummy indicating the presence of a sister at home (*FG*) and of *E(PE)* with individual parental education (*PE*), that we standardize to have zero mean and unit variance. These additional variables are further interacted with the gender dummy. We do not include *FG* and *PE* in (2), because they are absorbed by the inclusion of family fixed effects. We estimate Equations (1) and (2) by ordinary least squares using two-way clustered standard errors, by school and by family (see Cameron, Gelbach and Miller, 2011).

5. Results

We organize the presentation of results in four sub-sections. First, we discuss the estimated effects of peer characteristics at school on the selected outcomes. Second, we consider whether and how family interactions affect the impact of school interactions on these outcomes. Third, we focus on peers' parental background and examine the effects of both its average and its standard deviation. In the final sub-section, we briefly describe sensitivities.

5.1. Average gender and background peer effects

Table 2 reports parameter estimates for Equation (1), separately for males and females. For each gender, we show the estimated effects of the share of female schoolmates *SG* and of the peers' average parental background *E(PE)* in the 9^{th} grade of compulsory education - normalized to have zero mean and unit standard deviation - on individual years of education, average log earnings between age 31 and 40 and employment for at least five years between age 31 and 40.

We find that *SG* has no statistically significant effect (at the conventional 5 percent level of confidence) on the selected outcomes. As shown by Tables A3 and A4 in the Appendix, this result is not qualitatively altered by omitting family fixed effects or by changing the sample to include also single children - who are omitted when estimating with family fixed effects. Turning to peers' parental education, we estimate that a one standard deviation increase in *E(PE)* rises the years of education completed by males by

0.33 percent (0.043/12.944), a very small effect, and has a negative and imprecisely estimated effect on the years of education completed by females.

A higher value of *E(PE)* also generates higher earnings and employment, but these effects are imprecisely estimated. The comparison of estimates in Table 2 and Table A4 indicates that precision increases when we omit family fixed effects, as done for instance by Lavy and Schlosser, 2011, and Black, Devereux and Salvanes, 2013. When we do so, we estimate that a one standard deviation increase in *E(PE)* has a statistically significant and positive effect on earnings, ranging from 0.4 percent for males to 0.7 percent for females.

The share of girls *SG* does not affect also additional outcomes such as the type of the highest degree (vocational versus academic) or the college field - see Table A.5. On the other hand, a one standard deviation increase in *E(PE)* reduces the probability that the highest degree is vocational by 0.6% for males and by 0.4% for females. A higher *E(PE)* also increases the (conditional) probability that Law and Social Sciences are selected by males and Humanities are chosen by females going to college, and reduces the probability that males enroll in STEM fields and females enroll in Health related fields.[9]

While our findings that average parental background in the school increases the educational attainment of males and the earnings of both males and females are broadly consistent with Black, Devereux and Salvanes, 2013, we cannot confirm their findings that a higher share of female schoolmates reduces the educational attainment of males and increases female earnings. Overall, our evidence suggests that, independently of gender, peers' gender has little impact on education and labor market outcomes.

5.2. Do school peer effects vary with family characteristics?

Our estimates that consider family-school interactions are shown in Tables 3 and 4, where we separately present the results for the two peer characteristics (gender and

---

[9] We estimate the effects of peer characteristics on conditional probabilities as follows. Let F be the field in question, and C be college attendance. We are interested in $\partial P(F \mid C)/\partial SC$, C=G, H. Using the definition of conditional probability, P(F|C)= P(F,C)/P(C), we obtain $\partial P(F \mid C)/\partial SC = [\partial P(F,C)/\partial SC] * P(C) - [\partial P(C)/\partial SC] * P(F,C)/P(C)^2$. 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).

educational background) that are estimated jointly in Equation (2). Table 3 focuses on the interaction between *SG* and *FG* and is organized in six columns. Columns (1) and (2) for males and (4) and (5) for females show the estimated effects of the share of female schoolmates SG on the selected outcomes for individuals without and with at least one sister in the household. Column (3) for males and (6) for females present the differences between the coefficients reported in the previous two columns. Table 4 is devoted to the interaction between *E(PE)* and *PE* and consists of four columns. Columns (1) for males and (3) for females show the estimated effect of the peers' parental education, and columns (2) and (4) indicate the effect of the interaction between *E(PE)* and *PE*, that are both normalized to have zero mean and unit standard deviation.

Table 3 indicates that a higher share of girls at school significantly reduces (increases) employment opportunities between age 31 and 40 only for males (females) who have at least one sister. The impact of female schoolmates on males and females without any sister is virtually zero. A tentative explanation of these findings is that males with sisters may grow up not only as 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. Conversely, females growing up with sisters and interacting with females schoolmates may be less exposed to stereotyped behaviors and less inclined to acquire traditional gender roles, with positive employment effects.

Table 4 shows that the interaction of peers' and individual parental education always attracts a statistical significant coefficient, with the single exception of male employment. We illustrate how the marginal effect of *E(PE)* on years of education, log earnings and employment vary with individual *PE* in Figures 3 to 5 both for males and females.

In the case of males, assignment to a school where peers have a relatively high parental education improves educational attainment, earnings and employment for those with a disadvantaged parental background (*PE* negative) and reduces attainment and earnings for those from a privileged background (*PE* positive). In the case of females, assignment of the disadvantaged to a similar school has positive effects on earnings, negative effects on employment and virtually no effect on achievement. Privileged females, on the other hand, experience positive employment gains but negative effects on attainment and earnings.

These results indicate that de-segregation policies reassigning disadvantaged students to "better" schools benefit these students, especially if they are males. If de-segregation alters the characteristics of peers not only for reassigned students but also for receiving students, by reducing their *E(PE)*, privileged students may also benefit.[10]

5.3. Does the Dispersion of Peer Characteristics Affect Outcomes?

Following Booji, Leuven and Oosterbeek, 2017, we examine whether individual outcomes are affected not only by the average parental education of peers but also by its standard deviation (see also Lyle, 2009). Conditional on the mean, a higher dispersion of individual backgrounds in the class may loosen social ties or reduce teacher effectiveness that is typically higher in more homogeneous classes. Figure 6 shows that the relationship between *E(PE)* and the standard deviation of *PE* within the school, *SD(PE)*, is hump-shaped. With the exception of the extreme values of the mean, there is also substantial variation in *SD(PE)* for each selected *E(PE)*.

We estimate the following equation:

$$Y_i = \rho_1 E(PE)_i + \rho_2 SD(PE)_i + \rho_3 E(PE)_i * PE_i + \rho_4 SD(PE)_i * PE_i + X_i'\lambda + \\ + \alpha_{c(i)} + \beta_{s(i)} + \gamma_{f(i)} + \varepsilon_i \tag{3}$$

and report the results in Table 5. In the first panel of the table, we show the estimated effects of *E(PE)* and its interaction with *PE*. In the second panel, we present estimated coefficients of *SD(PE)* and its interaction with *PE*. As for Table 4, we illustrate both for males and for females how the marginal effect of *SD(PE)* on years of education, log earnings and employment vary with individual *PE* in Figures 7 to 9. *SD(PE)* is normalized to have zero mean and is divided by the standard deviation of *PE* - that sets the scale of the analysis.

Conditional on *E(PE)*, we find that an increase in the dispersion of peers' parental education improves the years of schooling, earnings and employment of males with a

---

[10] Tables A6 and A7 show the estimates of Eq. (3) for a few additional outcomes, including the type of highest degree (vocational or academic) and the field of study in college.

privileged background (*PE* positive) but has negative effects on disadvantaged males (*PE* negative). For females, higher dispersion has mixed effects: it reduces employment and slightly increases earnings for privileged females but reduces earnings and increases employment for the disadvantaged.

In the previous sub-section, we have shown that increasing the average background of peers can improve both educational attainment and the labor market performance of disadvantaged students, especially males. This sub-section indicates that improvements are less likely to occur when a higher average education is accompanied by higher dispersion within the school.

5.4. Sensitivities

We have carried out several sensitivities, none of which affects qualitatively our results. First, so far we have computed the dummy *FG* using information on completed fertility within a mother-father couple. However, younger siblings may not have yet been born when older ones attended grade 9 - when we measure the variable *SG*. To take this into account, we have re-defined the dummy *FG* using the gender composition of siblings when individuals were aged 15. As shown in Table A8, this correction is empirically negligible and does not affect our empirical results.

Second, we have replaced average parental education with the maximum level of education attained by the parents (in Table A9), or with the father's education for males and the mother's education for females (in Table A10), but these changes make no qualitative difference.

Third, to dispel concerns that - given residential-based school assignment mechanisms - our measures of school composition could be capturing neighborhood composition effects, we add to regressions the share of female schoolmates and average parental education in the parish where the individual was living at age 15 - a good approximation for neighborhood composition (see Bingley, Cappellari and Tatsiramos, 2015). Again, our results are not affected - see Table A11.

Finally, as done by Booij et al, 2017, we have estimated Equation (3) by including interactions between *E(PE)* and *SD(PE)*, as well as the triple interaction between *E(PE), SD(PE)* and *PE*, but found that these additions are never statistically significant.[11]

Conclusions

Using Danish register data, we have investigated whether and how the effects of peer characteristics - measured with the share of girls in the school and with the average parental education of schoolmates - on educational achievement and labor market outcomes vary with family characteristics, including the gender composition of siblings and parental education.

In contrast with previous literature in the area, and independently of whether we control or not for family fixed effects in our estimates, we have found that the share of female schoolmates has no statistically significant effect on individual educational attainment and labor market outcomes, measured by average earnings between age 31 and 40 and by employment for at least five years between age 31 and 40.

When family fixed effects are not included in our regressions, as done in previous literature, we have also found that individuals in schools where peers have higher average parental education complete more education and have higher lifetime earnings, similarly to the results by Black, Devereux and Salvanes, 2013, who measure parental background with maternal education or with average log father's earnings. As in Black, Devereux and Salvanes, 2013, the statistically significant effects that we estimate are very small.

One possible reason why measured peer characteristics appear to be ineffective is that they are treated as if they homogeneously affected everyone. Instead, their effect may vary significantly with family characteristics, such as the composition of siblings and average parental education. When the heterogeneity of effects is properly accounted for, we have shown that males (females) with at least one sister are less (more) likely to be employed after experiencing a higher share of female schoolmates. No such effect can be found for males and females without any sister.

[11] Results available from the authors.

We have also found that exposure to schoolmates having higher average parental education increases both attained years of education and earnings among "disadvantaged" males and earnings among "disadvantaged" females, while having effects of opposite sign on "privileged" males and females. Although the estimated effects are small, these results suggest that assigning "disadvantaged" students to schools with high parental background can contribute to reduce the inequality of outcomes.

We have shown that individuals are affected not only by the average parental education of their peers but also by its standard deviation. In particular, an increase in the former that is accompanied by an increase in the latter is less effective in improving the economic outcomes of "disadvantaged" students. Potential mechanisms driving these results include loosened social ties and reduced teacher effectiveness. With the data at hand, however, we are unable to distinguish between these alternatives.

On the one hand, our finding that the share of female schoolmates has virtually no effect on individual performance suggests that the gender composition of peers should not be the driving force behind the choice of single-sex schools. On the other hand, the evidence showing that average peers' parental background can partially compensate for the lack of own parental human capital points to the existence of intergenerational social returns of education and supports social mixing in schools. Our results also indicate that the effectiveness of de-segregation policies could be lower than expected if they increase within-school inequality.

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. (2015). *Family, Community and Long-Term Earnings Inequality*. IZA Discussion Paper 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.

Cunha, F., & Heckman, J. (2007). The technology of skill formation, American Economic Review, Papers and Proceedings, 97(2): 31-47

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.

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.

Lyle, D. S. (2009). The effects of peer group heterogeneity on the production of human capital at West Point. *American Economic Journal: Applied Economics*, *1*(4), 69-84.

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
a) Main Outcomes (by gender)				
Years of education	12.944	(2.28)	13.113	(2.20)
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)
b) *Additional Educational Outcomes (by gender)*				
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)
c) Other variables (full sample)				
At least one sister in the family	0.456	(0.49)		
Average years of education of mother and father (YE)	10.743	(2.79)		
Number of siblings	1.445	(0.95)		
First-born	0.499	(0.50)		
% of girls in school and cohort (SG)	0.489	(0.09)		
Average years of educations of peers' parents $\overline{YE}$	10.734	(1.23)		
SD of years of educations of peers' parents SD(YE)	2.511	(0.25)		
Enrollment in school and cohort	47.044	(21.28)		

Notes: 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 E(PE) on school and labor market outcomes. By gender. With school, cohort and family fixed effects.

	(1)	(2)	(3)	(4)
Peer variable	SG		E(PE)	
Gender	Male	Female	Male	Female
Dependent Variable				
Years of education	-0.083	0.035	0.043***	-0.022*
	(0.053)	(0.052)	(0.010)	(0.011)
Average log income 31-40	-0.007	0.009	0.000	0.004*
	(0.012)	(0.012)	(0.002)	(0.002)
Employed at least 5 years between age 31 and 40	-0.017*	0.009	0.004*	0.001
	(0.010)	(0.010)	(0.002)	(0.002)

Notes: each regression includes family and school dummies, linear school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, parental education, the number of siblings and a dummy for first-born. Peers' parental background E(PE) is standardized (zero mean and unit standard deviation). Standard errors clustered by school and family. Total number of observations: 725,722 (356,251 females). Observations for average log income 31-40: 565,284 (287,200 females). ***: p<.01; **: p<.05; *:p<.10

Table 3. Estimated effects on school and labor market outcomes of the interaction between SG and FG.

	(1)	(2)	(3)	(4)	(5)	(6)
Gender		Male			Female	
	FG=0	FG=1	Difference	FG=0	FG=1	Difference
Dependent variable			(2)-(1)			(2)-(1)
Years of education	-0.056	-0.099	-0.043	-0.008	0.057	0.065**
	(0.088)	(0.064)	(0.107)	(0.057)	(0.053)	(0.030)
Average log income, age 31-40	0.004	-0.011	-0.015	0.007	0.010	0.003
	(0.020)	(0.015)	(0.024)	(0.012)	(0.012)	(0.006)
Employed at least 5 years between age 31 and 40	-0.002	-0.025**	-0.023	-0.005	0.018*	0.023***
	(0.016)	(0.012)	(0.020)	(0.011)	(0.011)	(0.005)

Notes: each regression includes family and school dummies, linear school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, parental education, the number of siblings and a dummy for first-born. Standard errors clustered by school and family. Own parental background PE and peers' parental background E(PE) are standardized (zero mean and unit standard deviation). Total number of observations: 725,722 (356,251 females). Observations for average log income 31-40: 565,284 (287,200 females). ***: p<.01; **: p<.05; *:p<.10

Table 4. Estimated effects on school and labor market outcomes of the interaction between E(PE) and PE.

	(1)	(2)	(3)	(4)
Gender	Male		Female	
	E(PE)	E(PE)*PE	E(PE)	E(PE)
Dependent variable				
Years of education	0.044***	-0.042***	-0.025**	-0.016**
	(0.010)	(0.007)	(0.011)	(0.07)
Average log income, age 31-40	-0.001	-0.005***	0.004	-0.005***
	(0.012)	(0.001)	(0.002)	(0.001)
Employed at least 5 years between	0.004**	-0.002	0.000	0.003**
age 31 and 40	(0.002)	(0.002)	(0.002)	(0.001)

Notes: each regression includes family and school dummies, linear school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, parental education, the number of siblings and a dummy for first-born. Standard errors clustered by school and family. Own parental background PE and peers' parental background E(PE) are standardized (zero mean and unit standard deviation). Total number of observations: 725,722 (356,251 females). Observations for average log income 31-40: 565,284 (287,200 females). ***: p<.01; **: p<.05; *:p<.10

Table 5. Estimated heterogeneous effects of E(PE) and SD(PE) on school and labor market outcomes. With school and cohort dummies, family fixed effects and interactions of E(PE) and SD(PE) with PE.

a. E(PE)

	(1)	(2)	(3)	(4)
Gender	Male		Female	
	E(PE)	E(PE)*PE	E(PE)	E(PE)*PE
Dependent variable				
Years of education	0.052***	-0.046***	-0.023**	-0.016**
	(0.011)	(0.007)	(0.011)	(0.007)
Average log income, age 31-40	-0.001	-0.004***	0.004*	-0.006***
	(0.002)	(0.001)	(0.002)	(0.001)
Employed at least 5 years between age 31 and 40	0.004**	-0.002	0.000	0.003**
	(0.002)	(0.001)	(0.002)	(0.001)

b. SD(PE)

	(1)	(2)	(3)	(4)
Gender	Male		Female	
	SD(PE)	SD(PE)*PE	SD(PE)	SD(PE)*PE
Dependent variable				
Years of education	-0.072***	0.076***	0.014	-0.011
	(0.021)	(0.019)	(0.021)	(0.019)
Average log income, age 31-40	0.010**	0.007*	-0.005	0.005
	(0.005)	(0.004)	(0.004)	(0.004)
Employed at least 5 years between age 31 and 40	-0.004	0.007**	0.002	-0.009***
	(0.004)	(0.003)	(0.004)	(0.003)

Notes: each regression includes family and school dummies, linear school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, parental education, the number of siblings and a dummy for first-born. Standard errors clustered by school and family. Own parental background PE and peers' parental background E(PE) are standardized (zero mean and unit standard deviation).The standard deviation of peers' parental background SD(E(PE)) is also normalized to have 0 mean and standard deviation equal to E(PE). Total number of observations: 725,722 (356,251 females). Observations for average log income 31-40: 565,284 (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)

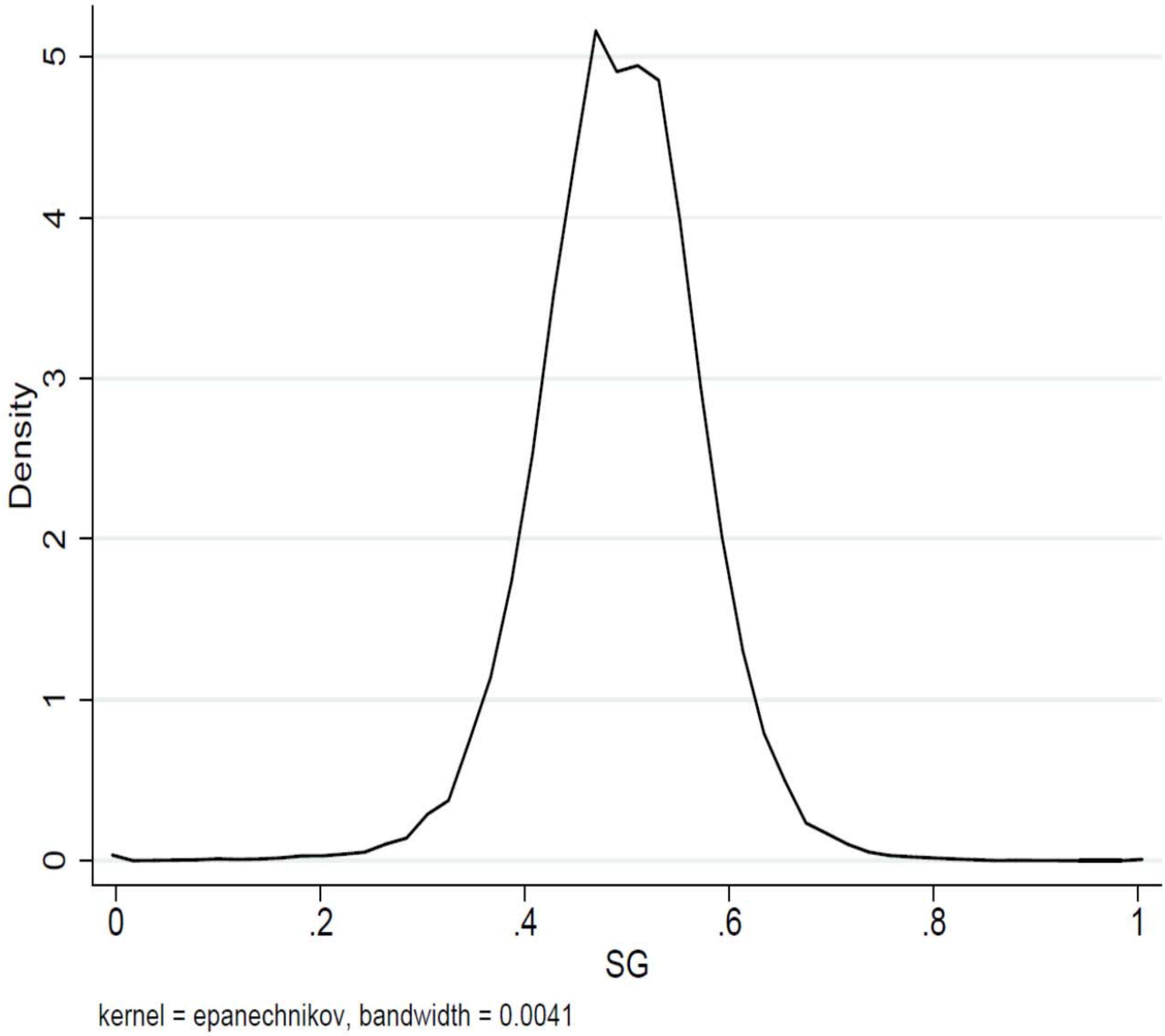


Figure 2. Kernel density estimate of the average parental education of peers E(PE)

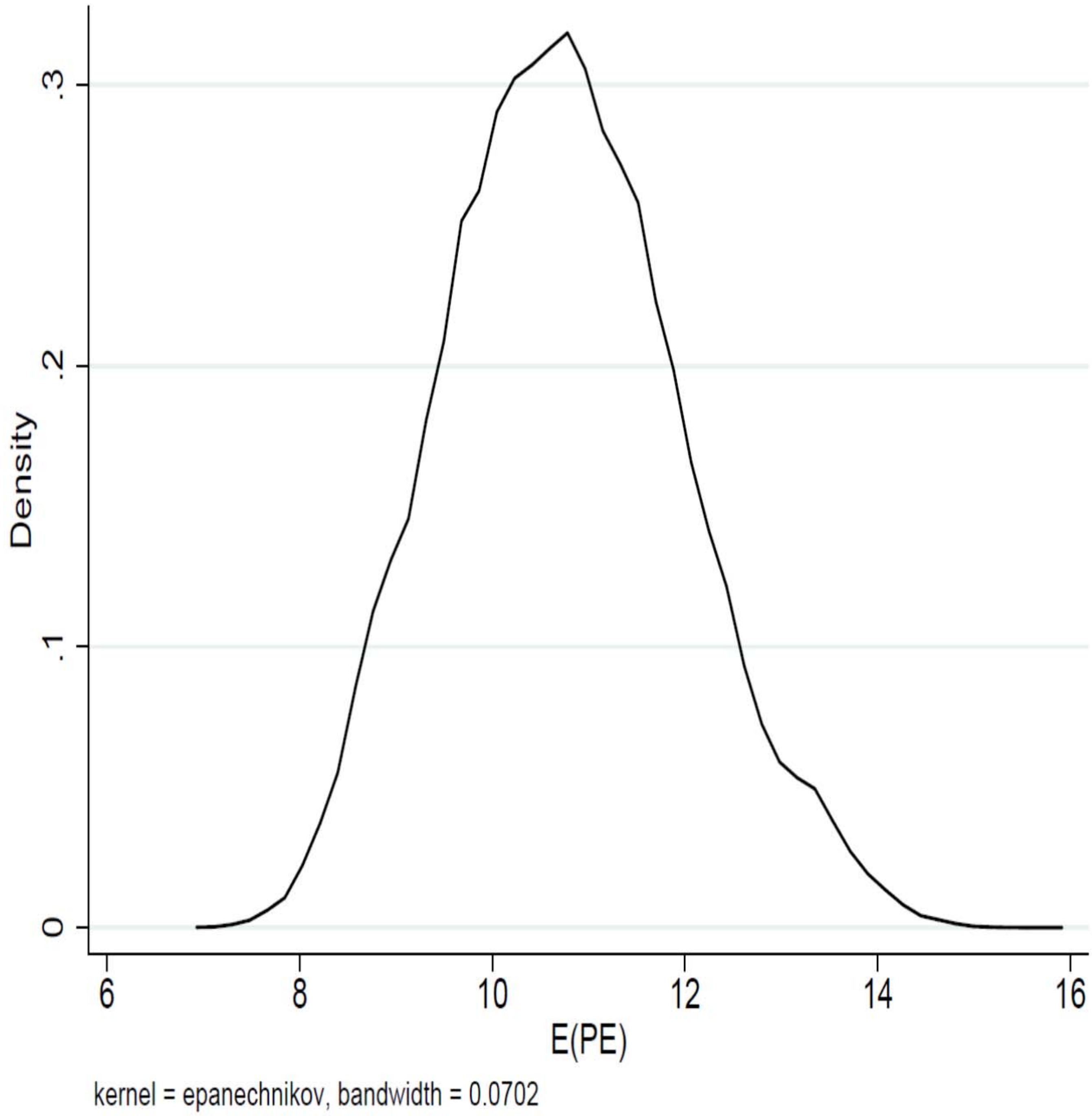


Figure 3. Marginal effect of E(PE) on years of education for different values of individual PE. By gender

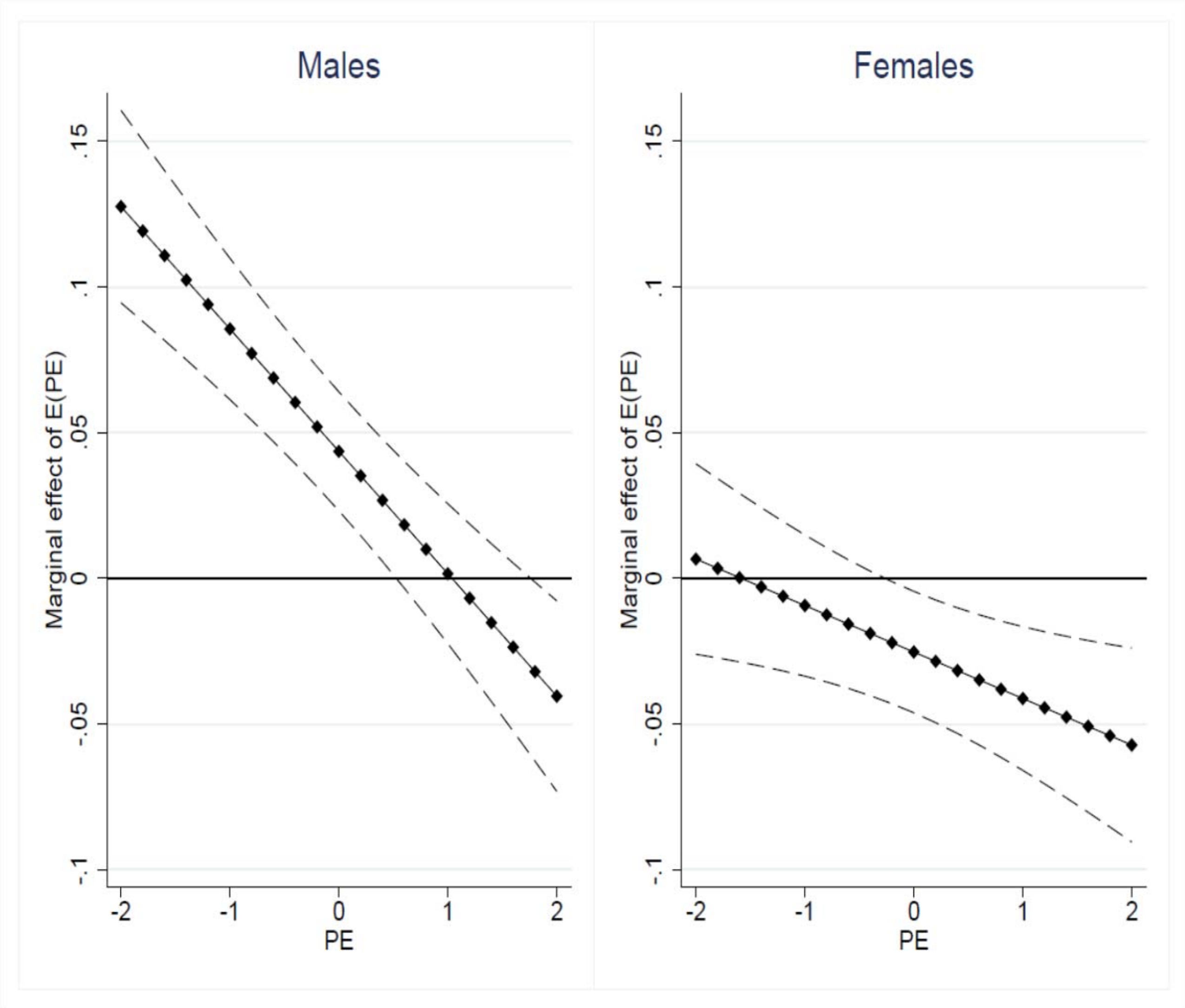


Figure 4. Marginal effect of E(PE) on average log earnings between age 31 and 40 for different values of individual PE. By gender

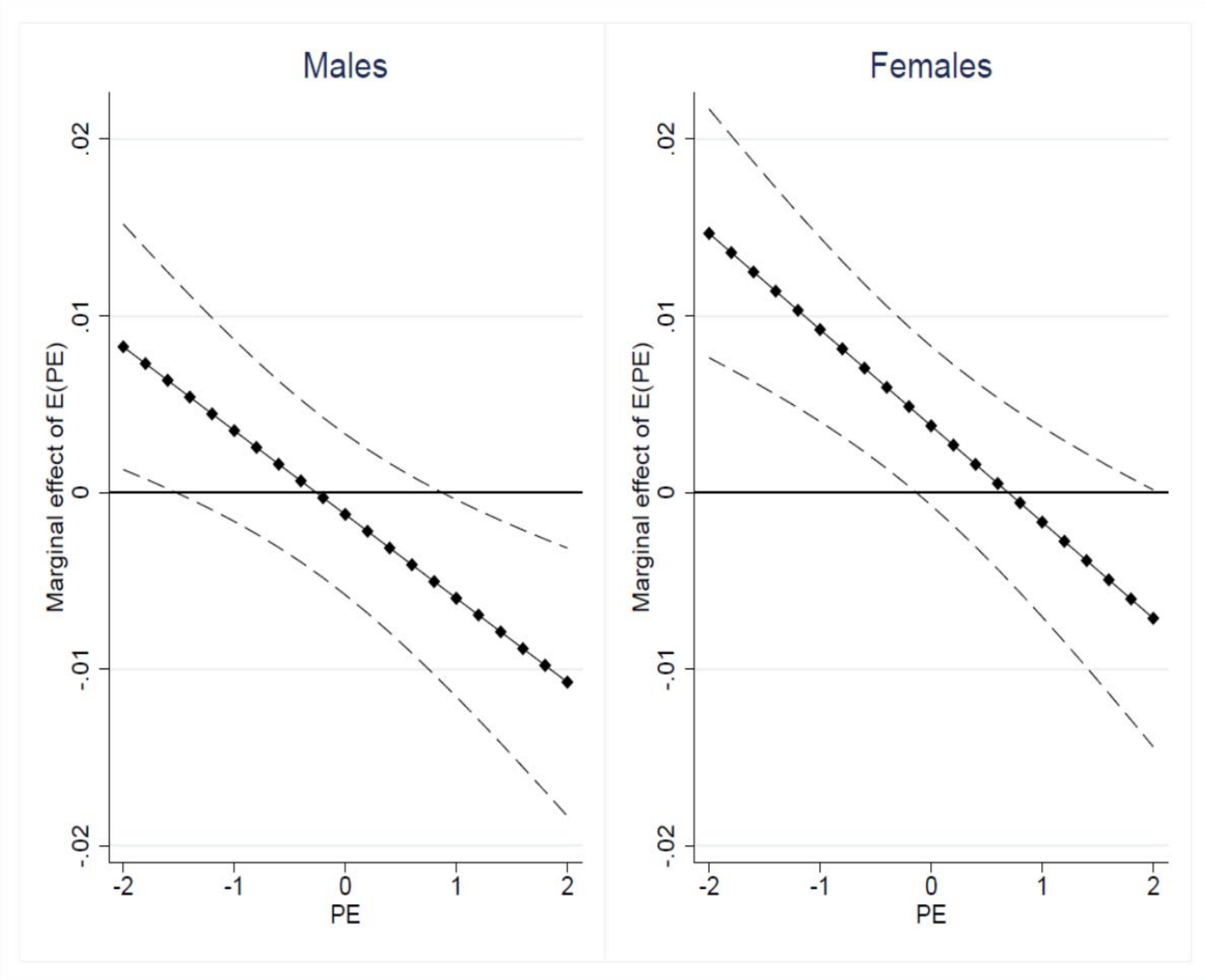


Figure 5. Marginal effect of E(PE) on employment at least five years between age 31 and 40 for different values of individual PE. By gender.

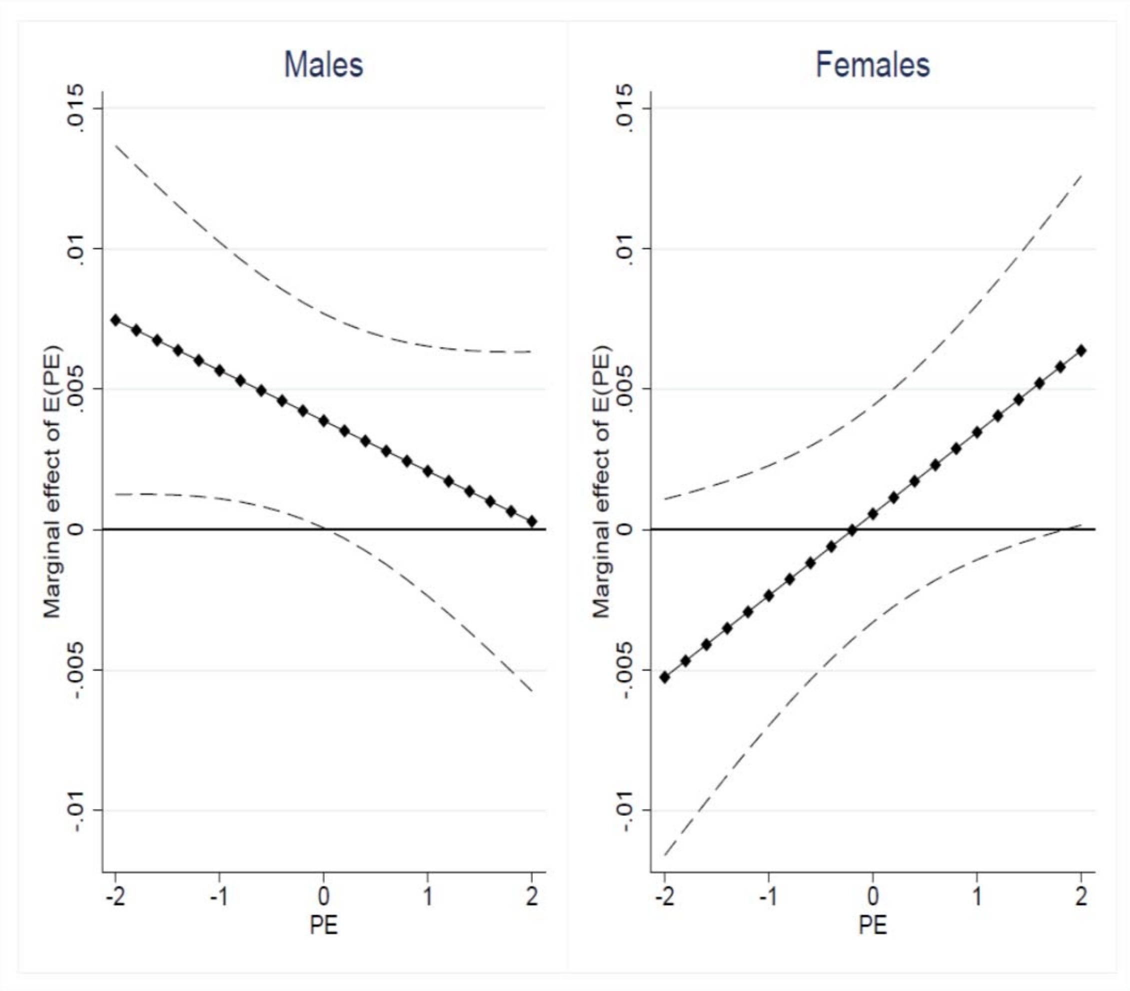


Figure 6. Mean and standard deviation of parental education in the school.

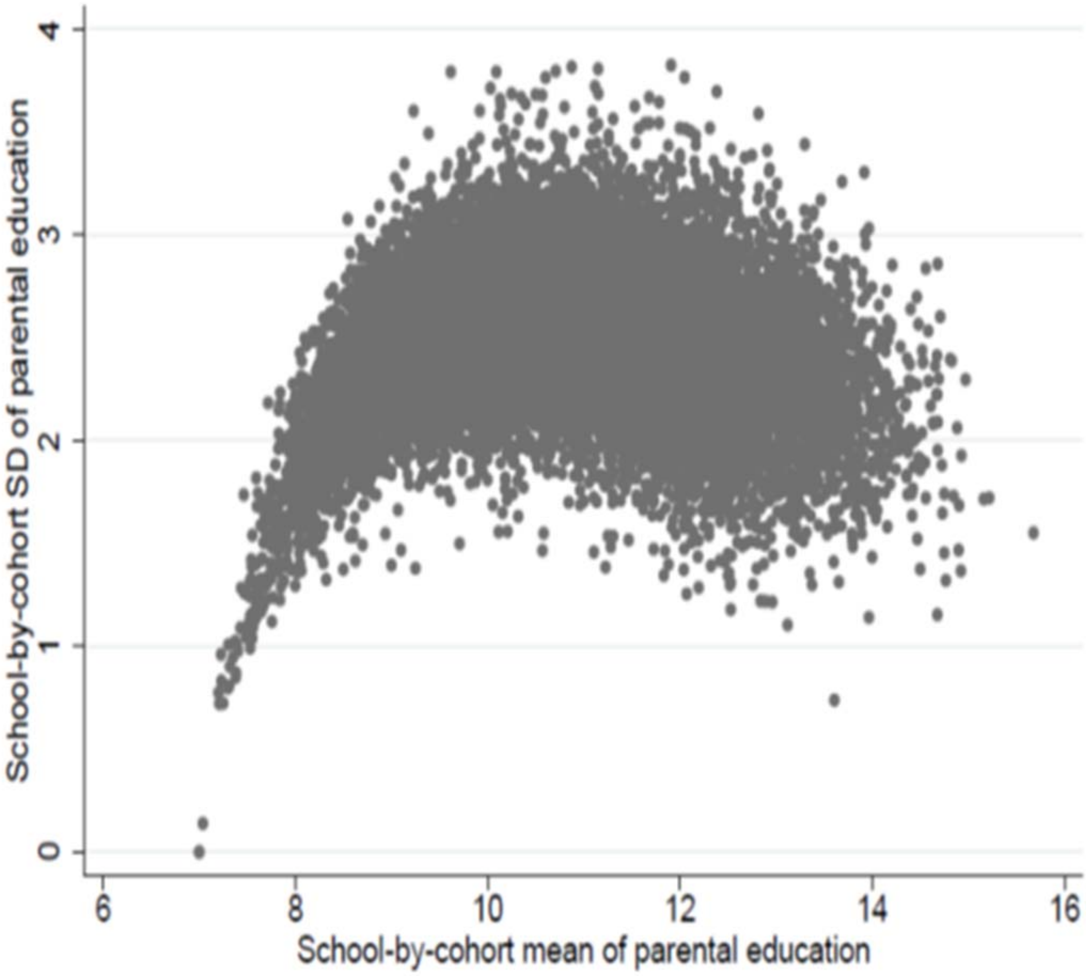


Figure 7. Marginal effect of SD(PE) on years of education for different values of individual PE. By gender.

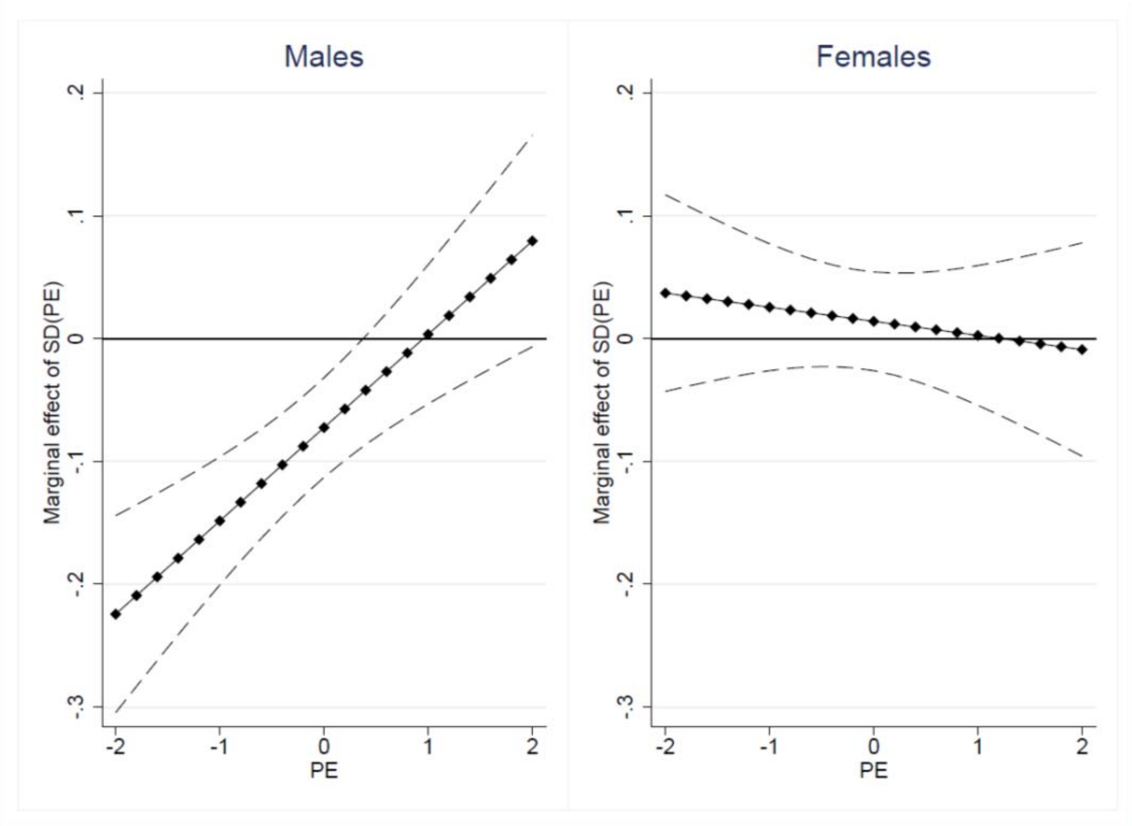


Figure 8. Marginal effect of SD(PE) on average log earnings between age 31 and age 40 for different values of individual PE. By gender.

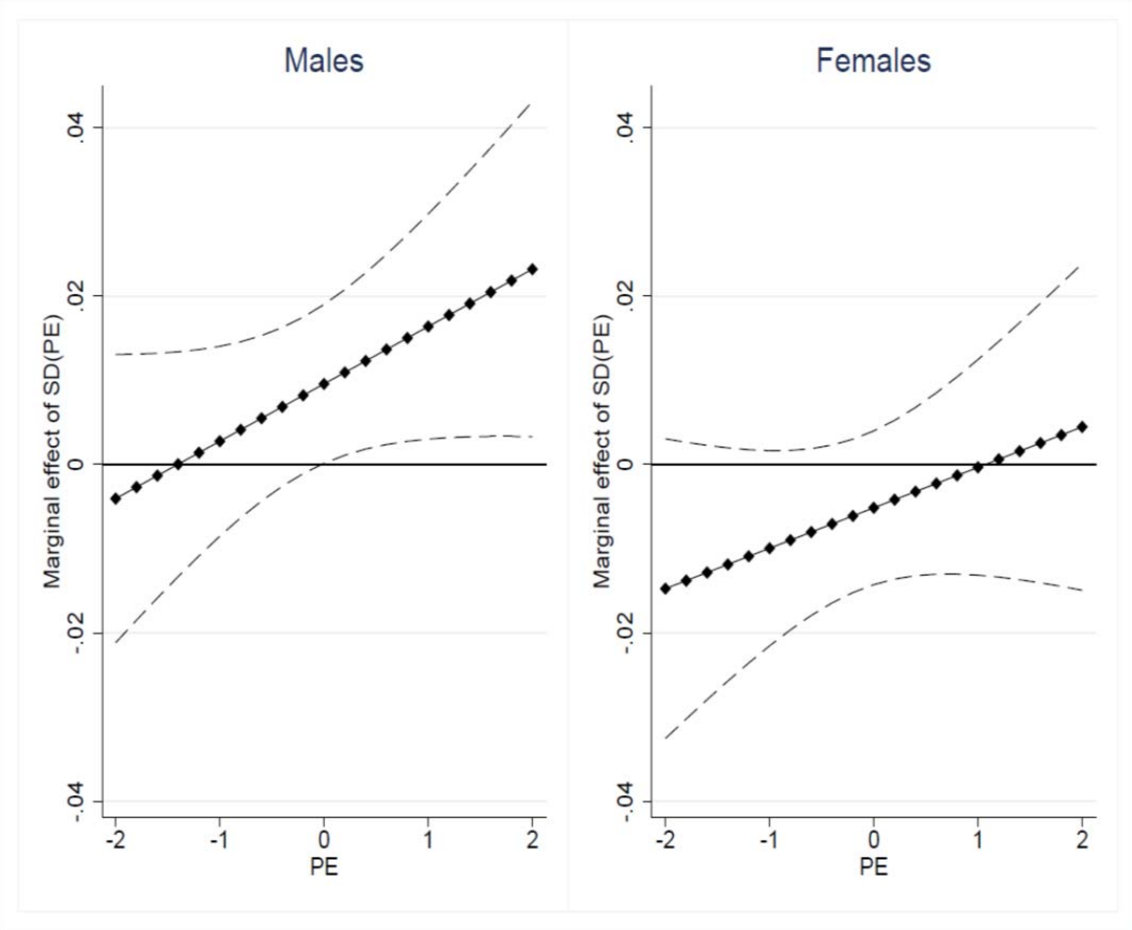
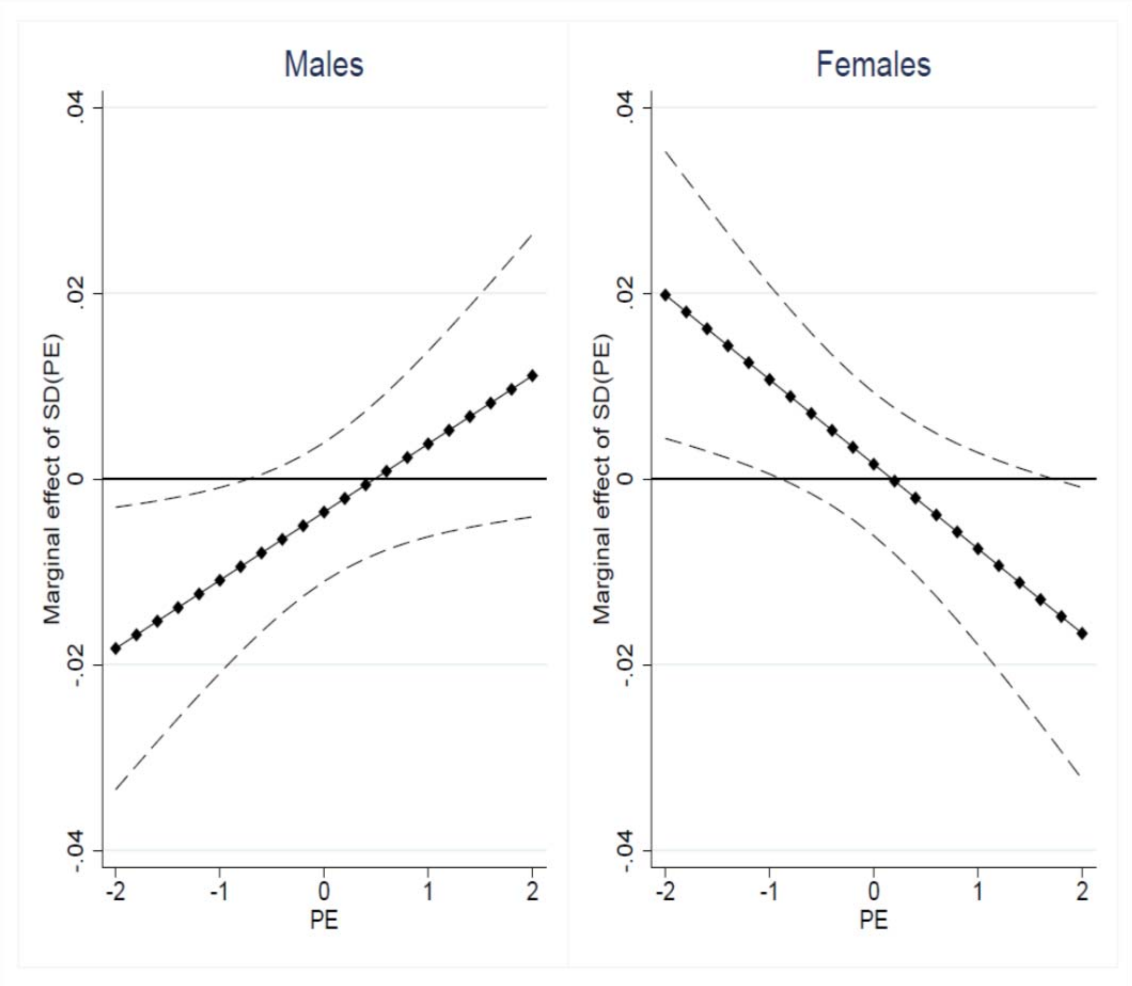


Figure 9. Marginal effect of SD(PE) on employment at least five years between age 31 and 40 for different values of individual PE. By gender.



Appendix

Table A1. Balancing tests. Reverse regressions of individual covariates on the peer variables SG and E(PE). With and without family fixed effects and single children.

	(1)	(2)	(3)	(4)	(5)	(6)
Peer variable	SG	E(PE)	SG	E(PE)	SG	E(PE)
Dependent variable						
Female	-0.100***	0.004**	-0.095***	0.002	-0.025**	0.000
	(0.011)	(0.002)	(0.012)	(0.002)	(0.013)	(0.002)
Average years of education of father and mother - PE	0.023	-0.037***	0.025	-0.053***	*Absorbed by family fixed effects*	
	(0.014)	(0.006)	(0.016)	(0.006)		
Number of siblings	0.014	-0.004	-0.011	0.003	*Absorbed by family fixed effects*	
	(0.013)	(0.003)	(0.014)	(0.003)		
Firstborn	0.009	0.004*	0.014*	0.004**	0.015*	0.001
	(0.007)	(0.002)	(0.008)	(0.002)	(0.009)	(0.013)
Family fixed effects	No	No	No	No	Yes	Yes
Includes single children	Yes	Yes	No	No	No	No

Notes: each regression includes school enrolment, cohort and school dummies and school-specific trends. Columns (1) and (2) consider the full sample. Columns (3) and (4) drop single children, excluded from estimation with family fixed effects. Columns (5) and (6) also include family fixed effects. Standard errors are clustered by school, for columns (5) and (6) by school and family. Peers' parental background E(PE) is standardized (zero mean and unit standard deviation). Total number of observations: 1,009,924. Observations for columns (3) to (6): 725,722. ***: p<.01; **: p<.05; *: p<.10

Table A2. Characteristics of Mover and Non-Mover Families.

	(1)	(2)	(3)
	Movers	Non-movers	Difference (1) - (2)
Number of children	2.72	2.55	.17***
Age of mother at first birth	23.11	23.86	.75***
Year of birth of first-born	1963.62	1964.67	-1.05***
Average years of schooling of the parents	10.73	10.78	-.05***
At least one parent with college degree	.235	.234	.001

Notes: the table reports the average of a set of observable characteristics for mover (Column 1) and non-mover (Column 2) families. Every observation is a family with more than one child. Column 3 reports the difference between the average for movers and non-movers, with its statistical significance. The number of observations is equal to 88,006 and 226,761 for movers and non-movers, respectively. ***: p<.01; **: p<.05; *: p<.10.

Table A3. Estimated effects of SG and E(PE) on school and labor market outcomes. With school and cohort dummies, full sample, without family fixed effects.

	(1)	(2)	(3)	(4)
Peer variable	SG		E(PE)	
Gender	Male	Female	Male	Female
Dependent Variable				
Years of education	0.035	0.079*	0.069***	0.022**
	(0.032)	(0.041)	(0.008)	(0.009)
Average log income 31-40	0.002	0.003	0.004***	0.007***
	(0.009)	(0.009)	(0.001)	(0.002)
Employed at least 5 years between age 31 and 40	-0.004	0.005	0.002	-0.001
	(0.007)	(0.007)	(0.001)	(0.001)

Notes: each regression includes family and school dummies, linear school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, parental education, the number of siblings and a dummy for first-born. Standard errors clustered by school and family. Peers' parental background E(PE) is standardized (zero mean and unit standard deviation). Total number of observations: 1,009,924 (496,439 females). Observations for average log income 31-40: 860,879 (425,162 females). ***: p<.01; **: p<.05; *: p<.10

Table A4. Estimated effects of SG and E(PE) on school and labor market outcomes. With school and cohort dummies, excluding single children, without family fixed effects.

	(1)	(2)	(3)	(4)
Peer variable	SG		E(PE)	
Gender	Male	Female	Male	Female
Dependent Variable				
Years of education	0.001	0.088*	0.071***	0.020**
	(0.047)	(0.047)	(0.010)	(0.010)
Average log income 31-40	0.002	0.005	0.004**	0.007***
	(0.010)	(0.010)	(0.002)	(0.002)
Employed at least 5 years between age 31 and 40	-0.004	0.006	0.003*	-0.000
	(0.008)	(0.008)	(0.002)	(0.002)

Notes: each regression includes family and school dummies, linear school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, parental education, the number of siblings and a dummy for first-born. Standard errors clustered by school and family. Peers' parental background E(PE) is standardized (zero mean and unit standard deviation). Total number of observations: 725,722 (356,251 females). Observations for average log income 31-40: 565,284 (287,200 females). ***: p<.01; **: p<.05; *: p<.10

Table A5. Estimated effects of SG and E(PE) on additional educational outcomes. With school, cohort and family fixed effects.

	(1)	(2)	(3)	(4)
Peer variable	SG		E(PE)	
Gender	Male	Female	Male	Female
Dependent Variable				
Highest degree is vocational	0.010	-0.008	-0.006**	-0.004*
	(0.013)	(0.013)	(0.003)	(0.002)
Tertiary STEM (conditional)	0.030	0.013	-0.015***	0.003
	(0.027)	(0.016)	(0.005)	(0.003)
Tertiary Humanities (conditional)	-0.003	-0.003	-0.002	0.010***
	(0.021)	(0.017)	(0.004)	(0.003)
Tertiary Health and Related Fields	0.000	-0.011	0.012*	-0.015**
(conditional)	(0.027)	(0.034)	(0.006)	(0.006)
Tertiary Law and Social Sciences	-0.024	0.001	0.011**	-0.003
(conditional)	(0.027)	(0.020)	(0.005)	(0.004)

Notes: each regression includes family and school dummies, linear school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, parental education, the number of siblings and a dummy for first-born. Standard errors clustered by school and family. Peers' parental background E(PE) is standardized (zero mean and unit standard deviation). Total number of observations: 725,722 (356,251 females). Observations for average log income 31-40: 565,284 (287,200 females). ***: p<.01; **: p<.05; *: p<.10

Table A6. Heterogeneous effects of SG and E(PE) on additional educational outcomes. With school and cohort dummies, family fixed effects and interactions of SG with FG and E(PE) with PE.

	(1)	(2)	(3)	(4)	(5)	(6)
Gender	Male			Female		
	FG=0	FG=1	Difference (2)-(1)	FG=0	FG=1	Difference (2)-(1)
Dependent variable						
Highest degree is vocational	0.033	-0.002	-0.035	-0.005	-0.009	-0.004
	(0.021)	(0.016)	(0.026)	(0.014)	(0.013)	(0.007)
STEM at tertiary (conditional)	0.005	0.043	0.038	0.023	0.007	-0.015
	(0.005)	(0.032)	(0.053)	(0.017)	(0.016)	(0.009)
Humanities at tertiary (conditional)	0.001	-0.006	-0.007	0.003	-0.005	-0.002
	(0.034)	(0.026)	(0.041)	(0.017)	(0.017)	(0.042)
Health at tertiary (conditional)	0.029	-0.016	-0.045	-0.036	0.002	0.038**
	(0.035)	(0.035)	(0.049)	(0.036)	(0.034)	(0.017)
Law and Social at tertiary (conditional)	-0.014	-0.028	-0.014	0.003	0.001	-0.002
	(0.045)	(0.033)	(0.054)	(0.021)	(0.021)	(0.011)

Notes: each regression includes family and school dummies, linear school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, parental education, the number of siblings and a dummy for first-born. Standard errors clustered by school and family. Own parental background PE and peers' parental background E(PE) are standardized (zero mean and unit standard deviation). Total number of observations: 725,722 (356,251 females). Observations for average log income 31-40: 565,284 (287,200 females). ***: p<.01; **: p<.05; *: p<.10

Table A7. Heterogeneous effects of E(PE) on additional educational outcomes. With school and cohort dummies, family fixed effects and interactions of E(PE) with PE.

	(1)	(2)	(3)	(4)
Gender	Male		Female	
	E(PE)	E(PE)*PE	E(PE)	E(PE)*PE
Dependent variable				
Highest degree is vocational	-0.006**	-0.004***	-0.005*	-0.002
	(0.003)	(0.002)	(0.003)	(0.002)
STEM at tertiary (conditional)	-0.018***	0.005*	0.002	-0.005***
	(0.006)	(0.003)	(0.004)	(0.002)
Humanities at tertiary (conditional)	-0.001	0.001	0.084***	0.002
	(0.004)	(0.002)	(0.031)	(0.002)
Health at tertiary (conditional)	0.012*	-0.005	-0.015**	0.000
	(0.007)	(0.003)	(0.007)	(0.035)
Law and Social at tertiary (conditional)	0.013**	-0.001	-0.003	0.004**
	(0.006)	(0.003)	(0.004)	(0.002)

Notes: each regression includes family and school dummies, linear school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, parental education, the number of siblings and a dummy for first-born. Standard errors clustered by school and family. Own parental background PE and peers' parental background E(PE) are standardized (zero mean and unit standard deviation). Total number of observations: 725,722 (356,251 females). Observations for average log income 31-40: 565,284 (287,200 females). ***: p<.01; **: p<.05; *: p<.10

Table A8. Robustness of the estimated heterogeneous effects of SG to considering any sisters at age 15 instead of the any sister ever. With school and cohort dummies, family fixed effects and interactions of SG with FG. Considering any sister at age 15 instead of any sister.

	(1)	(2)	(3)	(4)	(5)	(6)
Gender	Male			Female		
	FG=0	FG=1	Difference (2)-(1)	FG=0	FG=1	Difference (2)-(1)
Dependent variable						
Years of education	-0.107	-0.073	0.034	-0.004	0.055	0.059*
	(0.082)	(0.062)	(0.097)	(0.057)	(0.053)	(0.030)
Average log income, age 31-40	-0.002	-0.009	-0.007	0.007	0.010	0.003
	(0.013)	(0.014)	(0.022)	(0.012)	(0.012)	(0.007)
Employed at least 5 years between age 31 and 40	0.003	-0.028**	0.009	-0.005	0.018*	0.023***
	(0.015)	(0.012)	(0.009)	(0.011)	(0.011)	(0.005)

Notes: each regression includes family and school dummies, linear school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, parental education, the number of siblings and a dummy for first-born. Standard errors clustered by school and family. Own parental background PE and peers' parental background E(PE) are standardized (zero mean and unit standard deviation). Total number of observations: 725,722 (356,251 females). Observations for average log income 31-40: 565,284 (287,200 females). ***: p<.01; **: p<.05; *: p<.10

Table A9. Robustness of the estimated heterogeneous effects of E(PE) to considering maximum instead of average parental education. With school and cohort dummies, family fixed effects and interactions of E(PE) with PE.

	(1)	(2)	(3)	(4)
Gender	Male		Female	
	E(PE)	E(PE)*PE	E(PE)	E(PE)*PE
Dependent variable				
Years of education	0.041***	-0.037***	-0.027**	-0.015**
	(0.010)	(0.006)	(0.010)	(0.06)
Average log income, age 31-40	-0.001	-0.003**	0.003	-0.004***
	(0.012)	(0.001)	(0.002)	(0.001)
Employed at least 5 years between age 31 and 40	0.003*	-0.002**	-0.000	0.003**
	(0.002)	(0.002)	(0.002)	(0.001)

Notes: each regression includes family and school dummies, linear school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, parental education, the number of siblings and a dummy for first-born. Standard errors clustered by school and family. Own parental background PE and peers' parental background E(PE) are standardized (zero mean and unit standard deviation). Total number of observations: 725,722 (356,251 females). Observations for average log income 31-40: 565,284 (287,200 females). ***: p<.01; **: p<.05; *: p<.10

Table A10. Robustness of the estimated heterogeneous effects of E(PE) to considering parental education of father for males and mothers for females instead of average parental education. With school and cohort dummies, family fixed effects and interactions of E(PE) with PE.

	(1)	(2)	(3)	(4)
Gender	Male		Female	
	E(PE)	E(PE)*PE	E(PE)	E(PE)*PE
Dependent variable				
Years of education	0.059***	-0.043***	-0.036***	-0.014**
	(0.009)	(0.006)	(0.009)	(0.06)
Average log income, age 31-40	-0.000	-0.004***	0.002	-0.005***
	(0.002)	(0.001)	(0.002)	(0.001)
Employed at least 5 years between	0.004**	-0.005***	-0.001	0.000
age 31 and 40	(0.002)	(0.001)	(0.002)	(0.001)

Notes: each regression includes family and school dummies, linear school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, parental education, the number of siblings and a dummy for first-born. Standard errors clustered by school and family. Own parental background PE and peers' parental background E(PE) are standardized (zero mean and unit standard deviation). Total number of observations: 725,722 (356,251 females). Observations for average log income 31-40: 565,284 (287,200 females). ***: p<.01; **: p<.05; *: p<.10

Table A11. Robustness of the heterogeneous effects of SG and E(PE) to the inclusion of the share of girls and average parental education in the parish of residence as additional controls. With school and cohort dummies, family fixed effects and interactions of SG with FG and E(PE) with PE.

a. SG

	(1)	(2)	(3)	(4)	(5)	(6)
Gender	Male			Female		
	FG=0	FG=1	Difference (2)-(1)	FG=0	FG=1	Difference (2)-(1)
Dependent variable						
Years of education	-0.052	-0.092	-0.040	-0.045	0.020	0.065**
	(0.090)	(0.069)	(0.107)	(0.061)	(0.058)	(0.030)
Average log income, age 31-40	0.005	-0.011	-0.016	0.006	0.009	0.003
	(0.021)	(0.015)	(0.024)	(0.013)	(0.013)	(0.006)
Employed at least 5 years between age 31 and 40	-0.002	-0.025*	-0.023	-0.005	0.018	0.023***
	(0.017)	(0.013)	(0.020)	(0.011)	(0.011)	(0.005)

b. E(PE)

	(1)	(2)	(3)	(4)
Gender	Male		Female	
	E(PE)	E(PE)*PE	E(PE)	E(PE)*PE
Dependent variable				
Years of education	0.022**	-0.041***	-0.000	-0.017***
	(0.011)	(0.007)	(0.012)	(0.007)
Average log income, age 31-40	0.003	-0.005***	0.004	-0.005***
	(0.003)	(0.001)	(0.002)	(0.001)
Employed at least 5 years between age 31 and 40	0.000	-0.002	0.004**	0.003**
	(0.002)	(0.002)	(0.002)	(0.001)

Notes: each regression includes family and school dummies, linear school specific trends, and the following covariates interacted by gender: cohort dummies, school enrolment, parental education, the number of siblings and a dummy for first-born. Standard errors clustered by school and family. Own parental background PE and peers' parental background E(PE) are standardized (zero mean and unit standard deviation). Total number of observations: 725,722 (356,251 females). Observations for average log income 31-40: 565,284 (287,200 females). ***: p<.01; **: p<.05; *: p<.10. The models presented in this table include the share of girls and average parental education in the parish of residence as additional controls.