Early Socialization and the Gender Wage Gap

Demid Getik[‡] Armando N. Meier[⋆]

October, 2021

Lund Department of Economics WP No. 2021:13

— Click here for the latest version —

Abstract

We study the impact of early socialization on gender inequality in the labor market. To this end, we link the gender environment in the primary-school cohort to later occupations and wages. We find that women exposed to more girls at this critical age earn more later on, leading to a reduction in the gender wage gap. We explore mechanisms and find that women exposed to a more female-dominated environment select into less gender-stereotypical occupations with higher wage potential. The gender environment at an early age, therefore, shapes career trajectories and lifetime earnings.

JEL Classifications: D91, I24, I26, J16, J24, J70

Keywords: socialization, school environment, peers, occupational sorting, gender wage gap

[‡] Demid Getik, Department of Economics and Center for Economic Demography (CED), Lund University, email: demid.getik@nek.lu.se
[⋆] Armando N. Meier, University of Lausanne and University of Basel, email: armando.meier@unil.ch

We thank numerous scholars and seminar participants at the University College Dublin, Lund University, the University of Basel, as well as conference participants at the Copenhagen Education Workshop, the IZA Economics of Education Workshop, and the Workshop for Societal and Policy Problems for their comments.

1 Introduction

Women still earn less than men, with more than 70% of the gender wage gap unaccounted for by traditional explanations such as educational attainment (Card, Cardoso and Kline, 2016; Goldin et al., 2017; Blau and Kahn, 2017). The persistent and partly unexplained gap has spurred research on other potential explanations for inequality in labor market outcomes. Recent literature has proposed early-life gender socialization and related identity formation as one important explanation for the remaining inequality in occupational selection and wages (see, e.g., Bertrand, 2011; Bertrand, Kamenica and Pan, 2015; Olivetti and Petrongolo, 2016; Bursztyn, Fujiwara and Pallais, 2017; Blau and Kahn, 2017). Unfortunately, it has been difficult to identify the causal impact of this potential explanation on labor market trajectories (Bertrand, 2011). One reason is a lack of data linking key features of early-life environments, such as the gender environment in primary school, to long-run labor market outcomes.

In this paper, we study whether and how a female-dominated environment at the critical ages of 6 to 16 shapes career trajectories in the long-run. We overcome the challenge of linking early gender environment to labor market outcomes by using comprehensive Swedish register data ($N = 757,560$). We link the share of female peers in a student's primary school cohort to gender gaps in grades, post-primary educational choices, career choices, and wages.

We find that a more female-dominated environment in primary school leads to higher wages for women, lowering the gender wage gap. Changing from a 40% female to a 60% female cohort leads to a \$404 (SEK 3,434) increase in annual wages of women at the age of 30, corresponding to a \$14,173 (SEK 120,471) increase in lifetime earnings. The impact on women's wages is large when compared to teacher value-added estimates. Chetty et al. (2011) estimate that kindergarten students with a teacher who had more than 10 years of experience earn \$364 more per year at ages 25 to 27. In comparison, we find that a 10 percentage point shift in the gender composition of the cohort changes annual wages by \$202 at age 30. The effects are also large when compared to the association between parental education and child wages. A parental university degree translates into \$2,065 higher yearly wages. A 10 percentage point shift in the gender composition therefore amounts to almost 10% this relationship. The corresponding effect on the gender wage gap is substantial. A corresponding increase in the share of female peers, which is equivalent to 2-3 more girls in the classroom, decreases the gender wage gap by 2.7%.

We also examine the impact of a more female-dominated environment on occupational wage potential only relying on variation across occupations. We find that women exposed to more girls in primary school enter occupations with higher wages. A 60% female-dominated class, when compared to a 40% female class, leads to $9,091 higher occupational lifetime earnings and a reduction of the gender gap by $10,597.

A key reason for these findings is that women select into less gender-stereotypical occupations after being exposed to more girls in primary school. One reason for the change in occupational selection is that girls select into less gender-stereotypical high-school specializations and college majors. Girls with more girls in the cohort choose more male-dominated high-school tracks and choose more male-dominated jobs, leading to higher wages. We can attribute roughly 60% of the impact of the early gender environment on the gender wage gap to educational and occupational selection. While we do not find an impact of the gender environment on selection into more or less competitive educational tracks or into occupations with a higher variance in wages, the remaining unexplained variation suggests an additional role for other factors, such as commuting distance, within household work distribution, non-cognitive skills, mental health, competitiveness, negotiation behavior, and risk aversion (Bertrand, 2011; Buser, Niederle and Oosterbeek, 2014; Flory, Leibbrandt and List, 2015; Landaud, Ly and Maurin, 2018; Born, Ranehill and Sandberg, 2020; Getik and Meier, 2020; Biasi and Sarsons, 2021; Le Barbanchon, Rathelot and Roulet, 2021).

To estimate the causal impact of gender environment in early life, we use idiosyncratic variation in the gender composition across cohorts within schools (see also, e.g., Hoxby, 2000; Lavy and Schlosser, 2011; Carrell, Hoekstra and Kuka, 2018).[1] The extensive administrative data allows us to check in detail whether the share of girls in a cohort is arbitrary. First, we observe that cohorts with more boys or girls are comparable based on an extensive range of parental characteristics, including parental age, educational attainment, and wages. Second, we document that whether a girl or a boy ends up in a specific cohort is arbitrary: Student gender does not predict the leave-one-out share of female peers (following Guryan, Kroft and Notowidigdo, 2009), and school-by-cohort fixed effects do not predict student gender (following Chetty et al., 2011; Balestra, Eugster and Liebert, 2020). Third, we compare the distribution of the share of female peers to a simulated distribution based on random

[1] Data on classrooms is not available for our sample period. However, looking at more recent data containing classroom indicators (which we use in Getik and Meier, 2020 to examine the impact of class composition on mental health), we see that classroom gender composition is also arbitrary.

assignment to cohorts (following Bietenbeck, 2020). The actual distribution of the share of female peers looks equivalent to the distribution from simulations based on actual random assignment of students. Taken together, the gender environment in cohorts seems arbitrary.

We do several further plausibility checks. The data allows us to check the robustness of our results to different sets of identifying assumptions: i) We can include school and cohort fixed effects separately, and ii) cohort–by–school fixed effects, allowing us to identify the impact on the gender gap. With the second specification, we can examine how more girls in the cohort affect girls and boys differently (see, e.g., Brenøe and Zölitz, 2020). We therefore study whether the share of female peers differently affects boys and girls even within the same school-cohort, holding constant any potential selection into specific school-cohorts. Accordingly, including fixed effects for each cohort of each school controls for the exact level at which selection on time variant characteristics and unobservables would occur. The results are equivalent independent of the specification. Placebo checks confirm that only the cohort the student attends matters for outcomes. Moreover, we show that the results are similar when we drop students who move during the sample period or in the year before entering school.

The paper offers a comprehensive evaluation of the impact of early gender environment on labor market outcomes, contributing to two broad strands of literature. First, we offer evidence for how gender socialization affects the gender wage gap. We therefore add to a large literature on occupational sorting and the gender wage gap (see, e.g., Altonji and Blank, 1999; Mulligan and Rubinstein, 2008; Bertrand, 2011; Olivetti and Petrongolo, 2016; Lundborg, Plug and Rasmussen, 2017; Blau and Kahn, 2017; Kleven, Landais and Søgaard, 2019; Card, Colella and Lalive, 2021).

Previous literature has examined how educational attainment, occupation, and discrimination affect gender gaps, but has given less consideration to the role of early social environments and socialization:[2] "... we feel that much more validating empirical work will be needed in the near future for gender identity insights to have a long-lasting impact on how labor economists approach gender issues (Bertrand, 2011; p. 1545)." Findings from psychological surveys indicate that girls in female-dominated environments are more self-confident and less gender-conforming (Bertrand, 2011). This is consistent with our results, which show that women choose less gender-stereotypical jobs when exposed to more girls in primary school,

[2]One exception is Slotwinski and Stutzer (2018) who show that voting rights increased labor force participation among women.

suggesting that the early social environment shapes gender roles. While women attain a higher level of education in high-income countries, the gender wage gap persists (Olivetti and Petrongolo, 2016; Blau and Kahn, 2017). Our results suggest that socialization, even independent of educational attainment, is a key driver of occupational selection. Therefore persisting gender roles could play a critical part in explaining why the rise in educational attainment has not lead to a reduction in gender inequality.

Second, we complement the literature on how school inputs and childhood environments affect later labor market outcomes (see, e.g., Chetty et al., 2011; Eisenkopf et al., 2015; Feld and Zölitz, 2017; Carrell, Hoekstra and Kuka, 2018; Balestra, Eugster and Liebert, 2020; Balestra, Sallin and Wolter, 2021; Bietenbeck, 2020; Golsteyn, Non and Zölitz, 2021; Elsner, Isphording and Zölitz, 2021). Among others, previous research has examined the role of neighborhoods and the role of parents and siblings in shaping labor market ouctomes (see, e.g., Chetty et al., 2016; Almås et al., 2016; Brenøe, 2021; Brenøe and Lundberg, 2018).

Research examining the consequences of gender environment in school has focused on the more short-run impact on educational choices and outcomes (Lavy and Schlosser, 2011; Park, Behrman and Choi, 2012; Black, Devereux and Salvanes, 2013; Eisenkopf et al., 2015; Anelli and Peri, 2019; Giardili, 2020; Brenøe and Zölitz, 2020; Lu and Anderson, 2015; Borbely, Norris and Romiti, 2021). For instance, Brenøe and Zölitz (2020) document that girls are less likely to choose a STEM college major when exposed to a male-dominated environment in the math track in high school.[3] In contrast, Anelli and Peri (2019) find that boys graduating from 80% male high-school classes are more likely to choose gender stereotypical college majors, such as engineering, but there is no impact on girls.[4] We build on and contribute

[3]Evidence on the impact of the share of female peers on educational attainment and choices is mixed. Black, Devereux and Salvanes (2013) document lower educational attainment for boys with more girls in the cohort with no impact on girls, while Lavy and Schlosser (2011) document higher attainment for boys and girls with more girls in the cohort. The latter finding is consistent with findings from single-sex schooling (Park, Behrman and Choi, 2012; Giardili, 2020) or seating patterns within the classrom (Lu and Anderson, 2015), which show that girls do better with more girls around them. Regarding educational choices, Brenøe and Zölitz (2020) find that female high-school students who joined the math track are more likely to make gender-congruent educational choices when exposed to more boys. Anelli and Peri (2019) find no impact of gender composition on educational choices of girls, and Giardili (2020) finds that girls in single-sex schools make less gender-congruent educational choices.

[4]Anelli and Peri (2019) document that while the high-school gender environment affects college major choice, there is no impact on degree completion because of attrition. This may be one reason why there are no lasting effect on income of 80% male-dominated classes. Brenøe and Zölitz (2020) document a decrease in the percentile rank of women's incomes after being exposed to more girls in the math track in high school. In contrast, Black, Devereux and Salvanes (2013) find an increase in log incomes of women after exposure to more girls. We complement the different results in several ways. First, we provide a comprehensive assessment of the impact the gender wage gap. Second, we document effects on occupational wage potential,

to this literature by providing a comprehensive and unified evaluation of the impact of the gender environment at an earlier, more malleable age (Bertrand, 2011) on long-run labor market outcomes.

2 Institutional Background and Data

2.1 Swedish Educational System

All Swedish children have to complete nine years of primary schooling (grundskolan), usually from ages 6-7 to 15-16, with a standardized curriculum. Admissions to schools is based on residence, with the catchment area of a school district determined by traveling distance to the school.[5] After starting school in the year when they turn seven, they usually pass through three stages of primary schooling: grades 1-3 (low), grades 4-6 (middle), and grades 7-9 (high).[6]

The institutional context leads to a likely arbitrary share of female peers in one's school-cohort as the place of residence and the age determine the assignment to specific school-cohorts. The government implicitly requires educational facilities to provide equal access and uniform standards for students.[7] An overwhelming majority of students in our sample ($\approx$

---

which is key to better understand how lifetime earnings are affected. Third, we complement the previous evidence by providing detailed evidence on occupational sorting, particularly regarding gender conformity, and mechanisms (e.g., educational attainment, high-school track choices, gender congruence of occupational choices). Fourth, we focus on earlier student exposure of the whole student population in primary school, rather than on students in high-school when students have selected into specific tracks and schools and some important choices have already been made (Brown and Corcoran, 1997).

[5] Most municipalities set a maximal acceptable travel distance. Some municipalities apply a measure known as relative distance (relativ närhet). This metric involves comparing the relative distance between a school and the next best alternative across students. See, for example, the explanation by the schooling authority for Stockholm. In the 1990s, charter schools (friskolor), which could apply additional criteria for admission, were introduced into the Swedish system. However, we only use data from schools we consistently observe from 1989 to 2002. Accordingly, we do not have charter schools in our sample. (In the Swedish system, there is a distinction between charter schools (friskolor) and private schools. While charter schools can collect profits, they are still funded by the state. The number of actual private schools that operate on their own costs is exceedingly low.)

[6] Note that, at the beginning of each of these stages, students are assigned to classes in which they remain for the duration of the stage. Thereby, schools have some discretion over which stages to offer and when to reshuffle classes. However, there is no reshuffling after grade 7 (around age 13) and students remain in the same class. Across these stages, the cohort composition remains very stable. Reasons for changes would include grade repetition, which is very rare in the Swedish primary-school context (Collins and Lundstedt, 2021) and people moving. We address the concern of movers in Section 5.

[7] Swedish Primary School Regulation (Grundskoleförordning), SFS:1994:1194, 4 kap 4§.

96%) thus comply with the requirement and enter school in the year they turn seven. Previous research using Swedish data shows that for a different time window, classroom and cohort gender composition are indeed orthogonal to parental and student characteristics (Getik and Meier, 2020). A battery of balance and robustness checks detailed below corroborate previous findings and the adherence of schools, parents, and principals to institutional rules, suggesting that gender composition across school-cohorts is largely arbitrary.

In the last year of primary schooling, students can apply for high-school admission within their municipality. Just under 90% of students in our sample complete high-school, which takes three years. There are currently 18 high-school programs that students in Sweden can choose from: 6 academic, and 12 vocationally oriented.[8] Graduation from an academic program provides the necessary basic qualification to enter a university, with the Natural Sciences track allowing for admission into the widest range of university programs. In the vocational programs, there is a possibility to fulfil extra requirements to attend university if the student chooses a sufficient load of academic courses. Students can then apply for specific programs at a university of their choice. Admission is based on high-school grades or the results of a national test if the grades are insufficient.

2.2 Data and Descriptive Statistics

Sample. We use administrative data on primary-school cohort composition for the years 1989 to 2002, linked to labour-market outcomes at the age of 30 for the years 2003 to 2016. We link the data using an anonymized personal ID issued to Swedish residents. Data restrictions that we face come from two main sources. First, the data on primary-school cohort composition starts in 1989. Second, we want to observe people up to age 30, or 14 years after primary schooling. Income data is available until 2016, which means we can use primary school data up to 2002.

The sample includes all Swedish primary-school students who completed that stage of education between the years of 1989 and 2002 and whose school we observe consistently for

[8] The number and the list of available programs has changed over time. The academic programs available to students during the period covered by our data are: Natural Sciences; Social Sciences; Humanities; Arts; Healthcare; The Industrial Program. The vocationally oriented programs are: The Food Program; The Handicraft Program; Natural Resources; The Construction Program; Children Recreation; Electrical Engineering; Vehicle Engineering; The Energy Program; Business Administration, The Hotel, Restaurant and Catering Program; and Media.

the entire duration.[9] From this sample, we remove students for whom information about both parents is missing. Finally, following Brenøe and Zölitz (2020), we exclude cohorts with fewer than 10 students (around 1% of the sample). Those are excluded primarily to avoid conflation with non-regular education, such as evening classes. The final dataset consists of 752,560 primary-school students from 539 schools across Sweden.

Calculating the Share of Female Peers by School–cohort. To calculate the share of female peers in a school–cohort, we use data from the Primary Schooling Outcomes Register (Registret över ämnesprov i årskurs 9). The data for each student is recorded once upon completion of primary school. The register contains an indicator for which school an individual attended in their last stage of primary schooling and the year when they completed it.[10] Accordingly, we know each student's school and cohort based on the combination of the school attended and the year of degree completion.[11] Since we observe the gender of each student, we can then calculate the share of female peers for each school–cohort during primary school by counting the number of female students and dividing it by the total number of students.

Educational Outcomes: Grades, High-school Tracks, University Enrollment, College Degree. To construct these variables, we use the data from the Primary Schooling Exams Register, as well as the High-School Outcomes (Registret över slutbetyg från gymnasiet), University Registration (HregGrundRegistrerade), and the Accumulated Education (Utbildning Ackumulerad) registers. The Primary Schooling Exams Register records students' grades and outcomes in primary schooling. The High-School Outcomes register their study programs in high school, and the University Registration register provides information about university programs or stand-alone courses for which they registered. Finally, the Accumulated Education register records the highest level of education obtained by everyone

[9] Using a sample including also schools which we do not observe for all years yields similar results with the coefficient for female x share females being SEK 11,862 ($se = 3,861$) and the gap between the impact on males and females being SEK 17,600 ($se = 5,472$).

[10] Unfortunately, the data do not contain a classroom indicator. The classroom indicator variable is available only in the exams register, which starts in the year 2004, two years after the last cohort in our sample.

[11] Note that cohort composition at the end of primary school is a good proxy for cohort composition throughout primary school, because of low grade retention and little moving. Moreover, for a relatively high share of the schools, there is no division into three primary school stages, but only two or even one. According to Skolverket, in the academic year 2019/2020, approximately 26% of the schools did not subdivide into stages at all. This further limits changes in cohort composition.

in the population, which allows us to infer whether someone completed the program they started.

Labor Market Outcomes: Wages and Occupation. We use the Income and Taxation register (Registret över inkomster- och taxeringar) to link a student's subsequent earnings to their former school–cohort composition. Using the Occupation register (Yrkeregistret), we do the same for their subsequent chosen occupation. Note that occupation data is available for three years fewer than wage data, which reduces the sample size when we consider occupation-related outcomes. Our two main dependent variables are the former students' annual wage and the median wage in a former student's occupation, both measured at age 30. We choose age 30 for two main reasons. First, examining age 30 allows us to observe labor market outcomes for people who attend college and, on average, finish at age 26. Second, we do not lose too much power when we restrict the sample only to those we observe until age 30.[12]

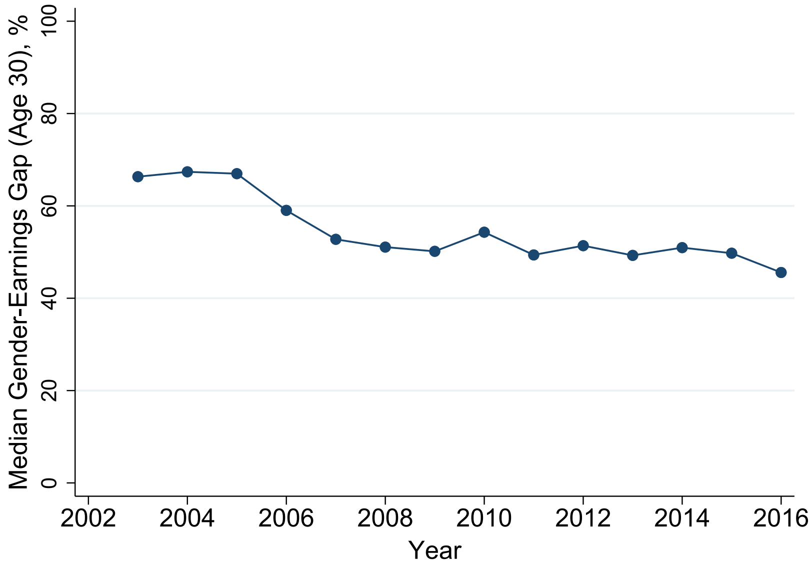
The annual individual wage is labor income as recorded in the income register. This data comes from incomes employers directly report to the tax office for taxation purposes, so there is little leeway, if any, for manipulation. Occupational information comes from the Occupation register and contains a three-digit code which identifies and individual's profession. Each digit represents a subsequent sub-division into categories of occupation, starting with eight main categories. We use these eight main categories to examine which types of occupations people select.

We use the more detailed account of 186 unique occupations to calculate a proxy for wage potential by taking the median wage for each occupation based on all Swedes aged 31 and older. To do this, we combine this information with the income register. We use this outcome to assess potential effects of gender composition on selection into better-paying occupations. We also use the fine-grained categorization to calculate the share of males in each occupation, again based on all Swedes aged 31 and older.

Descriptive Statistics. Figure 1 shows the development of the unadjusted relative gender gap in labor income over time. While the gap has also been decreasing in Sweden, the reduction of the gender gap has stagnated in recent years. Over the whole sample period,

[12]The findings are similar when using age 29 or 31 as the cutoff for wages, see Table D.2.

Figure 1: Raw Wage Gap at Age 30—Median Wages of Women Are Roughly 45% Lower than Men's



Note: The figure represents the gender gap in median annual earnings in relative terms at the age of 30 across the years for which we observe labour-market outcomes in our sample. The y-axis denotes the percentage difference between crude median male and female earnings in a given year. The underlying data is adjusted for inflation.

the average the raw difference between male and female wages corresponds to roughly SEK 80,000 at the age of 30.

Table A.1 presents more comprehensive summary statistics. The first panel shows background variables related to family characteristics. Boys comprise approximately 51% of the sample, and there is little difference between the genders with respect to family characteristics. The second section shows school-level variables. An average school-cohort consists of around 123 students. Approximately 10% of the students complete the natural sciences track in high school, and around a quarter of students proceed to university. There is no meaningful difference with respect to cohort size and composition between the genders. However, girls in the sample obtain higher grades and are more likely to enroll in college. The last section describes labor market outcomes, our main focus. These are recorded when the former students reach the age of 30, approximately 14 years after completion of primary schooling. For both annual and occupation wage, men out-earn women at that age, in spite of the previously higher educational attainments of girls. There also is gender homophily in occupational choice: Individuals at the age of 30 work in occupations comprised of more than 60% of their own gender.

3 Empirical Strategy and Plausibility Checks

3.1 Specifications

We estimate the effects of cohort gender composition on labor market outcomes using the following main specification:

$$\begin{aligned} Y_{isc} = {} & \beta_1 \times Female_i \times ShareFemPeers_{isc} + \beta_2 \times Male_i \times ShareFemPeers_{isc} + \\ & \beta_3 \times Female_i + \alpha_{school} + \delta_{cohort} + X_i\gamma' + e_{isc} \end{aligned} \tag{1}$$

Y_{isc} is the outcome of interest for student i in school s and cohort c. The explanatory variable is $ShareFemPeers_{isc}$, which represents the proportion of female students in a given school-cohort. It is calculated as $\frac{ngirls_c}{size_c}$, where $ngirls_c$ is the number of girls in a given cohort and $size_c$ is the cohort size. The estimate for β_1 describes the effect of the share of female peers on women, β_2 describes the effect on men, and β_3 represents the gender gap in Y_{isc}. α_{school} denotes school-fixed effects, while δ_{cohort} denotes cohort fixed effects. X_i is a vector of school trends, individual- and school-level controls. In our preferred specification, we combine school and cohort fixed effects with school-specific trends. The vector of controls includes parental education (measured by whether at least one of the parents went to college), log family income, and an indicator for a single-parent household. It also includes cohort size (Epple and Romano, 2011) and the number of schools available in a given municipality. We cluster standard errors on the school level, thus allowing students' outcomes to correlate within schools.

In the tables we also show $\beta_1 - \beta_2$ which we call "Gap". It denotes the gender gap in the impact of female peers on each outcome — impact on women minus impact on men — and therefore shows by how much the gender gap changes with more girls in the cohort.

When analyzing peer effects at the school-cohort level, the primary threat to identification lies in potential sorting of students. Here, the institutional context ensures that there is little selection apart from geographic location and age into schools and cohorts. The inclusion of school fixed effects absorbs static heterogeneity in selection into schools, and cohort fixed effects control for national level changes which affect all students in a cohort. A large literature shows that peer composition, conditional on school and cohort fixed effects, is arbitrary in many contexts (see, e.g, Hoxby, 2000; Black, Devereux and Salvanes, 2013; Helene et al., 2015; Carrell, Hoekstra and Kuka, 2018). While unlikely, dynamic selection across schools and cohorts may be a potential threat to identification.

We do two things to address potential dynamic selection into school–cohorts across schools, over time. First, we include school-specific time trends to analyze deviations from peer composition conditional on dynamic trends (see, e.g., Carrell, Hoekstra and Kuka, 2018). This allows to control parametrically for unobserved factors that may correlate with time trends in cohort composition. Second, we include fixed effects for each individual cohort in a given school (see, e.g., Brenøe and Zölitz, 2020). This more conservative specification including *school*×*cohort* fixed effects allows us to address remaining selection into different schools in different years. A drawback is that we are then only able to compute the impact on the gender gap in outcomes and not also on levels. This second specification looks as follows:

$$Y_{isc} = \beta_1 \times Female_i \times ShareFemPeers_{isc} + \beta_2 \times Female_i + \alpha_{school} \times \delta_{cohort} + X_i\gamma' + e_{isc} \quad (2)$$

The underlying assumption for a causal interpretation in the first, main specification including separate school and cohort fixed effects as well as school-specific trends, is that no omitted variable simultaneously satisfies the following conditions: (i) time-variant and cohort-specific; (ii) not captured by school fixed effects, cohort fixed effects, or linear time trends (iii) correlates with peer composition as well as labour-market outcomes, (iv) not included in our vector of controls. The existence of such a variable seems highly unlikely given the rate of compliance with the school starting age. Still, to assess the likelihood of such a variable existing, we examine the relationship between high-quality and detailed observable characteristics from administrative registers and gender peer composition in the cohort. In addition, we include school-by-cohort fixed effects, which fully accounts for potential changes in sorting to schools over time. Taken together, the results from a battery of checks which we detail below suggest no rejection of the main identifying assumptions.

3.2 Balance and Placebo Checks

While specification 1) with separate school and cohort fixed effects addresses static selection into schools or cohorts, we now check whether remaining variation is likely arbitrary. The following list summarizes the key tests:

1. Across 72 bivariate regressions we find no indication that maternal and paternal education and income or factors such as family size or immigrant status systematically predict the gender peer share.
2. We also find no indication that there is an imbalance in family characteristics across quintiles of the share of female peers.
3. The residualized share of female peers is normally distributed.
4. The distribution of the residualized share of female peers is similar to the simulated residualized share of female peers when we randomly assign students to cohorts.
5. Student gender does not predict the leave-one-out share of female peers.
6. School-by-cohort fixed effects do not jointly predict student gender.
7. Including school-by-cohort fixed effects does not markedly affect the coefficient size of estimated gender differences.
8. The share of females in previous and subsequent cohorts does not consistently affect individual wages or occupational wage potential.

Family Characteristics. We provide a series of balance checks for high-quality background variables on parental and family characteristics from administrative data.[13] The 24 variables include, among others, detailed education and labour-market outcomes for each parent, as well as multiple family composition variables.

For each variable, we examine whether there is a correlation with the share of female peers in the cohort across three specifications including: (1) separate school and cohort fixed effects, (2) fixed effects and school trends, and (3) fixed effects, trends, and controls. Table 1 shows the results of the 72 bivariate regressions.[14] If there was a consistent correlation between these variables and the share of female peers, this would indicate that parents may be able to select

[13] We measure those just before children start the last phase of primary school.

[14] Using school-by-cohort fixed effects would not allow us to do the balancing checks since the share of female peers is the same within each school specific cohort.

Table 1: Bivariate Balance Checks

Dependent Variable:	(1)	(2)	(3)
Mother High School	0.015	0.009	0.009
	(0.013)	(0.011)	(0.011)
Father High School	0.007	0.010	0.010
	(0.012)	(0.011)	(0.011)
Mother Vocational Degree	0.016*	0.012	0.013
	(0.008)	(0.008)	(0.008)
Father Vocational Degree	0.006	0.005	0.005
	(0.007)	(0.007)	(0.007)
Mother College Degree	0.015	0.009	0.009
	(0.009)	(0.009)	(0.010)
Father College Degree	0.013	0.009	0.009
	(0.009)	(0.009)	(0.009)
Mother STEM Degree	0.002	0.003	0.003
	(0.002)	(0.002)	(0.002)
Father STEM Degree	-0.005	-0.004	-0.004
	(0.004)	(0.004)	(0.004)
Log Family Income	0.013	-0.076	-0.082
	(0.118)	(0.081)	(0.081)
Wage Mother (1000 SEK)	1.993	-0.800	-0.800
	(3.463)	(2.446)	(2.446)
Wage Father (1000 SEK)	1.875	1.275	1.275
	(6.445)	(4.769)	(4.769)
Mother Unemployed	-0.292	0.192	0.240
	(0.950)	(0.809)	(0.806)
Father Unemployed	0.561	1.049	1.072
	(0.870)	(0.815)	(0.817)
First-Born Child	-0.009	-0.011	-0.011
	(0.012)	(0.012)	(0.012)
Number Siblings	0.006	0.019	0.019
	(0.033)	(0.028)	(0.028)
Immigrant	-0.001	0.003	0.004
	(0.011)	(0.007)	(0.007)
2nd Generation Immigrant	0.007	0.006	0.007
	(0.012)	(0.010)	(0.010)
Adopted	-0.003	-0.001	-0.001
	(0.003)	(0.003)	(0.003)
Age Mother	0.083	0.088	0.088
	(0.143)	(0.129)	(0.129)
Age Father	0.119	0.125	0.128
	(0.151)	(0.141)	(0.141)
Mother Unknown	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)
Father Unknown	0.001	-0.000	-0.000
	(0.003)	(0.003)	(0.003)
Single Mother	0.008	0.004	0.004
	(0.006)	(0.006)	(0.006)
Single Father	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)
School FE	X	X	X
Cohort FE	X	X	X
School Trends	-	X	X
Controls	-	-	X

Note: The table shows the estimated relationship between student family characteristics and the share of female peers in their cohort. The specifications in the table incrementally include school and cohort fixed effects, school trends, and school-level controls. Those controls include cohort size and the number of schools in the municipality. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

into specific school–cohorts. In the absence of systematic sorting, one would expect to have approximately 10%, 5%, and 1% of the coefficients to be significant at each corresponding level. Across all variables and specifications, we find only one coefficient that is statistically significant at the 10% level. This is below what one would expect by chance. The results are in line with previous findings of Getik and Meier (2020), who find that the share of female peers is arbitrary in Swedish primary schools at the classroom and at the cohort level.

Balance Across the Distribution of the Share of Female Peers. One concern may be that there is a balance in family characteristics when including the share of female peers linearly, but not when using more flexible functional forms. We check whether there is any indication that different quintiles of the share of female peers systematically relate to family characteristics in Table B.1. We run one regression including quintiles for each family characteristic. Across 96 coefficients, there is 1 coefficient statistically significant at the 1% level, 1 coefficient statistically significant at the 5% level, and 4 coefficients statistically significant a the 10% level. Accordingly, there is no systematic correlation of family characteristics across the distribution of the share of female peers. The results corroborate the findings from the linear specifications above: Cohorts with more girls are comparable to cohorts with fewer girls.

Gender and the Share of Female Peers. A violation of the identifying assumption would occur if there was gender-based selection into schools. To test this, we examine whether a student's own gender correlates with the leave-one-out share of girls in their cohort, following the methodology proposed by Guryan, Kroft and Notowidigdo (2009). Across all specifications, we control for the school-level leave-one-out cohort mean of the share of girls to account for the mechanical relationship between peer and own gender. That is, we control for the average share of female peers in the other cohorts of a student's school. There is no statistically significant correlation between own gender and the share of female peers (see Table B.2).

Distribution and Simulated Distribution of the Share of Girls. In another plausibility check of as-good-as-random assignment to school-cohorts, we examine the variation of the gender peer share that we eventually exploit. If gender peer share were as-good-as-randomly assigned at the school-cohort level, we would expect that the corresponding distribution of

Figure 2: Residual Share of Female Peers Across School-Cohorts

Note: The figure above represents the distribution of the residualized female peer share across school-cohorts, conditional on separate fixed effects for schools and cohorts. The overlaid curve represents normal distribution.

peer shares would look normally distributed, conditional on school and cohort fixed effects. Figure 2 suggests that the residual share of female peers is indeed well-behaved and follows a normal distribution, suggesting that the peer gender composition is likely arbitrary.

We further test whether peer-gender variation within schools is consistent with random assignment by comparing the actual distribution to a simulated distribution of the female peer share. To this end, we do Monte Carlo simulations in which we assign students randomly to cohorts within their schools. We take the number and size of cohorts from the actual data. Similarly to Bietenbeck (2020), we then regress the share of female peers on school and cohort fixed effects in the simulated data and collect the residuals. We plot the simulated residuals from random assignment alongside the residuals from the actual data in Figure B.2. The distributions look very similar, a result consistent with as-good-as-random assignment of the share of female peers.

Gender and Cohorts. Following Chetty et al. (2011), and Balestra, Eugster and Liebert (2020), we regress student gender on school-by-cohort fixed effects. The school-by-cohort fixed effects should be jointly insignificant if assignment to a school specific cohort is independent of student gender (Chetty et al., 2011).

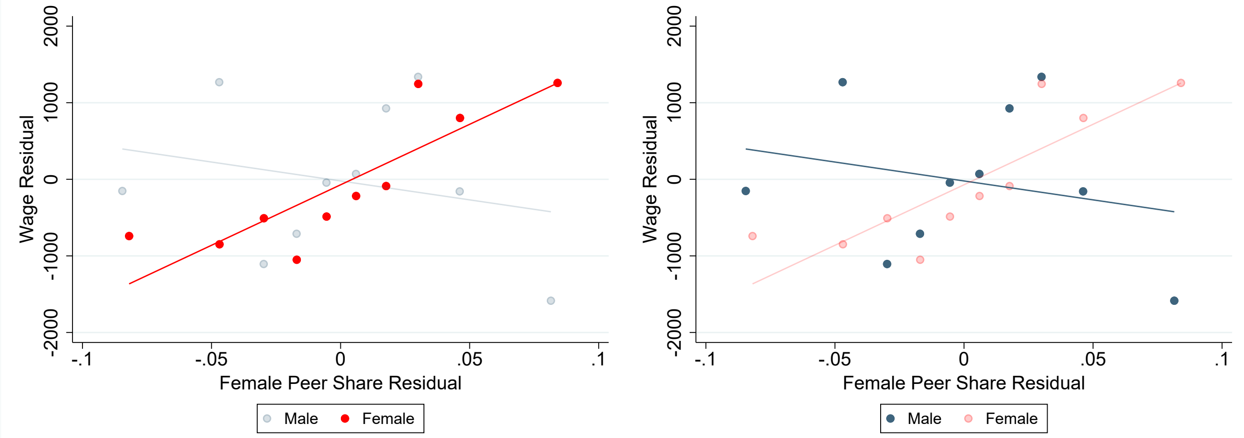
We proceed as follows: In the first step, we regress student gender on separate school and cohort fixed effects as well as on controls and we then retrieve the residuals from this regression. In the second step, we regress the residuals obtained in the prior regression on school-by-cohort fixed effects. We then do a joint F-test to determine whether the school-by-cohort fixed effects are jointly significant. Across three different specifications the F-statistics suggest no predictive power of school-by-cohort fixed effects for student gender: $F = 1.019$ without controls, $F = 1.020$ with school-level controls, and $F = 1.020$ with school-level and individual-level controls. The F-statistics are all not statistically significant at the 10% level.

Stability of Gender Differences to School-by-cohort Fixed Effects. While it is not possible to estimate the effect on levels of outcomes in specifications with school-by-cohort fixed effects (specification 2), it is possible to estimate how the share of girls shapes the difference between girls and boys.

Examining the stability of the gender gap to the inclusion of school-by-cohort fixed effects is informative: If including school-by-cohort fixed effects changes the estimated effect of the gender peer share on the gender gap in labor market outcomes, this may indicate potential selection. However, if the estimated impact on the gender gap remains similar when we include school-by-cohort fixed effects, it is unlikely that selection into specific school cohorts has a significant effect. The stable impact on gender gaps across specifications (see, e.g., Table 2), therefore, render it unlikely that selection into specific school-cohorts is a key driver of the main results.

Placebo Check: Previous and Subsequent Cohorts. Finally, we do a placebo check examining whether the share of female peers in other cohorts affects outcomes. More specifically, we examine whether the share of female peers in the previous or the past cohort affects labor-market outcomes. Table B2 shows that none of the corresponding regression coefficients are statistically significant at the 5% level. Across 24 coefficient estimates three coefficients are statistically significant at the 10% level. The coefficient estimates are substantially smaller than the impact of the current cohort and have opposing signs. Taken together, the results indicate that our estimates capture idiosyncratic variation coming from the current cohort rather than from previous or past cohorts.

Figure 3: Primary-School Gender Composition and Wages



Note: The figure shows the relationship between residualized wages measured at age 30 and the residuals of the share of female peers in a given school-cohort (N = 752,560). The residuals stem from regressions of the respective variables on school and cohort fixed effects, school-specific trends, and controls (specification 3 in Table 2). The dots show the binned averages across the deciles of the distribution. The left section of the figure highlights the linear fit from OLS regressions for females ($\beta = 17,211$, $se = 5,172$), whereas the right section does so for males ($\beta = -6,714$, $se = 5,674$).

4 Gender Composition, Wages, and Wage Potential

4.1 The Impact of Primary School Gender Composition on Wages

Figure 3 and Table 2 show that a higher share of female peers in primary school increases women's wages at age 30, reducing the gender wage gap. In Table 2, Female × Share Females shows the effects on the wage level of females, Male × Share Females shows the effects on the wage level of males. "Gap" shows the difference between the two effects and therefore indicates by how much an increase in the share of female peers reduces the gender gap. Finally, the female dummy shows the wage gap, conditional on controls.

The effects on female wage levels and the gender wage gap are precisely estimated across specifications and substantial. The coefficient estimate in column (3) suggests a SEK 3,442 or \$404 increase in annual female wages when changing from a 40% female to a 60% female cohort. Neglecting wage trajectories across age, this corresponds to a \$14,140 (35 years x \$404) lifetime earnings difference up to the retirement age of 65.

But even a 10 percentage point increase in the share of females (avg. = 49%) results in an SEK 1,721 or \$202 increase in annual wages for women at age of 30. Similarly, a 10 percentage point increase in the share of females, which corresponds to 12 more female students in a cohort, reduces the gender wage gap by 2,419 SEK. The effect is economically meaningful:

Table 2: Primary-School Gender Composition and Wages

	Annual Wage Mean: 218,380				
	(1)	(2)	(3)	(4)	(5)
Female × Share Females	16,822***	18,626***	17,211***		
	(5,062)	(5,199)	(5,172)		
Male × Share Females	-9,072	-7,704	-6,714		
	(5,912)	(5,776)	(5,674)		
Gap	26,098***	26,491***	24,197***	26,488***	23,552***
	(7,571)	(7,587)	(7,536)	(7,638)	(7,582)
Female	-90,886***	-91,077***	-90,002***	-91,072***	-89,681***
	(3,732)	(3,739)	(3,719)	(3,766)	(3,743)
School FE	X	X	X	-	-
Cohort FE	X	X	X	-	-
School Trends	-	X	X	-	-
Controls	-	-	X	-	X
School × Cohort FE	-	-	-	X	X
Observations	752,560	752,560	752,560	752,560	752,560
Schools	537	537	537	537	537
R-squared	0.09	0.09	0.10	0.10	0.11

Note: The table shows the estimated relationship between annual wage at age 30 and the share of female peers in one's cohort. The wage is recorded in Swedish crowns (SEK). The first row shows the coefficient estimates for women; the second row for men. The third row shows the difference in response to the share of female peers between the genders. Finally, the last row "Female" shows the gross difference in annual wage between the genders. The coefficients in the first three columns are based on the first specification that relies on school and cohort fixed effects. Columns (4) and (5) show estimates from our second specification, which include school-by-cohort fixed effects. Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are based on clustering at the school level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A 10 percentage point increase in the share of female peers reduces the raw gender wage gap by 2.7%.

The impact on female wages are also large when compared to teacher value-added estimates. For instance, Chetty, Friedman and Rockoff (2014) estimate that a one standard deviation improvement in teacher value-added leads to a $350 wage increase at age 28. Chetty et al. (2011) find that kindergarten students with teachers who had more than 10 years of experience earn $364 more per year between the ages of 25 and 27. In comparison, we find that modest changes in the gender composition, a 10 percentage point shift in the cohort, changes annual wages by $202 at age 30. The effects are also large when compared to the association of parental education and child wages. A parental university degree translates into

$2,065 higher yearly wages, a 10 percentage point shift in the gender composition therefore amounts to almost 10% this relationship. Taken together, the coefficient estimates indicate an important role of gender composition in primary school for female wage levels and the gender wage gap.

Based on related evidence on the impact of gender environments on outcomes during school years, we expect that the classroom composition drives most of the effect of the gender environment captured by the cohort composition.[15] The estimates here are, therefore, likely a lower bound of the true effect given that we examine cohorts rather than classes.

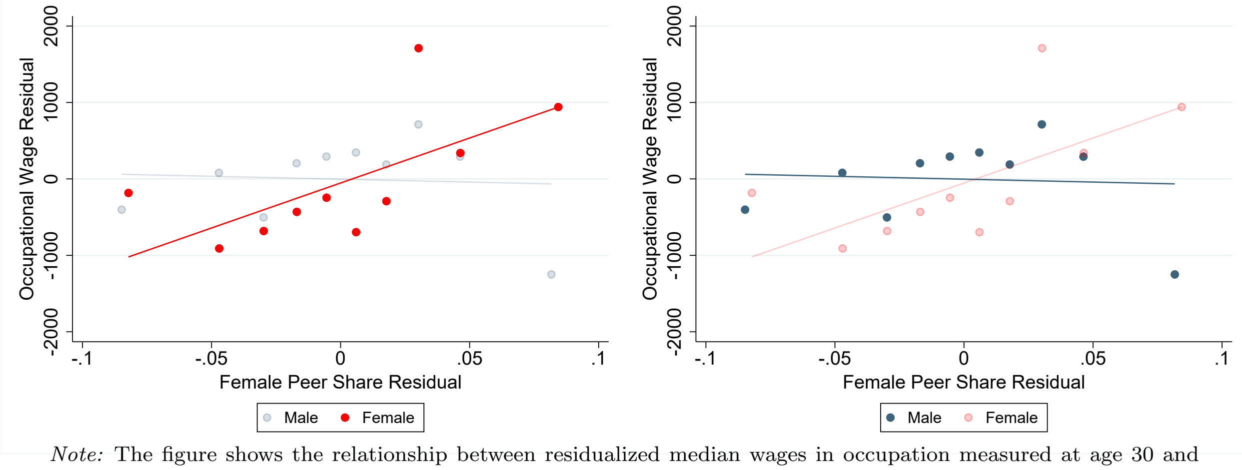
In contrast, men do not seem to benefit from more girls in their cohort in terms of later wages. Based on specification (3), a 10 percentage points higher share of female peers leads to a statistically insignificant reduction of SEK 671 or $79 annual wage among men.

The reduction in the gender wage gap is robust to the inclusion of school and cohort-fixed effects in column (1) in Table 2, school-specific time trends in column (2), as well as to parental and school-level controls in column (3). The gender gap estimates remain stable even when including fixed effects for each school specific cohort in column (4) without controls and in column (5) with controls. Specifications (4) and (5) allow us to address potential remaining concerns about selection on the school-cohort level. While we cannot observe gender-specific effects in these specifications, we can still observe the difference between the effects on males and females. The robustness to the different specifications suggests that the estimates are not the results of static or dynamic sorting into schools, cohorts, or school specific cohorts (Oster, 2019).

We also consider how the female wages shift across the wage distribution in Table C.1. We find that women with more females in the cohort are 4 percentage points less likely to end up in the lowest quintile of the wage distribution, but 4 percentage points more likely to end up in the highest quintile of the wage distribution. The point estimates suggest corresponding shifts across the wage distribution: Women are less likely to find themselves in the bottom half of the income distribution and substantially more likely to find themselves in the upper half when exposed to more girls in primary school. In summary, the results indicate that women earn higher wages after socialization among girls, putting a substantial dent in the gender wage gap.

[15] For instance, when studying the impact of the gender environment on mental health during school years using Swedish primary school data, we find that estimates on classroom level are approximately 60% larger than those on the cohort level (Getik and Meier, 2020).

Figure 4: Primary School Gender Composition and Occupational Wage Potential



Note: The figure shows the relationship between residualized median wages in occupation measured at age 30 and the residuals of the share of female peers in a given school-cohort (N = 652,115). The residuals stem from regressions of the respective variables on school and cohort fixed effects, school-specific trends, and controls (specification 3 in Table 3). The dots show the binned averages across the deciles of the distribution. The left section of the figure highlights the linear fit from OLS regressions for females ($\beta = 11,040$, $se = 3,394$), whereas the right section does so for males ($\beta = -1,600$, $se = 3,084$).

4.2 The Impact of Gender Composition on Wage Potential

Figure 4 and Table 3 indicate that a higher share of female peers in primary school not only decreases the gender wage gap at age 30, but also the gap in wage potential. Here, the dependent variable is the median wage in the occupation each individual works in. The results indicate that females select into higher wage occupations after attending primary school with more female peers.

Again, the estimated impact on female wage levels and on the change in the gender gap are stable and statistically significant across specifications. In specification (3), we find that a 10 percentage point increase in the share of female peers in primary school increases female occupational wage potential by 1,104 SEK per year. Neglecting wage trajectories across age, this corresponds to a remaining SEK 38,640 (35 years x 1,104 SEK) or $4,546 lifetime earnings difference up to the retirement age of 65. When changing from a 60% female cohort to a 40% female cohort, the lifetime foregone occupational wage potential could be up to $9,091. Correspondingly, a 10 percentage point increase in the share of female peers reduces the gender wage gap by SEK 1,286 which corresponds to a 3.2% of the raw gap in occupational wage potential between males and females. The results indicate that gender peer composition in primary school does not only affect current wages, but also has substantial effects on lifetime earnings potential.

Table 3: Primary School Gender Composition and Occupational Wage Potential

	Median Wage in Occupation Mean: 281,634				
	(1)	(2)	(3)	(4)	(5)
Female × Share Females	11,906***	12,661***	11,040***		
	(3,432)	(3,467)	(3,394)		
Male × Share Females	-2,476	-2,112	-1,600		
	(3,118)	(3,111)	(3,084)		
Gap	14,633***	15,025***	12,868***	14,845***	12,422***
	(4,693)	(4,698)	(4,647)	(4,711)	(4,676)
Female	-40,579***	-40,772***	-39,701***	-40,679***	-39,477***
	(2,296)	(2,299)	(2,278)	(2,309)	(2,295)
School FE	X	X	X	-	-
Cohort FE	X	X	X	-	-
School Trends	-	X	X	-	-
Controls	-	-	X	-	X
School × Cohort FE	-	-	-	X	X
Observations	652,115	652,115	652,115	652,115	652,115
School-Cohorts	537	537	537	537	537
R-squared	0.20	0.20	0.22	0.21	0.23

Note: The table shows the estimated relationship between median wage in a given individual's occupation at age 30 and the share of female peers in one's cohort. These wages are computed based on 186 unique occupations in our registers and are recorded in Swedish crowns (SEK). The first row shows the coefficient estimates for women; the second row for men. The third row shows the difference in response to the share of female peers between the genders. Finally, the last row "Female" shows the gross difference in annual wage between the genders. The coefficients in the first three columns are based on the first specification that relies on school and cohort fixed effects. Columns (4) and (5) show estimates from our second specification, which include school-by-cohort fixed effects. Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5 Robustness Checks

Estimates from Split Samples by Gender. In a first robustness check, we estimate the effect of the share of female peers on annual wages and occupational wage potential separately for women and men. A disadvantage of estimating the effects relying only on men or women is that it makes it impossible to estimate the impact on the gender gap using school-by-cohort fixed effects. An advantage of splitting the sample is that potential interactions of gender with fixed effects and controls are absorbed.

Table D.1 shows the estimated effects on wage levels and occupational wage potential separately by gender. The estimates for women's wages remain positive and statistically significant throughout. The coefficient sizes are not statistically distinguishable from the main estimates. Again, the point estimates for men are negative and statistically insignificant. Estimating the impact of the share of female peers with split samples therefore does not substantially affect the coefficient estimates.

Different Ages for Measuring Current Wages and Using Log Wages. In a second robustness check in Table D.2, we examine what happens when we use ages 29 and 31 as the ages at which we measure wages and occupational wage potential. Note that we lose some observations when examining wages at age 31, since we observe one cohort less up to this age. We observe no large differences across ages: Across ages 29, 30, and 31, we observe that wage levels of women are higher when they were exposed to more girls in their cohort, leading to a reduction in the gender wage gap. Accordingly, the results are not sensitive to using a specific cutoff age around age 30.

In Table D.3 we also examine the impact on log wages. We see robust and statistically significant effects of the share of female peers.

Movers. Individuals who stay in the same municipality do not have an institutional reason to change schools. However, could people who move from one municipality to another drive the results? To examine this concern, we split the sample into those who move across municipalities and those who do not. We define non-movers as students who lived in the same municipality throughout the entire nine-year period of primary schooling as well as the year before the start of primary school. Just over 83% of students do not move across municipalities and therefore likely remain in the same school. The coefficient estimates across the groups do not differ substantially (see Table D.4). This suggests that the results are not driven by students moving.

Outliers in Gender Peer Share and Small Cohorts. The results might be driven primarily by the tails of the distribution. To examine the possibility of outliers driving the results, we first estimate the main specification without the top and bottom 5% of cohorts with the highest share of female peers. In our data, these cut-offs roughly correspond to cohorts having 60% or more students of the same gender. We then also drop the smallest

10% of cohorts. Dropping extreme observations results in qualitatively equivalent coefficient estimates (see Table D.4).

Heterogeneous Effects. We also examine heterogeneous effects by students' socioeconomic background, cohort, or municipality size (parental unemployment, single parenthood, parental college degree, parental income, cohort size, and the number of schools in the municipality). Table C.2 shows the corresponding interactions. We do not observe any statistically significant heterogeneity in effects across socio-demographics, cohort, or municipality size for either gender. In sum, the results suggest that the effects are likely not driven by a particular socio-economic group, nor by smaller schools or municipalities.

6 Mechanisms: Occupational and Educational Selection

Why do we observe an increase in women's wages after they were exposed to more girls in primary school? Previous research highlights the impact of the gender composition in school on educational attainment and choices (see, e.g., Lavy and Schlosser, 2011; Park, Behrman and Choi, 2012; Black, Devereux and Salvanes, 2013; Eisenkopf et al., 2015; Anelli and Peri, 2019; Giardili, 2020; Brenøe and Zölitz, 2020). However, how gender composition affects later occupational selection remains an open question. Table 4 shows the impact throughout a students' educational and labor market career, its effects on educational attainment, educational selection, and labor market selection. This table guides our exploration of what drives the impact of female peers on labor market outcomes.

Educational Attainment, Networks, and Selection. There are two main educational stages between primary school and labor market outcomes that we can observe: high school (gymnaiset) and tertiary education in university or vocational training. We can trace each student's school path: primary- and high-school grades, entry into high school and the study track there, as well as tertiary education choices.

We first document pronounced effects on primary school grades. An increase in the share of female peers improves grade point averages for girls and worsens boys' grades. The esti-

Table 4: Summary of Main Mechanisms

Dependent variable:	**Female**	**Male**	**Gap**
	(1)	(2)	(3)
Primary-School Grades (.001)	0.082**	-0.141***	0.223***
	(0.038)	(0.038)	(0.053)
High-School Grades (.051)	0.011	-0.045	0.056
	(0.039)	(0.039)	(0.046)
Enter High School (.852)	0.037***	-0.002	0.040**
	(0.014)	(0.016)	(0.018)
Science Track High School (.112)	0.030**	0.004	0.026
	(0.013)	(0.014)	(0.018)
Cohort-Mates in High School (9.00)	1.638**	-1.335**	3.000***
	(0.709)	(0.675)	(0.688)
College Enrolment (.264)	0.007	0.006	0.002
	(0.017)	(0.017)	(0.025)
STEM Enrolment (.108)	0.007	0.004	0.004
	(0.009)	(0.013)	(0.016)
Female Share High-School Track (.497)	-0.037*	-0.027	-0.010
	(0.022)	(0.022)	(0.041)
Female Share College (.59)	-0.021**	-0.001	-0.021
	(0.011)	(0.016)	(0.023)
Female Share Occupation (.496)	-0.042**	0.008	-0.050
	(0.017)	(0.017)	(0.031)
Married by 30 (.225)	-0.020	0.006	-0.025
	(0.015)	(0.014)	(0.020)
Has Children by 30 (.299)	-0.038**	0.013	-0.051**
	(0.017)	(0.015)	(0.023)
Unemployed (.097)	0.000	0.005	-0.005
	(0.011)	(0.009)	(0.015)
School FE	X	X	X
Cohort FE	X	X	X
School Trends	X	X	X
Controls	X	X	X

Note: The table estimates the effect of the share of female peers in classroom on the observable intermediate outcomes that represent potential channels through which future income can be affected (population mean provided in parentheses next to the variable). The coefficient estimates below "Female" represent the interaction of the share of female peers with being a girl. The coefficient estimates below "Male" represents the impact of female peers on boys. "Gap" shows the difference in response to the share of female peers between the genders Both the primary- and high-school grades variables are standardized. The share variables represent the proportion of females, entry, enrolment and completion variables are binary. The fertility variable takes on a value of 1 if the student had a child prior to the age of 30. Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. We include all the observation for which we have data on. Standard errors (in parentheses) are clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

mate suggests that a 10 percentage point increase in the share of female peers results in an approximately 6% widening of the grade gap favoring girls.

We further study what happens to peer networks. We find that girls with more girls in their cohort during primary school later on go to high-school with more of their previous same-gender peers. The opposite happens for boys. Accordingly, more same-gender peers in primary school lead to stronger same-gender networks in high-school.

We also find that girls with more female peers are more likely to attend high school as well as select the natural sciences high school track. We do not observe direct effects of primary-school composition on grades in high school or effects on college enrollment for either gender in line with Anelli and Peri (2019).

Last, we examine whether girls with more girls in the classroom chose less gender stereotypical subjects as shown by Giardili (2020). Consistent with her findings, we observe that girls exposed to more female peers have a smaller share of girls in the track they choose in high school. There appears to be no corresponding selection effects on boys. We also observe a similar pattern with respect to university programs. In sum, girls with more female peers therefore have a higher educational attainment in primary school and are less likely to choose gender stereotypical high school or college tracks. This finding is interesting in light of the network formation of girls: The results suggest that girls form groups with other girls early on and then proceed with them to study less gender-stereotypical subjects.

Occupational Selection. While we find that girls exposed to more female peers select into better-paying occupations, we have not yet explored exactly what kind of jobs they select into. To study occupational selection we focus on four factors: 1) unemployment,[16] 2) the share of females in a given occupation, 3) selection into specific occupational groups, and 4) gender gaps in the selected occupation.

First, we observe no statistically discernible impact on unemployment for either men or women (the estimates are close to zero across specifications, see Table E.1). The result indicates that the wage effects we observe do not come from lower unemployment among women.

[16]The unemployment indicator takes value 1 if a person received any unemployment benefits during the year he or she turned 30.

Second, women who had more female peers in primary school seem to have a smaller share of women in the occupation that they pursue at the age of 30. Consistent with high-school track choices, women therefore sort into less gender-stereotypical jobs with higher wages.

Third, for a better understanding of the effects on occupational selection, we also present regressions on the 8 general categories of occupation available in the occupation register (Yrkeregistret). Table E.2 shows the results. A higher share of female peers results for women in a decreased likelihood of working in a service and care profession and an increased likelihood of having an occupation that requires academic education. Consistent with the results in Table 4, there are no statistically significant effects on boys.

Fourth, we observe that women end up sorting into occupations with a slightly higher gender wage gap (see Table E.5). Together with the finding that they earn more in those occupations, the results suggest that women exposed to more girls in their cohort may be trailblazers: They end up in more highly qualified, traditionally male-dominated jobs with higher gender gaps.

Fertility and Marriage. Additionally, register data allows us to observe fertility prior to age 30. Fertility reduces female wages (Lundborg, Nilsson and Rooth, 2014; Kleven et al., 2019). Using the register of newborns, we observe whether women have children prior to the age of 30. We find that women with more girls in the classroom have a lower likelihood of getting children. Women experience an approximately 2.5% reduction in the relative likelihood of having a child prior to the age of 30 with a 10 percentage point increase in the share of females in primary school (Black, Devereux and Salvanes (2013) find qualitatively similar results). Finally, there is no statistically significant impact on the likelihood of being married.

Can These Factors Explain the Impact on the Wage Gap? We explore the extent to which we can attribute the higher wages among women and the corresponding lower gender wage gap to the above mechanisms. To this end, we begin by naively controlling for each mechanism separately (Table E.3) and then successively add the mechanisms jointly depending on the timing during the educational and labor market career (Table E.4) . Note, our goal is to assess the extent to which these mechanisms can account for the effects of gender composition, and not to make adjustments to our estimates by controlling for endogenous variables.[17]

Accounting for primary-school and high-school grades, dummies for high-school study tracks and college tracks, eight occupation dummies, and fertility, reduces the size of the coefficient estimate capturing the impact of the share of females peers on the gender gap by 62% (Table E.4, column 5).[18] Comparing across the different mechanisms, including dummies for the occupations accounts for the largest reduction in coefficients (Table E.3, column 4). This specification also indicates that even within broad occupational categories, women earn higher wages after being exposed to more girls. Importantly, fertility does not account for a sizable portion of the impact of female peers on the gender gap.

The biggest factor explaining the effects on the wage gap is not fertility or grades, but occupational selection. While the observable mechanisms can account for roughly 60% of the impact of the gender environment on wages, there is still some remaining variation.

Other Potential Factors: Competitiveness and Risk Aversion. Preferences such such as competitiveness and risk aversion play an important role for educational attainment and occupational selection. The register data set clear limits on exploring these mechanisms, but we attempt to explore whether there are hints of a large impact of the share of female peers on competitiveness and risk aversion. Note that the data do not allow us to link levels of risk aversion or competitiveness to labor market outcomes, only to examine whether people behave consistent with changes in risk aversion or competitiveness due to a higher share of female peers.

First, we examine whether the average primary school grade is higher in high-school tracks that females choose after being exposed to more females as a proxy for competitiveness. We do not see such impact (Table E.6).[19] Second, we examine whether women choose more high-variance occupations. We do not see that females exposed to more females choose occupations with a higher variance in wages (Table E.7). Accordingly, we do not see concrete evidence

[17]We reduce the sample of these analyses to observations for which we observe all variables that we control for.

[18]Controlling for educational attainment and choices does not affect the estimated impact of the gender environment on the selection into non-gender stereotypical occupations.

[19]We also do not see an impact when examining college majors.

that females exposed to more females become more competitive or seek occupations with a higher wage variance.

The limits of administrative data do not allow us to explore other mechanisms which have been discussed in the literature. For instance, evidence from psychology suggests that girls in female-dominated environments are more self-confident, which could explain parts of the effects we document (Bertrand, 2011).

How Large Does the Share of Female Peers Need to Be? Exploring non-linearities. We examine the impact of different quintiles of the share of female peers on wages in Table E.8. Relative to the highest quintile of the share, we find that lower quintiles have lower wage gains. Balanced and female-dominated cohorts increase women's wages more than male-dominated cohorts. The highest quintile of 53% to 75% share of female peers increases wages the most. The finding that female-dominated cohorts improve girls outcomes most is in line with previous literature on the impact of single sex schooling on educational attainment (see, e.g., Park, Behrman and Choi, 2012). However, in female-dominated cohorts, men also experience the highest reduction in earnings.

Hence, policy implications depend on the welfare function. If a policy maker prefers to increase men's and women's wages, they should balance classrooms. However, if they aim primarily at reducing the gender wage gap, they should try to maximize the share of female peers girls are exposed to. In the latter case, and whilst keeping co-educational schooling, they might want to consider all-female study groups for some topics.

Summary of Mechanisms. Our exploration of the impact of female peers across educational stages and early career reveals six main findings:

1. Girls exposed to more girls have higher primary-school grades and are more likely to attend high school. Conversely, boys' primary-school grades worsen. Accordingly, a higher share of female peers widens the grade gap.

2. Women exposed to more girls do not have a lower likelihood of unemployment.

3. Women exposed to more girls choose less gender-stereotypical educational paths and jobs.

4. Women exposed to more girls have lower fertility by age 30.

5. The above factors account for roughly 60% of the reduction in the gender wage gap, occupational sorting playing the biggest role.

6. Female-dominated cohorts largely drive the overall effects on wages.

The results provide a fairly consistent picture. Early exposure to more girls has lasting effects on educational and career trajectories. Women socialized among girls seem to be less bound by gender norms and enter occupations with higher lifetime earnings.

7 Conclusion

We examine the impact of early gender socialization on the gender wage gap. To this end, we use unique register data that combines cohort composition during primary school with earnings and occupational data for up to 15 years after graduation. We exploit arbitrary variation in the share of girls in a student's cohort to estimate the long-run effects of gender environment at a critical age.

We find that a higher share of female peers in primary school increases women's later wages, thereby reducing the gender wage gap. Strikingly, we find that women exposed to more girls in primary school select into less gender-stereotypical jobs with higher lifetime earnings potential.

Because of a dearth of data, there is only little evidence on how gender socialization affects long-run labor market outcomes. Our evidence suggests that wages and occupational selection depend on the social environment in primary school: More girls in the cohort lead to less gender-congruent behavior. Gender roles likely play an important role for explaining why the gender wage gap has persisted in spite of advances in educational attainment. The evidence we present indicates that early gender environment shapes gender roles with a persistent impact on labor market outcomes.

References

Almås, Ingvild, Alexander W. Cappelen, Kjell G. Salvanes, Erik Sørensen, and Bertil Tungodden. 2016. "What explains the gender gap in college track dropout? Experimental and administrative evidence." Vol. 106, 296–302. American Economic Association.

Altonji, Joseph G., and Rebecca M. Blank. 1999. "Chapter 48 Race and gender in the labor market." In *Handbook of Labor Economics*. Vol. 3 PART, 3143–3259. Elsevier.

Anelli, Massimo, and Giovanni Peri. 2019. "The effects of high school peers' gender on college major, college performance and income." *Economic Journal*, 129(618): 553–602.

Balestra, Simone, Aurélien Sallin, and Stefan C. Wolter. 2021. "High-ability influencers? The heterogeneous effects of gifted classmates." *Journal of Human Resources*, forthcoming.

Balestra, Simone, Beatrix Eugster, and Helge Liebert. 2020. "Peers with Special Needs: Effects and Policies." *Review of Economics and Statistics*, forthcoming.

Bertrand, Marianne. 2011. "New perspectives on gender." In *Handbook of Labor Economics*. Vol. 4, 1543–1590. Elsevier.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan. 2015. "Gender Identity and Relative Income within Households." *The Quarterly Journal of Economics*, 130(2): 571–614.

Biasi, Barbara, and Heather Sarsons. 2021. "Flexible Wages, Bargaining, and the Gender Gap." *Quarterly Journal of Economics*, forthcoming.

Bietenbeck, Jan. 2020. "The long-term impacts of low-achieving childhood peers: Evidence from project STAR." *Journal of the European Economic Association*, 18(1): 392–426.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2013. "Under pressure? The effect of peers on outcomes of young adults." *Journal of Labor Economics*, 31(1): 119–153.

Blau, Francine D., and Lawrence M. Kahn. 2017. "The gender wage gap: Extent, trends, & explanations." *Journal of Economic Literature*, 55(3): 789–865.

Borbely, Daniel, Jonathan Norris, and Agnese Romiti. 2021. "Peer Gender and Schooling: Evidence from Ethiopia." Mimeo, University of Strathclyde.

Born, Andreas, Eva Ranehill, and Anna Sandberg. 2020. "Gender and Willingness to Lead: Does the Gender Composition of Teams Matter?" *The Review of Economics and Statistics*, forthcoming.

Brenøe, Anne. 2021. "Brothers Increase Women's Gender Conformity." *Journal of Population Economics*, forthcoming.

Brenøe, Anne A, and Ulf Zölitz. 2020. "Exposure to more female peers widens the gender gap in STEM participation." *Journal of Labor Economics*, 38(4): 1009–1054.

Brenøe, Anne Ardila, and Shelly Lundberg. 2018. "Gender gaps in the effects of childhood family environment: Do they persist into adulthood?" *European Economic Review*, 109: 42–62.

Brown, Charles, and Mary Corcoran. 1997. "Sex-based differences in school content and the male-female wage gap." *Journal of Labor Economics*, 15(3): 431–465.

Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais. 2017. "Acting wife: Marriage market incentives and labor market investments." *American Economic Review*, 107(11): 3288–3319.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2014. "Gender, Competitiveness, and Career Choices." *The Quarterly Journal of Economics*, 129(3): 1409–1447.

Card, David, Ana Rute Cardoso, and Patrick Kline. 2016. "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women." *Quarterly Journal of Economics*, 131(2): 633–686.

Card, David, Fabrizio Colella, and Rafael Lalive. 2021. "Gender Preferences in Job Vacancies and Workplace Gender Diversity." NBER WP No. 29350.

Carrell, Scott E., Mark Hoekstra, and Elira Kuka. 2018. "The long-run effects of disruptive peers." *American Economic Review*, 108(11): 3377–3415.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review*, 104(9): 2633–2679.

Chetty, Raj, John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How does your kindergarten classroom affect your earnings? Evidence from project star." *Quarterly Journal of Economics*, 126(4): 1593–1660.

Chetty, Raj, Nathaniel Hendren, Frina Lin, Jeremy Majerovitz, and Benjamin Scuderi. 2016. "Childhood environment and gender gaps in adulthood." *American Economic Review*, 106(5): 282–288.

Collins, Matthew, and Jonathan Lundstedt. 2021. "The Long-term Consequences of More Informative Grading itle." Lund Papers in Econ. Dem. 2021:1.

Eisenkopf, Gerald, Zohal Hessami, Urs Fischbacher, and Heinrich W. Ursprung. 2015. "Academic performance and single-sex schooling: Evidence from a natural experiment in Switzerland." *Journal of Economic Behavior and Organization*, 115: 123–143.

Elsner, Benjamin, Ingo E. Isphording, and Ulf Zölitz. 2021. "Achievement Rank Affects Performance and Major Choices in College." *Economic Journal*, forthcoming.

Epple, Dennis, and Richard E. Romano. 2011. "Chapter 20 - Peer Effects in Education: A Survey of the Theory and Evidence." In *Handbook of Social Economics*. 1053–1163.

Feld, Jan, and Ulf Zölitz. 2017. "Understanding peer effects: On the nature, estimation, and channels of peer effects." *Journal of Labor Economics*, 35(2): 387–428.

Flory, J. A., A. Leibbrandt, and J. A. List. 2015. "Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job Entry Decisions." *The Review of Economic Studies*, 82(1): 122–155.

Getik, Demid, and Armando N. Meier. 2020. "Peer Gender and Mental Health." WWZ Paper No. 2020/15.

Giardili, Soledad. 2020. "Single-sex primary schools and student achievement: evidence from admission lotteries." Mimeo, University of Edinburgh.

Goldin, Claudia, Sari Pekkala Kerr, Claudia Olivetti, and Erling Barth. 2017. "The expanding gender earnings gap: Evidence from the LEHD-2000 census." *American Economic Review*, 107(5): 110–114.

Golsteyn, Bart H. H., Arjan Non, and Ulf Zölitz. 2021. "The impact of peer personality on academic achievement." *Journal of Political Economy*, 129(4): 1052–1097.

Guryan, Jonathan, Kory Kroft, and Matthew J. Notowidigdo. 2009. "Peer effects in the workplace: Evidence from random groupings in professional golf tournaments." *American Economic Journal: Applied Economics*, 1(4): 34–68.

Helene, Jannie, Grøne Kristoffersen, Morten Visby, Helena Skyt, and Marianne Simonsen. 2015. "Disruptive school peers and student outcomes." *Economics of Education Review*, 45: 1–13.

Hoxby, Caroline M. 2000. "Peer effects in the classroom: Learning from gender and race Variation." NBER WP No. 7867.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2019. "Children and gender inequality: Evidence from Denmark." *American Economic Journal: Applied Economics*, 11(4): 181–209.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. 2019. "Child Penalties Across Countries: Evidence and Explanations."

Landaud, Fanny, Son-Thierry Ly, and Éric Maurin. 2018. "Competitive Schools and the Gender Gap in the Choice of Field of Study." *Journal of Human Resources*, 55(1): 0617–8864R.

Lavy, Victor, and Analía Schlosser. 2011. "Mechanisms and impacts of gender peer effects at school." *American Economic Journal: Applied Economics*, 3(2): 1–33.

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet. 2021. "Gender Differences in Job Search: Trading off Commute against Wage." *The Quarterly Journal of Economics*, 136(1): 381–426.

Lu, Fangwen, and Michael L. Anderson. 2015. "Peer effects in microenvironments: The benefits of homogeneous classroom groups." *Journal of Labor Economics*, 33(1): 91–122.

Lundborg, Petter, Anton Nilsson, and Dan Olof Rooth. 2014. "Adolescent health and adult labor market outcomes." *Journal of Health Economics*, 37(1): 25–40.

Lundborg, Petter, Erik Plug, and Astrid Würtz Rasmussen. 2017. "Can women have children and a career? IV evidence from IVF treatments." *American Economic Review*, 107(6): 1611–1637.

Mulligan, Casey B., and Yona Rubinstein. 2008. "Selection, Investment, and Women's Relative Wages Over Time." *Quarterly Journal of Economics*, 123(3): 1061–1110.

Olivetti, Claudia, and Barbara Petrongolo. 2016. "The Evolution of Gender Gaps in Industrialized Countries." *Annual Review of Economics*, 8(1): 405–434.

Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Validation." *Journal of Business Economics and Statistics*, 37(2): 187–204.

Park, Hyunjoon, Jere R. Behrman, and Jaesung Choi. 2012. "Single-sex education: Positive effects." *Science*, 335(6065): 165–166.

Slotwinski, Michaela, and Alois Stutzer. 2018. "Women Leaving the Playpen: The Emancipating Role of Female Suffrage." CESifo Working Paper No. 7002, Munich.

Early Socialization and the Gender Wage Gap

Demid Getik[‡] Armando N. Meier[⋆]

Online Appendix

Table of Contents

A Summary Statistics **1**

B Balance and Placebo Checks **2**

B.1 Female Peer Share Distribution 2

B.2 Balance Checks 4

B.3 Student Gender and the Gender Peer Share 5

B.4 Placebo Check: Previous and Past Cohorts 6

C Gender Peer Share and Wages **7**

C.1 Impact on Wage Quintile 7

C.2 Heterogeneities 8

D Robustness and Sample Checks **9**

D.1 Sample Splits 9

D.2 Measurement Age 10

D.3 Log Wages 11

D.4 Outliers and Movers 12

E Occupational Selection and Other Mechanisms **13**

E.1 Occupational Selection 13

E.2 Attribution of Effects to Mechanisms 15

E.3 Other Potential Mechanisms: Gender Inequality, Risk Aversion and Competitiveness 17

E.4 Nonlinearities 20

[‡] Demid Getik, Department of Economics, Lund University, email: demid.getik@nek.lu.se

[⋆] Armando N. Meier, University of Lausanne and University of Basel, email: armando.meier@unil.ch

A Summary Statistics

Table A.1: Summary Statistics

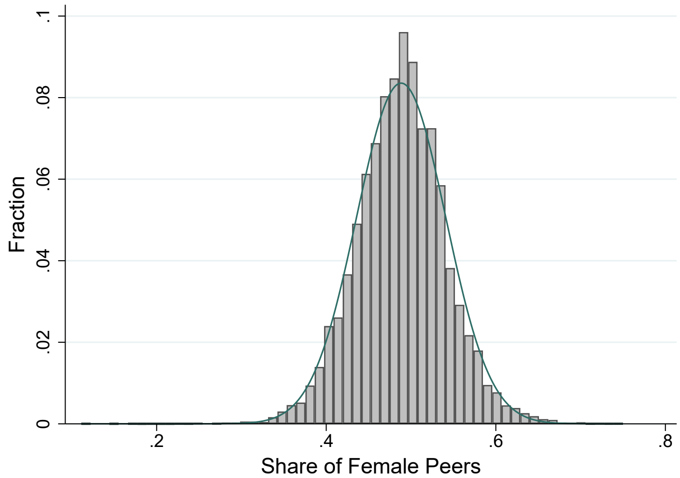
Variable	N	Mean	SD	N	Mean	SD
Background Variables		**Men**			**Women**	
Mother High School	385,574	0.69	0.46	366,986	0.69	0.46
Father High School	385,574	0.71	0.45	366,986	0.71	0.45
Mother Vocational Degree	385,574	0.14	0.35	366,986	0.14	0.35
Father Vocational Degree	385,574	0.11	0.32	366,986	0.11	0.32
Mother College Degree	385,574	0.19	0.39	366,986	0.19	0.39
Father College Degree	385,574	0.19	0.40	366,986	0.19	0.40
Mother STEM Degree	385,574	0.01	0.09	366,986	0.01	0.09
Father STEM Degree	385,574	0.04	0.20	366,986	0.04	0.20
Log Family Income	385,574	11.95	2.79	366,986	11.97	2.76
Wage Mother (1000 SEK)	381,217	136.99	95.50	363,193	137.52	96.02
Wage Father (1000 SEK)	368,791	210.92	185.20	350,543	211.55	186.41
Mother Unemployed	385,574	10.57	30.74	366,986	10.44	30.58
Father Unemployed	385,574	10.96	31.24	366,986	11.06	31.36
First-Born Child	385,574	0.66	0.47	366,986	0.66	0.47
Number Siblings	385,574	1.40	1.14	366,986	1.39	1.15
Immigrant	385,574	0.08	0.27	366,986	0.08	0.27
2nd Generation Immigrant	385,574	0.21	0.41	366,986	0.21	0.41
Adopted	385,574	0.01	0.11	366,986	0.02	0.13
Age Mother	384,736	43.61	5.09	366,442	43.62	5.11
Age Father	379,537	46.50	5.79	360,898	46.50	5.80
Father Unknown	385,574	0.02	0.12	366,986	0.02	0.13
Mother Unknown	385,574	0.00	0.05	366,986	0.00	0.04
Single Mother	385,574	0.07	0.25	366,986	0.07	0.26
Single Father	385,574	0.01	0.10	366,986	0.01	0.09
Schooling Variables						
Primary-School Grade	385,574	-0.17	0.96	366,986	0.18	0.95
High-School Grade	308,584	-0.09	0.89	304,446	0.19	0.86
High-School Attendance	385,574	0.84	0.36	366,986	0.86	0.35
High-School Science Track	385,574	0.12	0.33	366,986	0.10	0.30
Share Females HS Track	325,678	0.38	0.26	316,015	0.62	0.15
Cohort Size	385,574	122.94	41.54	366,986	123.30	41.51
Schools in Municipality	385,574	42.16	57.80	366,986	42.28	57.93
Share Females	385,574	0.49	0.05	366,986	0.49	0.05
College Enrolment	385,574	0.25	0.43	366,986	0.28	0.45
College Degree	385,574	0.18	0.38	366,986	0.30	0.46
Share Females College	168,494	0.50	0.22	220,660	0.66	0.15
Labour-Market Variables						
Share Females Occupation	333,413	0.34	0.26	318,702	0.66	0.23
Mean Occupation Wage	333,413	266.72	101.06	318,702	234.25	995.13
Log Occupation Wage	333,413	12.43	0.37	318,702	12.28	0.39
Wage (1000 SEK)	385,574	256.58	169.39	366,986	178.30	138.51
Log Income	342,678	12.34	0.94	323,448	11.84	1.17
Unemployed	385,574	8.44	27.79	366,986	11.13	31.45

Note: This table presents summary statistics for the key variables in the paper. The grades variables in both high- and primary school are standardized average grades. We do not observe all variables for the same number of people. Data for parental occupational wage potential is not available to us at the time of school attendance as it starts after our observation period.

B Balance and Placebo Checks

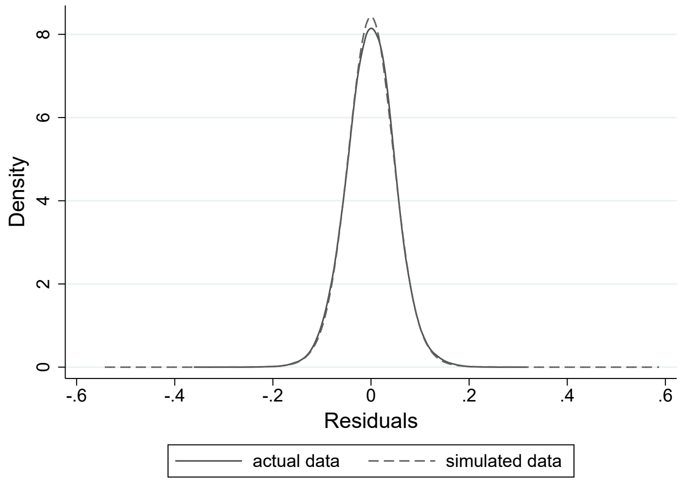
B.1 Female Peer Share Distribution

Figure B.1: Share of Female Peers Across School-Cohorts



Note: The figure above represents the distribution of the share of female peers (average ≈ 0.48). The overlaid curve represents normal distribution.

Figure B.2: Simulated and Actual Residual Share of Female Peers Across School-Cohorts



Note: The figure above represents the actual and simulated distribution of the residualized female peer share across school-cohorts, conditional on school and cohort fixed effects.

B.2 Balance Checks

Table B.1: Balance Checks Across the Distribution of the Share of Female Peers by Quintile

Peer Quintile:	**1**	**2**	**3**	**4**	**5**
Mother High School	0.002	-0.003**	-0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father High School	-0.001	-0.002	0.003*	-0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mother Vocational Degree	0.000	-0.001	0.001	-0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father Vocational Degree	-0.001	0.001	0.000	-0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mother College Degree	0.001	-0.001	-0.002	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father College Degree	0.001	-0.004***	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mother STEM Degree	-0.000	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Father STEM Degree	0.001	-0.001	-0.000	-0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log Family Income	-0.007	0.008	0.016	-0.007	-0.011
	(0.010)	(0.011)	(0.010)	(0.011)	(0.011)
Log Wage Mother	-0.000	0.001	-0.008	0.007	0.001
	(0.013)	(0.014)	(0.013)	(0.013)	(0.014)
Log Wage Father	-0.029**	0.020	0.029**	-0.015	-0.007
	(0.014)	(0.015)	(0.014)	(0.015)	(0.015)
Mother Unemployed	-0.039	0.010	-0.017	0.035	0.012
	(0.108)	(0.107)	(0.114)	(0.107)	(0.118)
Father Unemployed	-0.092	-0.075	-0.023	0.088	0.110
	(0.114)	(0.105)	(0.104)	(0.111)	(0.112)
First-Born Child	0.001	0.001	0.001	-0.002	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number Siblings	0.003	0.001	-0.006*	-0.003	0.006
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Immigrant	0.000	-0.000	-0.001	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
2nd Generation Immigrant	-0.000	0.001	-0.002*	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Adopted	-0.000	-0.000	0.001*	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age Mother	0.003	-0.010	0.000	-0.006	0.013
	(0.016)	(0.015)	(0.016)	(0.016)	(0.017)
Age Father	-0.007	0.011	-0.020	0.018	-0.002
	(0.019)	(0.018)	(0.019)	(0.018)	(0.019)
Mother Unknown	-0.000	-0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Father Unknown	-0.000	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Single Mother	-0.001	-0.000	0.001	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Single Father	0.000	0.000	0.000	0.000	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note: The table shows the estimated relationship between student family characteristics and different quintiles of the share of female peers in one's cohort (each quintile is regressed separately on each variable separately indicated in the first column). The specifications in the table incrementally include school and cohort fixed effects, school trends, and school-level controls. Those controls include cohort size and the number of schools in the municipality. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B.3 Student Gender and the Gender Peer Share

Table B.2: Effects of Own Gender on the Share of Female Peers

	Share of Female Peers		
	(1)	(2)	(3)
Female	-0.0006	-0.0007	-0.0006
	(0.0007)	(0.0006)	(0.0006)
School FE	X	X	X
Cohort FE	X	X	X
School Trends	-	X	X
Controls	-	-	X
Observations	752,560	752,560	752,560
Schools	537	537	537
R-squared	0.12	0.20	0.20

Note: The table shows the estimated relationship between the share of female peers in a cohort and a student's own gender. Following Guryan, Kroft and Notowidigdo (2009), we control for the school-level leave-one-out cohort share of females. That is, we control for the share of females in the rest of the school leaving out the cohort of the student under consideration. Controls include parental education, income, and mental health as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B.4 Placebo Check: Previous and Past Cohorts

Table B2: Gender Composition in Other Cohorts, Wages, and Occupational Wage Potential

	Annual Wage			**Occupation Wage**		
	(1)	(2)	(3)	(4)	(5)	(6)
Males (t-1)	-9,757	-7,554	-7,120	2,630	4,075	3,375
	(5,961)	(6,070)	(5,971)	(3,459)	(3,562)	(3,450)
Females (t-1)	-8,903*	-8,808	-9,063*	3,743	2,001	1,562
	(4,871)	(5,154)	(5,134)	(3,495)	(3,630)	(3,552)
Males (t+1)	5,513	8,897	9,886*	-2,184	-1,120	-818
	(5,311)	(5,538)	(5,507)	(3,252)	(3,366)	(3,343)
Females (t+1)	-1,467	-490	790	3,240	2,545	3,274
	(4,569)	(4,831)	(4,814)	(3,543)	(3,603)	(3,605)
School FE	X	X	X	X	X	X
Cohort FE	X	X	X	X	X	X
School Trends	-	X	X	-	X	X
Controls	-	-	X	-	-	X

Note: The table shows the estimated relationship between annual wage at 30 and the share of female peers in the previous (t-1) and the subsequent cohorts (t+1). Males refers to the sample only consisting of boys, females refers to the sample only consisting of girls. The first three columns present the relationship for annual wage; the subsequent three columns for occupation wage. Occupational wages are computed based on 186 unique occupations. The specifications for each of the variables incrementally include school and cohort fixed effects, school trends, and a vector of controls. This vector includes parental education, income, and mental health as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C Gender Peer Share and Wages

C.1 Impact on Wage Quintile

Table C.1: Gender Peer Share Effects by Wage Quintile

Income Quintile:	**Male**	**Female**
	(1)	(2)
Quintile 1 (1 - 112632)	0.002	-0.040**
	(0.013)	(0.017)
Quintile 2 (112634 - 216219)	0.021	-0.026
	(0.015)	(0.017)
Quintile 3 (216220 - 281551)	0.016	-0.000
	(0.017)	(0.016)
Quintile 4 (281552 - 353564)	-0.007	0.026
	(0.019)	(0.017)
Quintile 5 (353565 - 9650019)	-0.032*	0.041***
	(0.018)	(0.014)
School FE	X	X
Cohort FE	X	X
School Trends	X	X
Controls	X	X

Note: The table shows the estimated relationship between annual wage at age 30 expressed as 5 bins (quintiles) of earnings and the share of female peers in one's cohort. The boundaries of a given bin in terms of the respective annual wages are recorded in parentheses. Each row illustrates the effect of the female peer share in one's cohort on the likelihood of being in a given income quintile. The outcomes are estimated separately for males and females. All outcomes are estimated using the preferred specification from column (3) in the main tables. Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C.2 Heterogeneities

Table C.2: Heterogeneities

Dependent variable:	**Female**	**Male**	**Gap**
	(1)	(2)	(3)
Female Share	18,626***	-7,704	26,491***
	(5,199)	(5,776)	(7,587)
Above-Median Income	5,968	-14,276	20,679
	(9,413)	(10,851)	(13,712)
Parent Went to College	-2,747	-7,528	4,911
	(10,652)	(12,932)	(15,703)
Parent Unemployed	-4,694	4,923	-10,018
	(11,320)	(12,001)	(16,956)
Single Parent	-20,148	26,530	-46,377*
	(15,984)	(18,304)	(24,283)
Larger Cohort	7,031	653	5,678
	(10,781)	(12,470)	(16,024)
Larger Municipality	5,756	-4,904	10,960
	(10,006)	(12,028)	(15,229)
School FE	X	X	X
Cohort FE	X	X	X
School Trends	X	X	X
Observations	756,560	756,560	756,560
Schools	539	539	539

Note: The table presents heterogeneous effect of the share of female peers in the cohort on annual wage at age 30. The first row replicates the main results of the paper presented in column (2) of Table 2. The next three rows represents interacted effects with the female peer share of variables coded as a binary with 1 for "yes" and 0 for "no". For the remaining four variables, 0 indicates a below-median and 1 an above-median value. For each interaction, we run the baseline regression once–so each row presents estimates from one regression. Column (1) presents the interaction effects for females, and column (2) for males. Column (3) estimates the gender difference in the interacted effects. Standard errors (in parentheses) are clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D Robustness and Sample Checks

D.1 Sample Splits

Table D.1: Effects on Wages and Occupational Wage Potential when Splitting the Sample by Gender

	Annual Wage			Occupation Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Females	13,088***	14,918***	13,568***	10,941***	11,401***	9,782***
	(4,941)	(5,123)	(5,101)	(3,529)	(3,586)	(3,535)
Males	-6,854	-5,860	-4,428	-2,200	-1,517	-1,169
	(5,982)	(6,056)	(5,951)	(3,393)	(3,417)	(3,381)
School FE	X	X	X	X	X	X
Cohort FE	X	X	X	X	X	X
School Trends	-	X	X	-	X	X
Controls	-	-	X	-	-	X
Observations	752,560	752,560	752,560	652,115	652,115	652,115
Schools	537	537	537	537	537	537
R-squared	0.03	0.04	0.05	0.16	0.16	0.18

Note: The table shows the estimated relationship of the share of female peers with annual wage at age 30 and occupational wage potential, with the sample split by gender. The outcomes are recorded in Swedish crowns (SEK). The first row shows the results for women; the second row for men. The first three columns show the relationship for the annual wage; the last three columns for median occupation wage. Occupational wages are computed based on 186 unique occupations. Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D.2 Measurement Age

Table D.2: Effects on Labour-Market Outcomes at Different Age Cut-Offs

	Annual Wage			Occupation Wage		
Age:	29	30	31	29	30	31
Female × Share Females	17,781***	17,211***	15,396***	11,416***	11,866***	12,509***
	(5,057)	(5,172)	(5,264)	(3,422)	(3,608)	(3,821)
Male × Share Females	-5,247	-6,714	-4,991	-831	-2,198	-3,392
	(5,439)	(5,674)	(6,715)	(3,305)	(3,331)	(3,591)
Female × Gap	23,138***	24,197***	20,556**	12,417***	14,305***	16,210***
	(7,433)	(7,536)	(8,179)	(4,635)	(4,937)	(5,228)
Female	-80,827***	-90,002***	-95,994***	-36,871***	-39,198***	-41,382***
	(3,712)	(3,719)	(4,068)	(2,280)	(2,419)	(2,563)
School FE	X	X	X	X	X	X
Cohort FE	X	X	X	X	X	X
School Trends	X	X	X	X	X	X
Controls	X	X	X	X	X	X
Observations	751,550	752,560	698,950	651,224	652,115	609,003
Schools	537	537	537	537	537	537
R-squared	0.09	0.10	0.10	0.18	0.21	0.20

Note: The table shows the estimated relationship between annual wages at ages 29-31/occupational wage potential at ages 29-31 and the share of female peers in one's cohort. Occupational wages are computed based on 186 unique occupations. The outcomes are recorded in Swedish crowns (SEK). The first row shows the results for women; the second row for men. The third row shows the difference in response to the female share between the genders. Finally, the last row shows the gross difference between the genders. Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are based on clustering at the school level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D.3 Log Wages

Table D.3: Effects of the Gender Peer Share on the Annual Log Wage

	Log Annual Wage Mean: 12.1				
	(1)	(2)	(3)	(4)	(5)
Female × Share Females	0.11**	0.13***	0.12***		
	(0.04)	(0.04)	(0.04)		
Male × Share Females	-0.01	0.00	0.00		
	(0.04)	(0.04)	(0.04)		
Gap × Share Females	0.12**	0.13**	0.12**	0.13**	0.12*
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Female	-0.56***	-0.56***	-0.56***	-0.56***	-0.56***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
School FE	X	X	X	-	-
Cohort FE	X	X	X	-	-
School Trends	-	X	X	-	-
Controls	-	-	X	-	X
School × Cohort FE	-	-	-	X	X
Observations	666,126	666,126	666,126	666,126	666,126
School-Cohorts	537	537	537	537	537
R-squared	0.07	0.07	0.08	0.08	0.08

Note: The table shows the estimated relationship between log annual wage at age 30 and the share of female peers in one's cohort. The first row shows the results for women; the second row for men. The third row shows the difference in response to the female share between the genders. Finally, the last row shows the gross difference in annual wage between the genders. Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D.4 Outliers and Movers

Table D.4: Robustness to Excluding Extreme Observations of the Share of Female Peers or Movers

	Dropping:			
	None	**Share Outliers**	**Size Outliers**	**Movers**
Females × Share Females	16,481***	19,616***	14,804**	17,113***
	(5,250)	(6,989)	(6,300)	(5,748)
Male × Share Females	-5,631	-3,357	-4,840	-5,216
	(5,714)	(7,735)	(7,100)	(6,366)
Female × Share Females	22,357***	23,281**	20,251**	22,330***
	(7,596)	(10,425)	(9,191)	(8,455)
Female	-89,558***	-90,185***	-87,984***	-91,002***
	(3,746)	(5,114)	(4,516)	(4,185)
School FE	X	X	X	X
Cohort FE	X	X	X	X
School Trends	X	X	X	X
Controls	X	X	X	X
Observations	742,833	681,621	592,172	616,827
School-Cohorts	537	537	478	537
R-squared	0.10	0.10	0.10	0.11

Note: The table shows the estimated relationship between annual wage at age 30 and the share of female peers in one's cohort. The outcomes are recorded in Swedish crowns (SEK). The first row shows the results for women; the second row for men. The third row shows the difference in response to the female share between the genders. Finally, the last row shows the gross difference in annual wage between the genders. The estimates in column (1) correspond to the estimates of our main results. In column (2), we exclude individuals who come from a cohort from an extreme in the distribution of the female peer share (top or bottom 5%). In column (2), we exclude individuals who come from a cohort from a cohort that lies in the top or the bottom 10% of the cohort size distribution. In column (4), we only include non-movers. We define those to be individuals who resided in one municipality throughout the entirety of primary schooling (9 years). Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

E Occupational Selection and Other Mechanisms

E.1 Occupational Selection

Table E.1: Gender Peer Share and Unemployment

	Unemployed at 30 Mean: 9.75				
	(1)	(2)	(3)	(4)	(5)
Female × Share Females	0.02	-0.04	0.08		
	(1.13)	(1.11)	(1.11)		
Male × Share Females	0.71	0.71	0.70		
	(0.98)	(0.93)	(0.93)		
Gap	-0.68	-0.74	-0.62	-0.68	-0.49
	(1.46)	(1.47)	(1.47)	(1.48)	(1.48)
Female	3.00***	3.03***	2.97***	3.00***	2.91***
	(0.72)	(0.73)	(0.73)	(0.73)	(0.73)
School FE	X	X	X	-	-
Cohort FE	X	X	X	-	-
School Trends	-	X	X	-	-
Controls	-	-	X	-	X
School × Cohort FE	-	-	-	X	X
Observations	752,560	752,560	752,560	752,560	752,560
School-Cohorts	537	537	537	537	537
R-squared	0.03	0.03	0.04	0.04	0.04

Note: The table shows the estimated relationship between having been unemployed in the year when turning 30 and the share of female peers in one's cohort. The variable is 100 for unemployment and 0 otherwise. We classify someone as having been unemployed in that year if they received unemployment benefits at any point during the year. The first row shows the results for women; the second row for men. The third row shows the difference in response to the female share between the genders. Finally, the last row shows the gross difference between the genders. Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table E.2: Effects of Female Peer Share on Occupational Group

Dependent variable	Female	Male	Gap
	(1)	(2)	(3)
Leading Role	0.006	0.004	0.003
	(0.006)	(0.006)	(0.008)
Requires Degree	0.013	0.017	-0.004
	(0.016)	(0.015)	(0.022)
Specialised Work	0.030*	0.005	0.025
	(0.016)	(0.015)	(0.021)
Office and Customer Work	0.001	-0.000	0.001
	(0.012)	(0.010)	(0.015)
Service/Care	-0.039*	0.014	-0.053
	(0.022)	(0.019)	(0.033)
Nature-related	-0.003	-0.001	-0.002
	(0.003)	(0.004)	(0.005)
Craft and Building	0.008	-0.016	0.024
	(0.012)	(0.017)	(0.026)
Unqualified Work	-0.008	0.007	-0.014
	(0.008)	(0.008)	(0.012)
School FE	X	X	X
Cohort FE	X	X	X
School Trends	X	X	X
Controls	X	X	X

Note: The table estimates the relationship between the share of classroom female peers and subsequent selection into different categories of occupation. This classification is based on the eight primary occupation categories in the Occupation register (Yrkeregistret). "Leading Role" corresponds to the "Ledningsarbete" category; "Requires Theoretical Competence" refers to occupations requiring longer academic training (Arbete som kräver teoretisk specialkompetens); "Requires Shorter Degree" refers to occupations with shorter academic trainings (Arbete som kräver kortare högskoleutbildning eller motsvarande kunskaper); "Office and Customer Work" corresponds to "Kontors- och kundservicearbete". "Care and Service" to "Service-, Omsorg-, och Försäljningsarbete"; "Nature-related" to "Arbete inom jordbruk trädgård, skogsbruk och fiske". "Craft and Building" to "Hantverksarbete inom byggverksamhet och tillverkning". "Unqualified Work" refers to work that does not require special qualifications (Arbete utan krav på särskild yrkesutbildning). The dependent variable is a dummy for working in a given occupation. Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are clustered at the school level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

E.2 Attribution of Effects to Mechanisms

Table E.3: Effects on Wages, Including Intermediate Stage Controls

	Annual Wage Mean: 246,128				
	(1)	(2)	(3)	(4)	(5)
Female × Share Females	17,798***	17,569***	13,676**	12,159**	16,687***
	(5,737)	(5,680)	(5,425)	(4,918)	(5,535)
Male × Share Females	-6,560	-1,165	-6,460	-5,210	-6,328
	(6,006)	(5,873)	(5,793)	(5,428)	(6,007)
Female × Gap	24,686***	19,056**	20,465***	17,634**	23,361***
	(7,948)	(7,915)	(7,440)	(6,986)	(7,809)
Grades	-	X	-	-	-
Study Tracks	-	-	X	-	-
Occupation	-	-	-	X	-
Fertility	-	-	-	-	X
School FE	X	X	X	X	X
Cohort FE	X	X	X	X	X
School Trends	X	X	X	X	X
Controls	X	X	X	X	X
Observations	538,099	538,099	538,099	538,099	538,099
Schools	537	537	537	537	537
R-squared	0.15	0.19	0.20	0.30	0.16

Note: The table shows the estimated relationship between annual wage at age 30 and the share of female peers in one's cohort including intermediate controls discussed in the mechanisms section. The first row shows the results for women; the second row for men. The third row shows the difference in response to the female share between the genders. The first column represents the main set of results. The second column includes primary- and high-school grades. The specification in column (3) includes dummies for high-school tracks and university programs. Column (4) includes dummies for the 8 occupational categories shown in Table E.2. Column (5) also includes a dummy for giving birth. The sample size in each specification is reduced to a sub-sample for which we can observe all the relevant variables. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table E.4: Effects on Wages, Including All Intermediate Stage Controls

	Annual Wage Mean: 246,128				
	(1)	(2)	(3)	(4)	(5)
Female × Share Females	17,798***	18,528***	14,726***	11,540**	11,050**
	(5,737)	(5,641)	(5,362)	(4,829)	(4,746)
Male × Share Females	-6,560	-236	-3,049	-3,860	-3,891
	(6,006)	(5,962)	(5,783)	(5,413)	(5,442)
Female × Gap	24,686***	19,071**	18,084**	15,655**	15,211**
	(7,948)	(7,914)	(7,362)	(6,819)	(6,812)
Grades	-	X	X	X	X
Study Tracks	-	-	X	X	X
Occupation	-	-	-	X	X
Fertility	-	-	-	-	X
School FE	X	X	X	X	X
Cohort FE	X	X	X	X	X
School Trends	X	X	X	X	X
Controls	X	X	X	X	X
Observations	538,099	538,099	538,099	538,099	538,099
Schools	537	537	537	537	537
R-squared	0.15	0.19	0.22	0.32	0.32

Note: The table shows the estimated relationship between annual wage at age 30 and the share of female peers in one's cohort including the main intermediate controls discussed in the mechanisms section. The first row shows the results for women; the second row for men. The third row shows the difference in response to the female share between the genders. The first column represents the main set of results. Column (2) additionally includes primary- and high-school grades. The specification in column (3) additionally includes dummies for high-school tracks and university programs. Column (4) also includes dummies for the 8 occupational categories shown in Table E.2. Column (5) also includes a dummy for giving birth. The sample size in each specification is reduced to a sub-sample for which we can observe all the relevant variables. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

E.3 Other Potential Mechanisms: Gender Inequality, Risk Aversion and Competitiveness

Table E.5: Gender Gap in Median Wage in Occupation at Age 30

	Earnings Gap in Chosen Occupation Mean: 48410				
	(1)	(2)	(3)	(4)	(5)
Female × Share Females	2,909**	2,968**	2,721*		
	(1,414)	(1,508)	(1,504)		
Male × Share Females	-1,279	-1,511	-1,415		
	(1,542)	(1,532)	(1,537)		
Female × Gap	4,218*	4,502*	4,159*	4,590*	4,215*
	(2,268)	(2,329)	(2,334)	(2,342)	(2,346)
Female	-15,736***	-15,877***	-15,706***	-15,928***	-15,742***
	(1,109)	(1,138)	(1,141)	(1,146)	(1,148)
School FE	X	X	X	-	-
Cohort FE	X	X	X	-	-
School Trends	-	X	X	-	-
Controls	-	-	X	-	X
School × Cohort FE	-	-	-	X	X
Observations	652,115	652,115	652,115	652,115	652,115
School-Cohorts	537	537	537	537	537
R-squared	0.05	0.05	0.06	0.06	0.07

Note: The table shows the estimated relationship between the gender gap in median earnings in one's chosen occupation at the age of 30 and the share of female peers in one's cohort. The first row shows the results for women; the second row for men. The third row shows the difference in response to the female share between the genders. Finally, the last row shows the gross annual occupational wage gap between the genders. These wages are computed based on 186 unique occupations in our registers and are recorded in Swedish crowns (SEK). The coefficients in the first three columns are based on the first specification that relies on school and cohort fixed effects. Columns (4) and (5) record the estimates produced by our second specification, which include school-by-cohort fixed effects. Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table E.6: Gender Peer Share and Mean Grade in High-School program

	Mean Grade in High-School Track Mean: .04				
	(1)	(2)	(3)	(4)	(5)
Female × Share Females	0.02	0.02	0.01		
	(0.02)	(0.02)	(0.02)		
Male × Share Females	0.03	0.02	0.02		
	(0.02)	(0.02)	(0.02)		
Gap	-0.01	-0.01	-0.01	-0.01	-0.02
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Female	0.09***	0.09***	0.10***	0.09***	0.10***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
School FE	X	X	X	-	-
Cohort FE	X	X	X	-	-
School Trends	-	X	X	-	-
Controls	-	-	X	-	X
School × Cohort FE	-	-	-	X	X
Observations	641,467	641,467	641,467	641,467	641,467
School-Cohorts	537	537	537	537	537
R-squared	0.07	0.08	0.10	0.09	0.12

Note: The table shows the estimated relationship between the average grade in one's chosen high-school program and the share of female peers in one's cohort. The first row shows the results for women; the second row for men. The third row shows the difference in wage response to the female share between the genders. Finally, the last row shows the gross difference between the genders. The coefficients in the first three columns are based on school and cohort fixed effects. Columns (4) and (5) include school-by-cohort fixed effects. Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. The sample size is smaller than the one for the main results due to a fraction of students ($\approx$ 10%) in our sample not proceeding to high school. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table E.7: Gender Peer Share and Earnings Variance in the Chosen Occupation

	Variance in Occupational Wages Mean: 160,064				
	(1)	(2)	(3)	(4)	(5)
Female × Share Females	4,182	4,700	3,647		
	(2,986)	(3,004)	(2,972)		
Male × Share Females	-1,837	-1,616	-1,370		
	(3,009)	(3,011)	(2,960)		
Gap	6,371	6,663	5,342	6,073	4,630
	(4,143)	(4,168)	(4,111)	(4,182)	(4,126)
Female	-18,238***	-18,383***	-17,739***	-18,097***	-17,393***
	(2,018)	(2,029)	(2,003)	(2,039)	(2,013)
School FE	X	X	X	-	-
Cohort FE	X	X	X	-	-
School Trends	-	X	X	-	-
Controls	-	-	X	-	X
School × Cohort FE	-	-	-	X	X
Observations	652,115	652,115	652,115	652,115	652,115
School-Cohorts	537	537	537	537	537
R-squared	0.07	0.08	0.09	0.09	0.10

Note: The table shows the estimated relationship between the wage variance in one's chosen occupation at the age of 30 and the share of female peers in one's cohort. Occupational wages are computed based on 186 unique occupations. The first row shows the results for women; the second row for men. The third row shows the difference in response to the female share between the genders. Finally, the last row shows the gross difference between the genders. The coefficients in the first three columns are based on the first specification that relies on school and cohort fixed effects. Columns (4) and (5) record the estimates produced by our second specification, which include school-by-cohort fixed effects. Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are based on clustering at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

E.4 Nonlinearities

Table E.8: Annual Wage Effects by Gender Peer Share Quintile

Peer Share Quintile:	**Male**	**Female**
	(1)	(2)
Quintile 1 (.42)	1554.90	-1740.82**
	(981.57)	(863.82)
Quintile 2 (.46)	972.11	-3197.99***
	(1029.30)	(817.77)
Quintile 3 (.49)	1654.07*	-1347.55*
	(958.19)	(804.76)
Quintile 4 (.52)	2285.97**	-1197.37
	(1025.39)	(797.63)
School FE	X	X
Cohort FE	X	X
School Trends	X	X
Controls	X	X

Note: The table shows the estimated relationship between annual wage at age 30 and the share of female peers in one's cohort expressed as 5 bins (quintiles) of that share, with quintile 5 as the reference category. The mean female share in a given bin is shown in parentheses. The outcome is recorded in Swedish crowns (SEK). Each row represent the corresponding quintile of the female share in cohort. All outcomes are estimated using the preferred specification from column (3) in the main tables. Controls include parental education, income and family composition as well as class size, cohort size, and the number of schools in the municipality. Standard errors (in parentheses) are clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$