The long-run effects of peer gender
on occupational sorting and the wage gap

Demid Getik[‡]   Armando N. Meier[⋆]

Online Appendix

Table of Contents

[‡] Demid Getik, Durham University and Center for Economic Demography at Lund University (CED), email: demid.getik@durham.ac.uk

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

A   Balance

A.1   Linear Balance Checks

### Table A.1: Linear, 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)
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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.2 Distribution of the Peer Share

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 A.1. The distributions look very similar, a result consistent with as-good-as-random assignment of the share of female peers.

Figure A.1: Simulated and Actual Residual Share of Female Peers Across School-Cohorts
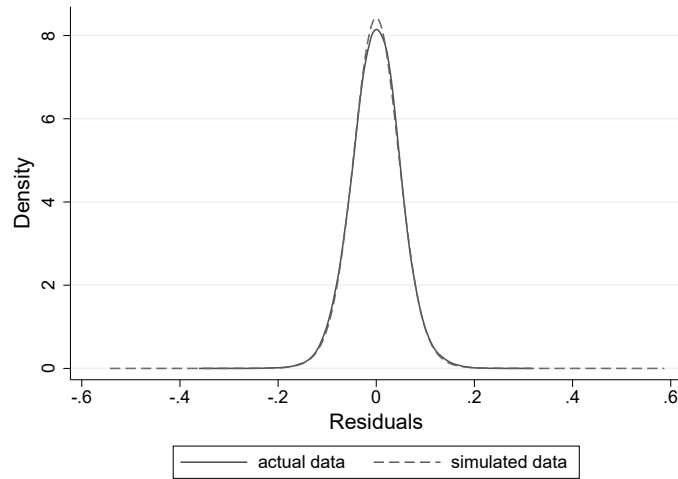
Note: The figure above represents the actual and simulated distribution of the residualized female peer share across school-cohorts, conditional on school and cohort fixed effects.

A.3 Student Gender and the Gender Peer Share

Table A.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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Gender and Cohorts. Following Chetty et al. (2011), and Balestra, Eugster and Liebert (2022) we do another check 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.

B Effects on Earnings

B.1 Effects on Earnings Quintile

Table B.1: Gender Peer Share Effects by Earnings Quintile

Income Quintile:	Male	Female
	(1)	(2)
Quintile 1 (0 - 47778)	-0.000	-0.017
	(0.013)	(0.014)
Quintile 2 (47779 - 182656)	0.024*	-0.027
	(0.014)	(0.017)
Quintile 3 (182657 - 265422)	0.014	-0.016
	(0.014)	(0.015)
Quintile 4 (265423 - 342073)	0.005	0.020
	(0.019)	(0.016)
Quintile 5 (342074 - 9650019)	-0.042**	0.040***
	(0.016)	(0.013)

Note: The table shows the estimated relationship between annual earnings 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 earnings 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B.2 Heterogeneity

Table B.2: Heterogeneity by SES Variables

SES variable:	Female (1)	Male (2)
Female Share	18,626***	-7,704
	(5,199)	(5,776)
Parental Education	739	378
	(1,243)	(1,624)
Parent Went to College	-2,748	-7,529
	(10,652)	(12,932)
Parent Unemployed	-4,694	4,923
	(11,320)	(12,001)
Parental Unemployment Benefits	361	-572
	(1,337)	(1,421)
Log Family Income	3,647**	2,259
	(1,579)	(1,926)
Above-Median Family Income	5,250	-17,006
	(9,736)	(10,633)
Cohort Size	84	24
	(113)	(121)
Above-Median-Size Cohort	6,914	-450
	(10,731)	(12,251)
Schools in Municipality	17	-120
	(122)	(109)
Above-Median-Size Municipality	5,760	-4,897
	(10,006)	(12,031)
Observations	753,131	752,561
Schools	537	537

Note: The table presents heterogeneous effect of the share of female peers in the cohort on annual earnings at age 30. The first row replicates the main results of the paper presented in column (2) of Table 2. Column (1) presents the interaction effects for females, and column (2) for males. Parental Education variable is calculated as the total number of years outside of compulsory education; parental unemployment benefits is the amount of unemployment benefits received. Standard errors (in parentheses) are clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C   Robustness Checks

C.1   Sample Splits

Table C.1: Effects on Earnings 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 earnings at age 30 and occupational earning 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 earnings; the last three columns for median occupation earnings. Occupational earnings 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C.2 Measurement Age

Table C.2: Effects on Labor-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	-80,827***	-90,002***	-95,994***	-36,871***	-39,198***	-41,382***
	(3,712)	(3,719)	(4,068)	(2,280)	(2,419)	(2,563)
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)
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 earnings at ages 29-31/occupational earning potential at ages 29-31 and the share of female peers in one's cohort. Occupational earnings 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 row "Female" shows the gross difference in annual earnings between the genders. The "Gap" row shows the difference in response to the share of female peers 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.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C.3 Extreme Observations and Movers

Table C.3: Robustness to Excluding Extreme Observations or Movers

	Full Sample	Extreme 10%	Small Cohorts	Movers
	(1)	(2)	(3)	(4)
Females × Share Females	16,480***	19,615***	14,804**	17,320***
	(5,250)	(6,989)	(6,300)	(5,725)
Male × Share Females	-5,631	-3,358	-4,840	-4,871
	(5,714)	(7,735)	(7,100)	(6,359)
Female	-89,558***	-90,185***	-87,983***	-90,838***
	(3,746)	(5,114)	(4,516)	(4,180)
Gap	22,356***	23,281**	20,251**	22,192***
	(7,596)	(10,425)	(9,191)	(8,447)
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	622,924
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 earnings 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 row "Female" shows the gross difference in annual earnings between the genders. The "Gap" row shows the difference in response to the share of female peers 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 (4), we exclude individuals who come from a cohort from a cohort that lies in the bottom 10% of the cohort size distribution. In column (5), we only include non-movers. We define those to be individuals who resided in one municipality throughout the entirety of primary school (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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C.4 Log Earnings

Table C.4: Effects of the Gender Peer Share on the Annual Log Earnings

	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)		
Female	-0.56***	-0.56***	-0.56***	-0.56***	-0.56***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Gap	0.12**	0.13**	0.12**	0.13**	0.12*
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
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 earnings 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 row "Female" shows the gross difference in annual earnings between the genders. The "Gap" row shows the difference in response to the share of female peers 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C.5 Placebo Check

Table C.5: Effect of Gender Composition in Other Cohorts on Earnings

	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 earnings 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 earnings; the subsequent three columns for occupational earnings. Occupational earnings 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D Mechanisms

D.1 Occupational Sorting

Table D.1: Effects of Female Peer Share on General Occupational Groups

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)
Specialized Work	0.030*	0.005	0.025
	(0.016)	(0.015)	(0.021)
Office/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 cohort 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 Degree" refers to occupations requiring longer academic training (Arbete som kräver teoretisk specialkompetens); "Specialized Work" refers to occupations with shorter academic trainings (Arbete som kräver kortare högskoleutbildning eller motsvarande kunskaper); "Office/Customer Work" corresponds to "Kontors- och kundservicearbete". "Service/Care" 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.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D.2: Effects of Female Peer Share on Occupational Choice, Most Common Occupations

Group:	Female	Male
	(1)	(2)
Technology/Data (.09)	0.019**	0.004
	(0.008)	(0.006)
Biology/Medicine (.04)	-0.007	-0.002
	(0.009)	(0.005)
Teaching (.08)	0.004	0.009
	(0.011)	(0.008)
Other Requiring Higher Education (.14)	0.001	-0.016
	(0.013)	(0.011)
Office Work (.05)	0.006	-0.002
	(0.008)	(0.007)
Service (.12)	-0.001	0.018
	(0.017)	(0.013)
Sales (.06)	-0.012	-0.008
	(0.009)	(0.008)
Building/Construction (.07)	0.004	-0.008
	(0.011)	(0.015)
Machinery/Transport (.05)	-0.001	-0.011
	(0.009)	(0.011)

Note: The table estimates the relationship between the share of cohort female peers and subsequent selection into different categories of occupation. This classification is based on the occupation categories in the Occupation register (Yrkeregistret) which comprise more than 4% of the sample. "Technology/Data" corresponds to the "Arbete som kräver teoretisk specialkompetens inom teknik och datavetenskap" category; "Biology/Medicine" to "[...] inom biologi, hälso- och sjukvård"; "Teaching" to "Lärararbete". 'Other Requiring Higher Education" to "Annat arbete som kräver teoretisk specialkompetens"; "Service" to "Kontors- och Kundservicearbete" "Sales" to "Service-, Omsorg-, och Försäljningsarbete"; "Building/Construction" to "Hantverksarbete inom byggverksamhet och tillverkning"; "Machinery/Transport" to "Process- och Maskinsoperatörsarbete, Transportarbete". 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.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D.3: Gender Gap in Median Earnings 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	-15,736***	-15,877***	-15,706***	-15,928***	-15,742***
	(1,109)	(1,138)	(1,141)	(1,146)	(1,148)
Gap	4,218*	4,502*	4,159*	4,590*	4,215*
	(2,268)	(2,329)	(2,334)	(2,342)	(2,346)
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 row "Female" shows the gross difference in annual earnings between the genders. The "Gap" row shows the difference in response to the share of female peers between the genders. These earnings 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D.2 Unemployment and Indicator for Positive Income

Table D.4: 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)		
Female	3.00***	3.03***	2.97***	3.00***	2.91***
	(0.72)	(0.73)	(0.73)	(0.73)	(0.73)
Gap	-0.68	-0.74	-0.62	-0.68	-0.49
	(1.46)	(1.47)	(1.47)	(1.48)	(1.48)
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 row "Female" shows the gross difference in annual earnings between the genders. The "Gap" row shows the difference in response to the share of female peers 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 D.5: Gender Peer Share and Positive Income in Given Year

	Positive Income Mean: .89				
	(1)	(2)	(3)	(4)	(5)
Female × Share Females	-0.01	-0.01	-0.01		
	(0.01)	(0.01)	(0.01)		
Male × Share Females	-0.01	-0.01	-0.00		
	(0.01)	(0.01)	(0.01)		
Female	-0.01	-0.01	-0.00	-0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Gap	-0.00	-0.00	-0.01	-0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
School FE	X	X	X	-	-
Cohort FE	X	X	X	-	-
School Trends	-	X	X	-	-
Controls	-	-	X	-	X
School × Cohort FE	-	-	-	X	X
Observations	752,561	752,561	752,561	752,561	752,561
School-Cohorts	537	537	537	537	537
R-squared	0.01	0.01	0.02	0.02	0.03

Note: The table shows the estimated relationship between having been having a positive income in the year when they turn 30 and the share of female peers in one's cohort. The variable is 1 for positive income and 0 otherwise. The first row shows the results for women; the second row for men. The row "Female" shows the gross difference in annual earnings between the genders. The "Gap" row shows the difference in response to the share of female peers 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 Attribution of Effects to Mechanisms

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.

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 about 40% (Table D.7, column 5).[22] Comparing across the different mechanisms, including dummies for the occupations accounts for the largest reduction in coefficients (Table D.6, column 4). This specification also indicates that even within broad occupational categories, women earn more 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.[23]

---

[22]Controlling for educational attainment and choices does not affect the estimated impact of the gender environment on the selection into non-gender stereotypical occupations. Note also that controlling for study track dummies accounts for differential competitiveness across tracks.

[23]For all analysis, we reduce the sample of these analyses to observations for which we observe all variables that we control for.

Table D.6: Effects on Earnings, 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)
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 earnings 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 "Gap" row shows the difference in response to the share of female peers 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 D.1. 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 D.7: Effects on Earnings, 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)
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 earnings 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 "Gap" row shows the difference in response to the share of female peers 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 D.1. 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D.4 Non-linearities

Table D.8: Are There Important Non-linearities? Residual Annual Earnings Effects by Gender Peer Share Quintile

Peer Share Quintile:	Male	Female
	(1)	(2)
Quintile 1 (.11-.45)	-515.65	801.73
	(934.91)	(846.53)
Quintile 2 (.45-.48)	-831.75	-1399.98*
	(911.19)	(791.90)
Quintile 3 (.48-.5)	ref.	ref.
Quintile 4 (.5-.53)	778.88	-272.46
	(863.74)	(804.12)
Quintile 5 (.53-.75)	-1248.02	163.53
	(957.81)	(803.73)

Note: The table shows the estimated relationship between annual earnings residuals taking out the linear relationship between the annual earnings and gender peer share and the share of female peers in one's cohort expressed as 5 bins (quintiles) of that share, with quintile 3 as the reference category. The mean female share and the range for 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$