MONASH University

MONASH BUSINESS SCHOOL

Department of Economics
ISSN number 1441-5429

Female Neighbors, Test Scores, and Careers

Discussion Paper no. 2023-06

Sofoklis Goulas, Rigissa Megalokonomou and Yi Zhang

Abstract:

How much does your neighbor impact your test scores and career? In this paper, we examine how an observable characteristic of same-age neighbors—their gender—affects a variety of high school and university outcomes. We exploit randomness in the gender composition of local cohorts at birth from one year to the next. Using new administrative data for the universe of students in consecutive cohorts in Greece, we find that a higher share of female neighbors improves both male and female students' high school and university outcomes. We also find that female students are more likely to enroll in STEM degrees and target more lucrative occupations when they are exposed to a higher share of female neighbors. We collect rich qualitative geographic data on communal spaces (e.g., churches, libraries, parks, Scouts and sports fields) to understand whether access to spaces of social interaction drives neighbor effects. We find that communal facilities amplify neighbor effects among females.

Keywords: neighbor gender peer effects, cohort-to-cohort random variat, birth gender composition, geodata, STEM university degrees

JEL Classification: J16

Sofoklis Goulas: Hoover Institution, Stanford University (email: sofoklis.goulas@hoover.org); Rigissa Megalokonomou: Department of Economics, Monash University (email: rigissa.megalokonomou@monash.edu); Yi Zhang: School of Economics, University of Queensland (email: yi.zhang1@uq.edu.au).

Female Neighbors, Test Scores, and Careers

March 31, 2023

Sofoklis Goulas
Hoover Institution
Stanford University

Rigissa Megalokonomou
Department of Economics
Monash University

Yi Zhang
School of Economics
University of Queensland

ABSTRACT

How much does your neighbor impact your test scores and career? In this paper, we examine how an observable characteristic of same-age neighbors—their gender—affects a variety of high school and university outcomes. We exploit randomness in the gender composition of local cohorts at birth from one year to the next. Using new administrative data for the universe of students in consecutive cohorts in Greece, we find that a higher share of female neighbors improves both male and female students' high school and university outcomes. We also find that female students are more likely to enroll in STEM degrees and target more lucrative occupations when they are exposed to a higher share of female neighbors. We collect rich qualitative geographic data on communal spaces (e.g., churches, libraries, parks, Scouts and sports fields) to understand whether access to spaces of social interaction drives neighbor effects. We find that communal facilities amplify neighbor effects among females.

Keywords: neighbor gender peer effects, cohort-to-cohort random variation, birth gender composition, geodata, STEM university degrees

JEL Classification: J16, J24, I24, I26

[1]Corresponding author: Rigissa Megalokonomou, Monash University, Department of Economics and University of Queensland, IZA, and CESifo. Email: rigissa.megalokonomou@monash.edu. We are grateful to Sascha Becker, Dan Bernhardt, Clement de Chaisemartin, David Deming, Victor Lavy, and Fabian Waldinger for insightful discussions and helpful comments. We are indebted to Robert Akerlof, Samia Badji, Simon Burgess, Vasilakis Chrysovalantis, Luke Chu, Anne Cumpston, Jan Feld, Thiemo Fetzer, Alessandro Iaria, Asad Islam, Jan Kabatek, Claudio Labanca, Pushkar Maitra, Andrew McLennan, Robin Naylor, Laura Puzzello, Tina Rampino, Heather Royer, Haishan Yuan, Chris Woodruff, and three anonymous referees for valuable suggestions. We thank seminar participants at the University of Warwick, University of Queensland, Queensland University Technology, University of Macedonia, Monash University and the Athens University of Economics and Business. We also thank participants at the Economics and Econometrics conference at the Athens University of Economics and Business, XXVII Meeting of the Economics of Education Association at the University of Barcelona, the Natural Experiments in History: Development, Health, and Labor Workshop at Deakin University, Annual American Economic Association Meeting in Atlanta, 15th North American Meeting of the Urban Economics Association and the CSDA Workshop in Auckland. We are grateful to Athina Skiadopoulou for superb research assistance and for helping to compile the qualitative geographic data. Rigissa Megalokonomou acknowledges research support from the University of Queensland BEL Early Career Grant (No: UQECR1833757).

1 Introduction

The question of which peers are relevant and how they affect one's education outcomes has long been a concern for academics, social scientists, and policymakers (Foster, 2006; Carrell, Fullerton, and West, 2009). Networks between same-age individuals exist not only within school but also outside of school. Peers in broader peer environments—*beyond* school walls—are important for student performance and university attendance decisions (Barrios-Fernández, 2022; Deutscher, 2020; Goux and Maurin, 2007). Unlike college students who might live close to each other for a predetermined, and typically short, period of time, high school students are likely to live in a specific area for many years. Physical proximity peers are particularly important during adolescence, given the limited geographic mobility at that age (Buu, DiPiazza, Puttler, Fitzgerald, and Zucker, 2009). Thus, adolescents are likely to form natural reference groups with similar-age peers with whom they interact daily for years within a relatively contained geographic area, such as their neighborhood.

Physical proximity may affect a student's education outcomes for several reasons. For example, adolescents might have a desire to conform with others in their proximity as a result of peer pressure, role models, or social norms (Akerlof and Kranton, 2002; Mota, Patacchini, and Rosenthal, 2016). Moreover, students might gain from information transmission (Bikhchandani, Hirschleifer, and Welch, 1992; Xiong, Payne, and Kinsella, 2016) in networks broader than their schoolmates. A higher number of communal spaces may amplify peer interactions and social integration and encourage the development of non-cognitive skills for young people (Majee and Anakwe, 2020). Thus, high school students are not only likely to be influenced by their schoolmates, but they are also likely to be affected by their neighbors.

In this paper, we study the impact of peers in physical proximity on student outcomes and decisions in a non-selective setting. Measuring peer effects outside the school can be challenging, because key student characteristics may be unobservable and nonrandom. We use an observable, random peer characteristic to study social interactions between peers in physical proximity: gender. Student academic achievement has been found to be associated with the gender of peers in the classroom or school (Hoxby, 2000; Lavy and Schlosser, 2011; Goulas, Megalokonomou, and Zhang, 2020),[1] but gender peer effects in education at neighborhood level have received recent attention (Barrios-Fernández, 2022; Deutscher, 2020).[2] Using new data for the universe of students in public high schools in Greece in consecutive cohorts and in combination with rich geographic data, we are able to identify reference groups broader than schoolmates: same-cohort neighbors. Students are assigned to public schools based on the school's proximity to the student's residential address. We exploit the institutional setting to define neighborhood and neighbors in an innovative way. In this institutional setting, schools are built very close to each other. Same-cohort peers who reside in the same residential building or complex are schoolmates. Same-cohort peers who reside in close proximity to a student—but not as close as her schoolmates—attend by law "next

[1] A strand of the literature examines the effect of single-sex schooling and single-sex classes on test scores (Lee, Turner, Woo, and Kim, 2015; Jackson, 2016; Park, Behrman, and Choi, 2018; Lee and Nakazawa, 2022).

[2] The impact of gender composition on other, non-education-related, outcomes has been studied in the literature. In particular, studies have examined how changes in gender composition affect one's propensity to engage in alcohol consumption, smoking and drinking behavior, and teenage pregnancy (Waddell, 2012; Black, Devereux, and Salvanes, 2013; Fletcher and Ross, 2018).

door" schools. These students who attend neighboring schools are still very likely to interact with each other in the local community. We define as *neighbors* all same-cohort peers who attend any *other* public neighboring school in the close vicinity (within 1 mile) of one's school. Identification relies heavily on the relative absence of school choice in this context. We examine how idiosyncratic cohort-to-cohort changes in the proportion of females within the neighborhood affect a student's academic performance, university enrollment, quality of higher education, field of study, and expected wage.[3] We show that these cohort-to-cohort changes in the female share can be attributed to randomness in the gender composition of local cohorts at birth.

How can same-age peers who live close to each other but attend different schools interact with each other? Young people are likely to interact socially in their local community. Neighbors are likely to play together in their spare time at the park, interact in communal spaces, play sports together, attend church, attend a library, attend a Scout facility (i.e., facilities such as a clubhouse), take part in group activities, etc. To examine whether the presence of those communal spaces interacts with gender neighbor effects, we obtain rich new geographic information about the intensity of communal facilities in each neighborhood. In particular, we study how the gender neighbor effects vary in neighborhoods depending on the intensity of communal resources and public facilities. Motivated by the literature, we include churches, libraries, parks or squares, Scouts, sport fields and an *Overall Intensity* index for all facilities generated by principal component analysis (PCA) (Majee and Anakwe, 2020; Zeldin, Christens, and Powers, 2013; Goulding, 2009). These channels may foster social interactions and encourage learning engagement.

To overcome selection and reflection problems (Manski, 1993) in the identification of gender neighbor effects, we rely on within-neighborhood variation to control for unobserved characteristics of neighborhoods and families that might be correlated with the proportion of females and could also affect the outcomes. While there is a large literature documenting peer effects in schools using cohort-to-cohort variation, the use of cohort-to-cohort variation to identify neighborhood effects is rare. We also control for neighborhood time-invariant unobserved factors that might confound our identification. The basic idea is to compare the outcomes and choices of students from consecutive cohorts who are exposed to the same neighborhood environment and have similar characteristics, except for the fact that one cohort has a higher share of female neighbors than the other for idiosyncratic reasons (i.e., randomness in birth rates). Then we simultaneously control for gender composition in the school and neighborhood to isolate neighborhood peer effects, while we control for peer effects generated within a school. This allows us to mitigate any potential mechanical bias arising from the fact that each student's peers are drawn from a population with a different mean gender composition at school level (Guryan, Kroft, and Notowidigdo, 2009). We use variance decomposition to confirm that there is sufficient cohort-to-cohort variation in gender composition within neighborhoods. Our balancing tests also provide evidence that this variation is not associated with within-neighborhood variation in students' background characteristics. Finally, we perform a Monte Carlo simulation that verifies that the variation in the proportion of female students can be generated by a random process.

[3]This cohort-to-cohort identification method has also been used to examine peer effects along other dimensions, such as race (Hanushek and Rivkin, Hanushek and Rivkin); parental characteristics (Bifulco, Fletcher, and Ross, 2011); family violence at home (Carrell and Hoekstra, 2010); and home language (Friesen and Krauth, 2011).

We find that an increase in the proportion of females in the neighborhood has positive and significant effects on academic performance, college admission, and quality of postsecondary degree for both genders, but especially for females. We highlight a few key results. A 10-percentage-point increase in the share of females in the neighborhood (1) increases the national exam score of males and females by 2.9 and 3.3 percentage points, respectively; (2) causes males and females to be more likely to enroll in an academic university by 1.6 and 1 percentage points, respectively; (3) increases the likelihood of enrollment in a STEM (science, technology, engineering, and mathematics) degree by 0.5 percentage points for females only;[4] and (4) increases the expected wage for females by 0.2 standard deviations. Estimates on field of study and expected wage are less precise for males. Our results suggest that exposure to a 10-percentage-points higher share of female neighborhood peers is associated with a 2-percentage-point drop in the gender gap in STEM enrollment in college. Our more precise female peer effects impacts are in line with the strand of literature that suggests that females are more likely to develop confidence and leadership skills when they are surrounded by a higher share of females (Schneeweis and Zweimuller, 2012). These skills are likely to help females perform better in school, invest more human capital in math and science, and target more male-dominated fields and occupations, such as STEM (Lavy and Schlosser, 2011).

A higher share of female students may also lower disruption and violence, improve adult-student and inter-student relationships, and increase student experiences and satisfaction (Lavy and Schlosser, 2011; Schone, von Simson, and Strom, 2017; Goulas, Megalokonomou, and Zhang, 2020). Individuals may then become less disruptive, more focused, and better behaved (Figlio, 2007). Lower neighborhood disruption increases trust, discipline and feeling safe among individuals in the local community and improves student behaviour and engagement in learning activities (Burdick-Will, 2018).

We then explore nonlinearities and heterogeneous effects. First, we find that our effects are larger when the proportion of female neighbors is higher (i.e., over 56%). Females are 1 percentage point more likely to study STEM and are expected to earn a higher wage by 0.4 standard deviations when they have a female share in the neighborhood above 61% relative to having a female share below 53%. When testing for nonlinearities, we allow for multiple hypothesis testing using the procedure of Romano and Wolf (2005) and find that neighbor effects on academic outcomes and education choices are stronger in neighborhoods with high student density. Across outcomes, gender peer effects in high-density neighborhoods are as strong as 10 times the corresponding effects in low-density neighborhoods. This finding underscores the value of peer networks outside the school in high-density environments such as urban areas.

We conduct falsification and robustness exercises to support the causal interpretation of our results. We first construct a group of false peers based on distance, with whom students are unlikely to interact, and find no gender peer effects. In the second falsification exercise, we consider the impact of false peers (neighbors in preceding and succeeding cohorts) in time. These falsification estimates show no effects. This indicates that our results are driven by interactions among students of the same cohort and reassures us that our estimates do not pick up any confounding factors. We perform a series of robustness exercises to ensure that our results are not driven by time-varying confounding factors (e.g.,

[4] Policymakers are often interested in increasing female representation in STEM degrees and occupations, because gender differences in STEM participation have been linked to inequalities in earnings (OECD, 2017); STEM degrees are associated with higher lifetime earnings (Oreopoulos and Petronijevic, 2013).

opening or closing of private or single-sex schools). We also provide evidence that changes in a school's female share are uncorrelated with changes in dropout and intake rates. Our results remain unchanged when we expand our specification to control for the proportion of females in the preceding, current, and succeeding cohorts, which mitigates concerns about serial correlation within neighborhoods across time. Overall, our identification assumptions and results remain unaffected in an extensive battery of robustness exercises.

In the last part of the paper, we investigate how these gender neighbor effects interact with specific neighborhood environment characteristics. We view those interactions as the channels behind gender neighbor effects. We obtain new qualitative geographic information about communal facilities for each neighborhood in the administrative data. We hypothesize that peer social interactions are amplified in local communities with more public facilities. We focus on five facilities that are discussed in the literature: churches, libraries, parks or squares, Scouts, and sports fields, and also construct an overall index using PCA of all facilities. We find that gender neighbor effects are more pronounced in neighborhoods that have a higher intensity of communal facilities for females. The intensity of facilities in the local community does not seem to affect males' outcomes.

Our study extends beyond prior literature in several important ways. First, we contribute to the economics and education literature on nontraditional peer effects. We are the first to identify the impact of gender composition in neighborhoods on academic performance, STEM participation, and career outcomes; most peer effects studies focus on schoolmates (Hoxby, 2000; Lavy and Schlosser, 2011). The effects of other unconventional peer groups have also been examined. For instance, Foster (2006) defines a peer group as all college students residing in rooms that are in the same wing of a residence hall as a given student, and finds that the room proximity of college students alone yields no evidence of contextual peer effects. Carrell, Fullerton, and West (2009) use the random assignment of college students to squadrons at the US Air Force Academy and find positive peer effects when a college student is assigned to small groups of students who spend most of their time working together. Similarly, Sacerdote (2001); Zimmerman (2003); and Stinebrickner and Stinebrickner (2006) use the random assignment of students to college dormitories and find only moderate college roommate peer effects. Halliday and Kwak (2012) study the role of friends and find large peer effects on educational attainment. Hill (2015) uses survey data to study the effect of opposite-gender school friends on high school GPA.[5] Finally, there is recent evidence that neighbors affect college attendance rates and family decisions (Deutscher, 2020; Barrios-Fernández, 2022). We contribute to the literature by examining the effect of same-age neighbors on a variety of high school, university admissions and career-related outcomes. Understanding the connection between gender composition in the neighborhood and education outcomes can help stakeholders identify areas that need support from other programs and policies to ensure equitable access to school and career opportunities.

Second, we study whether specific neighborhood environment characteristics interact with gender neighbor effects. We focus on public facilities in the local community. Access to community spaces is important for young people, because they promote youth engagement and social integration. We collect

[5] In particular, Hill (2015) uses the share of opposite-gender schoolmates in one's neighborhood as an instrument for the gender composition of a student's self-reported friendship network. He finds that a student's share of opposite-gender school friends negatively affects scholastic outcomes in high school.

new information on the number of communal spaces for each neighborhood, which we combine with our administrative data. To the best of our knowledge, we are the first to combine administrative data on school performance and university admissions for neighbors with qualitative geographic data about the facilities in their local communities and study gender neighbor effects. We show that gender neighbor effects are amplified in local communities with more public facilities for females. This suggests that a higher share of female neighbors improves females' outcomes in high facility intensive neighborhoods. This is in line with prior work showing that peer effects in neighborhoods are mainly driven by same-sex peers, based on Australian data Deutscher (2020). There does not seem to be an effect on males. This is also in line with studies that show that males' education choices are not affected by peer or other school attributes Goulas, Griselda, and Megalokonomou (2022a); Lavy and Megalokonomou (2017). There is a growing body of literature that studies how the neighborhood environment affects college attendance rates and earnings (Chetty, Hendren, and Katz, 2016; Deutscher, 2020; Barrios-Fernández, 2022); crime behavior (Billings, Deming, and Ross, 2019); and mental health (Kling, Liebman, and Katz, 2007).[6] Together, these findings demonstrate that important social costs are related to geographic segregation and provide further evidence that school assignment policies may have unintended effects on academic, crime, and health outcomes.

Third, we contribute to the economics and education literature on field specialization and career path decisions. We provide previously undocumented evidence on the long-run occupational consequences of neighborhood peers. The literature has considered the effect, mainly on shorter-term outcomes, of gender composition in peer groups. At primary school level, a positive effect on both genders' test scores is found when there is a higher proportion of females in the classroom (Lavy and Schlosser, 2011; Hoxby, 2000). Lu and Anderson (2015) find that if a female student is surrounded by five females in the class rather than five males, this increases her performance by 0.2-0.3 s.d., but has no significant effects on males' test scores. At the high school level, Lavy and Schlosser (2011) find that a higher proportion of females in the school setting positively influences academic scores for both genders, but it is males who are more likely to enroll in advanced math and science classes in high school. Hu (2015) finds that a higher share of female peers in the classroom improves males' academic performance. Schone, von Simson, and Strom (2017) show that having more female peers in lower secondary grades renders females more likely to choose male-dominated courses in upper secondary grades.[7] Taken together, these studies have largely focused on student performance and enrollment in advanced school courses. More recently, Anelli and Peri (2019) show that males become more likely to major in engineering, economics, or business when they are exposed to a classroom in high school with more than 80% of male students using data from a small number of schools. Brenøe and Zölitz (2020) find that a larger proportion of female peers reduces females' probability of enrolling in and graduating from STEM programs at college and having more children. We contribute to the literature by exploring gender peer effects on short-

[6]There are also studies employing natural experiments of public housing opportunities to quantify neighborhood effects on a variety of economic outcomes in large North American cities (Oreopoulos, 2003; Jacob, 2004; Jacob and Ludwig, 2012).

[7]At university level, some studies have used the random assignment of students to groups and obtained mixed results. Other studies have found that the college peer environment has important effects on academic performance and/or students' choices (Carrell, Fullerton, and West, 2009; De Giorgi and Woolston, 2012; Booth, Sosa-Cardona, and Nolen, 2014; Hill, 2015; Zolitz and Feld, 2017).

and longer-term outcomes on a broader level of social networks: the neighborhood. Our longer-term outcomes include university enrollment, quality of postsecondary degree, enrollment in a STEM degree, and expected wage.

The fourth contribution of our study relates to the broad external validity of our findings. This is the first study to examine neighborhood gender peer effects using the full support of student characteristics. Unlike Anelli and Peri (2019), who use data from a relatively limited number schools in Italy to study gender peer effects on students' choices of university study, we provide evidence from a non-selective setting. Our results draw on a broader range of student backgrounds. Zölitz and Feld (2021) study the effects of gender composition in a business school setting and find that a higher share of female peers is positively associated with additional courses in marketing. Our study expands on the literature by investigating the effects of gender composition in neighborhoods in an entire country. Neighbor effects may depend on urban development. Using data from an entire country with patches of high urban density provides rich variation in neighbor effects. Estimates of the entire distribution of gender composition effects in a neighborhood may be of particular policy relevance. Overall, this paper contributes to a better understanding of the origins of gender differences in educational choices and labor market outcomes. The paper also shows that the gender composition of neighborhood peers is an important aspect of the social environment that shapes an individual's outcomes and choice of field specialization, quality of higher education, occupation, and consequently wages.

2 Institutional Setting and Data

2.1 Institutional Setting

Students are assigned to public schools based on proximity to their residential address. Most students (92%) in Greece attend public schools (OECD, 2018b). Parents are offered the opportunity to enroll their child to a unique public high school. Local school authorities inform parents which public school their child has been assigned to, and parents must provide proof of their residential address and utility bills to register their child in that school.[8] Assignment of students to high schools takes place at the beginning of the 10$^{\text{th}}$ grade.[9] Once assigned to a school, students are by law alphabetically assigned to classes for their general education courses (Goulas, Griselda, and Megalokonomou, 2022a,b; Goulas, Megalokonomou, and Zhang, 2020).[10] In addition to general education classes, students are required to choose a track for specialization in the 11$^{\text{th}}$ and 12$^{\text{th}}$ grades.[11] Usually, students choose a track based on their desired field of study at university level.[12] There are three options: (1) classics or humanities, (2) science, and (3) exact science. Students take school-level exams throughout their senior year and national exams at the end of the last senior high school grade—i.e., 12$^{\text{th}}$ grade.

[8]If the family pays rent, the school requires proof of rent payments and official documents from the Tax Authority that certify their residential address.

[9]Thus, school authorities have control of schools' enrollment and avoid over/undersubscription in a given school. Unfortunately, the Ministry does not collect data on 10$^{\text{th}}$-grade students.

[10]Teachers are also randomly assigned to classrooms within schools (Lavy and Megalokonomou, 2017, 2022).

[11]All schools provide these three tracks. The data confirm that there are students in all tracks in each high school.

[12]This track decision may be based on several factors, including comparative advantage in a subject, preferences over a specific major, gender composition, etc. (Goulas, Griselda, and Megalokonomou, 2022a).

The transition from high school to postsecondary education in Greece is based on a centralized and transparent allocation of students to university departments. Many countries have a similar university admission system; for instance, Chile, China, Korea, Taiwan, and Turkey. Students are compared with each other based on their admissions grade,[13] which is the only criterion for university admissions. Following their university admissions exams, students compile a list of ranked choices of specific departments in universities (degree programs), which they submit to the Ministry of Education. A computerized central system at the Ministry of Education ranks all students by their admissions score and assigns the highest ranked student to her preferred degree choice. It then moves to the next student and assigns her to the first degree on her list for which there is an available place, and so on (Goulas and Megalokonomou, 2018; Bizopoulou, Megalokonomou, and Simion, 2022). There are two main postsecondary types of institutions: academic universities and technical schools. Academic universities have, on average, higher admissions criteria, and thus are considered to be more prestigious. The duration of studies in both types of degrees is four years. On average, technical schools have a more applied focus.

2.2 Data Description

To study neighborhood gender peer effects, we use data from all high schools in Greece for six consecutive 12$^{\text{th}}$-grade cohorts of age-17 students in academic years from 2003/2004 through 2008/2009. For our empirical analysis, we construct a unique dataset of all students who take the end-of-secondary-education national exams and link this information to their postsecondary enrollment data. We obtain this information from various sources:

1. Administrative data from the Hellenic Ministry of Education, which contain information on all 12$^{\text{th}}$-grade students and all schools in the country. For each student, we have information on the high school they attend, their baseline exam scores, end of grade 12 national exam scores, and the track/specialization they choose at the beginning of 12$^{\text{th}}$ grade. We also have information on their gender, age, graduation year, and an indicator for students born in the first quarter of the calendar year. We also obtained information on all students' postsecondary education choices from the Hellenic Ministry of Education, which we link to the main dataset described above using a unique identification code. This dataset contains information on students' exact postsecondary matriculation grade, an indicator for whether they enrolled in a university department or a technical school, the exact university/technical school enrolled identifier, a degree program identifier, and information on whether a student enrolled in (1) a science or mathematics degree or (2) a STEM degree. We also obtained information on each degree's *admissions threshold* or *cutoff grade* for each year. A degree's threshold or cutoff grade is the grade of the last student admitted in that year, which reflects how selective or prestigious a degree program is. More selective/prestigious degrees have higher admission thresholds or cutoff grades.

[13]National end-of-high-school exams are administered by the Ministry of Education. They take place once every year and last from late May to early June. The exams are graded blindly and externally. Students usually take exams in six subjects, with a combination of common subjects (language, mathematics, physics, biology or history), four compulsory track-specific subjects, and one elective exam. The postsecondary matriculation grade is the average of all the exams a student takes. Track-related subjects are assigned a greater weight in calculating the admissions grade, and special weight is assigned to the track subjects most closely related to the university departments a student plans to apply to.

2. Labor Force Survey data for the year 2003 from the National Statistical Authority. We use this quarterly data to map each college major onto the most closely related occupation identifier and then onto the annual earnings reported in 2003 for each occupation category. Respondents have 209 occupation categories available to them, and they select their occupation with high precision. Then the earnings data are grouped into 10 bins that represent the 10 national deciles with the highest frequency. For each bin, we use the lowest bound to construct a proxy for the minimum expected annual salary earnings from each occupation (*"Expected Wage"*). We consider this to be a proxy for students' expected annual earnings after graduation from each university degree (in euros). In other words, this is a proxy for how lucrative, on average, each occupation is.
3. Geographic coordinates (latitude and longitude) for each high school using Google Maps.
4. The Ministry of Economy and Finance provided us with average net income information for each postcode in the country in 2009 in euros.
5. Detailed qualitative geographic information about the intensity of public facilities in the local communities that correspond to specific geographic areas using Google Maps. These facilities include the number of churches, libraries, parks, Scouts and sports facilities within each geographic area.

We restrict the data to obtain the final panel of the students and schools we use in our analysis.[14] We drop neighborhoods that have only one school because, according to our definition of neighbors, these students do not have any neighbors. These areas include, in particular, islands or other remote areas in which only one high school operates (38,333 students). The final sample includes 283,730 students and 222 neighborhoods for the period 2004-2009.

2.2.1 Construction of Neighborhoods

To study gender peer effects outside school walls, we examine gender peer effects of same-cohort peers in the local residential area (excluding school peers). Same-cohort peers who reside in the same residential building or complex with student i are likely to be her schoolmates. We identify school peers in the data, as they have the same school identifier. We exploit the institutional rules and define neighborhoods and neighbors in an innovative way. We consider broader neighbors who live one residential block or complex away. We define neighborhoods in relation to the school address, since we have no information on students' residential addresses. We exploit an institutional feature, according to which students who attend neighboring schools reside close to each other. In particular, same-cohort peers who reside in close proximity to student i—but not as close as her schoolmates—by law attend a neighboring school. Thus, we define as *neighbors* all same-cohort peers who attend neighboring schools (within a 1-mile radius of

[14] We exclude private (27,784 students) and evening (2,741 students) schools from our analysis to avoid the endogenous selection of students into those schools, and thus endogeneity in the proportion of females. For the same reason, we also remove single-sex schools (427 students in the country attend female single-sex schools and 363 students attend male single-sex schools) from the main analysis. We also drop school cohorts or neighborhood cohorts for which the proportion of females was either 0 (150 students) or 100% (656 students). Finally, we exclude students who are observed in a school or neighborhood whose cohort size in a particular year is smaller than 10 because an additional female (or male) student in these small schools or neighborhoods may cause a substantial change in the proportion of females. When included, the estimations produce similar results that only vary at the second decimal place.

their own school). In practice, these are same-cohort peers who attend "next door" schools and thus reside in the local area, but are not schoolmates.[15] Our definition of neighbors exploits two features: First, schools are built very close to each other in this context, and thus, students who attend neighboring schools also live very close to each other. Second, there is relative absence of school choice in this context, and thus, students can only enroll to their assigned school in the local area, as discussed in Section 2.1.

Figure 1 provides a real-life example of school locations in a neighborhood in Piraeus, a dense region in the southwest of Athens. We show the location of schools that operated in the sample period in two neighborhoods. In each neighborhood, schools all located within an 1-mile distance. Schools in one neighborhood are denoted by red circles (neighborhood 1) and schools in the other neighborhood are denoted by blue rectangles (neighborhood 2). Seven schools operate in neighborhood 1 and six schools operate in neighborhood 2.

We then illustrate how close students who attend the same school and neighboring schools reside. We present again a real example of how we construct neighborhoods from the same area in Athens, as in Figure 1. In Figure 2 we focus on three schools located in neighborhood 1 in Figure 1. These schools are represented by the three large shapes (triangular, rectangle, and circle) and are located within 500 meters, as shown in Figure 2. We show with same shapes of smaller size and same colour where same-age students who attend each school reside. Students who reside in the small green triangular/blue rectangle/red circle shapes attend the large green triangular/blue rectangle/red circle school. In other words, students who reside around each school in their local area attend the same school, and are thus, schoolmates.

Students who attend neighboring schools live close to each other and are considered to be a student's neighbors. For illustration purposes, let's denote school 1 the one shown by the large (green) triangular, school 2 the one shown by the large (blue) rectangle, and school 3 the one shown by the large (red) circle. In our design, for students who attend school 1 (2, or 3), their neighbors are considered to be students attending schools 2 and 3 (1 and 3, or 1 and 2). Neighbors are likely to socially interact with each other in the local community. The same logic applies if we consider all schools in neighborhood 1 (instead of only 3). For students in each school in neighborhood 1, same-age students who attend all other neighboring schools in the same neighborhood are considered to be their neighbors.

In Table 1, Panels A and B, we provide some summary statistics at student and school level, respectively. On average, each school is attended by 54 students (s.d.=32) in grade 12. In Table 1, Panel C, we see that each student's neighborhood same-age peers come from three schools different from her own. Each student has on average 162 same-age peers in the neighborhood. We limit our neighborhoods to a 1-mile distance so that they are big enough to allow for school diversity but also compact enough to capture common behavioral patterns or synergies in the local area.

We use a geographic grouping algorithm to define and construct neighborhoods within 1 mile of each school. This algorithm calculates the Euclidean distance between every two schools for all schools in the

[15]The median distance of a school from each nearest neighboring school in the sample is 0.32 miles.

sample.[16] Then, the algorithm uses the cutoff rule of 1 mile to define neighborhoods.[17] We construct 392 neighborhoods that cover the whole country. We remove 142 neighborhoods that contain only one school, one neighborhood that contains only single-sex schools, and 16 neighborhoods that contain only schools with fewer than 10 students. We further remove 11 neighborhoods that contain exactly two schools and at least one of the schools has fewer than 10 students. Our final sample contains 222 neighborhoods. Every neighborhood includes all 12^{th}-grade students who attend any other high school within a 1-mile distance of a student's high school.[18]

2.3 Descriptive Statistics

Table 1 describes our pooled data across cohorts. Panel A presents descriptive statistics at student level. Students are on average 17 years old and 57% are females.[19] The beginning-of-the-year score (our baseline performance measure) on school exams is roughly 87% with a standard deviation equal to 10. We calculate the ordinal rank of students within a neighborhood, based on the beginning-of-the-year score. The mean ordinal rank of students within a neighborhood is around 236, with a standard deviation equal to 283. The high school exact science track is the most popular (almost 49% of students enroll in this track), and the second most popular high school track is classics or humanities. Only 15% of students enroll in the science track. Almost 81% of students enroll in some postsecondary institution. Panel B uses school-level variables. All schools in the sample are public, and 82% are located in urban areas, which are defined as those with more than 20,000 inhabitants.[20] On average, 54 students are enrolled in 12_{th} grade in each school. Panel C uses neighborhood-level variables. Each neighborhood contains, on average, three other schools, and each student in the neighborhood has, on average, 162 neighbors.

Table A.1 reports various demographics for our sample separately for each year from 2003-2004 to

[16]We apply a nearest-neighbor clustering process that calculates the distance between every pair of schools and uses the cutoff rule of 1 mile to assign schools to groupings. The maximum distance between any two schools in the same grouping is 1 mile. Thus, 1 mile represents the distance that defines a grouping of schools. We apply this grouping algorithm without replacement.

[17]An alternative way to construct our neighborhoods would be to group all schools around each school using a different grouping formula than the clustering algorithm. The clustering algorithm we use allows us to construct neighborhoods without recounting the same students multiple times. Each school and student is assigned to a unique neighborhood. One limitation of the clustering algorithm is that schools very close to several neighborhood borders get categorized in one neighborhood, but not in any other. This mis-measurement of the clustering algorithm is orthogonal to any student characteristics and would not bias the regression estimates.

[18]Schools are built very close to each other in most urban settings in Greece. To give an example, Figure A.2 depicts the school complex of Grava in Athens, where six high schools and several elementary and middle schools form a massive complex and share various facilities. Another example is the school complex of Kaisariani in Athens, where two high schools share a yard and a basketball court (Figure A.3).

[19]In the past, men were overrepresented in different levels of education, but this trend has been changing in the last decades. Since the 1990s, in many countries around the world, females have started to become the dominant gender in terms of education participation at various levels of education. In 2005, women represented 55% of the higher education student population in the OECD area (OECD, 2018a). If these trends continue, in countries such as Austria, Canada, Iceland, Norway, and the United Kingdom, there will be almost twice as many female students as males in higher education in 2025 (OECD, 2018a). In 2005, 62% of higher education degrees in Greece were awarded to females, while the OECD average is 57%. Other countries with similar female shares are Canada, Finland, Ireland, Italy, Spain and New Zealand (OECD, 2018a). In Greece, we find that the proportion of females who graduate from high schools and take the national exams—a prerequisite for university admissions—is on average 57%. This is just 2 percentage points above the average OECD higher education student population.

[20]This definition is used by the Hellenic Ministry of Internal Affairs.

2008-2009. The last row reports the average of these demographics across all years. In a typical year, the average proportion of females within a neighborhood is 57%. We also observe that in every year, female students on average outperform male students in all academic outcomes. In particular, females' national exam scores, matriculation status, and matriculation scores are higher than those of males. Also, females are on average 5.5 percentage points (63.7-58.2) more likely than males to enroll in an academic university rather than a technical school. What is interesting here is that although females on average outperform males in all academic outcomes (columns 5-12), males pursue more lucrative occupations and thus their expected wage (column 13) is higher than that of females (column 14) in each year.

3 Empirical Strategy

Our goal is to estimate the effect of neighborhood peers' gender on a variety of outcomes, such as test scores, matriculation, college study, and expected wage. To mitigate endogeneity issues related to the sorting of students across neighborhoods, or unobserved correlated factors (Manski, 1993), we rely on within-neighborhood variation. In particular, we exploit within-neighborhood variation in the proportion of females in 12th grade across consecutive cohorts to examine the degree to which changes in students' outcomes can be attributed to an observable same-age neighbors' characteristic, such as gender. We use the following specification to estimate gender peer effects for males and females separately:

$$Outcome_{i,u,T} = \alpha_u + \beta_u year_{u,t} + \gamma X_i + \delta Prop.Females_{u,t} + \zeta Z_{u,t} + \psi_t + \epsilon_{i,u,t}, \tag{1}$$

$Outcome_{i,u,T}$ is the outcome variable of student i, in neighborhood u and time T. T is either measured in t or t+1.[21]. Outcomes measured in year t include the national exams score, postsecondary matriculation status (takes the value of 1 if a student enrolls in a postsecondary institution), and postsecondary matriculation score. Our outcomes measured in year t+1 include enrollment in an academic university (takes the value of 1 if a student enrolls in an academic university vs a technical school), quality of postsecondary degree, enrollment in a mathematics or science[22] degree at university level (takes the value of 1 if a student enrolls in a mathematics or science degree), enrollment in a STEM[23] university degree (takes the value of 1 if a student enrolls in a STEM degree), and expected wage. The quality of the postsecondary degree reflects the prestige/selectiveness of a degree, is increasing in quality, and is calculated as a degree's ranking based on the average admission threshold of each university department across sample years.[24]

Neighborhood fixed effects are captured in α_u. Year fixed effects are reflected in ψ_t. X_i includes student controls for age, a dummy for quarter of birth, indicators for the track chosen at the end of 11th grade, and baseline performance measured by school exam scores shortly after the end of 11th grade across all subjects derived at neighborhood level. $Z_{u,t}$ is a vector of neighborhood-by-year characteristics of neighborhood u at time t, which contains (a) the number of 12th-grade students in the neighborhood

[21] t is practically the beginning of grade 12 and T is the end of grade 12

[22] Science degrees are all degrees offered by physical and earth sciences, biology, veterinary science, medicine, and pharmacy departments.

[23] STEM degrees include all degrees in science, technology, engineering, and mathematics.

[24] The quality of a postsecondary degree takes values from 0 to 100 with a mean of 50.

(neighborhood-level enrollment) and (b) a set of leave-out mean student controls at neighborhood and year level to account for the mechanical relationship between a student's and their peers' group variables (Guryan, Kroft, and Notowidigdo, 2009). These leave-out means exclude all students in student i's school. The variable of interest, $Prop.Females_{u,t}$, is the leave-out mean of the female indicator across all students in each neighborhood, excluding students in student i's school. The coefficient of interest, δ, captures the effect of having a higher share of females in 12$^{\text{th}}$ grade in a neighborhood on a student's educational outcomes and choices. We cluster standard errors at school level to allow for heteroskedasticity and serial correlation among students within each school. Specification (1) is also estimated at school level (with u reflecting the school level). The full analysis at the school level can be found in the appendix.

Given that our identification relies on cohort-to-cohort variation, it is important to control for any unobserved neighborhood-specific time-varying heterogeneity that might be correlated with the share of females. Thus, we include a neighborhood-specific linear time trend, $year_{u,t}$.[25] We augment specification (1) with school-level mean regressors to exclude the possibility of confounding between school and neighborhood influences.

4 Validity of the Identification Strategy

As discussed earlier, we rely on cohort-to-cohort variation in the share of female students within neighborhoods to estimate gender peer effects in space. In this section, we provide evidence of the sufficiency and randomness of variation in the share of female students from one cohort to the next to identify neighbor gender effects.

4.1 Variation Sufficiency and Balancing Tests

To test whether there is sufficient variation in the proportion of female neighbors to identify gender neighbor effects, we adopt two approaches. First, we decompose the variation in the proportion of female students in the sample into within-unit (neighborhood/school) variation and between-unit variation. Results are reported in Table A.2. We find substantial within-neighborhood variation in the female share of students (sum of squares: 2.99), which represents 63% of the total variation in gender composition at neighborhood level. Between-neighborhood variation (sum of squares: 1.74) is lower than the within-neighborhood variation. The school-level variation decomposition pattern is similar to that at neighborhood level. Figure 3 provides visual evidence of the level of variation used for the identification of gender peer effects. The top (bottom) panel shows that there is significnat variation in the distribution of the proportion of female neighbors (schoolmates).

To provide further evidence of sufficient variation, we calculate the percentage difference in the share of female neighbors between consecutive cohorts. We draw the distributions for each pair of cohorts separately in Figure 4. The top (bottom) panel presents variation in the proportion of females within neighborhoods (schools) for each pair of consecutive cohorts. We notice that most of the variation lies between a -20% and +20% difference in the share of female students from one cohort to the next. Figures 3

[25] This is a common technique in studies that use cohort-to-cohort variation and has also been employed by Hoxby (2000); Lavy and Schlosser (2011); and Bifulco, Fletcher, and Ross (2011).

and 4 reassure us that there is substantial cohort-to-cohort variation in the proportion of females within neighborhoods.

We then test whether the variation in gender composition in the neighborhood from one cohort to the next is random using two approaches. First, we perform balancing tests to examine whether cohort-to-cohort changes in the share of female students are systematically associated with cohort-to-cohort changes in students' predetermined characteristics, annual neighborhood enrollment, and other characteristics. In particular, we regress students' characteristics, annual neighborhood enrollment, and other characteristics on the share of females in the neighborhood. Table 2 shows the estimated effects for the full sample (columns 1-2), for males (columns 3-4), and for females (columns 5-6) separately. These results indicate an insignificant association between the proportion of females and student characteristics, annual neighborhood enrollment, and neighborhood postcode income for the full sample and by gender.[26]

4.2 Simulated Gender Composition

To provide further evidence that the variation in the proportion of females is consistent with a random process, we employ Monte Carlo simulations to randomly assign gender to students in the neighborhood. We present the distribution of the actual and the simulated within-neighborhood standard deviation of the proportion of female students in Figure 6. The simulations work as follows: For each neighborhood and cohort, we assign each student a random gender using a binomial distribution function with probability (p) equal to 0.5—the average proportion of female students in the sample—and compute the within-neighborhood standard deviation of the artificially generated proportion of female neighbors. Figure 6 shows the Epanechnikov kernel density plot of the simulated standard deviation of the proportion of female students. The pattern between the actual standard deviation of female share within a neighborhood and the simulated standard deviation is similar, which indicates that the actual within-neighborhood gender composition may be generated by a random process.

4.3 Source of Variation in Gender Composition

In this section we explore the source of variation in the proportion of females from one year to the next. We provide evidence that the cohort-to-cohort variation in gender composition may be attributable to randomness in the gender composition of local cohorts at birth. In particular, we show that changes in the gender composition of high school-age students from one year to the next may be generated by fluctuations in the number of males and females being born and living in the area around the schools. We compare the variation in the proportion of females within schools to the average proportion of females in the very small geographic area around each school. We obtain data from the 1991 and 2001 censuses from the Hellenic National Statistical Authority.[27] The census data contain the population of females (and the total population) in different age groups and geographic areas. These age groups are 0-4, 5-9, 10-14, and 15-19 years. We use available census waves to show that the variation in the proportion of

[26]Robustness checks show that excluding controls for student characteristics from our main specification (1) produces similar results.

[27]The census is conducted every 10 years in Greece.

females in high school may be generated by variation in the proportion of females in the population in the local area. The students we use in our main analysis are 0-4 years old in the 1991 census and 10-14 years old in the 2001 census. These data indicate that the annual gender share at birth for the whole country between 1991 and 2001 is constant.[28] Figure 5 shows that there is variation in the local share of female births across cohorts. This translates into variation from one year to the next in the proportion of females in local areas around schools.

To investigate whether the variation in the proportion of female high school students in the neighborhood is associated with variation in the birth rates, we use the smallest possible geographic structures the census provides. Specifically, there are 5,947 distinct low-level geographic units (census blocks), and the average population (of males and females across all age groups) in these geographic units is 1,806 residents. For consistency with our main analysis, we drop all geographic areas that have a female share equal to 0 or 100. We calculate the percentage change in the proportion of females in different age groups within each census block between 1991 and 2001. Figure 5 shows that the variation in the proportion of females within schools is visually very similar to the variation in the proportion of females within census blocks. The densities are centered around zero, which indicates that for several schools and census blocks the changes in gender share from one cohort to the next are small. The smallest and largest deviations in gender share are also similar in these two figures. Overall, this figure indicates that the variation in the proportion of females we observe within schools could be generated by randomness in the gender composition of local cohorts.

4.4 Non-sorting to Schools

It would confound our gender neighbor effects if students move across schools in a response to the share of female students in their schools or school quality. However, the institutional setting here does not leave a lot of space for this kind of behavior, for two main reasons. First, students are not allowed to enroll in a public school based on their preferences. Second, a school's gender composition or quality measures are not publicly available, and thus it would be difficult for parents and students to predict the gender composition of a cohort that enters the school in the next year. Therefore, parents or students are unlikely to select schools based on gender composition.[29] An additional empirical exercise is presented in Appendix Table B.1, which shows that student enrollments are uncorrelated with the proportion of female schoolmates or school quality.

5 Main Results

In this section, we present our main results for the effects of neighbor gender composition on different students' outcomes in high school and university. In Subsection 5.1 we show how neighbor gender interactions affect high school and university admissions outcomes. In Subsection 5.2 we examine the effects on the choice of university study and expected wages. In Subsection 5.3 we identify neighbor gender effects while controlling for schoolmate gender effects and in Subsection 5.4 we consider nonlinearities.

[28] The average change in gender share from one year to the next is 0.5%.

[29] Funding and resources per student are the same across all public schools in Greece. Ninety-five percent of schools are public. In this setting, school quality is not a common driver for families who move between neighborhoods.

5.1 Academic Performance and University Admissions Outcomes

Table 3 shows the estimated effects of the share of female neighbors on student scholastic and matriculation outcomes, university major, and labor market outcome for males and females, separately. It also reports the outcome means (columns 1 and 4). In columns 2 and 5, we control for year fixed effects, neighborhood fixed effects, neighborhood-specific linear trends, student-level controls (age, quarter of birth, student's choice of track at the beginning of 12$^{\text{th}}$ grade, and baseline performance) and neighborhood income; in columns 3 and 6, we add controls for neighborhood-by-year characteristics.

Our results in Table 3 show that when there is a higher proportion of female peers in a neighborhood, both male and female students (a) perform better on subsequent national examinations and are more likely to (b) enroll in a postsecondary institution, (c) obtain a higher postsecondary matriculation score, (d) enroll in an academic university department (rather than a technical school), and (e) enroll in a more selective or prestigious university department. In particular, we find that a 10-percentage-point increase in the share of female neighbors increases the national exam score of males and females by 0.29 and 0.34 percentage points, respectively. The same increase in the share of female neighors would increase the matriculation likelihood for females by 0.4 percentage points; for males, the effect is smaller and statistically insignificant. Our findings are slightly smaller than those of Lavy and Schlosser (2011) at school level, who find that a 10-percentage-point increase in the proportion of females increases the likelihood of matriculation by 1 percentage point for females. We also find that a 10-percentage-point increase in the share of female neighbors increases the matriculation score by 3.4% and 2.2% of a s.d. for males and females, respectively. Our estimates are slightly larger than those of Hoxby (2000) at school level, who finds that a 10-percentage-point increase in the proportion of females increases students' mathematics scores by 1%-2% of a s.d. in Texas elementary schools. We also find that a 10-percentage-point increase in the female share in the neighborhood renders males and females more likely to enroll in an academic university, rather than a technical school, by 1.6 and 1 percentage points, respectively. Additionally, a 10-percentage-point increase in the female share in the neighborhood increases the quality of enrolled university degree for males and females by 1.1 and 0.7 percentage points, respectively.

5.2 Postsecondary Degree Choices

We next examine whether a higher share of female neighbors impacts student postsecondary decisions. In Table 3 we present the estimated effects of a higher share of female neighbors on the following postsecondary choices: enrollment in a math or science degree, enrollment in a STEM degree, and expected wage.

We find that in a neighborhood-cohort that has—for idiosyncratic reasons—a higher share of female neighbors, there are long-reaching effects on students' university and career choices, especially for females. In particular, in Table 3, we present the estimated effects of the proportion of female neighbors on the likelihood of enrollment in a science and mathematics degree and in a STEM university degree. An increase in the share of females in the neighborhood positively and significantly impacts the likelihood that females pursue a STEM degree. Estimates for males exhibit less precision than those for females. Enrollment rates for females in math and science degrees increase by 0.4 percentage points when they are exposed to a 10-percentage-point increase in the share of female neighbors. This is a sizeable effect, since on average only 8% of females enroll in math and science degrees. Our estimates remain qualitatively similar when we add neighborhood-by-year-characteristics as controls (columns 3 and 6).

Moreover, we find that a 10-percentage-point increase in the share of female neighbors increases the proportion of females who enroll in STEM postsecondary degrees by 0.5 percentage points. Given the low percentage of females who enroll in STEM, a 10-percentage-point increase in the share of females in the neighborhood increases females' likelihood of enrolling in STEM postsecondary degrees by 2.2% [(0.048/0.219)*10]. Our findings suggest that increasing the share of females within the neighborhood in a given cohort may reduce the gender gap in STEM degree enrollment, since the effects are positive and statistically significant only for females, and are not statistically different from zero for males. Considering the significant peer effects on females and based on our back-of-the-envelope calculation, a 10-percentage-point increase in the proportion of females in the neighborhood could potentially reduce the gender gap in STEM enrollments by roughly 2%.

Our finding on the impact of increasing the share of female neighbors on the drop in gender gap in STEM enrollment extends to wages. In particular, we find that females are more likely to choose occupations with higher pay when they are surrounded by a higher share of females.[30] In particular, a 10-percentage-point increase in the proportion of female neighbors increases females' expected wage by 1.3% and 2% of a s.d., respectively. We find no statistically significant impact on males' wages. Females' average standardized wage is -0.101, and males' average standardized wage is 0.136. This implies that a higher share of females within a neighborhood may reduce gender wage differences. In particular, based on our back-of-the-envelope calculations, a 10-percentage-point increase in the proportion of female neighbors reduces the occupation-related expected wage gap by 5% and 9%, respectively.

5.3 Simultaneous Controls for Gender Peer Effects at School Level

Neighbors' characteristics may be associated with those of schoolmates. We isolate the distinct influences of gender neighbor share by simultaneously controlling for the shares of females in both the neighborhood and school. This is important for two reasons. First, we mitigate any potential bias arising from the fact that each student's peers are drawn from a population with a different mean gender composition at school level (Guryan, Kroft, and Notowidigdo, 2009). Second, we want to ensure that we do not allow the exact same variation to drive the variation in our treatment variable of interest.[31] Specifically, we exploit the within-neighborhood variation in gender composition, while comparing students who have a similar gender share in their own schools. The following specification is used to isolate the neighbor gender composition effects while we control for the school gender composition effects:

$$Outcome_{i,n,T} = \alpha_{1n} + \beta_{1n} year_{n,t} + \gamma_1 X_i + \delta_1 Prop.Females_{n,t} + \theta_1 Prop.Females_{s,t} + \zeta_1 Z_{n,t} + \psi_{1t} + \epsilon_{1i,n,t}, \quad (2)$$

[30]Our estimates remain almost unchanged when we control for a binary indicator that indicates whether a student graduated from an academic university (versus a technical school), as well as a variable that indicates how male-dominated the degree is. The latter is calculated based on the proportion of male students who enroll in each degree every year. If a degree is male-dominated, this variable takes a value closer to 1.

[31]The simple correlation coefficient between the gender composition of a school and neighborhood is 0.133.

where n denotes neighborhood and s school. The coefficient of interest remains δ_1. In Table 4 columns 1 and 3 we present the main neighborhood estimates for males and females (same as in Table 3), respectively, while in columns 2 and 4 we also control for the proportion of female schoolmates in students' own schools. All estimated effects follow the same pattern and the magnitude of those effects drops only slightly when we use specification (2) instead of (1).[32]

5.4 Nonlinear Neighbor Gender Effects

Our specifications thus far have assumed that the effects are linear, but it is possible that the effects are higher when the share of females is larger. In this subsection, we examine nonlinear effects by splitting our variable of interest at neighborhood level into five quintiles. The first quintile indicator (for the lowest proportion of females) is our reference group. In Table A.4 we show that neighborhoods may switch between quintiles multiple times throughout the years (Lavy and Schlosser, 2011).[33] In a within-neighborhood regression model, this transition between quintiles within sample years offers variation for the identification of nonlinear neighbor gender effects.

Table 5 presents the estimated gender neighbor effects of switching from an environment with only a very low proportion of females (quintile 1) to an environment with an increasingly higher proportion of females (quintiles 2, 3, 4, and 5). The results largely indicate a positive relationship between academic outcomes and the share of female neighbors, with students performing better in neighborhood cohorts with a higher share of female neighbors. In particular, effects become more evident at the third, fourth and fifth quintiles, in which the proportion of female students exceeds 56%, 58%, and 61%, respectively. We also allow for multiple hypothesis testing using the procedure of Romano and Wolf (2005). We compute P-values using the Romano-Wolf step-down method and report those in square brackets. The statistical significance of our estimated results is generally robust to allowing for multiple hypothesis testing.

6 Falsification Exercises

In this section, we construct false groups of peers with whom students are unlikely to interact and check whether they generate results similar to those in the main analysis.

[32]In Table A.3 (columns 2-4 for males and 6-8 for females), we examine gender peer effects at school level. Columns 1 and 5 show the mean of each outcome variable for males and females in a school. Columns 2 and 6 present the main estimates for males and females, respectively, while in columns 3 and 7 we also control for school-by-year characteristics. In columns 4 and 8 we simultaneously control for the proportion of females in the neighborhood. The estimates of interest remain positive and change only slightly across different specifications. A higher proportion of female schoolmates increases both genders' scholastic and matriculation outcomes. This finding is established in the literature (Hoxby, 2000; Lavy and Schlosser, 2011; Lu and Anderson, 2015; Schone, von Simson, and Strom, 2017), and we provide evidence that this pattern is also evident in our setting.

[33]Cells on the diagonal of the matrix show the number of neighborhoods whose proportion of females remains in the same quintile across the sample period. This is the case for only 4 out of 222 neighborhoods. Each cell in the off-diagonal of the matrix displays the number of neighborhoods that are observed in two different quintiles during the 6 sampled years. For example, the first row shows that 80 neighborhoods moved between quintile 1 and quintile 2 (in both directions); 53 neighborhoods moved between quintile 1 and quintile 3 (in both directions); 53 neighborhoods moved between quintile 1 and quintile 4 (in both directions); and 73 neighborhoods between quintile 1 and quintile 5 (in both directions). We conclude that there is a substantial amount of quintile movement in each neighborhood.

6.1 False Peers in Distance

Our first falsification exercise demonstrates that students' geographic proximity alone does not generate positive gender peer effects. It also establishes that identifying the relevant peer group is important. We replace our gender composition variable within a neighborhood with a *false gender composition*. To test this, we construct a false peer group: all same-cohort students in one's periphery, excluding one's own neighborhood.[34] These peripheries are geographically larger than neighborhoods and thus, students within a periphery—who reside in different neighborhoods—are less likely to interact with each other.

Table 6 presents the estimated effects for the false peer group—the proportion of same-cohort females in one's periphery, except for their own neighborhood.[35] Of the 16 estimates, only two are statistically significant for males—at the 5% and 10% level—and are both negative. All remaining estimates are statistically insignificant and negative (except three). These findings indicate that geographic proximity alone does not generate the pattern we observe in our main analysis and that more distant peers do not seem to matter. Our findings confirm the result in the literature whereby social interaction effects at neighborhood level may decay rapidly over space (Billings, Deming, and Ross, 2019).

6.2 False Peers in Time

We then examine whether our main results can be explained by changes in the proportion of female neighbors in the younger (t-1) or older cohort (t+1). To study this, we replace the actual measure of treatment (proportion of female students at time t) within a neighborhood with the proportion of females within the same neighborhood in the previous (t-1) or following (t+1) year.

Results are presented in Table 7. In columns 1-2 and 3-4 we report estimates of the proportion of female neighbors in cohort t-1 and t+1 on various education outcomes, respectively.We report estimates for male (columns 1 and 3) and female (columns 2 and 4) neighbors, separately. All estimated effects are statistically insignificant and are all quantitatively much smaller than the corresponding effects in the main analysis. These results show that neighbor gender effects operate mainly at grade level and that male and female neighbors are not much affected by a higher share of females in the younger or older cohort. These results confirm the finding in the literature that social interaction effects are stronger for same-grade individuals (Lavy and Sand, 2015; Billings, Deming, and Ross, 2019).

7 Heterogeneity by Neighborhood Density

In this section, we investigate whether gender peer effects vary by neighborhood density. To do so, we run our main specification separately for neighborhoods with density above or below the average. Table 8 presents the estimated effects of the gender composition of neighbors on scholastic outcomes and education decisions for samples stratified by neighborhood density: low (below-average neighborhood-cohort size of 162 students) and high (above average neighborhood-cohort size of 162 students). The estimated effects

[34] There are 13 peripheries in total.

[35] We follow Guryan, Kroft, and Notowidigdo (2009) and additionally control for the cohort-to-cohort proportion of females in one's school in columns 1 and 2. We do so to eliminate potential bias that stems from the fact that each student is drawn from a population that has a different proportion of females at school level. Our results remain very similar, but are not reported here due to space constraints. Results will be provided upon request.

are more precise and stronger in neighborhoods with higher student density. Our findings indicate that in high-density neighborhoods, a 10-percentage-point increase in the female share of neighbors increases the matriculation score of male and female neighbors by 7.2% and 7.1% of a s.d., respectively, while it increases females' expected wage by 4.5% of a standard deviation. Overall, the main estimates we present in Table 3 seem to be driven by high-density neighborhoods. This finding emphasizes the value of peer networks outside of school in high-density environments such as urban areas.

8 Robustness Checks

In this section, we present a set of robustness exercises that support the causal interpretation of our findings. In general, any factor that may influence the proportion of females in a neighborhood in a time-invariant fashion may be less of a threat to our empirical strategy, since identification comes from cohort-to-cohort changes in the share of female neighbors. Potential identification challenges may primarily arise in situations in which omitted factors may influence education and career outcomes, as well as the proportion of females in some cohorts but not in others.

8.1 Single-sex and/or Private Schools in the Neighborhood

A potential concern is the presence of single-sex or private schools in the area. In some areas, single-sex or coeducational private schools operate and that may affect the share of females attending public schools. For example, if a single-sex school for males (or females) is in close proximity to a public coeducational school, that could result in a smaller share of males (or females) attending public coeducational schools in that neighborhood. However, any selection bias resulting from single-sex schools may not be substantial in this setting for two reasons. First, to the extent that single-sex schools are present in some neighborhoods for the entire sample period, their influence would be picked up by neighborhood fixed effects. Almost all single-sex and private schools in the sample operate for the whole sample period. Second, there are only a few single-sex schools in Greece, so they are an outside option for only a small sample of students.[36]

To address this concern we perform several robustness exercises. First, we control for neighborhood-specific linear time trends in all specifications to capture any unobserved factors that might confound neighborhood gender peer effects. Second, we include controls for the cohort-to-cohort enrollment of males and females in single-sex or private schools in each neighborhood. In Table 9, columns 1 and 3, we present the main estimated coefficients (from Table 5), and in columns 2 and 4 we add controls for the cohort-to-cohort enrollments of males and females in single-sex or private schools in the neighborhood. Estimates with and without these controls are very similar and vary only at the second decimal place. Third, we check whether the share of female students in public schools is affected by enrollments enrollments in single-sex or private schools in the nieghborhood. Estimates are shown in Table A.7. We restrict the analysis to neighborhoods that have at least one single-sex or private school. The outcome variable is the total number of students in the nieghborhood who attend single-sex or coeducational private schools (column 1), the total number of students who attend only coeducational private schools (column 2), the total number of all students who attend only single-sex schools (column 3), the number of female

[36]There are only 3 single-sex schools for females and 11 single-sex schools for males.

students who attend only single-sex schools (column 4), and the number of male students who attend only single-sex schools (column 5). All coefficients are statistically insignificant, which indicates that the proportion of females in public schools is unaffected by enrollments in single-sex or private schools in the neighborhoods. Fourth, we re-run our main specification (1), excluding areas in which single-sex schools (or private schools) operate. Our estimates change only in the second decimal place, and the overall pattern remains unaffected.[37]

8.2 Dropout and Intake Rates

One might worry that students who are—for idiosyncratic reasons—exposed to a lower (or higher) female share might respond by dropping out of (or enrolling in) a school. We examine the possibility that students might leave their initial school during the academic year (or enroll the following year) after they are exposed to a different proportion of females. This would also affect the gender composition of neighbors. Students have no public school choice. However, they could potentially switch to (or away from) a private school if they realize that the proportion of females in their initial school is low (high).

To empirically address these concerns, we use a sample of 134 representative schools[38] for which we have information about whether a student dropped out during the academic year, as well as subsequent enrollment. We first notice that drop out and intake rates are very low. As we can see in Table A.6, drop out rates and student intake rates are very low (drop out: 3% and 1% of males and females, respectively, Panel A, first row, intakes: 7% and 6% for males and females, respectively, Panel B, first row). We then examine whether changes in the current dropout or subsequent enrollment rates are related to changes in the proportion of current female schoolmates. We estimate specifications similar to (1), using as the dependent variable an indicator variable for whether a student left his/her initial school or re-enrolled in the same school a year later. Estimated effects of the share of female schoolmates on the likelihood of leaving the school during the year or enrolling in the school a year later are small and insignificant. These findings are in line with our results in Table B.1 and provide additional evidence that there may be limited mobility across schools.

8.3 Measurement Error in the Proportion of Female Neighbors

The proportion of same-age female students in neighboring schools may be an imperfect representation of the actual share of same-age females in one's neighborhood. Sources of measurement error may include the misclassification of same-age individuals who dropped out of school or peers who managed to enroll in a school located outside their neighborhood. We consider two types of measurement error: systematic error, which affects all individuals' measured share of females in their neighborhood the same, regardless of the actual share of females in their neighborhood, and non-systematic error, which may differ across individuals with different levels of actual neighborhood female share.

We perform a data-driven bounding exercise to assess the degree to which our neighbor gender effect estimates could be diluted when artificial additional measurement error of different sizes is introduced.

[37] Results are not reported here in the interest of space, but will be provided upon request.

[38] We provide evidence that this sample of 134 schools is a random sample in terms of the share of females and other characteristics in Table A.5 and also in Goulas and Megalokonomou (2021).

This involves a series of Monte Carlo simulations in which we add additional measurement error drawn from a normal distribution (with μ=0) to the measured proportion of females in the neighborhood and re-estimate our main specification (1) at ever-increasing levels of standard deviations of the measurement error. Through this process, we gauge the direction and approximate magnitude of any measurement error-induced bias. The standard deviation of the error distribution increases from 1% of the standard deviation in the female share in the neighborhood up to 40%. In terms of the proportion of females in the neighborhood, this represents an increase in the standard deviation from 0.056% to 2.24%. For each measurement error distribution, we simulate the data 1,000 times and estimate the neighbor gender effect coefficient. Figure 7 and Panel A shows the simulated estimates of the mean and the 2.5th and 97.5th percentiles from the sampling distribution of the estimated neighbor gender effect on national exam scores for each level of measurement error. We see that as measurement error increases, there is a downward (attenuation) bias. This bias grows slightly non-linearly with measurement error. Small additional measurement error has little impact on the results. The amount of downward bias from increasing the additional error from 1% to 25% of a standard deviation amounts to the same level of bias as increasing the error from 25% to 40%. Figure 7 and Panel B repeats this process with multiplicative measurement error (rather than additive measurement error). Specifically, we consider measurement error that increases as students' measured proportion of females is further from 1%. The patterns of potential bias under multiplicative error are similar to those under additive error. Even at high levels of measurement error, either additive or multiplicative, neighbor gender effects seem to have substantial impact on student performance.

8.4 Simultaneous Controls for the Proportion of Females in t, t-1 and t+1

We examine whether there is serial correlation in the proportion of female students within neighborhoods from one year to another. Our results in Table 7 suggest that gender peer effects operate among peers within the same cohort rather than across consecutive cohorts, so we do not expect to find a significant effect from the proportion of females in other cohorts. In the absence of serial correlation in gender composition within the same neighborhood, we do not expect our results to be driven by future or past values in the share of females. In Table 10 we focus on our estimates of the female share of neighbors in cohort t, while we simultaneously controlling, in the same regression, for the proportion of female neighbors in cohort t-1, cohort t and cohort t+1. We present estimated effects of the lagged (column 1 for males and 4 for females), current (column 2 for males and 5 for females), and the lead value (columns 3 for males and 6 for females) of the proportion of female neighbors. The main gender peer effect estimates remain positive and significant in almost all cases, while the estimated effects of future and lagged values are only occasionally positive and rarely significant. Overall, controlling for the future and lagged values of the proportion of females within a neighborhood does not seem to affect our actual treatment variable much.[39]

[39]This is not surprising given the results in Table 6, where we found no effects when we replaced our main treatment variable with the proportion of females in the same neighborhood in year t-1 or t+1.

8.5 Different Definitions of Neighborhoods

In this section, we examine whether our estimated effects are sensitive to how central the schools in each neighborhood are. To address this, we focus on those neighborhoods that contain at least 3 schools and examine the sensitivity of estimated effects when we drop (1) the farthest school—the school with the largest average distance to all other schools within the neighborhood; and (2) the closest school—the school with the smallest average distance to all other schools within the neighborhood, separately.[40] This exercise is equivalent to using different definitions of neighbors. Columns 1 and 4 in Table 12 replicate the main results for neighborhoods that have at least 3 schools for males and females, respectively. Columns 2 and 5 show the estimated female share effect when we drop the school with the highest average distance to all other schools (i.e., the most distant school) within the neighborhood. Columns 3 and 6 show the estimated female share effects when we drop the school with the lowest average distance to all other schools (i.e., the most central school) in each neighborhood. We find that dropping the farthest school generates, on average, slightly larger neighbor effects for female students across outcomes. This would be equivalent to focusing on slightly more distant neighbors. On the other hand, dropping the closest school produces, on average, slightly smaller estimated effects across outcomes. This would be equivalent to focusing on slightly closer neighbors. This pattern is more salient among females than males.

8.6 Addressing Potential Mechanical Relationship

A last concern would be that the characteristics of school peers may influence both the characteristics of neighborhood peers and outcomes and potentially bias the estimated neighbor effects. We investigate this hypothesis by augmenting our main specification to simultaneously control for school-by-year and neighborhood-by-year characteristics. Following Guryan, Kroft, and Notowidigdo (2009), we use the school- and neighborhood-level leave-out means of student characteristics to account for the mechanical relationship between a student's and both peer groups' characteristics. The estimated effects are presented in Table 11 and show that an increase in the proportion of female neighbors produces effects on males' and females' outcomes similar to those in our main specification (columns 9 and 12 in Table 3). Specifically, almost all estimates remain positive and statistically significant for both males and females. The additional inclusion of school-by-year characteristics leaves the standard errors virtually unchanged, but slightly reduces the magnitude of the estimates. For example, a 10-percentage-point increase in the share of female neighbors increases females' matriculation score by 0.208 s.d. (instead of 0.217 s.d.) and increases STEM enrollment by 0.46 (instead of 0.48) percentage points.

9 Mechanisms

Understanding the mechanisms through which gender peer effects at neighborhood level operate is as important as identifying the magnitude of those effects. In this section, we attempt to understand which neighbrohoods' characteristics drive our effects. We provide a brief review of the related literature for gender peer effects and empirically investigate those channels.

[40]The average distance of one school to all other schools is computed as the mean of the Euclidean distance between that school and all other schools in each neighborhood. The farthest (closest) school is the school with the highest (lowest) mean distance from all other schools.

9.1 Motivation

Community spaces have been found to be important for young people, because they promote youth engagement and social integration. Neighborhoods with more communal resources may amplify gender peer effects, because more communal resources (e.g. churches, parks, libraries, sports teams) help young individuals become more self confident and socially active (Majee and Anakwe, 2020). Other communal spaces (e.g., churches) may help youth take on organizational or community leadership roles (Zeldin, Christens, and Powers, 2013). Libraries have been leveraged in some communities to increase community engagement by providing spaces for community activities, creating opportunities for volunteer activities and building partnerships among young people (Goulding, 2009). Participating in Scouts, sport clubs, and other communal groups has been found to be positively associated with self-reported human capital, and particularly higher self-reported knowledge and confidence in science and technology. This may work through fostering a sense of brotherhood, community, and working together for a common goal—traits that are important in developing leaders (Falk and Needham, 2013). Those channels are more relevant for younger individuals.

In-class learning environment is crucial for student human capital development. There is evidence that a higher female share of students in a class lowers disruption and violence levels, improves adult-student and inter-student relationships, and increases student experiences and satisfaction (Lavy and Schlosser, 2011; Schone, von Simson, and Strom, 2017; Goulas, Megalokonomou, and Zhang, 2020). Individuals are less disruptive, more focused, and better behaved in environments in which females outnumber males (Figlio, 2007). Similar to the in-class behavioral effects of females on disruption and violence, out-of-class positive ramifications may also exist. A lower level of disruption or violence in a neighborhood increases trust, discipline and feeling safe among individuals in the local community and improves student behavior and engagement in learning activities (Burdick-Will, 2018). These positive externalities of a higher share of females may be more relevant in neighborhoods in which young individuals have more interaction channels and interact more with each other. These interactions may encourage communal engagement and inter-youth relationships, and can include after-school learning experiences, social gatherings, involvement in sports, etc.[41]

9.2 Empirical Investigation

We obtain new qualitative geographic information about neighborhood-specific communal spaces to investigate whether neighbor effects are more pronounced in areas with more communal resources and communal spaces. We hypothesize that peers' communication and interaction is amplified in public facilities within the local community. Thus, we use the intensity of public facilities as a proxy for the degree

[41] Our finding of higher precision in the estimated neighbor effects on the STEM enrollment of females compared with males may speak to plausible channels. Goulas, Griselda, and Megalokonomou (2022a) find that social interactions matter more for females' educational decisions and career choices, while males are not influenced by the gender composition of their peer group. Gneezy, Niedelre, and Rustichini (2003) and Niedelre and Vesterlund (2007) find that gender composition in a group impacts the level of peer competition. Males can display dominating behavior in groups (Karpowitz, Mendelberg, and Shaker, 2012).[42] Consequently, females might be more likely to end up in male-dominated fields, such as STEM, when surrounded by a higher share of females, potentially due to the development of noncognitive skills such as self-confidence and leadership skills (Schneeweis and Zweimuller, 2012). Thus, the finding of more consistent neighbor effects for females than males fits the narrative whereby neighbor effects may operate through peer comparison.

of neighbor interactions. Motivated by the literature discussed in the previous section, we focus on five communal facilities: (1) Church; (2) Library; (3) Parks and Squares; (4) Scouts; and (5) Sports Fields. Table A.8 shows the summary statistics of these neighborhood-specific facilities. The median of those facilities is around 2 churches, 1 library, 2 parks or squares, 1 Scout facility, and 3 sports fields or clubs in each neighborhood. We also construct an *Overall Intensity* index using the first component from a PCA that includes of all communal facilities. This method collapses the several dimensions (facilities) into a single dimension (index) and would represent the *Overall Intensity* of the reported facilities in the local community.

Figure 8 plots the locations of schools in neighborhood 1 from Figure 1 and the available communal facilities in that neighborhood. There is a significant number of churches, parks or squares, and sports fields for young people in this neighborhood. There are also 2 Scouts and libraries in which students are likely to interact with each other.

We then examine whether gender neighbor effects are more pronounced in neighborhood that have a higher intensity of communal facilities. For each type of facility, we create a binary indicator that takes the value of one if a neighborhood has an above-median intensity of the corresponding facility and zero otherwise. We add an interaction term between the neighorhood's female share and each of the binary indicators that correspond to each facility in our baseline regression (1). The coefficient of the interaction captures the differential effects of female neighbor share on the outcomes for neighborhoods that have an above-median density and those with have a below-median density in the corresponding facility. Figure 9 presents the coefficient plots of the interaction term between female neighbor share and a binary that acts as an indicator for whether the neighborhood has an above-median *Overall Intensity* index. We have grouped outcomes into panels depending on their value ranges. Panel A shows the estimated female neighbor effects on national exam scores and the quality of the enrolled postsecondary degree. Panels B and C plot the estimated female neighbor effects on all outcomes that are binary indicators and all standardized outcomes, respectively. Most estimated effects for males are around zero, but are positive for females. The positive estimated effects for females are also statistically significant for most outcomes (5/8). The overall pattern suggests that a higher share of female neighbors has positive and more pronounced effects on females' outcomes in neighborhoods with higher intensity (compared with those with a lower density) of communal spaces. This suggests that females benefit more from social interactions in neighborhoods that have a higher number of communal spaces. Figure A.1 presents the related coefficient plots for each facility and outcome, separately. Libraries and Scouts seem to be the most frequent drivers of the very positive effects on females, but all facilities contributes to the overall positive effect. Since all of our analyses are performed separately by gender, we are not as interested in the comparison between estimated coefficients for females and males, but rather the importance of the facility intensity for each of the genders.

Increasing social interactions for females in neighborhoods with more communal resources amplify neighbor effects for females and improve their test scores and aspirations to study for a STEM degree and pursue a more lucrative occupation. However, we do not find evidence of positive externalities on males' outcomes by facility intensity. Figures 9 and A.1 indicate no difference between high and low facility-intensive neighorhoods for males. These findings are also reflected in Table A.9 in which we report estimated neighbor effects for above-median and below-median facility-intensive neighborhoods and for each facility, gender, and outcome, separately. Effects on females are much more pronounced in neighborhoods that have a high intensity of facilities, while for males there is no distinguishable difference. For instance, the estimated neighbor effects are similar for males below and above the median of the *Overall Intensity* of facilities (6.320 and 4.639) when the outcome is the national exam score (Panel A). For females, the effect is 4 times larger when the *Overall Intensity* is above median compared with below median (3.259 compared with 12.206). This pattern remains the same across all facilities, with the estimated effects being 2-12 times larger for females when the intensity of the facility is above the median. This is not completely surprising. Exploiting cross-cohort variation within neighorhoods, Deutscher (2020) also finds that peer effects appear to be driven by same-sex peers.

10 Conclusion

In this paper, we study nontraditional student peer effects and examine whether social networks outside of school matter for students' learning and postsecondary choices. A common threat in the literature is the inclusion of irrelevant peers. We believe that a relevant peer group for high school students is their same-cohort peers who reside near them but attend different schools in the local area. High school students, unlike college students, are likely to interact with their physically proximal peers (i.e., neighbors). There are frequent occasions in adolescents' lives in which they participate in activities with their neighbors (e.g., play sports together, participate in after-school learning activities, participate in local community activities, etc.). To showcase neighbor effects, we examine the effect of an observable non-serially correlated same-age neighbor characteristic, such as gender, on academic performance, postsecondary admission, degree choice, and expected wage.

We combine unique administrative data for the universe of students in Greece and geographic school data to uncover the effect of peer networks outside of school and understand what drives these effects. The data allow us to identify all same-cohort students who attend any other school within 1 mile of one's school, whom we define as one's neighbors. We exploit a unique institutional setting in which (1) students are assigned to schools based on the proximity of their residential address and (2) schools are built very close to each other. Identification relies heavily on the relative absence of school choice in this context. We define neighborhoods and neighbors in an innovative way. In particular, we define as neighbors all same-cohort peers who attend other public schools within a short distance (1 mile) from one's school. A student's neighbors live nearby (e.g., a block away) and attend neighboring schools. We exploit cohort-to-cohort variation in the gender composition of 12^{th}-grade students within neighborhoods to deal with the usual sorting and endogeneity problems. We compare the outcomes and choices of students from consecutive cohorts that have similar characteristics and face the same environment, except for the fact that one cohort has a higher share of female neighbors than the other.

We show that changes in the share of female neighbors can be attributed to randomness in the gender composition of local cohorts at birth. We provide evidence that our cohort-to-cohort variation in gender composition is uncorrelated with cohort-to-cohort changes in students' characteristics or dropout or intake rates. We also perform a Monte Carlo simulation that corroborates that our cohort-to-cohort variation in gender composition results from a random process. Two falsification exercises reassure that our estimates do not pick up any confounding factors at neighborhood level. In the first exercise, we do not find effects generated by more distant peers. In the second exercise, we do not find effects from younger or older cohorts, which increases our confidence in neighbor effects within the same cohort.

Our results reveal significant neighbor effects in a series of outcomes. We find that having a higher share of female neighbors increases both genders' national exam scores, matriculation rates, and matriculation scores and renders both genders more likely to enroll in an academic university (versus a technical school) and a more selective/prestigious university department. Neighbors' gender is also found to impact enrollment in STEM degrees for females but not for males. We show that student peer networks outside of the school environment have long-lasting impacts, and affect expected wage of females. Our effects are larger for higher proportions of females in the neighborhood—i.e., above 56%. We also show that our results overall are driven by neighborhoods with high student density. This result underscores the value of peer networks outside of school in high-density environments, such as urban areas.

To understand whether access to communal spaces of social interaction drives neighbor effects, we examine how gender neighbor effects interact with specific local community characteristics. Access to those local facilities may foster social interaction, social integration and encourage learning engagement. We collect rich qualitative geographic information about communal facilities for each neighborhood that we observe in the administrative data. We focus on five public facilities that are discussed in the literature: churches, libraries, parks or squares, Scouts, and sports fields. We also construct an overall index using PCA of all facilities. We find that gender neighbor effects are amplified for females in local communities that have a higher intensity of public facilities. Females are found to improve their scholastic outcomes, be more likely to study for a STEM degree, and pursue a more lucrative occupation when there are more public resources in their local community. The intensity of facilities in the local community does not seem to affect males' outcomes.

References

Akerlof, G. and R. Kranton (2002). Identity and Schooling: Some Lessons for the Economics of Education. *Journal of Economic Literature, 40*(4), 1167–1201.

Anelli, M. and G. Peri (2019). The Effects of High School Peers' Gender on College Major College Performance and Income. *Economic Journal 129*(618), 553–602.

Barrios-Fernández, A. (2022). Neighbors' Effects on University Enrollment. *American Economic Journal: Applied Economics 14*(3), 30–60.

Bifulco, R., J. M. Fletcher, and S. L. Ross (2011). The Effect of Classmate Characteristics on Post-secondary Outcomes: Evidence from the Add Health. *American Economic Journal: Economic Policy 3*(1), 25–53.

Bikhchandani, S., D. Hirschleifer, and I. Welch (1992). A Theory of Fads, Fashion, Custom, Cultural Changes as Informational Cascades. *Journal of Political Economy, 100*(4), 992–1026.

Billings, S. B., D. J. Deming, and S. L. Ross (2019). Partners in Crime. *American Economic Journal: Applied Economics 11*(1), 126–50.

Bizopoulou, A., R. Megalokonomou, and S. Simion (2022). Do Second Chances Pay Off? *CESifo Working Paper No. 9620*.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2013). Under Pressure? The Effect of Peers on Outcomes of Young Adults. *Journal of Labor Economics 31*(1), 119–153.

Booth, A., L. Sosa-Cardona, and P. Nolen (2014). Gender Differences in Risk Aversion: Do Single-Sex Environments Affect their Development? *Journal of Economic Behavior and Organizations, 99*, 126–154.

Brenøe, A. A. and U. Zölitz (2020). Exposure to More Female Peers Widens the Gender Gap in STEM Participation. *Journal of Labor Economics 38*(4), 1009–1054.

Burdick-Will, J. (2018). Neighborhood Violence, Peer Effects, and Academic Achievement in Chicago. *Sociology of Education 91*(3), 205–223.

Buu, A., C. DiPiazza, L. Puttler, H. Fitzgerald, and R. Zucker (2009). Parent, Family, and Neighborhood Effects on the Development of Child Substance Use and Other Psychopathology from Preschool to the Start of Adulthood. *Journal of Studies on Alcohol and Drugs, 70*(4), 489–498.

Carrell, S., R. Fullerton, and J. West (2009). Does Your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics, 27*(3), 439–464.

Carrell, S. E. and M. L. Hoekstra (2010). Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids. *American Economic Journal: Applied Economics 2*(1), 211–28.

Chetty, R., N. Hendren, and L. F. Katz (2016). The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review 106*(4), 855–902.

De Giorgi, Pellizzari, M. and W. Woolston (2012). Class Size and Class Heterogeneity. *Journal of the European Economic Association, 10*(4), 795–830.

Deutscher, N. (2020). Place, Peers, and the Teenage Years: Long-Run Neighborhood Effects in Australia. *American Economic Journal: Applied Economics 12*(2), 220–49.

Falk, J. H. and M. D. Needham (2013). Factors Contributing to Adult Knowledge of Science and Technology. *Journal of Research in Science Teaching 50*(4), 431–452.

Figlio, D. (2007). Boys Named Sue: Disruptive Children and Their Peers. *Education Finance and Policy 2*(4), 376–394.

Fletcher, J. M. and S. L. Ross (2018). Estimating the Effects of Friends on Health Behaviors of Adolescents. *Health Economics 27*(10), 1450–1483.

Foster, G. (2006). It's Not Your Peers, and It's Not Your Friends: Some Progress Toward Understanding the Educational Peer Effect Mechanism. *Journal of Public Economics, 90*(8-9), 1455–1475.

Friesen, J. and B. Krauth (2011). Ethnic Enclaves in the Classroom. *Labour Economics 18*(5), 656–663.

Gneezy, U., M. Niedelre, and A. Rustichini (2003). Performance in Competitive Environments: Gender Differences. *Quarterly Journal of Economics, 118*, 1049–1074.

Goulas, S., S. Griselda, and R. Megalokonomou (2022a). Comparative Advantage and Gender Gap in STEM. *Journal of Human Resources*, 0320–10781R2.

Goulas, S., S. Griselda, and R. Megalokonomou (2022b). Compulsory Class Attendance versus Autonomy. *IZA Working Paper No. 14559*.

Goulas, S. and R. Megalokonomou (2018). Which Degrees Do Students Prefer During Recessions? *Empirical Economics*, 1–33.

Goulas, S. and R. Megalokonomou (2021). Knowing who you actually are: The effect of feedback on short- and longer-term outcomes. *Journal of Economic Behavior and Organization 183*, 589–615.

Goulas, S., R. Megalokonomou, and Y. Zhang (2020). Females, Classroom Disruption and STEM Careers. *Working Paper*.

Goulding, A. (2009). Engaging with Community Engagement: Public Libraries and Citizen Involvement. *New Library World 110*(1/2), 37–51.

Goux, D. and E. Maurin (2007). Close neighbours matter: Neighbourhood effects on early performance at school. *The Economic Journal 117*(523), 1193–1215.

Guryan, J., K. Kroft, and M. 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.

Halliday, T. and S. Kwak (2012). What Is a Peer? The Role of Network Definitions in Estimation of Peer Effects. *Applied Economics, 44*(3), 289–302.

Hanushek, E. A. and S. G. Rivkin. Harming the Best: How Schools Affect the Black-White Achievement Gap. *NBER Working Paper No. 14211*.

Hill, A. J. (2015). The Girl Next Door: The Effect of Opposite Gender Friends on High School Achievement. *American Economic Journal: Applied Economics 7*(3), 147–77.

Hoxby, C. (2000). Peer Effects in the Classroom: Learning from Gender and Race Variation. *NBER Working Paper No. 7867*.

Hu, F. (2015). Do Girl Peers Improve Your Academic Performance? *Economic Letters, 137*, 54–58.

Jackson, C. K. (2016). The Effect of Single-Sex Education on Test Scores, School Completion, Arrests, and Teen Motherhood: Evidence from School Transitions. *NBER Working Paper No. w22222*.

Jacob, B. A. (2004). Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing Demolitions in Chicago. *American Economic Review 94*(1), 233–258.

Jacob, B. A. and J. Ludwig (2012). The Effects of Housing Assistance on Labor Supply: Evidence from a Voucher Lottery. *American Economic Review 102*(1), 272–304.

Karpowitz, C. F., T. Mendelberg, and L. Shaker (2012). Gender Inequality in Deliberative Participation. *American Political Science Review*, 533–547.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental Analysis of Neighborhood Effects. *Econometrica 75*(1), 83–119.

Lavy, V. and R. Megalokonomou (2017). Persistency in Teachers' Grading Biases and Effect on Longer Term Outcomes: University Admission Exams and Choice of Field of Study. *Working Paper*.

Lavy, V. and R. Megalokonomou (2022). Longer- Term Effects of Teachers: Evidence from a Random Assignment of Teachers to Students. *Working Paper*.

Lavy, V. and E. Sand (2015). On the Origins of Gender Human Capital Gaps: Short and Long Term Consequences of Teachers? Stereotypical Biases. *NBER Working Paper (w20909).*

Lavy, V. and A. Schlosser (2011). Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics, 3*(2), 1–33.

Lee, S., L. J. Turner, S. Woo, and K. Kim (2015). The Impact of School and Classroom Gender Composition on Educational Achievement. *Working Paper*.

Lee, Y. and N. Nakazawa (2022). Does Single-Sex Schooling Help or Hurt Labor Market Outcomes? Evidence from a Natural Experiment in South Korea. *Journal of Public Economics 214*, 104729.

Lu, F. and M. Anderson (2015). Peer Effects in Microenvironments: The Benefits of Homogeneous Classroom Groups. *Journal of Labor Economics, 33*(1), 91–122.

Majee, W. and A. Anakwe (2020). Youth Engagement: A Mixed Method Investigation of Adult and Youth Perceptions of Community Resources in Rural America. *Community Development 51*(2), 140–156.

Manski, C. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies, 60*(3), 531–542.

Mota, N., E. Patacchini, and S. Rosenthal (2016). Neighborhood Effects, Peer Classification, and the Decision of Women to Work. *IZA Working Paper, No. 9985*.

Niedelre, M. and L. Vesterlund (2007). Do Women Shy away from Competition? Do Men Compete Too Much? *Quarterly Journal of Economics, 112*(3), 1067–1101.

OECD (2017). Report on the Implementation of the OECD Gender Recommendations. *OECD Publishing, Paris.*

OECD (2018a). OECD Employment Outlook 2018. *OECD Publishing, Paris.*

OECD (2018b). *PISA Results: Greece.*

Oreopoulos, P. (2003). The Long-Run Consequences of Living in a Poor Neighborhood. *Quarterly Journal of Economics 118*(02), 1533–1575.

Oreopoulos, P. and U. Petronijevic (2013). Making College Worth It: A Review of the Returns to Higher Education. *Future of Children, 23*(1), 41–65.

Park, H., J. R. Behrman, and J. Choi (2018). Do single-sex schools enhance students' stem (science, technology, engineering, and mathematics) outcomes? *Economics of Education Review 62*, 35–47.

Romano, J. P. and M. Wolf (2005). Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica 73*(4), 1237–1282.

Sacerdote, B. L. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *Quarterly Journal of Economics, 116*(2), 681–704.

Schneeweis, N. and M. Zweimuller (2012). Girls, Girls, Girls: Gender Composition and Female School Choice. *Economics of Education Review, 31*(4), 482–500.

Schone, P., K. von Simson, and M. Strom (2017). Girls Helping Girls: The Impact of Female Peers on Grades and Educational Choices. *IZA Working Paper, No. 10586*.

Stinebrickner, R. and T. Stinebrickner (2006). What Can be Learned about Peer Effects Using College Roommates? Evidence from New Survey Data and Students from Disadvantaged Backgrounds. *Journal of Public Economics, 90*(8-9), 1435–54.

Waddell, G. (2012). Gender and the Influence of Peer Alcohol Consumption on Adolescent Sexual Activity. *Economic Inquiry 50*(1), 248–63.

Xiong, H., D. Payne, and S. Kinsella (2016). Peer Effects in the Diffusion of Innovations: Theory and Simulation. *Journal of Behavioral and Experimental Economics, 63*, 1–13.

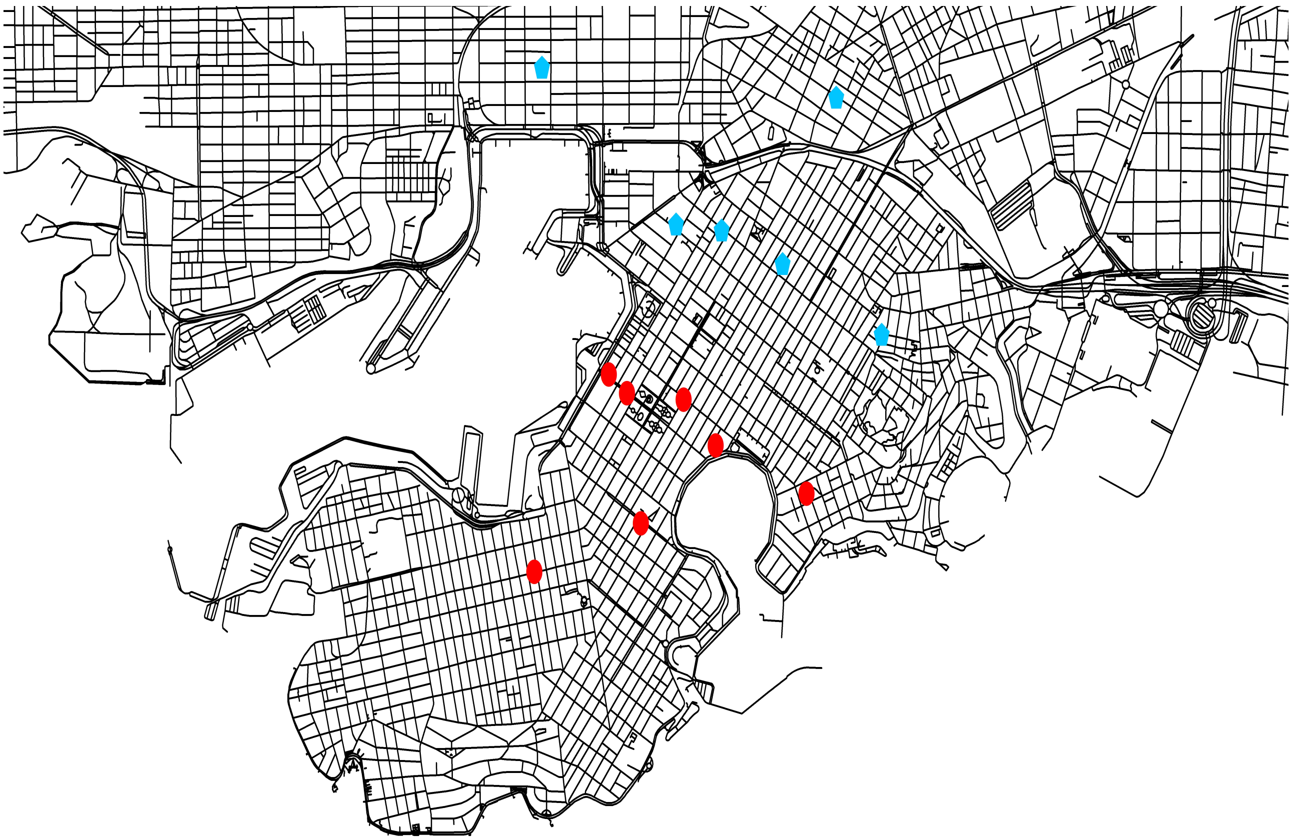
Zeldin, S., B. D. Christens, and J. L. Powers (2013). The psychology and practice of youth-adult partnership: Bridging generations for youth development and community change. *American journal of community psychology 51*(3), 385–397.

Zimmerman, D. (2003). Peer Effects in Higher Education: Evidence from a Natural Experiment. *Review of Economics and Statistics, 85*(1), 9–23.

Zolitz, U. and J. Feld (2017). The Effect of Peer Gender on Major Choice. *University of Zurich, Department of Economics, Working Paper No. 270*.

Zölitz, U. and J. Feld (2021). The Effect of Peer Gender on Major Choice in Business School. *Management Science 67*(11), 6963–6979.

Figure 1:
Location of Schools in Neighborhoods

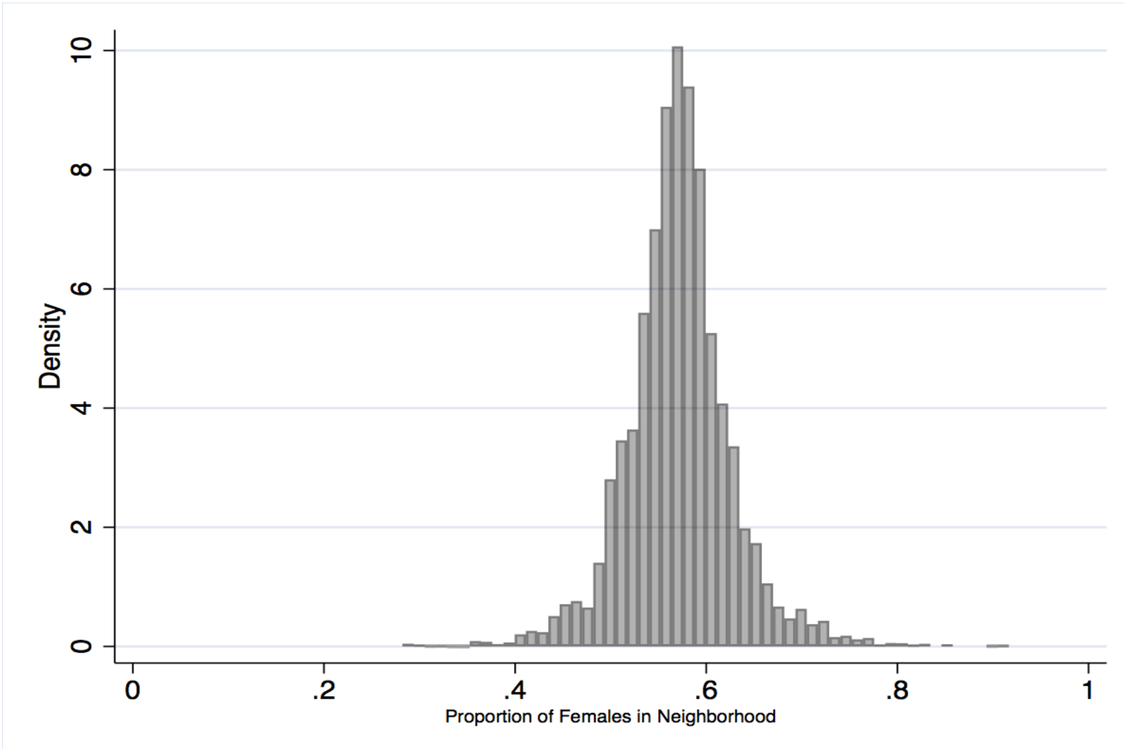


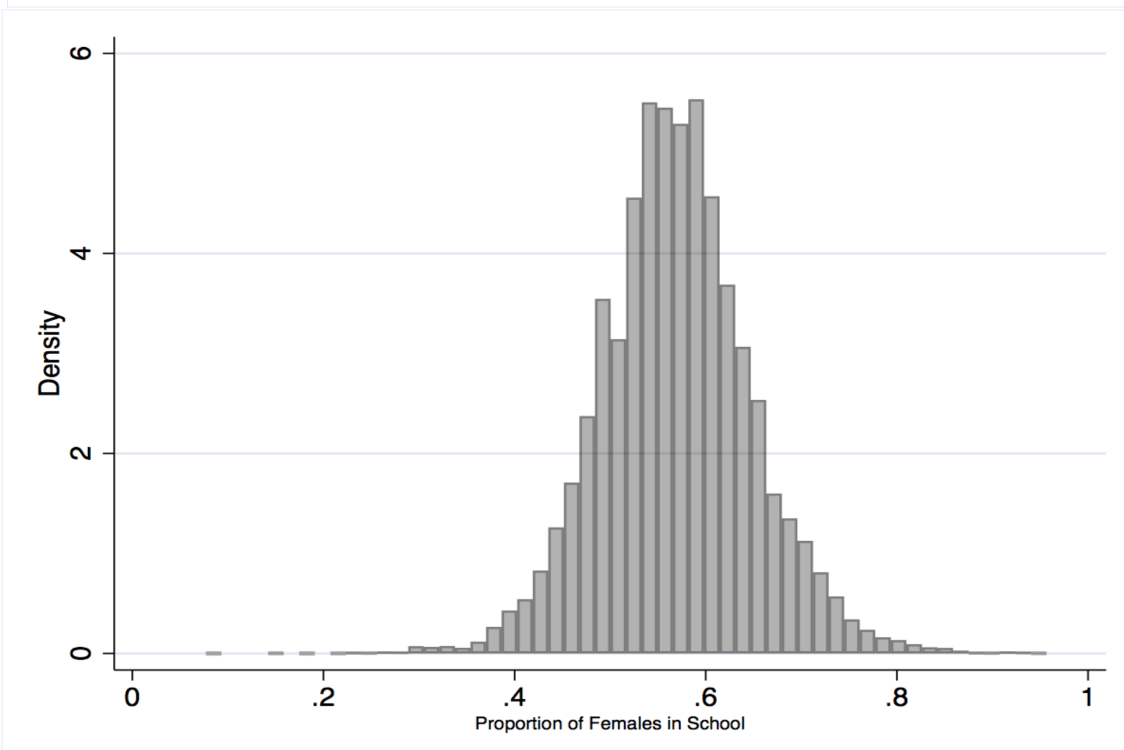
Notes: This figure shows the location of schools in two neighborhoods in Piraeus, a dense region in Athens: Schools in one neighborhood are denoted by red circles (neighborhood 1) and schools in the other neighborhood are denoted by blue rectangles (neighborhood 2). Schools within neighborhoods are located very close to each other (within 1 mile). In neighborhood 1 (2), we show the location of 7 (6) schools that operate in the local area in the sample period. Students who attend neighboring schools are considered to be a student's neighbors—in this example, a student's neighbors come from the remaining 6 schools in neighborhood 1 and from the remaining 5 schools in neighborhood 2. This is because students are assigned to schools based on geographic proximity to their residential address. Thus, students who attend neighboring schools—especially in such a dense geographic area—are very likely to live close to each other.

Figure 2:
CONSTRUCTION OF NEIGHBORHOODS

Notes: We focus on three schools that are located in neighborhood 1 in Figure 1 and are located very close to each other (within 500 metres distance). These three schools are denoted by the large shapes of a (green) triangular (school 1), a (blue) rectangle (school 2), and a (red) circle (school 3). We show with smaller shapes where same-age students who attend these three schools reside. Students who reside very close to a school in their local area are assigned to this school, as shown by the smaller shapes of the same colour in this map. In particular, we show where students who attend school 1 reside with smaller (green) triangular shapes. We show where students who attend school 2 reside with smaller (blue) rectangle shapes. We show where students who attend school 3 reside with smaller (red) circle shapes. The idea is that students who attend neighboring schools live close to each other and are considered to be a student's neighbors. Neighbors are likely to interact with each other in the local community, since they reside so close to each other. For students who attend school 1, same-age students who attend schools 2 and 3 are considered to be their neighbors. For students who attend school 2, same-age students who attend schools 1 and 3 are considered to be their neighbors. For students who attend school 3, same-age students who attend schools 1 and 2 are considered to be their neighbors.

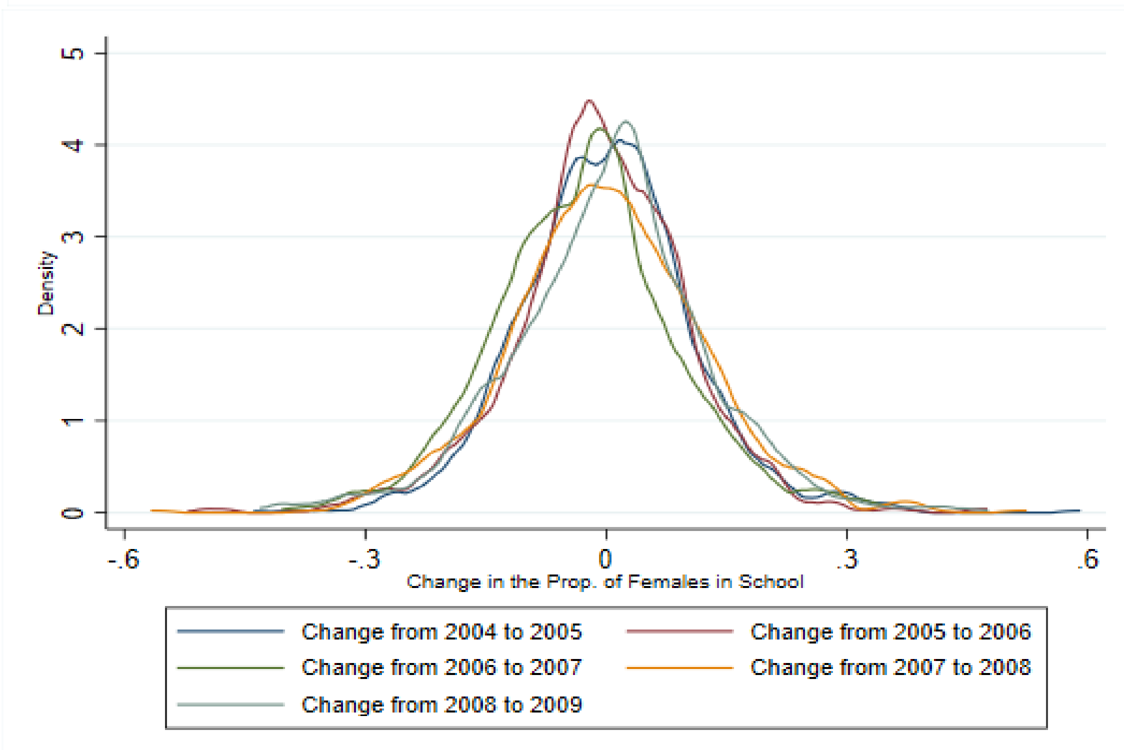
Figure 3:
HISTOGRAMS OF THE PROPORTION OF FEMALES IN SCHOOLS AND NEIGHBORHOODS





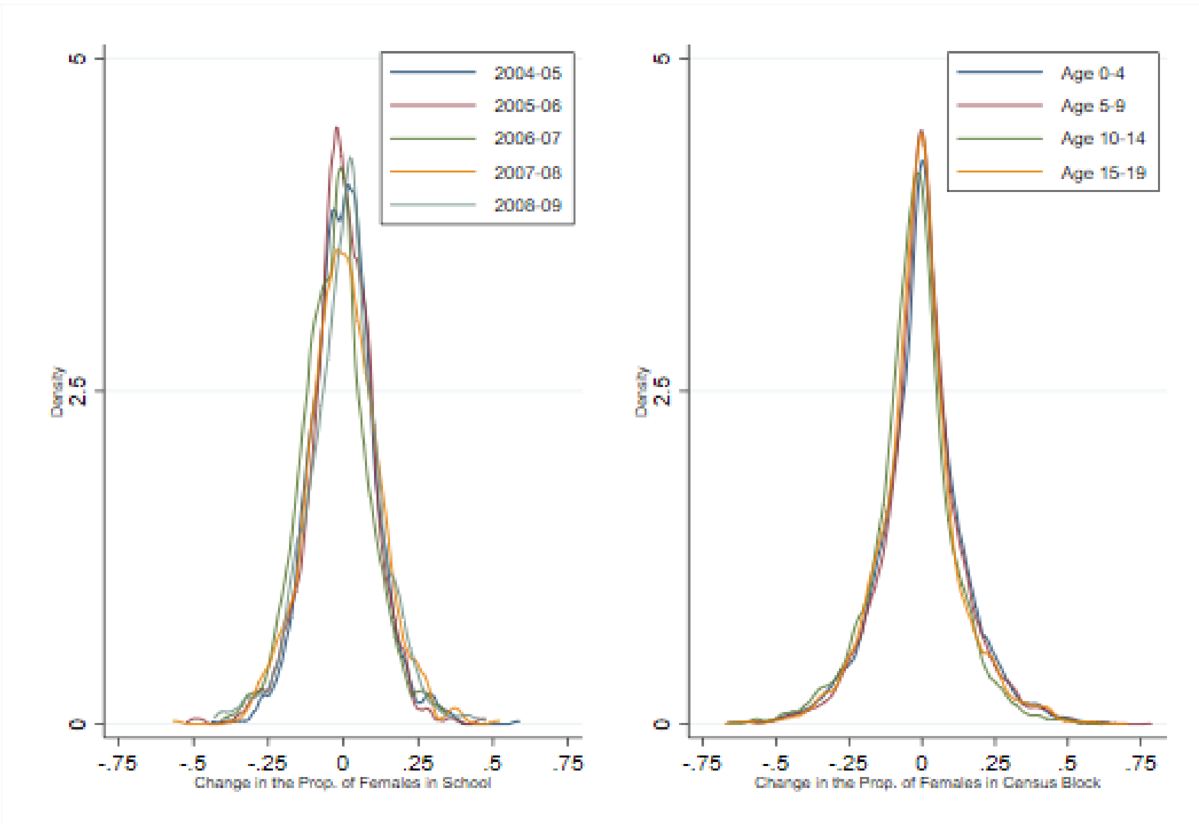
Notes: The top histogram presents the distribution of the proportion of female students in neighborhoods across cohorts and the bottom histogram presents the distribution of the proportion of female students in schools across cohorts.

Figure 4:
Densities of Changes in the Proportion of Females between Consecutive Years



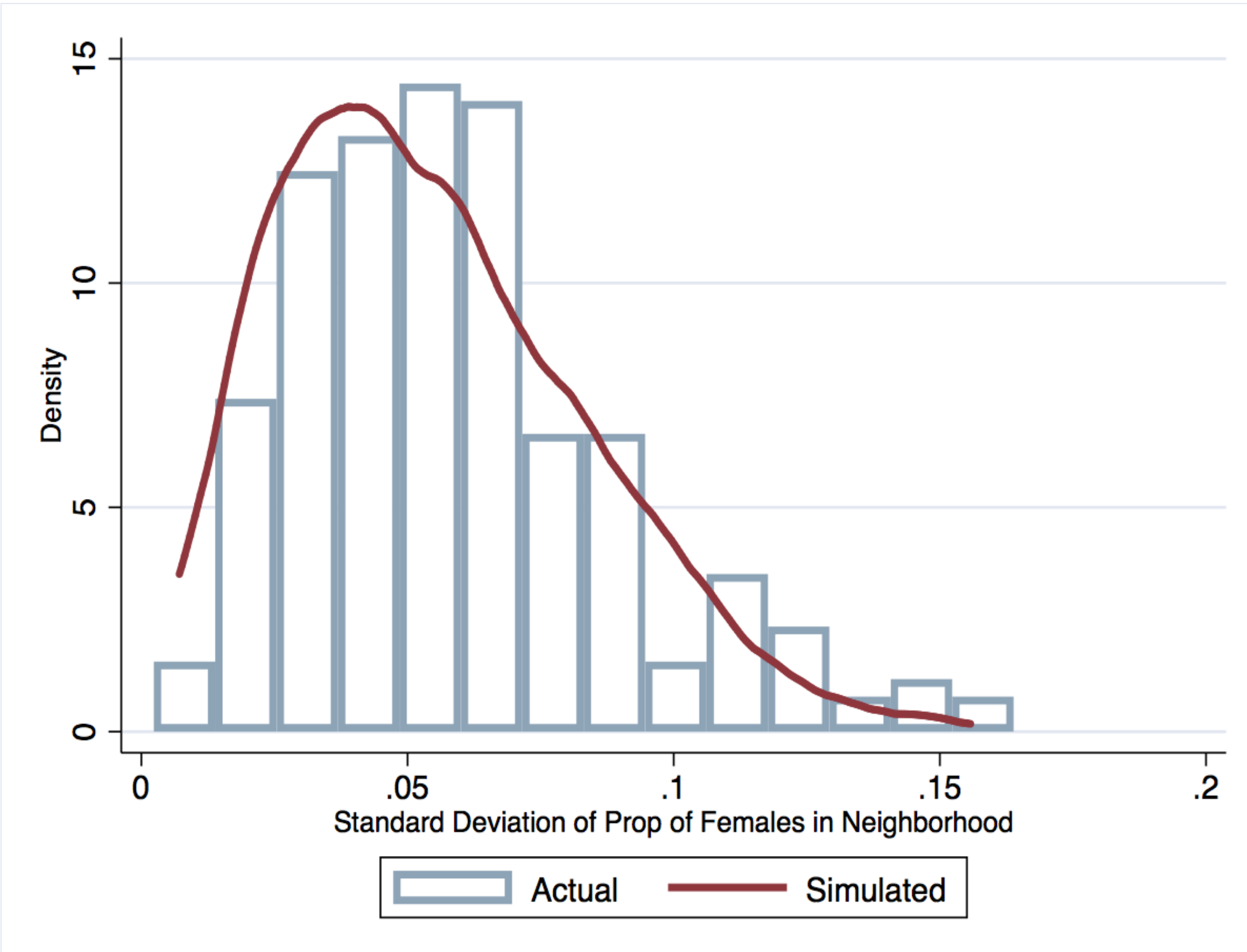
Notes: The top figure presents the density of the change in the proportion of females in neighborhoods for each pair of consecutive years from 2004 to 2009. The bottom figure presents the density of the change in the proportion of females in schools for each pair of consecutive years from 2004 to 2009.

Figure 5:
VARIATION IN THE PROPORTION OF FEMALES WITHIN SCHOOLS AND CENSUS BLOCKS



Notes: The left figure presents the distribution of the change in the proportion of females in schools for each pair of consecutive cohorts between 2004 and 2009. The proportion of females in these figures ranges within (0,100). The sample includes all students who attend public schools between 2004 and 2009. The right figure presents the distributions of the change in the proportion of females in the population within very small geographic areas (equivalent to census blocks) for 1991 and 2001, separately for different age groups. Specifically, separate distributions are drawn for the population in the age group between 0-4, 5-9, 10-14, and 15-19 years. Census data were obtained from the National Statistical Authority. The total number of distinct geographic areas is 5,947 and the average population of any age in these areas is 1,806 residents. These are the smallest geographic units available in the census and are equivalent to census blocks. We exclude areas that have a female share equal to 0 or 100 from both samples: the high-school sample and the census sample.

Figure 6:
Actual And Simulated Standard Deviation of Proportion of Females

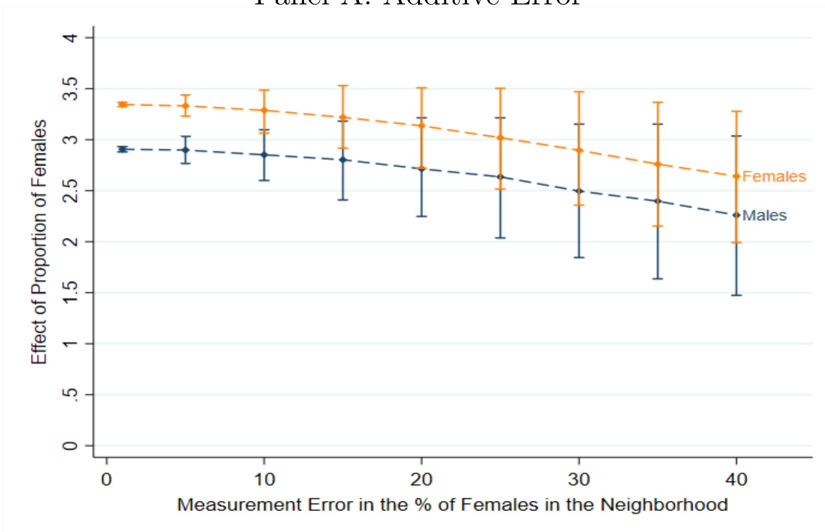


Notes: This figure shows the actual and simulated Monte Carlo standard deviation of the change in the proportion of females within neighborhoods. For each neighborhood, we randomly generate the gender of students in each cohort using a binomial distribution function, with p equal to the average proportion of females in the neighborhood across all years. We then compute the within-neighborhood standard deviation of the proportion of females. Details on the Monte Carlo simulation are provided in the text.

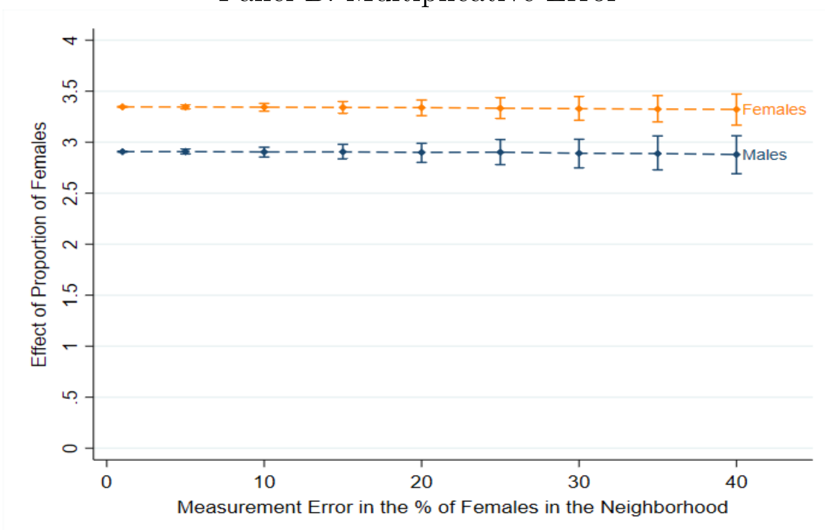
Figure 7:

ESTIMATES FROM MONTE CARLO SIMULATIONS WITH MEASUREMENT ERROR IN THE PROPORTION OF FEMALES IN THE NEIGHBORHOOD

Panel A: Additive Error

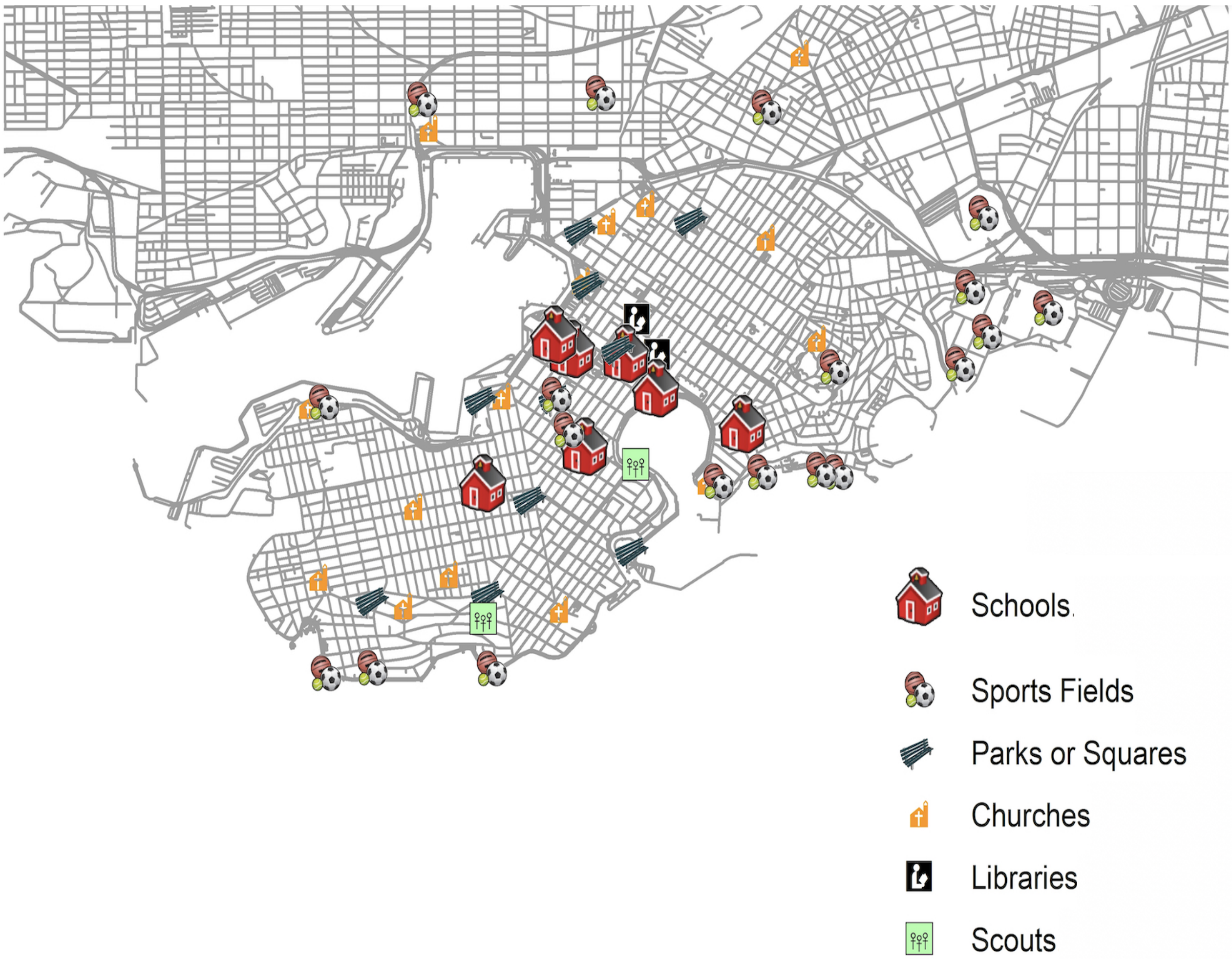


Panel B: Multiplicative Error



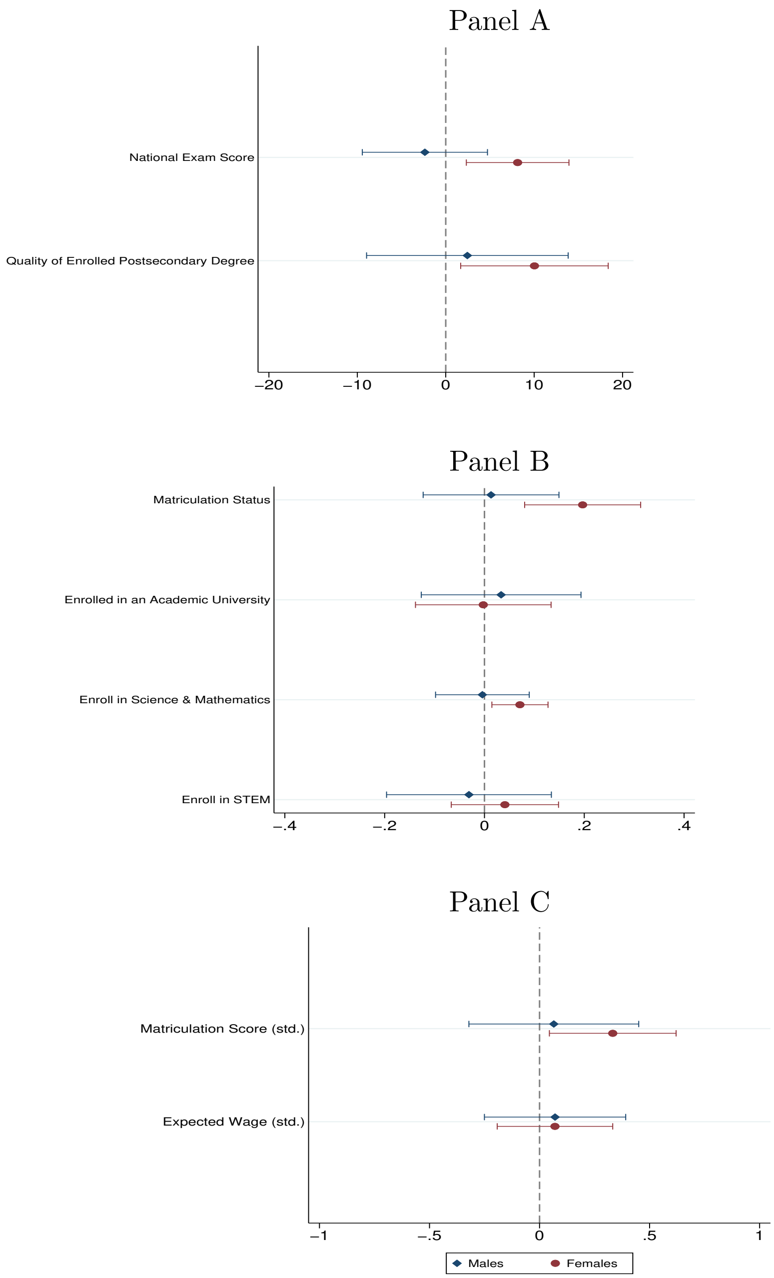
Notes: Panels A and B plot the mean neighbor gender effect on national exam scores for males and females from 1,000 simulations of specification (1), with increasing additional measurement error added to and multiplied by the proportion of females in the neighborhood, respectively. The measurement error for each student is independently drawn from a normal distribution with mean zero and a standard deviation proportional to the standard deviation of the neighborhood female share distribution. Error bars show the 2.5^{th} and 97.5^{th} percentiles from the sampling distribution of the estimated coefficient for each measurement error level. In Panel B, the size of the measurement error is increasing linearly in distance from 1%, so that students with balanced gender composition in the neighborhood experience no measurement error. Students with extreme neighborhood gender composition values experience additional measurement error drawn from a normal distribution with a mean zero and standard deviation equal to the proportion of the standard deviation in the measured neighborhood female share denoted on the horizontal axis.

Figure 8:
Locations of Schools and Facilities in a Neighborhood



Notes: We focus on neighborhood 1 from Figure 1 (schools denoted by red circles) and also show the location of the public facilities that are located in the local area. We also show the location of schools that operate in this neighborhood during the sample period. Additionally, we show the location of the following facilities of interest in this neighborhood: sports fields, parks or squares, churches, libraries, and Scout facilities. Each type of facility is denoted by a different shape. Students are likely to socially interact with each other in these public facilities in the local community.

Figure 9: Estimated Effects of Female Neighors by Intensity of Communal Facilities



Notes: These Panels show the estimated effects of the proportion of female neighbors by the intensity of communal facilities for different outcomes. The estimated coefficients plotted are generated by OLS regressions in which we use specification (1) and add an interaction term between the share of female neighbors and a binary indicator if a neighborhood has above-median *Overall Intensity* of communal facilities. The *Overall Intensity* of communal facilities in each neighborhood is an index generated by using the first component from a PCA consisting of all communal facilities: church, library, parks or squares, Scout facilities, and sports fields. A positive coefficient suggests a more pronounced gender neighbor effect when there is a higher density of communal facilities in the neighborhood compared with a lower density of communal facilities. Estimated effects report the interaction coefficients when different outcomes are used. Outcomes are listed vertically. For each outcome, we show bar plots for females and males separately in which bars denote 90% confidence intervals.

Table 1: Descriptive Statistics

	Mean	Std. Dev.	Min.	Max.	N
Panel A: Individual Level					
Age	16.891	0.536	14	58	283,730
Female	0.569	0.495	0	1	283,730
Born in First Quarter of Birth Year	0.161	0.367	0	1	283,730
National Exams Score (out of 100)	65.099	20.266	2.6	99.65	283,730
Baseline Test Score (out of 100)	87.195	10.429	0	100	283,730
Neighborhood Rank (based on Baseline Test Score)	236.381	283.304	1	1,950	283,730
Specialty in Classics	0.364	0.481	0	1	283,730
Specialty in Science	0.150	0.357	0	1	283,730
Specialty in Exact Science	0.486	0.500	0	1	283,730
Matriculation Status	0.809	0.393	0	1	283,730
Matriculation Score (std.)	0.008	1.001	-2.876	3.083	229,475
Postcode Income (Euro, 2009)	20,709	5,688	11,028	74,798	283,730
Panel B: School Level					
Age	16.897	0.095	16.462	18.103	940
Prop. of Females	0.577	0.059	0.143	0.882	940
Prop. of Students Born in First Quarter of Birth Year	0.166	0.048	0	0.538	940
Urban	0.823	0.382	0	1	940
Postcode Income (Euro, 2009)	19,699	5,758	11,028	74,798	940
No. of Students in Each School	54	32	11	190	940
Panel C: Neighborhood Level					
Age	16.897	0.055	16.774	17.186	222
Prop. of Females	0.574	0.041	0.382	0.749	222
Prop. of Students Born in First Quarter of Birth Year	0.164	0.029	0	0.302	222
Postcode Income (Euro, 2009)	19,978	5,911	6,500	46,058	222
No. of Other Schools in Each Neighborhood	3	4	1	31	222
Total Enrollment in Other Schools in Each Neighborhood	162	220	14	1,650	222

Notes: Data span six cohorts, 2004-2009. The indicator variable "Born in First Quarter of Birth Year" takes the value 1 if a student is born in the first quarter of the birth year. The indicator variable "Matriculation Status" takes the value 1 if a student enrolls in a postsecondary institution. Students have to enroll in a track at the beginning of the 12$^{\text{th}}$ grade, and that is their track or *specialty*. Students have three options: Classics, Science, or Exact Science. All schools offer these tracks. "Baseline Test Score" indicates students' performance on an exam taken shortly after the beginning of 12$^{\text{th}}$ grade. All schools are public. The matriculation score is observed only for students who enroll in a postsecondary institution. "Total Enrollment in Other Schools in Each Neighborhood" is equivalent to the number of neighbors a student has.

Table 2: Balancing Tests for the Proportion of Females in Neighborhood

	Full Sample		Males		Females	
	(1)	(2)	(3)	(4)	(5)	(6)
Age	-0.021	-0.006	-0.007	-0.007	-0.031	-0.001
	(0.025)	(0.026)	(0.036)	(0.037)	(0.031)	(0.032)
Born in First Quarter of Birth Year	-0.008	-0.025	-0.020	-0.033	-0.001	-0.019
	(0.015)	(0.015)	(0.020)	(0.022)	(0.020)	(0.021)
Annual Neighborhood Enrollment	15.301	0.937	16.145	4.926	14.651	-1.609
	(16.081)	(10.687)	(16.400)	(10.920)	(16.277)	(10.806)
Annual Neighborhood Average Age	0.014	0.036	0.012	0.029	0.018	0.039
	(0.024)	(0.025)	(0.024)	(0.025)	(0.024)	(0.026)
Annual Neighborhood Percent of Being Born in First Quarter of Birth Year	0.002	-0.013	0.004	-0.010	0.000	-0.015
	(0.016)	(0.016)	(0.015)	(0.016)	(0.016)	(0.017)
Log Neighborhood Postcode Income	-0.012	-0.023	-0.014	-0.024	-0.011	-0.023
	(0.016)	(0.015)	(0.015)	(0.014)*	(0.018)	(0.016)
Year FE	✓	✓	✓	✓	✓	✓
Neighborhood FE	✓	✓	✓	✓	✓	✓
Neighborhood-Specific Linear Trends		✓		✓		✓

Notes: The dependent variable is students' baseline characteristics (age or an indicator variable for being born in first quarter of birth year), neighborhood enrollment/size, and neighborhood-by-year characteristics (students' leave-out baseline characteristics averaged by neighborhood, year, and neighborhood annual enrollment). The variable of interest is the proportion of females in the neighborhood. Odd columns report estimates of the proportion of females at neighborhood level when year and neighborhood fixed effects are included. In all even columns, we also add neighborhood-specific linear trends. Standard errors are clustered at school level. Baseline test scores and track indicators are not included in the balancing tests under the assumption of strict exogeneity, since students' baseline test scores and their track specialization decision are realized at the beginning of 12^{th} grade. The sample size contains 283,730 students. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table 3: Estimated Effect of Proportion of Females in the Neighborhood on School Outcomes and Choice of University Major

	Mean	Males		Mean	Females	
	(1)	(2)	(3)	(4)	(5)	(6)
Scholastic and Matriculation Outcomes						
National Exam Score	63.554	3.137	2.907	66.267	3.592	3.346
		(1.532)**	(1.554)*		(1.327)***	(1.311)**
N		122,191	122,191		161,539	161,539
Matriculation Status	0.800	-0.024	-0.028	0.816	0.046	0.042
		(0.030)	(0.030)		(0.025)*	(0.025)*
N		122,191	122,191		161,539	161,539
Matriculation Score (std.)	-0.083	0.353	0.342	0.076	0.228	0.217
		(0.084)***	(0.085)***		(0.073)***	(0.073)***
N		97,734	97,734		131,741	131,741
Enrolled in an Academic University	0.582	0.159	0.156	0.637	0.107	0.104
		(0.037)***	(0.038)***		(0.033)***	(0.033)***
N		97,734	97,734		131,741	131,741
Quality of Enrolled Postsecondary Degree	47.541	11.166	10.871	51.923	7.675	7.395
		(2.446)***	(2.478)***		(2.091)**	(2.083)**
N		97,734	97,734		131,741	131,741
University Major						
Enroll in Science & Mathematics	0.083	0.030	0.029	0.076	0.040	0.039
		(0.019)	(0.019)		(0.014)***	(0.014)***
N		97,734	97,734		131,741	131,741
Enroll in STEM	0.408	0.053	0.052	0.219	0.049	0.048
		(0.036)	(0.036)		(0.025)**	(0.025)*
N		97,734	97,734		131,741	131,741
Labor Market Outcome						
Expected Wage (std.)	0.136	0.092	0.090	-0.100	0.200	0.201
		(0.069)	(0.069)		(0.065)***	(0.065)***
N		97,734	97,734		131,741	131,741
Year FE & Neighborhood FE		✓	✓		✓	✓
Neighborhood-Specific Linear Trends		✓	✓		✓	✓
Student-Level Controls		✓	✓		✓	✓
Neighborhood-by-Year Characteristics			✓			✓

Notes: Columns 1 and 4 report the mean of each outcome variable for males and females in the neighborhood. Outcome variables are listed vertically. Columns 2, 3, 5, and 6 present estimates from neighborhood fixed-effects regressions for males (columns 2, 3) and females (columns 5, 6) separately. All regressions include student-level controls (baseline performance; indicators for track choices in classics, science, or exact science; age; and an indicator for being born in the first quarter of the calendar year), year fixed effects, neighborhood fixed effects, neighborhood-level leave-out mean of the postcode income, and neighborhood-specific linear time trends. Columns 3 and 6 add neighborhood-by-year characteristics (students' leave-out characteristics averaged by neighborhood and year and neighborhood annual enrollment). Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Standard errors clustered at school level are reported in parentheses.

Table 4: Neighbor Effects while Simultaneously Controlling for the Percentage of Females in the Two Geographic Units (Neighborhood and School)

	Males		Females	
	(1)	(2)	(3)	(4)
Scholastic and Matriculation Outcomes				
National Exam Score	2.907	2.364	3.346	2.485
	(1.554)*	(1.540)	(1.311)**	(1.285)*
N	122,191	122,191	161,539	161,539
Matriculation Status	-0.028	-0.029	0.042	0.031
	(0.030)	(0.031)	(0.025)*	(0.025)
N	122,191	122,191	161,539	161,539
Matriculation Score (std.)	0.342	0.278	0.217	0.161
	(0.085)***	(0.085)***	(0.073)***	(0.072)**
N	97,734	97,734	131,741	131,741
Enroll in an Academic University	0.156	0.144	0.104	0.086
	(0.038)***	(0.039)***	(0.033)***	(0.033)***
N	97,734	97,734	131,741	131,741
Quality of Postsecondary Degree	10.871	8.968	7.395	6.147
	(2.478)***	(2.481)***	(2.083)***	(2.064)***
N	97,734	97,734	131,741	131,741
University Major				
Enroll in Science & Mathematics	0.029	0.035	0.039	0.035
	(0.019)	(0.020)*	(0.014)***	(0.014)**
N	97,734	97,734	131,741	131,741
Enroll in STEM	0.052	0.048	0.048	0.049
	(0.036)	(0.038)	(0.025)*	(0.026)*
N	97,734	97,734	131,741	131,741
Labor Market Outcome				
Expected Wage (std.)	0.090	0.041	0.201	0.184
	(0.069)	(0.070)	(0.065)***	(0.067)***
N	97,734	97,734	131,741	131,741
Year FE & Neighborhood FE	✓	✓	✓	✓
Neighborhood-Specific Linear Trends	✓	✓	✓	✓
Student-Level Controls	✓	✓	✓	✓
Neighborhood-by-Year Characteristics	✓	✓	✓	✓
Proportion of Females in School		✓		✓

Notes: Columns 1 and 3 report the main estimates of gender neighbor effects (the same as in columns 3 and 6 in Table 3). In columns 2 and 4, we add controls for the proportion of females in the school each student attends. All regressions include student-level controls (baseline performance; indicators for track choices in classics, science, or exact science; age; and an indicator for being born in the first quarter of the calendar year), neighborhood-by-year characteristics (students' leave-out characteristics averaged by neighborhood and year and neighborhood annual enrollment), year indicators, neighborhood-level leave-out mean of the postcode income, neighborhood fixed effects, and neighborhood-specific linear time trends. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Standard errors clustered at school level are reported in parentheses.

Table 5: Nonlinear Effects of Proportion of Females in the Neighborhood on School Outcomes and Choice of University Major

	Males				Females			
	Quintile 2 (1)	Quintile 3 (2)	Quintile 4 (3)	Quintile 5 (4)	Quintile 2 (5)	Quintile 3 (6)	Quintile 4 (7)	Quintile 5 (8)
Scholastic and Matriculation Outcomes								
National Exam Score	-0.125	0.412	0.368	0.392	-0.027	0.528	0.634	0.467
	(0.250)	(0.292)	(0.311)	(0.277)	(0.226)	(0.239)**	(0.266)**	(0.231)**
	[0.673]	[0.208]	[0.297]	[0.257]	[0.901]	[0.080]	[0.030]	[0.099]
Matriculation Status	-0.002	0.000	-0.001	-0.006	0.000	0.009	0.010	0.005
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)**	(0.005)**	(0.004)
	[0.663]	[0.970]	[0.901]	[0.317]	[0.970]	[0.060]	[0.089]	[0.178]
Matriculation Score (std.)	-0.007	0.041	0.049	0.055	-0.005	0.032	0.046	0.035
	(0.012)	(0.014)***	(0.015)***	(0.015)***	(0.012)	(0.012)***	(0.014)***	(0.013)***
	[0.604]	[0.010]	[0.010]	[0.010]	[0.693]	[0.020]	[0.010]	[0.020]
Enroll in an Academic University	0.002	0.018	0.027	0.026	0.001	0.013	0.020	0.018
	(0.006)	(0.006)***	(0.007)***	(0.007)***	(0.005)	(0.006)**	(0.006)***	(0.006)***
	[0.753]	[0.020]	[0.010]	[0.010]	[0.832]	[0.040]	[0.010]	[0.020]
Quality of Postsecondary Degree	-0.237	1.127	1.345	1.729	-0.106	0.817	1.301	1.088
	(0.355)	(0.393)***	(0.432)***	(0.420)***	(0.343)	(0.354)**	(0.396)***	(0.367)***
	[0.446]	[0.030]	[0.020]	[0.010]	[0.753]	[0.020]	[0.010]	[0.010]
University Major								
Enroll in Science & Mathematics	-0.000	0.005	0.006	0.003	0.006	0.009	0.006	0.008
	(0.003)	(0.003)	(0.003)*	(0.003)	(0.002)***	(0.002)***	(0.003)**	(0.002)***
	[0.990]	[0.168]	[0.080]	[0.317]	[0.010]	[0.010]	[0.020]	[0.010]
Enroll in STEM	0.001	-0.000	0.002	0.008	0.010	0.004	0.006	0.010
	(0.006)	(0.006)	(0.006)	(0.006)	(0.004)***	(0.004)	(0.004)	(0.004)**
	[0.842]	[0.951]	[0.683]	[0.238]	[0.020]	[0.356]	[0.267]	[0.089]
Labor Market Outcome								
Expected Wages (std.)	-0.008	0.008	-0.004	0.010	0.019	0.028	0.024	0.040
	(0.011)	(0.012)	(0.012)	(0.012)	(0.010)*	(0.010)***	(0.011)**	(0.011)***
	[0.436]	[0.505]	[0.743]	[0.495]	[0.040]	[0.010]	[0.040]	[0.010]
Neighborhood FE. & Year FE.	✓	✓	✓	✓	✓	✓	✓	✓
Neighborhood-Specific Linear Trends	✓	✓	✓	✓	✓	✓	✓	✓
Student-Level Controls	✓	✓	✓	✓	✓	✓	✓	✓
Student Characteristics	✓	✓	✓	✓	✓	✓	✓	✓
Neighborhood-By-Year Characteristics	✓	✓	✓	✓	✓	✓	✓	✓

Notes: This table reports nonlinear effects of the proportion of female students in the neighborhood on males' and females' academic outcomes and choices. The model replaces the single treatment variable with a set of quintile indicators for different quintiles for the proportion of female students in the neighborhood. The omitted category is quintile 1, where the share of females is between 0.282 and 0.532. Estimates in each row by gender are generated from the same regression. We control for the neighborhood-level leave-out mean of the postcode income. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Standard errors clustered at the school level are reported in parentheses. Tests of statistical significance were simultaneously performed using the Romano and Wolf (2005) procedure. Retrieved p-values are reported in brackets.

Table 6: Falsification Exercise, False Peers in Space

	False Peers in Neighborhood	
	Males	Females
	(1)	(2)
Scholastic and Matriculation Outcomes		
National Exam Score	-0.663	3.747
	(7.653)	(6.765)
N	121,770	160,823
Matriculation Status	-0.075	-0.179
	(0.161)	(0.137)
N	121,770	160,823
Matriculation Score (std.)	-0.663	-0.112
	(0.402)*	(0.365)
N	97,377	131,156
Enroll in an Academic University	-0.272	-0.010
	(0.193)	(0.158)
N	97,377	131,156
Quality of Postsecondary Degree	-17.107	-7.736
	(11.479)	(10.319)
N	97,377	131,156
University Major		
Enroll in Science & Mathematics	-0.164	0.090
	(0.101)	(0.092)
N	97,377	131,156
Enroll in STEM	-0.496	-0.102
	(0.193)**	(0.142)
N	97,377	131,156
Labor Market Outcome		
Expected Wage (std.)	0.044	-0.194
	(0.405)	(0.340)
N	97,377	131,156
Year FE & Neighborhood FE	✓	✓
Neighborhood-Specific Linear Trends	✓	✓
Student-Level Controls	✓	✓
Neighborhood-by-Year Characteristics	✓	✓

Notes: We replace the actual variable of interest (proportion of female students in neighborhood u) with the proportion of females in all other neighborhoods in a student's periphery, except their own neighborhood. Students are unlikely to interact with students in this broader geographic group. All regressions control for year fixed effects, neighborhood fixed effects, neighborhood-specific linear time trends, student-level controls (baseline performance; indicators for track choices in classics, science, or exact science; age; indicators for being born in the first quarter of the calendar year), and neighborhood-by-year characteristics. We also control for the neighborhood-level leave-out mean of the postcode income. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Standard errors clustered at school level are reported in parentheses.

Table 7: Falsification Exercise, False Peers in Time

	False Peers in Neighborhood			
	in cohort t-1		in cohort t+1	
	Males (1)	Girls (2)	Males (3)	Females (4)
Scholastic and Matriculation Outcomes				
National Exam Score	-2.114	0.026	1.025	0.935
	(2.592)	(2.163)	(2.634)	(2.186)
Matriculation Status	-0.090	0.019	0.020	-0.026
	(0.061)	(0.050)	(0.055)	(0.045)
Matriculation Score (std.)	-0.129	-0.116	0.008	0.210
	(0.152)	(0.119)	(0.139)	(0.113)
Enroll in an Academic University	-0.128	-0.081	-0.039	0.023
	(0.073)*	(0.055)	(0.069)	(0.057)
Quality of Postsecondary Degree	-3.421	-2.755	-1.622	1.559
	(4.397)	(3.418)	(4.043)	(3.319)
University Major				
Enroll in Science & Mathematics	0.027	-0.025	-0.025	-0.025
	(0.040)	(0.029)	(0.035)	(0.025)
Enroll in STEM	0.061	-0.051	0.014	0.064
	(0.070)	(0.045)	(0.067)	(0.044)
Labor Market Outcomes				
Expected Wage (std.)	0.132	-0.191	-0.008	0.017
	(0.134)	(0.125)	(0.132)	(0.114)
Year FE & Neighborhood FE	✓	✓	✓	✓
Neighborhood-Specific Linear Trends	✓	✓	✓	✓
Student-Level Controls	✓	✓	✓	✓
Neighborhood-by-Year Characteristics	✓	✓	✓	✓

Notes: We replace the actual variable of interest (proportion of female students in year t) with the proportion of female students in the younger (t-1) or older (t+1) cohort within the same neighborhood. All regressions control for year fixed effects, neighborhood fixed effects, neighborhood-specific linear time trends, student-level controls (baseline performance; indicators for track choices in classics, science, or exact science; age; indicators for being born in the first quarter of the calendar year), and neighborhood-by-year characteristics. We also control for the neighborhood-level leave-out mean of the postcode income. Each estimate is generated from a different regression. *, **, and *** denotes significance at the 10%, 5%, and 1% level, respectively. Standard errors are clustered at the school level.

Table 8: Heterogeneous Effects of Neighbors on Academic Outcomes and Decisions by Neighborhood Size

	Males		Females	
	Neighborhood Density		Neighborhood Density	
	Low	High	Low	High
	(1)	(2)	(3)	(4)
Scholastic and Matriculation Outcomes				
National Exam Score	0.617	4.160	-0.452	10.519
	(1.435)	(3.828)	(1.278)	(3.391)***
N	45,300	76,891	60,398	101,141
Matriculation Status	-0.059	-0.005	-0.029	0.197
	(0.033)*	(0.068)	(0.026)	(0.066)***
N	45,300	76,891	60,398	101,141
Matriculation Score (std.)	0.097	0.721	-0.026	0.707
	(0.086)	(0.174)***	(0.081)	(0.156)***
N	36,227	61,507	48,962	82,779
Enroll in an Academic University	0.056	0.360	0.037	0.207
	(0.042)	(0.085)***	(0.036)	(0.067)***
N	36,227	61,507	48,962	82,779
Quality of Postsecondary Degree	4.048	20.992	1.624	17.572
	(2.506)	(5.140)***	(2.292)	(4.363)***
N	36,227	61,507	48,962	82,779
University Major				
Enroll in Science & Mathematics	0.026	0.003	0.022	0.103
	(0.022)	(0.041)	(0.015)	(0.036)***
N	36,227	61,507	48,962	82,779
Enroll in STEM	0.059	-0.026	0.036	0.087
	(0.041)	(0.084)	(0.028)	(0.055)
N	36,227	61,507	48,962	82,779
Labor Market Outcome				
Expected Wage (std.)	0.086	0.009	0.115	0.451
	(0.078)	(0.158)	(0.074)	(0.137)***
N	36,227	61,507	48,962	82,779
Year FE & Neighborhood FE	✓	✓	✓	✓
Neighborhood-Specific Linear Trends	✓	✓	✓	✓
Student-Level Controls	✓	✓	✓	✓
Neighborhood-by-Year Characteristics	✓	✓	✓	✓

Notes: The table reports estimates of the proportion of females by neighborhood density on males' and females' academic outcomes and decisions. Results are reported separately for neighborhoods that have an enrollment size below (low density) and above the average (high density), for a given geographic size. Neighborhood average enrollment size is 162 students. All regressions control for year fixed effects, neighborhood fixed effects, neighborhood-specific linear time trends, student-level controls (baseline performance; indicators for track choices in classics, science or exact science; age; indicators for being born in the first quarter of the calendar year), and neighborhood-by-year characteristics. We also control for the neighborhood-level leave-out mean of the postcode income. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Standard errors are clustered at school level.

Table 9: Robustness Check: Controls for the Existence of Single-Sex and Private Schools

	Males		Females	
	(1)	(2)	(3)	(4)
Scholastic and Matriculation Outcomes				
National Exam Score	2.907	2.937	3.346	3.378
	(1.554)*	(1.558)*	(1.311)**	(1.313)**
N	122,191	122,191	161,539	161,539
Matriculation Status	-0.028	-0.025	0.042	0.044
	(0.030)	(0.030)	(0.025)*	(0.025)*
N	122,191	122,191	161,539	161,539
Matriculation Score (std.)	0.342	0.334	0.217	0.208
	(0.085)***	(0.085)***	(0.073)***	(0.073)***
N	97,734	97,734	131,741	131,741
Enroll in an Academic University	0.156	0.153	0.104	0.104
	(0.038)***	(0.038)***	(0.033)***	(0.033)***
N	97,734	97,734	131,741	131,741
Quality of Enrolled Postsecondary Degree	10.871	10.626	7.395	7.084
	(2.478)***	(2.481)***	(2.083)***	(2.082)***
N	97,734	97,734	131,741	131,741
University Major				
Enroll in Science & Mathematics	0.029	0.028	0.039	0.039
	(0.019)	(0.019)	(0.014)***	(0.014)***
N	97,734	97,734	131,741	131,741
Enroll in STEM	0.052	0.048	0.048	0.048
	(0.036)	(0.036)	(0.025)*	(0.025)*
N	97,734	97,734	131,741	131,741
Labor Market Outcome				
Expected Wage (std.)	0.090	0.084	0.201	0.193
	(0.069)	(0.069)	(0.065)***	(0.065)**
N	97,734	97,734	131,741	131,741
Year FE & Neighborhood FE	✓	✓	✓	✓
Neighborhood-Specific Linear Trends	✓	✓	✓	✓
Student-Level Controls	✓	✓	✓	✓
Neighborhood-by-Year Characteristics	✓	✓	✓	✓
No. of Males and Females Attending Private or Single-Sex Schools		✓		✓

Notes: In columns 2 and 4 we control for the number of males and females who attend single-sex or private schools in the neighborhood. We also report estimates from the main analysis (columns 1 and 3) for comparison purposes. All regressions control for year fixed effects, neighborhood fixed effects, neighborhood-specific linear time trends, student-level controls (baseline performance; indicators for track choices in classics, science, or exact science; age; indicators for being born in the first quarter of the calendar year), and neighborhood-by-year characteristics. We also control for the neighborhood-level leave-out mean of the postcode income. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Standard errors are clustered at school level.

Table 10: Robustness Exercise: Controls for Proportion of Females in Previous and Following Cohorts

	Males			Females		
	t-1	t	t+1	t-1	t	t+1
	(1)	(2)	(3)	(4)	(5)	(6)
Scholastic and Matriculation Outcomes						
National Exam Score	-1.262	2.464	0.677	2.306	5.392	3.393
	(3.518)	(2.651)	(3.564)	(2.912)	(2.141)**	(2.823)
Matriculation Status	-0.099	-0.056	-0.048	0.086	0.089	-0.008
	(0.087)	(0.052)	(0.077)	(0.068)	(0.041)**	(0.057)
Matriculation Score (std.)	0.032	0.331	0.163	-0.081	0.318	0.192
	(0.203)	(0.143)***	(0.191)	(0.172)	(0.119)***	(0.161)
Enroll in an Academic University	-0.102	0.113	0.003	-0.108	0.126	0.097
	(0.098)	(0.058)**	(0.092)	(0.081)	(0.052)**	(0.075)
Quality of Postsecondary Degree	0.818	10.004	3.312	-1.594	9.629	6.115
	(5.822)	(4.114)**	(5.508)	(4.983)	(3.315)***	(4.660)
University Major						
Enroll in Science & Mathematics	0.086	0.059	-0.012	-0.045	0.038	-0.019
	(0.052)	(0.029)**	(0.046)	(0.042)	(0.021)*	(0.034)
Enroll in STEM	0.194	0.126	0.053	0.003	0.058	0.120
	(0.093)**	(0.054)**	(0.086)	(0.063)	(0.035)*	(0.058)**
Labor Market Outcome						
Expected Wage (std.)	0.298	0.024	0.028	0.039	0.251	0.106
	(0.175)*	(0.113)	(0.146)	(0.159)	(0.100)**	(0.146)
Year FE & Neighborhood FE	✓	✓	✓	✓	✓	✓
Neighborhood-Specific Linear Trends	✓	✓	✓	✓	✓	✓
Student-Level Controls	✓	✓	✓	✓	✓	✓
Neighborhood-by-Year Characteristics	✓	✓	✓	✓	✓	✓

Notes: This table reports estimates of the proportion of females in a neighborhood in cohort t, cohort t-1, and cohort t+1 on a variety of outcomes for boys (columns 1-3) and females (columns 4-6) separately. In all regressions we simultaneously control for the current, lagged, and lead proportion of females in the same neighborhood. All regressions control for year fixed effects, neighborhood fixed effects, neighborhood-specific linear time trends, student-level controls (baseline performance; indicators for track choices in classics, science, or exact science; age; indicators for being born in the first quarter of the calendar year), and neighborhood-by-year characteristics. We also control for the neighborhood-level leave-out mean of the postcode income. *, **, and *** denotes significance at the 10%, 5%, and 1% level, respectively. Standard errors are clustered at school level.

Table 11: Robustness Check: Simultaneous Controls for Neighborhood-by-Cohort and School-by-Cohort peer controls in the Neighborhood Regression

	Males	Females
	(1)	(2)
Scholastic and Matriculation Outcomes		
National Exam Score	2.618	3.075
	(1.523)*	(1.302)**
N	122,191	161,539
Matriculation Status	-0.031	0.038
	(0.030)	(0.025)
N	122,191	161,539
Matriculation Score (std.)	0.333	0.208
	(0.084)***	(0.072)***
N	97,734	131,741
Enroll in an Academic University	0.154	0.101
	(0.037)***	(0.033)***
N	97,734	131,741
Quality of Enrolled Postsecondary Degree	10.612	7.129
	(2.462)***	(2.076)***
N	97,734	131,741
University Major		
Enroll in Science & Mathematics	0.028	0.037
	(0.019)	(0.014)***
N	97,734	131,741
Enroll in STEM	0.050	0.046
	(0.036)	(0.025)*
N	97,734	131,741
Labor Market Outcome		
Expected Wage (std.)	0.084	0.197
	(0.069)	(0.065)***
N	97,734	131,741
Year FE & Neighborhood FE	✓	✓
Neighborhood-Specific Linear Trends	✓	✓
Student-Level Controls	✓	✓
Neighborhood-by-Year Characteristics	✓	✓
School-by-Year Characteristics	✓	✓

Notes: This table shows the results when we simultaneously control for year-specific peer characteristics at school and neighborhood level. All regressions also control for year fixed effects, neighborhood fixed effects, neighborhood-specific linear time trends, student-level controls (baseline performance; indicators for track choices in classics, science, or exact science; age; indicators for being born in the first quarter of the calendar year), and neighborhood-by-year characteristics. We also control for the neighborhood-level leave-out mean of the postcode income. Standard errors are clustered at school level. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table 12: Robustness Check: Different Definitions of Neighborhoods

	Males			Females		
	No Drop	Drop Farthest	Drop Closest	No Drop	Drop Farthest	Drop Closest
	(1)	(2)	(3)	(4)	(5)	(6)
Scholastic and Matriculation Outcomes						
National Exam Score	3.227	4.541**	2.435	4.988**	6.001***	5.185**
	(2.350)	(2.240)	(2.261)	(1.984)	(2.001)	(2.040)
N	97,015	82,521	74,557	128,701	108,745	98,352
Matriculation Status	-0.030	-0.010	-0.053	0.078**	0.116***	0.091**
	(0.046)	(0.043)	(0.042)	(0.038)	(0.038)	(0.040)
N	97,015	82,521	74,557	128,701	108,745	98,352
Matriculation Score (std.)	0.481***	0.517***	0.463***	0.336***	0.405***	0.292***
	(0.119)	(0.113)	(0.112)	(0.102)	(0.098)	(0.098)
N	77,747	66,346	59,673	105,232	89,344	80,440
Enrolled in an Academic University	0.224***	0.183***	0.198***	0.149***	0.114**	0.125***
	(0.058)	(0.053)	(0.053)	(0.046)	(0.047)	(0.045)
N	77,747	66,346	59,673	105,232	89,344	80,440
Quality of Enrolled Postsecondary Degree	14.527***	14.087***	14.997***	9.472***	11.936***	7.334***
	(3.400)	(3.217)	(3.241)	(2.920)	(2.788)	(2.827)
N	77,747	66,346	59,673	105,232	89,344	80,440
University Major						
Enroll in Science & Mathematics	0.024	0.021	0.016	0.060***	0.057***	0.023
	(0.025)	(0.025)	(0.026)	(0.022)	(0.020)	(0.019)
N	77,747	66,346	59,673	105,232	89,344	80,440
Enroll in STEM	0.069	0.011	0.088*	0.068*	0.072**	0.009
	(0.056)	(0.050)	(0.051)	(0.036)	(0.035)	(0.033)
N	77,747	66,346	59,673	105,232	89,344	80,440
Labor Market Outcome						
Expected Wage (std.)	-0.008	0.097	-0.012	0.325***	0.303***	0.189**
	(0.100)	(0.093)	(0.094)	(0.094)	(0.090)	(0.078)
N	77,747	66,346	59,673	105,232	89,344	80,440
Year FE & Neighborhood FE	✓	✓	✓	✓	✓	✓
Neighborhood-Specific Linear Trends	✓	✓	✓	✓	✓	✓
Student-Level Controls	✓	✓	✓	✓	✓	✓
Neighborhood-by-Year Characteristics	✓	✓	✓	✓	✓	✓

Notes: This table replicates the main results when we use different definitions of neighborhoods. In particular, we drop the school with the highest or lowest distance to any other schools within the initial neighborhoods and then re-construct the female share of neighbors in the newly constructed neighborhoods. To conduct this exercise, we focus on neighborhoods that have at least 3 schools. Columns 1 and 3 show the estimated effects when we do not drop any schools (main results). Columns 2 and 5 show the estimated effects when we drop the school with the highest average distance to any other schools (i.e., the school that is farthest from other schools) within the neighborhood. Columns 3 and 6 show the estimated effects when we drop the school with the smallest average distance to other schools (i.e., the school that is closest to other schools) within the neighborhood. All regressions control for year fixed effects, neighborhood fixed effects, neighborhood-specific linear time trends, student-level controls (baseline performance; indicators for track choices in classics, science, or exact science; age; indicators for being born in the first quarter of the calendar year), and neighborhood-by-year characteristics. We also control for the neighborhood-level leave-out mean of the postcode income. Standard errors are clustered at school level. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Online Appendix A
Not For Publication

Table A.1: Descriptive Statistics by Year

Year	No. of Schools	No. of Neigh/hoods	Proportion of Females in School (sd.)	Proportion of Females in Neigh/hood (sd.)	National Exam Score	
					Males	Females
	(1)	(2)	(3)	(4)	(5)	(6)
2003-2004	864	213	0.557 (0.077)	0.558 (0.051)	59.0	60.7
2004-2005	877	213	0.557 (0.074)	0.558 (0.052)	57.2	59.8
2005-2006	858	210	0.562 (0.073)	0.563 (0.050)	59.4	62.7
2006-2007	809	201	0.586 (0.082)	0.588 (0.055)	69.6	71.2
2007-2008	787	198	0.586 (0.081)	0.590 (0.057)	70.2	73.1
2008-2009	784	198	0.578 (0.087)	0.582 (0.062)	72.8	74.9
All	940	222	0.569 (0.080)	0.571 (0.056)	63.6	66.3

Year	Matriculation Status		Matriculation Score		Enrolled in University vs Technical School		Expected Wage (std.)	
	Males	Females	Males	Females	Males	Females	Males	Females
	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
2003-2004	0.804	0.815	-0.060	0.060	0.502	0.569	0.082	-0.111
2004-2005	0.801	0.823	-0.084	0.073	0.496	0.552	0.059	-0.112
2005-2006	0.635	0.682	-0.061	0.062	0.602	0.669	0.200	-0.062
2006-2007	0.843	0.825	-0.079	0.073	0.595	0.649	0.199	-0.072
2007-2008	0.869	0.876	-0.128	0.106	0.652	0.702	0.174	-0.116
2008-2009	0.913	0.911	-0.093	0.083	0.700	0.713	0.146	-0.119
All	0.800	0.816	-0.083	0.076	0.582	0.637	0.136	-0.100

Notes: This table presents descriptive statistics of the variable of interest (i.e., proportion of females), the number of neighborhoods/schools by year. We show statistics for each of the outcome variables by year and gender.

Table A.2: Variance Decomposition

	Neighborhood			**School**		
	Sum of Squares	Share of Total	DF	Sum of Squares	Share of Total	DF
	(1)	(2)	(3)	(4)	(5)	(6)
Within Neighborhood/School	2.99	63%	1,011	35.97	69%	4,731
Between Neighborhood/School	1.74	37%	221	15.97	31%	1,096
Total	4.74		1,397	51.94		5,827

Notes: The variance decomposition of the proportion of female students is done at the neighborhood (columns 1-3) and school level (columns 4-6). We show the sum of squares, the share of total and the degrees of freedom (DF) for each unit.

Table A.3: Estimated Effect of Proportion of Females in the School on School Outcomes and Choice of University Major

	Mean	Males			Mean	Females		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Scholastic and Matriculation Outcomes								
National Exam Score	63.554	3.056	3.070	3.069	66.267	3.255	3.279	3.263
		(0.989)***	(0.987)***	(0.987)***		(0.849)***	(0.847)***	(0.846)***
Matriculation Status	0.800	0.045	0.045	0.045	0.816	0.033	0.034	0.034
		(0.023)*	(0.023)**	(0.023)*		(0.018)*	(0.018)*	(0.018)*
Matriculation Score (std.)	-0.083	0.159	0.159	0.159	0.076	0.180	0.182	0.180
		(0.057)***	(0.057)***	(0.057)***		(0.052)***	(0.052)***	(0.052)***
Enrolled in an Academic University	0.582	0.071	0.070	0.071	0.637	0.075	0.076	0.075
		(0.029)**	(0.029)**	(0.029)**		(0.025)***	(0.024)***	(0.024)***
Quality of Enrolled Postsecondary Degree	47.541	4.984	4.975	4.979	51.923	6.304	6.357	6.315
		(1.591)***	(1.593)***	(1.592)***		(1.500)***	(1.499)***	(1.496)***
University Major								
Enroll in Science & Mathematics	0.083	0.050	0.050	0.050	0.076	0.016	0.016	0.017
		(0.015)***	(0.015)***	(0.015)***		(0.012)	(0.012)	(0.012)
Enroll in STEM	0.408	0.038	0.038	0.038	0.219	0.034	0.034	0.034
		(0.028)	(0.028)	(0.028)		(0.020)*	(0.020)*	(0.020)*
Labor Market Outcome								
Expected Wage (std.)	0.136	0.014	0.013	0.013	-0.100	0.086	0.087	0.088
		(0.055)	(0.055)	(0.055)		(0.048)*	(0.048)*	(0.048)*
Year FE & School FE		✓	✓	✓		✓	✓	✓
School-Specific Linear Trends		✓	✓	✓		✓	✓	✓
Student-Level Controls		✓	✓	✓		✓	✓	✓
School-by-Year Characteristics			✓	✓			✓	✓
Proportion of Females in Neighborhood				✓				✓

Notes: Columns 1 and 5 report the mean of each outcome variable for males and females in the school. Columns 2, 3, 4, 6, 7, and 8 present estimates from school fixed-effects regressions for males (columns 2, 3, 4) and females (columns 6, 7, 8) separately. All regressions include student-level controls (baseline performance; indicators for track choices in classics, science, or exact science; age; and an indicator for being born in the first quarter of the calendar year), year fixed effects, school fixed effects, school-level leave-out mean of the postcode income, and school-specific linear time trends. Columns 3 and 7 add school-by-year characteristics (students' leave-out characteristics averaged by neighborhood and year and neighborhood annual enrollment). Columns 4 and 8 also control for the proportion of females in the neighborhood. Each estimate is generated from a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Standard errors clustered at school level are reported in parentheses.

Table A.4: Transition of the Proportion of Females between Quintiles

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Range:	0.282-0.532	0.532-0.561	0.561-0.581	0.581-0.607	0.608-0.917
Mean:	0.497	0.548	0.571	0.592	0.648
Quintile 1	2	80	53	54	73
Quintile 2		0	54	57	65
Quintile 3			0	54	58
Quintile 4				0	61
Quintile 5					1

Notes: This table presents the frequency of neighborhoods that hit different pairs of quintiles in the sampled years. Quintiles are formed based on the distribution of the proportion of females across the entire sample. The range and mean proportion of females in each quintile are reported in the table. Each neighborhood may have multiple transitions between quintiles in the sample years. The diagonal reports the number of neighborhoods with no transition. Only 3 neighborhoods remain in the same quintile of the proportion of females during the whole sample period (2 in quintile 1 and 1 in quintile 5). 219 neighborhoods have hit more than one quintile.

Table A.5: Sample Study and Population

	Sample (134 Schools)	**Population** (Remaining Schools)	**Difference**
	Mean/(s.d.)	**Mean/(s.d.)**	**b/(s.e.)**
	(1)	(2)	(3)
Female	0.571	0.580	-0.010
	(0.057)	(0.070)	(0.006)
Age	17.914	17.905	0.009
	(0.071)	(0.128)	(0.011)
Born in First Quarter of Birth Year	0.158	0.162	-0.004
	(0.035)	(0.036)	(0.003)

Notes: Column 1 presents descriptive statistics for the sample of 134 schools and column 2 presents the same descriptive statistics for the remaining public schools in the country. Column 3 shows the differences between columns 1 and 2 and the corresponding standard errors.

Table A.6: Dropout Rates, Intakes, and Proportion of Females

	Males	Females
	(1)	(2)
Panel A		
Outcome mean	0.028	0.014
Left the school in year *t*	-0.032	-0.003
	(0.026)	(0.019)
N	24,114	29,399
Panel B		
Outcome mean	0.074	0.063
Enrolled in school in year *t+1*	-0.007	-0.004
	(0.066)	(0.065)
N	24,114	29,399

Notes: We use data from 134 public schools to construct this table. We show the mean of the dependent variables by gender and the estimated effects of the proportion of females on the likelihood of dropping out (Panel A) and enrolling in 12^{th} grade (Panel B). All regressions include school fixed effects, year fixed effects, and school linear time trends. Standard errors clustered at school level are reported in parentheses.

Table A.7: Single-sex, Private schools, and Proportion of Females in Public Schools in the area

	No. of Students Attending Single-Sex and Coed. Private Schools in the Area	**No. of Students Attending Coeducational Private Schools in the Area**	**No. of Students Attending Single-Sex Schools in the Area**	**No. of Female Students Attending Single-Sex Schools in the Area**	**No. of Male Students Attending Single-Sex Schools in the Area**
	(1)	(2)	(3)	(4)	(5)
Proportion of Females	-11.665	-10.585	-1.080	-0.740	-0.339
(in public schools)	(10.075)	(9.936)	(1.615)	(0.981)	(1.243)
School FE and Year FE	✓	✓	✓	✓	✓
School-Specific Linear Trends	✓	✓	✓	✓	✓

Notes: We restrict the analysis to areas that have at least one single-sex or one private school. The variable of interest is the proportion of females in public schools in those areas. The outcome variable is the total number of students who attend single-sex or coeducational private schools in an area (column 1), total number of students who attend coeducational private schools in an area (column 2), total number of students who attend single-sex schools in an area (column 3), the number of female students who attend single-sex schools in the area (colum 4), and the number of male students who attend single-sex schools in the area (column 5). We control for the public school's annual enrollment size in all regressions. The number of observations is 293 and refers to school-year combinations. All regressions include school fixed effects, year effects, and school linear time trends. Standard errors are clustered at school level.

Table A.8: Summary Statistics for Communal Facilities in Each Neighborhood

	Median	SD	p10	p90
Church	2	1.118	0	3
Library	1	1.466	0	3
Park or Square	2	1.965	0	5
Scout	1	1.109	0	2.5
Sports Field	3	2.702	1	8

Notes: The table presents descriptive statistics (median, standard deviation, the 10^{th} and 90^{th} percentile) for each type of public facility we consider in each neighborhood: (1) Parks or Squares; (2) Sports Fields; (3) Library; (4) Church; and (5) Scout facilities. We also construct an *Overall Intensity* index, which is generated using the first component from a PCA consisting of all communal facilities. This index has a median of -0.205, and standard deviation of 1.673.

Table A.9: Gender Neighbor Effects by Facility Intensity

Panel A	**National Exam Score**				**Matriculation Status**			
	Males		**Females**		**Males**		**Females**	
	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median
Overall Intensity	6.320***	4.639	3.259*	12.206***	0.039	0.055	0.003	0.211***
	(2.236)	(3.390)	(1.898)	(2.799)	(0.056)	(0.057)	(0.042)	(0.055)
N	18,183	66,930	25,739	89,363	18,183	66,930	25,739	89,363
Church	6.178**	5.482*	4.060*	10.627***	0.004	0.089*	0.039	0.159***
	(3.072)	(2.941)	(2.126)	(2.607)	(0.068)	(0.050)	(0.044)	(0.050)
N	26,187	58,926	36,163	78,939	26,187	58,926	36,163	78,939
Library	3.945	5.874**	-1.261	11.933***	0.020	0.051	-0.067	0.181***
	(2.609)	(2.769)	(2.490)	(2.059)	(0.074)	(0.049)	(0.053)	(0.042)
N	8,108	77,005	11,754	103,348	8,108	77,005	11,754	103,348
Park or Square	9.834***	1.438	3.801	9.187***	0.085	0.014	0.017	0.149***
	(2.665)	(2.756)	(2.381)	(2.329)	(0.068)	(0.050)	(0.055)	(0.045)
N	13,771	71,342	19,638	95,464	13,771	71,342	19,638	95,464
Scout	6.185**	5.699**	1.527	12.295***	0.060	0.050	-0.020	0.188***
	(2.411)	(2.838)	(2.147)	(2.278)	(0.064)	(0.050)	(0.051)	(0.045)
N	8,393	76,720	12,301	102,801	8,393	76,720	12,301	102,801
Sports Field	5.554**	6.167**	3.626*	10.865***	0.067	0.055	0.036	0.155***
	(2.646)	(2.999)	(2.137)	(2.563)	(0.066)	(0.050)	(0.049)	(0.048)
N	18,062	67,051	25,403	89,699	18,062	67,051	25,403	89,699

Panel B	**Matriculation Score (std.)**				**Enrolled in an Academic University**			
	Males		**Females**		**Males**		**Females**	
	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median
Overall Intensity	0.428***	0.478***	0.342***	0.647***	0.182***	0.213***	0.175***	0.159**
	(0.153)	(0.164)	(0.114)	(0.127)	(0.068)	(0.069)	(0.050)	(0.065)
N	14,221	53,862	20,643	73,363	14,221	53,862	20,643	73,363
Church	0.520***	0.345**	0.366***	0.540***	0.203**	0.188***	0.189***	0.141**
	(0.172)	(0.164)	(0.132)	(0.118)	(0.080)	(0.061)	(0.050)	(0.062)
N	20,518	47,565	28,996	65,010	20,518	47,565	28,996	65,010
Library	0.203	0.519***	0.065	0.660***	0.176*	0.196***	0.169***	0.166***
	(0.183)	(0.139)	(0.149)	(0.103)	(0.101)	(0.056)	(0.062)	(0.052)
N	6,102	61,981	9,178	84,828	6,102	61,981	9,178	84,828
Park or Square	0.584***	0.269*	0.340**	0.507***	0.214***	0.148**	0.180***	0.136**
	(0.170)	(0.148)	(0.141)	(0.111)	(0.076)	(0.062)	(0.058)	(0.056)
N	10,932	57,151	15,798	78,208	10,932	57,151	15,798	78,208
Scout	0.392**	0.497***	0.248*	0.639***	0.157*	0.217***	0.192***	0.162***
	(0.161)	(0.144)	(0.137)	(0.109)	(0.083)	(0.061)	(0.061)	(0.054)
N	6,386	61,697	9,758	84,248	6,386	61,697	9,758	84,248
Sports Field	0.241	0.481***	0.281**	0.583***	0.102	0.227***	0.126**	0.179***
	(0.162)	(0.167)	(0.125)	(0.121)	(0.080)	(0.061)	(0.052)	(0.060)
N	13,814	54,269	20,095	73,911	13,814	54,269	20,095	73,911

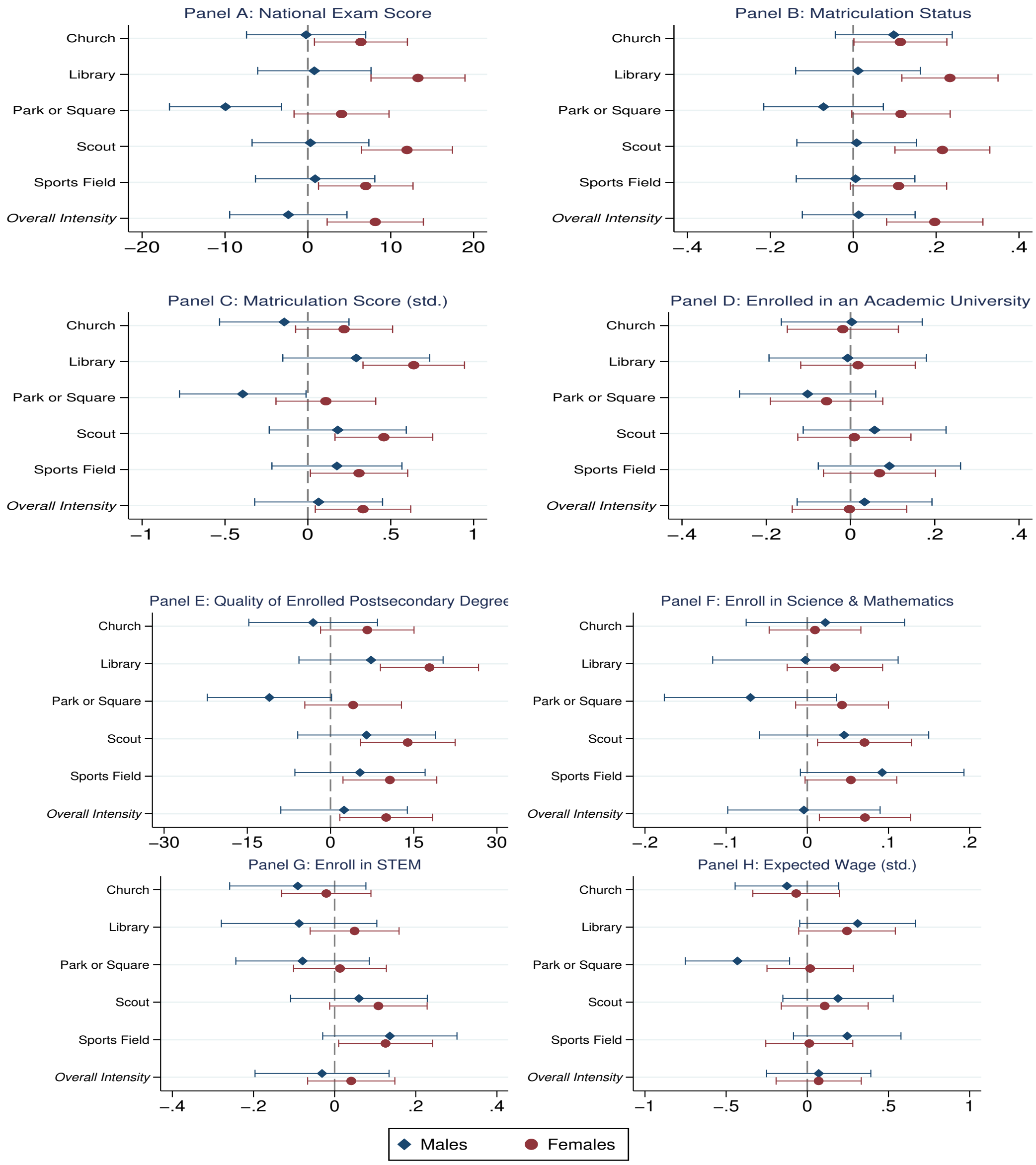
Gender Neighbor Effects by Facility Intensity (continuing)

Panel C	Quality of Enrolled Postsecondary Degree				Enroll in Science & Mathematics			
	Males		Females		Males		Females	
	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median
Overall Intensity	12.665***	14.483***	11.249***	20.624***	0.007	0.004	0.013	0.072***
	(4.638)	(4.719)	(3.184)	(3.803)	(0.046)	(0.032)	(0.022)	(0.027)
N	14,221	53,862	20,643	73,363	14,221	53,862	20,643	73,363
Church	15.172***	10.583**	12.100***	17.289***	-0.001	0.011	0.040	0.040
	(5.064)	(4.808)	(3.603)	(3.640)	(0.049)	(0.033)	(0.025)	(0.024)
N	20,518	47,565	28,996	65,010	20,518	47,565	28,996	65,010
Library	6.976	15.089***	3.991	20.588***	0.017	0.003	0.015	0.047**
	(5.242)	(4.152)	(4.210)	(2.963)	(0.058)	(0.031)	(0.028)	(0.022)
N	6,102	61,981	9,178	84,828	6,102	61,981	9,178	84,828
Park or Square	17.362***	8.465*	11.178***	16.632***	0.053	-0.024	0.009	0.057**
	(4.875)	(4.428)	(3.974)	(3.312)	(0.055)	(0.031)	(0.027)	(0.023)
N	10,932	57,151	15,798	78,208	10,932	57,151	15,798	78,208
Scout	10.438**	15.284***	8.203**	20.450***	-0.025	0.019	0.009	0.063***
	(4.924)	(4.204)	(3.909)	(3.131)	(0.054)	(0.032)	(0.028)	(0.023)
N	6,386	61,697	9,758	84,248	6,386	61,697	9,758	84,248
Sports Field	7.530	14.394***	8.563**	19.278***	-0.042	0.036	0.011	0.061**
	(4.835)	(4.887)	(3.482)	(3.665)	(0.054)	(0.029)	(0.025)	(0.024)
N	13,814	54,269	20,095	73,911	13,814	54,269	20,095	73,911

Panel D	Enroll in STEM				Expected Wage (std.)			
	Males		Females		Males		Females	
	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median
Overall Intensity	0.103	0.043	0.011	0.051	0.200	0.174	0.238**	0.289**
	(0.070)	(0.070)	(0.049)	(0.044)	(0.142)	(0.130)	(0.106)	(0.119)
N	14,221	53,862	20,643	73,363	14,221	53,862	20,643	73,363
Church	0.127*	0.016	0.048	0.016	0.241*	0.083	0.318***	0.204*
	(0.077)	(0.067)	(0.051)	(0.044)	(0.144)	(0.137)	(0.122)	(0.105)
N	20,518	47,565	28,996	65,010	20,518	47,565	28,996	65,010
Library	0.141	0.035	-0.005	0.041	-0.025	0.245**	0.093	0.333***
	(0.093)	(0.058)	(0.050)	(0.043)	(0.178)	(0.112)	(0.151)	(0.093)
N	6,102	61,981	9,178	84,828	6,102	61,981	9,178	84,828
Park or Square	0.138*	0.031	0.020	0.032	0.449***	-0.008	0.225*	0.247**
	(0.075)	(0.067)	(0.057)	(0.041)	(0.150)	(0.126)	(0.117)	(0.107)
N	10,932	57,151	15,798	78,208	10,932	57,151	15,798	78,208
Scout	0.065	0.081	-0.037	0.068*	0.123	0.218*	0.156	0.299***
	(0.081)	(0.064)	(0.065)	(0.038)	(0.169)	(0.119)	(0.126)	(0.103)
N	6,386	61,697	9,758	84,248	6,386	61,697	9,758	84,248
Sports Field	0.017	0.119*	-0.037	0.086**	-0.012	0.245*	0.244**	0.256**
	(0.075)	(0.069)	(0.060)	(0.038)	(0.154)	(0.136)	(0.121)	(0.106)
N	13,814	54,269	20,095	73,911	13,814	54,269	20,095	73,911

Notes: The table reports estimated effects of the proportion of females in the neighborhood by the intensity of each facility in the neighborhood. Results are reported separately for neighborhoods that have a facility intensity below the median (low density) and above the median (high density) and by gender. The top section in each panel shows the estimated effects of the proportion of female neighbors for the *Overall Intensity* of all facilities generated by the PCA. All regressions include controls for year fixed effects, neighborhood fixed effects, neighborhood-specific linear time trends, student-level controls (baseline performance; indicators for track choices in classics, science, or exact science; age; indicators for being born in the first quarter of the calendar year), neighborhood-by-year characteristics, and neighborhood-level leave-out mean of the postcode income. Each estimate is generated by a different regression. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Standard errors are clustered at school level.

Figure A.1: Estimated Neighbor Gender Effects on Outcomes by Density of Each Communal Facility



Notes: Each figure plots the estimated effects of the proportion of female neighbors by the intensity of communal facilities on the scholastic or university decision outcomes for males and females, separately. We estimate those effects by using specification (2) and adding an interaction term between the proportion of female neighbors and a binary indicator for whether this neighborhood has a higher than average intensity of each communal facility. We consider the following facilities: churches, libraries, parks or squares, Scouts, and sports fields. *Overall Intensity* is an index generated using the first component from the PCA that consists of all communal facilities. Coefficient bars represent 90% confidence interval.

Figure A.2:
INTERACTIONS BETWEEN NEIGHBORS: THE CASE OF A SCHOOL COMPLEX

Figure A.3:
INTERACTIONS BETWEEN NEIGHBORS: THE CASE OF A SMALLER SCHOOL COMPLEX

Appendix B

Additional Evidence of Non-sorting into Schools

To lend additional credibility to the validity of our identification strategy, we empirically investigate whether there is potential sorting of students across schools based on gender composition or school quality. In particular, we examine whether a school's subsequent enrollment is likely to change in response to changes in the current gender composition or school quality, and run the following regressions:

$$Enrollment_{s,t+1} = \alpha_2 + \zeta_s + \beta_t + +\gamma_s year + \delta_1 Prop.Females_{s,t} + \theta_2 Enrollment_{s,t}$$
$$+NeighborhoodCharacteristics_t + e_{s,t+1} \qquad (2)$$

$$Enrollment_{s,t+1} = \alpha_1 + \lambda_s + \kappa_t + +\gamma_s year + \delta_2 SchoolQualityRank_{s,t} + \theta_1 Enrollment_{s,t}$$
$$+NeighborhoodCharacteristics_t + u_{s,t+1} \qquad (3)$$

where s denotes schools and t denotes time. $Enrollment_{s,t+1}$ is the enrollment size of school s in time *t+1*. The coefficients of interest δ_1 and δ_2 indicate whether a school's subsequent enrollement is affected by the current proportion of females or a school's current performance ranking, respectively.[43] If students respond to changes in gender composition or school rankings, we would expect δ_1 and δ_2 to be statistically different from zero.

Table B.1 reports estimates of the share of females in school s in year t (columns 1 and 2) and a school's quality (columns 3 and 4) on subsequent enrollment. All estimates appear to be statistically insignificant, regardless of whether we control for year and school fixed effects (columns 2 and 4). These results suggest that changes in a school's proportion of females or a school's quality are uncorrelated with changes in subsequent enrollment. Therefore, changes in the gender composition of schools and neighborhoods are unlikely to be driven by the sorting of families based on prior school quality or school gender composition. In Section 8.2, we also show that changes in the share of female schoolmates are also uncorrelated with changes in dropout and intake rates.

[43] To construct a measure for school quality, we rank schools based on their average performance on the national exams across all nationally tested subjects. The school-ranking variable takes values from 0 to 100 and is increasing in performance, where the highest performing school has a ranking equal to 100 and the lowest performing school a ranking equal to 0.

Table B.1: Mobility of Students across Schools

Dependent Variable: School's Enrollment in Year t+1	(1)	(2)	(3)	(4)
Proportion of Females in School s in Year t	-3.814	-3.421		
	(2.591)	(2.165)		
School Quality Rank in Year t			-1.073	-0.357
			(1.190)	(1.098)
Current and Last Year's School Enrollment	✓	✓	✓	✓
Neighborhood Characteristics	✓	✓	✓	✓
School-Specific Linear Trends	✓	✓	✓	✓
School FE and Year FE		✓		✓
Observations	3,040	3,040	3,040	3,040

Notes: The table reports estimates for the effect of a school's proportion of females and school quality rank on a school's future enrollment. In all specifications, we control for a school's current and past year's enrollment, neighborhood characteristics, and a school-specific linear trend. *Neighborhood characteristics* include neighborhood average income and neighborhood average unemployment. Specifications in columns 2 and 4 also include school and year fixed effects. *School Quality Rank* in Year t is calculated based on the average school national exam performance across all subjects and takes values from 0 to 1 (increasing in performance); the mean rank equals 0.5. Standard errors are clustered at school level.