Mean and distributional impact of single-sex high schools on students' cognitive achievement, major choice, and test-taking behavior: Evidence from a random assignment policy in Seoul, Korea

Hosung Sohn*

Center for Policy Research, Maxwell School of Citizenship and Public Affairs, Syracuse University, 426 Eggers Hall, Syracuse, NY 13244-1020, USA

A R T I C L E   I N F O

Article history:
Received 25 May 2015
Revised 20 January 2016
Accepted 23 February 2016
Available online 2 March 2016

JEL classification:
I21
J16

Keywords:
Single-sex school
Major choice
Random assignment
Mean and distributional impact

A B S T R A C T

Single-sex schooling has been considered in many countries as a way to promote student achievement. This paper estimates the mean and distributional impact of single-sex high schools on students’ cognitive achievement, major choice, and test-taking behavior—by exploiting the random assignment policy adopted in Seoul, Korea. Based on administrative data for a period of seven years, I find that, on average, the positive effects of single-sex schooling on test scores are small, especially when the parental and teacher sorting are accounted for. Although the magnitude of the estimated effects is small, I find that the effect is relatively larger for students in quantiles 0.5–0.8 of the distribution of test scores. The impact is trivial, on the other hand, for students located at the very bottom and the very top quantiles. Moreover, I do not find any differences, both practically and statistically, in major choice and test-taking behavior.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Parents put much consideration into the choice of educational inputs such as teacher quality insomuch that their decisions affect their child’s academic achievement. Among these inputs, parents are selective in determining their child’s peer group, believing that peer groups are an important factor for child’s academic and non-academic achievements. One of the peer effects that parents and education policy makers pay attention to is the gender peer effect. This is due, in part, to a pervasive view that students (especially female students) learn better when they are with same-sex peers. Accordingly, implementation of single-sex education has been considered in many countries and has also been popular in the United States. As such, credible evidence on the efficacy of single-sex education is necessary.

As demonstrated by Manski (1993), estimating the effect of attending a single-sex school is difficult because, in general, it is an endogenous consequence of individual choice. As such, unobservable attributes that determine the selection of a single-sex school and unobservable factors that affect student performance may exist, introducing bias to estimates of single-sex schooling. For this reason, there is relatively little convincing evidence on the effects of single-sex education despite a long literature on single-sex schooling. The most convincing method for addressing the endogeneity problem is to randomly assign students to single-sex and coeducational schools. Middle school students in Seoul, the capital of South Korea, are randomly assigned to either a single-sex or coeducational high school.
upon graduation. Hence, in this study, I exploit this random assignment policy to analyze the effect of attending single-sex school.

Ideally, one needs information on the initial assignment of students to fully leverage the benefits of a randomized experiment. The data used in this study, however, do not contain this information. As documented by Krueger (1999) and Krueger and Zhu (2004), there may be potential biases associated with not complying with the initial assignment as well as differential sample dropout rates. In the results to follow, I present some facts and statistical tests to argue that these may not be a large source of bias in this study (see Section 5.2 and 5.3). I further note that the data used in the analyses are limited because some information is lacking at the individual level, and these limitations do not allow me to conduct ideal tests of balance in baseline student characteristics. I argue, however, that the baseline student covariates are balanced across the two types of schools by analyzing school-level data that are derived from student-level data.

Note that random assignment of students to either a single-sex or coeducational school solves the endogenous selection of students into either of these schools. And comparing student achievement in single-sex schools with that of coeducational schools will provide an effect of “attending” single-sex schools. This comparison, however, provides less insight unless one speaks to underlying mechanisms. That is, because we are making comparisons across schools, it is important to identify the differences between the two types of schools that might give rise to the estimated effects. Moreover, to draw policy implications from analyzing the effect of single-sex education, it is desirable to derive an effect estimate that is not driven by many kinds of factors—the estimated effects are likely to be driven by differences in school-, teacher-, and student-related characteristics. In Korea, school characteristics (e.g., school resources and curricula) are extremely homogeneous across high schools because the government is strongly committed to establishing homogeneity in these characteristics. On the other hand, the case may not be true for teacher quality. If we believe that teachers may prefer to teach in single-sex schools, it is likely that such schools may have more qualified teachers because they will be the ones, among their competitors, to get the job. Provided that single-sex schools have better teachers, estimated effects on single-sex schooling will reflect both the teacher-related and gender peer-related effects, and we cannot determine the magnitude of the effect of each factor. Hence, it is necessary to control for endogenous sorting of teachers, especially for the case of Korea, so that the estimated effects are not driven by teacher-related factors.

In South Korea, teachers cannot select the public school they wish to work for. Furthermore, neither can a public school choose its teachers. Rather, the Office of Education in each city assigns newly hired teachers to each of their public schools and rotates existing teachers every five years within the school district. As a consequence of this rigid nature of teacher assignment policy, I find that observable teacher characteristics are equally balanced across both types of public schools. Focusing on public high schools, therefore, is quite favorable for solving the bias associated with endogenous teacher sorting. I note, however, that although the rigid nature of assignment policy adopted in public schools will reduce the amount of teacher sorting that takes place, the estimated effect could still be driven by teacher quality in dimensions that are not measured in the data.

Exploiting the aforementioned setting—random assignment of students, the way in which teachers are assigned to a public school, and homogeneity in school characteristics—I compare test scores of students in single-sex public high schools with those of coeducational public high schools and attempt to derive single-sex schooling effects that are not driven by the differences in teacher- and school-related characteristics. It is important to note, however, that the estimated effects may still reflect other unobservable mechanisms other than the above factors.

In this study, because prior research focuses only on mean effects, I delineate distributional impact of single-sex schooling by conducting an unconditional quantile regression method. Furthermore, I analyze not only the achievement effect, but also look into whether attending a single-sex school affects students’ major choices and test-taking behavior.

The results of the analysis show that, on average, both boys and girls in boys- and girls-only schools scored about two percentile points higher in the reading and English sections of the college entrance exams compared with boys and girls in coeducational schools. Once endogenous sorting of parents and teachers are accounted for, however, the estimated mean effect decreases to around one percentile point for female students, and is close to zero for male students. Moreover, unconditional quantile regressions show heterogeneity in the estimated treatment effects for English exams. Single-sex education is favorable for students in quantiles 0.5–0.8 of the distribution of percentile ranks; the estimated treatment effects for these students are higher than the mean effects. On the other hand, the magnitude of the treatment effect is inconsequential for students at the very bottom and the very top quantiles of the distribution. Hence, although single-sex education may not benefit students, in expectation, it may be desirable for the above-average students.

I also find that single-sex schooling does not affect a student’s major choice and test-taking behavior. When endogenous sorting of parents and teachers are not accounted for, I find that male and female students are 1.2 and 2.9 percentage points more likely to choose a natural science major during high school. Moreover, male students are 2.1 and 3.2 percentage points more likely to take the advanced version of math and science college entrance exams. When such sorting is controlled for, A teacher’s school district will be assigned according to his or her home address. Also, in general, a teacher may not request for an exception to this assignment policy.
however, I do not find any statistically or practically significant differences in a student’s major choice and test-taking behavior.

The remainder of the paper is organized as follows. Section 2 provides an overview of research on single-sex schools. Section 3 illustrates the student and teacher assignment policy adopted in Seoul, followed by a detailed explanation of data and empirical strategies in Section 4. In Section 5, I present results on the validity of the research design. Section 6 presents estimates of single-sex schooling. Section 7 discusses some implications from this study’s results, and Section 8 concludes.

2. Prior literature on single-sex education

There are many arguments for and against single-sex schooling (Mael, Smith, Alonso, Rogers, & Gibson, 2004). Some of the pros raised in favor of single-sex settings are as follows:

- Female students are often distracted by the presence of male students, and vice versa;
- Students are more likely to pursue non-stereotypical curricular and courses.

Cons often cited against single-sex education are the following:

- Single-sex schools (especially female-only schools) may have fewer resources;
- A coeducational setting reduces stereotypes through familiarity.

The impact of single-sex schooling, therefore, should be empirically tested given the conflicting views.

Prior research on estimating the effects of single-sex school abounds, and researchers across many disciplines engage in examining the effects (e.g., Halpern et al., 2011). Rather than examining each and every study on single-sex schools, I briefly discuss results from a recent meta-analysis on single-sex schools (Pahlke, Hyde, & Allison, 2014). The meta-analysis is conducted using studies published in English, and the authors conduct the analysis by dividing the studies into uncontrolled and controlled studies (i.e., controls for selection bias). The results show a modest positive effect of single-schooling based on uncontrolled studies; a trivial difference based on controlled studies.

What is noteworthy about the meta-analysis is the existence of tremendous variation in the effect estimates across studies. For the uncontrolled studies, varying results are well expected for many reasons, but the variation in the estimated effects in the controlled studies needs some attention. As a matter of course, the variation in the effect estimates is understandable because, in general, controlling for selection bias is difficult. But the most important factor that drives the conflicting results is that schools are different. That is, given that schools are different in many aspects, there may be many potential channels through which single-sex education affect students. And it is likely that effect estimates vary to a large extent depending on the school sample used in a study.

When analyzing the effects of single-sex schooling, therefore, it is crucial to investigate the possible mechanisms that drive the single-sex school effect, if there is one. Suppose students are randomly assigned to either a single-sex or coeducational school, and assume that teachers in single-sex schools are more qualified than teachers in coeducational schools. Under these circumstances, it is likely that we obtain a positive effect estimate of single-sex schooling. Suppose this effect is driven mostly by teacher-related factors. Then changing only the gender composition of students would not lead to an increase in student achievement. Hence, examining the mechanism that drives an effect, if any, is important.3

On the other hand, examination of possible channels may not provide adequate policy implications if many kinds of mechanisms are revealed. For instance, suppose we observed that students in single-sex schools perform better than those in coeducational schools, and we elicited three mechanisms that contributed to the positive effect. In this case, it is difficult to judge the magnitude of the contribution of each mechanism. And without knowing the magnitude of each contribution, it provides less policy implications. A decomposition exercise, therefore, would be beneficial in this case, but conducting a credible decomposition is a challenging task.

Moreover, out of the many reasons for sending a child to a single-sex school, the most significant rationale for parents is their belief in gender peer effects. From a policy perspective, therefore, it is desirable to isolate the single-sex effects that are not driven by school- or teacher-related factors. Even though controlling for these two factors does not necessarily lead to estimating an effect that is entirely driven by gender peer effects, one may reasonably attribute the estimated effect to gender segregation when such factors are effectively controlled for.4

There are no studies, to my knowledge, that have attempted to isolate single-sex schooling effects that are not driven by school- and teacher-level characteristics. In this paper, therefore, I attempt to isolate such effects—although not conclusively, for there may still exist unobservable school- or teacher-level factors that affect student achievement—by using the institutional setting in Seoul. In Seoul, students are randomly assigned to high schools within some school districts, and as such, the inclusion of district fixed effects will eliminate the possibility of selection bias that may result from student sorting. Moreover, public school teachers do not have control over where to teach, because public school teachers are assigned to high schools within the district by the Office of Education. This rigid nature of teacher assignment policy, therefore, is likely to reduce the amount of teacher sorting. Furthermore, the Office of Education puts much effort into homogenizing school- and teacher-related

3 Jackson (2012), for example, finds a modest single-sex schooling effect. The effect, however, is mostly driven by students who gained admission to a preferred school. This suggests that the effect of single-sex education per se is trivial.

characteristics to enhance fairness and transparency of the random assignment policy. For example, teacher salary is controlled by the government, and each salary level is determined solely by a pay step based on the length of one’s service. Hence, monetary incentives faced by teachers are homogeneous across schools. Furthermore, all schools in Seoul follow the same governmental regulations and all students basically go by the same curriculum and textbooks. Accordingly, other than school types, school and teacher-level characteristics are similar—in important aspects—across school types. Hence, although a single-sex setting may promote student achievement through a variety of channels, the institutional setting in Seoul is advantageous as to ruling out, to some extent, the three noteworthy explanations: student and teacher sorting, and differences in observable school- and teacher-level factors.5

There are a few studies that have analyzed the effects of single-sex schools in Seoul. As with this study, Park, Behrman, and Choi (2013) have analyzed the effects of attending single-sex high schools on student achievement in Seoul. According to their analysis, attending single-sex schools significantly improves students’ test scores on college entrance exams. Note however, that there are some more steps that can be done to isolate single-sex schooling effects that are not driven mainly by parental and teacher sorting. First and foremost, the paper assumes that students are randomly assigned within a “school” district. A school district in Seoul consists of two or three “administrative” districts, and the majority of the students are assigned within their administrative district because the distance from a student’s home to the school is a determining factor in assignment. This is clearly stated in the Education Act. This implies that for a student in a school district with two administrative districts, the likelihood of being assigned to a school in one administrative district is substantially higher than being assigned to a school in the other administrative district—i.e., even within a school district, school assignment may not be fully random, which leads to the possibility of selection bias. If one administrative district has better schools and households with higher socioeconomic status, parents may move from one administrative district to another within the same school district; hence, selection within a school district is feasible. Indeed, a variation in the socioeconomic status exists even within a school district. Anecdotal evidence also shows that parents move within a school district—from one administrative district to another—in order to increase the chance of being assigned to a preferred school. This fact requires a comparison at least within administrative districts rather than within school districts.

Second, the paper does not address a teacher’s endogenous selection into a high school. Rather than restricting the analysis to students in public high schools, the paper controls for school type. Accordingly, it is difficult to determine factors that are giving rise to the estimated effects (i.e., whether it is derived by teacher-level factors, gender segregation, or both).6 Third, some schools are included in their analysis even though such schools should be discarded, for students self-select into these schools.7 Fourth, in order to check the balance in baseline characteristics of students, the authors use a dataset based on a survey which contains only about 3% of the sample size that they use. Furthermore, when testing for the balance in baseline covariates using this survey, the authors use students in South Korea rather than those in Seoul. These students are clearly not representative of the sample.8

3. Student and teacher assignment mechanism

Three types of high schools exist in Korea: special purpose, vocational, and general high schools. Special purpose high schools serve students who intend to major in arts, music, or physical education. Some of these schools also specialize in science or foreign languages. Students who graduate from vocational high schools enter the job market right after graduation or go on to two-year vocational colleges. The majority of middle school graduates enter a general high school. Students are not randomly assigned to the first two types of high schools. Hence, in this paper, I use data pertaining to students attending general high schools.9

The assignment of students to general high schools is conducted as follows.10 First, when students graduate from middle school, the Office of Education assigns a high school district for each student based on his or her residence. Fig. 1 shows the number of schools and administrative districts in Seoul for 2009. It also displays the number of coeducational, boys-only, and girls-only schools in each district, and as can be seen from the figure, a variation of school types exists for some administrative districts.11 Next, within the assigned high school district, students are classified into one of three blocks based on their middle school graduate standing percentile rank. Fig. 2 provides information on how students are randomly assigned to high schools based on these three blocks. The

5 The main reason for the coexistence of single-sex and coeducational schools is that the Office of Education determined the school type of public schools based on the number of each type of schools existed before public schools are established. For example, if the Office decides to open a public school in a certain district, the Office examines the ratio of each type of schools in the district. And if, say, there are more coeducational schools in the district, then the Office opens single-sex schools. And a difference in the number of each type of school within a district exists when a new public school opens because of private schools. Because private schools can choose the type of schools, it led to variation in the ratio of each type of school within a district.

6 For instance, Jackson (2009) documents such endogenous teacher sorting by student characteristics such as race. Furthermore, Dee (2007) finds that single-sex schools have more same-gender teachers.

7 The authors argue, in a footnote, that the estimated effects do not change when they include schools where students are not randomly assigned.

8 There are some studies that focus on middle school students in Korea (Lee, Turner, Woo, & Kim, 2014; Link, 2012; Pahlke, Hyde, & Mertz, 2013).

9 Note that random assignment of students was conducted in Seoul until 2009. From 2010, the Office of Education changed its policy and students are not longer randomly assigned to general high schools.

10 All information is based on the annual official documents published by the Office of Education.

11 Note that the variation of school types differs by year, because some schools changed school types. The analysis has been carefully conducted by accounting for such changes.
Fig. 1. School and administrative districts in Seoul (2009).

Fig. 2. Student random assignment mechanisms.
first block includes students whose middle school graduate standing percentile rank falls between 0.001 and 9.999 (upper-ranked). The second block consists of students with a graduate standing percentile rank between 10.000 and 49.999 (middle-ranked). Lastly, students with a percentile rank of 50.000 or above (lower-ranked) are placed in the third block. Once classified, the Office of Education randomly assigns each student from each block to a particular school using a computer-assisted lottery system (see Fig. 2 for a hypothetical case with four schools in one school district). Note that the distance between a student’s residence and a school is accounted for during this assignment process. To put it differently, a student’s administrative district is taken into account during the allocation process. Accordingly, many students are assigned to a school within their administrative district because a school district in Seoul typically consists of two or three administrative districts.

The Office of Education uses a method similar to a randomized block design to assign students to high schools. By implementing a randomized block design-type assignment, the Office of Education ascertains that upper-, middle-, and lower-ranked students are equally distributed among schools.

In Fig. 3, I describe the hypothetical teacher assignment process in public high schools (assuming there are three districts). To become a public school teacher, individuals need to initially pass the annual teacher recruitment test administered by the Office of Education. This test is highly competitive, and typically, one out of thirty applicants will pass. Second, those who pass the test go through two weeks of training. Upon completion of the training, teachers are assigned to a high school based on their home address. An important point to note from the teacher assignment policy is that although teachers can choose where to live, teachers cannot choose where they would like to teach. To be more specific, suppose a teacher’s home address is located in District A. Even though this teacher is assigned to a school in District A, the teacher cannot choose a school of his or her choice from this district. Furthermore, neither can public schools in District A choose its teachers.

4. Data and empirical strategy

4.1. Data

I use several sources of data to draw a causal effect of single-sex schooling on student performance. For student-level data, I use administrative records of the College Scholastic Aptitude Test (CSAT) maintained by the Ministry of Education, Science and Technology in Korea. For school-level data, I use data that are available at the School Information Website (SIW) and the EduData Service System (EDSS). Furthermore, I use school-level information that I obtained from both the Ministry of Education and the Seoul Metropolitan Office of Education. Lastly, I use the Statistical Yearbook of Seoul Education.

The CSAT is similar to the SAT in the United States. Unlike the SAT, however, the CSAT is a high-stakes test which is offered to students only once at the end of their third year of high school or after they have graduated. For example, students who entered high school in 2009 took the CSAT at the end of 2011. The CSAT held in 2011 is called CSAT 2012, rather than CSAT 2011. Although the proportion varies by year and by university, CSAT scores typically determine 50–100% of college admissions. Students are tested on five sections; reading, English, mathematics, social studies, and science. The CSAT has been conducted since 1993, and with the exception of the test held in 1993, it has been conducted only once every year.

Administrative records of the exam were not made public until 2009 due to concerns that the results might reveal educational gaps among schools and promote a sense of incongruity. After several years of administrative litigation filed by the members of the National Assembly, researchers, and parents, the Supreme Court ruled in favor of the disclosure and ordered the government to publicize the test scores solely for the purpose of scientific research. As a consequence, the government has allowed researchers to apply for the data starting from 2010.

I received data for the CSAT from 2002 to 2004 and from 2009 to 2012. I applied for these particular sets of data for two reasons. First, note that for the school-level variables, the CSAT contains only the name of each school and the city in which it is located. Student-level variables include gender, diploma types, major, scores on each subject, and names of the test districts in which the students took the test. As a result, I do not have data on school type. By using the school names, however, I was able to impute school type by consulting the Statistical Yearbook of Seoul Education. When imputing the school types for schools in CSAT 2002, for instance, I consult the Statistical Yearbook of 1999 because students who took CSAT 2002 were randomly assigned in 1999. A problem with using the information in the Yearbook, however, is that prior to Yearbook 1999, some school types have been miscoded. Using miscoded school types creates a serious threat to the validity of the analysis. What is assuring, however, is that starting from the Yearbook of 1999, two variables have been reported; school type and the share of female students. As a consequence, I am able to double check the

12 Information on the teacher assignment process is obtained from an official information disclosure request to the Office of Education.
13 Although the actual teacher assignment is more complicated, the assignment process basically follows Fig. 3.
14 www.schoolinfo.go.kr. The website is operated by the Ministry of Education, Science and Technology.
15 In order to obtain the data, one has to submit a research plan to the Ministry of Education. Once the Ministry of Education receives the research plans, an external committee reviews the appropriateness and feasibility of the research plan. Once the committee approves the plan, the Ministry of Education organizes screening committees, in this case, with inside members, to decide upon whether to approve access to the data. When the final approval has been made, researchers submit a written pledge regarding the data usage followed by several other administrative procedures. Once all necessary processes have been made, researchers visit the Ministry to retrieve the data stored on a compact disk.
16 Note that CSAT 2002 was held at the end of 2001, and students spend three years in high school.
17 I verified that in the Yearbook of 1996, for example, school types of five schools are miscoded. Note, furthermore, that this does not necessarily imply that other schools are correctly specified.
types of schools by examining the share of female students. Because of the presence of data issue in Yearbook 1998 or before, therefore, I did not apply for data prior to CSAT 2002. Second, test formats are identical for CSAT 2002 to 2004 and CSAT 2009 to 2012. During CSAT 2005 to 2008, however, the format was changed and for CSAT 2008, only the discrete rank was reported for each subject. Accordingly, I did not apply for CSAT 2005 to CSAT 2008.

To investigate the treatment effect, I generated three important variables. The first variable is the school type for each school mentioned above. The second variable is the school district indicator for each school. The CSAT data has information on the test districts—rather than school districts—of each student. For most students, test districts are equivalent to school districts. Some students, however, take the CSAT in districts outside of their school districts. Moreover, even if the test district corresponds with the school district, one cannot use a test district to determine the school district because the test district in the CSAT data corresponds to the school in CSAT year “minus” one. To reiterate, because students are randomly assigned three years before the CSAT year, it is critical to use the school district “three years” before the CSAT year, given the fact that some schools have moved to other school districts during the three-year period. Hence, using a test district as a school district is inappropriate because the test district in the CSAT data may not match the school district. Because the Yearbook contains an address of each school, however, I was able to impute the school district for each school as school districts are determined by their address. The third variable is the administrative districts of each school. As mentioned in Section 2, it is important to control for the administrative district as parents may move to a given administrative district to increase their chances of going to a school within that district. To identify the administrative districts for each school, I consulted the address information presented in the Yearbook.

I also make step-by-step restrictions to the initial sample for each of the CSAT tests. This is because students are randomly assigned to a partial set of schools that are located in Seoul only. First, the initial sample consists of students in Seoul. Second, because randomization has only been conducted on students in general high schools, I drop students in special purpose high schools and vocational high schools. Third, among general high schools in Seoul, some schools located at the center of the city (circled area

---

18 The share of female students in coeducational schools should be greater than zero. Contrarily, the share should be equal to zero for boys-only schools and 100% for girls-only schools. Using this argument, I verified that the school type of only one school was miscoded in the Yearbook between years 1999 and 2009.

19 I verified this fact by matching the test districts of students and school districts of schools after imputing the school district for each school.

20 I also confirmed that some schools, indeed, have moved to other districts.

21 For example, for CSAT 2002, I impute school districts for each school by retrieving the address of each school specified in the Yearbook of 1999 and matching it with the corresponding school districts.
in Fig. 1) are also excluded because the Office of Education allows students, who reside in Seoul, to apply to the schools in this area. While most students are admitted to the school of their choice, lottery assignment is used for oversubscribed schools. The number of schools in the circled area varies by year, and I obtained the list of schools in this area by year, from the Office of Education. Sample restrictions for each year are conducted based on this information. This policy has been adopted by the government since 1996 to palliate the resentment of parents who were deprived of their right of school choices for their children. I do not use student data from these schools because these students were not randomly assigned.

Fourth, some general high schools are operated both during the day and at night. These schools are consisted of both general high school as well as vocational high school students. I exclude students in these schools because I cannot determine whether these students are general high school or vocational high school students. Fifth, there are cases in which some single-sex schools have transitioned to coeducational schools during the three-year high school period. I remove these schools from the sample.第六, I eliminate students who have graduated from high school (i.e., re-takers). Including these students in the estimation process is inappropriate because it will introduce an interpolation bias. In Korea, there are many students who retake the college entrance exam one or more years after graduating from high school (in 2012, approximately 33% of test takers from Seoul were graduates). The exact graduation dates of these students are not specified in the dataset, and as a consequence, the time spent after the treatment may differ greatly between these students and those who are in their final year of high school. Moreover, if the retaking behavior is influenced by whether students attended single-sex or coeducational schools, then including these students in the analysis will lead to sample selection bias. To gauge the magnitude of the difference in the share of on-time takers between the two types of schools, I run a regression of an indicator for on-time takers on the single-sex school dummy using the CSAT 2009 to 2012 data. The results show that for female sample, the estimated difference in the share of on-time takers is 1.9 percentage points less for single-sex schools that is statistically significant at the 1% level. For male sample, the estimated difference is 0.02 percentage points less for single-sex schools that is also significant at the 1% level (results are not shown, but available upon request). Therefore, the significant difference in this share suggests that the school type may have differential effect on students’ retaking behavior. Final 1 sample consists of students in private and public high schools, and Final 2 sample consists of students in public high schools. The series of aforementioned sample restrictions and the resulting sample size are summarized in Table 1.

Another set of data I use for testing the validity of randomization is administrative records of school-level data stored in the SIW and EDSS. The websites were created in 2008 as part of the “Act on Special Cases Concerning the Disclosure of Information by Education-Related Institutions,” and include a rich set of information on school-level data such as the state of students, teachers, school activities, school conditions, and budget and account for each school in Korea. For the most part, data are covered from 2006. Moreover, I obtained information from the Office of Education regarding the number of students that were supported by the government or third parties during their first year of high school (from 2007 to 2009). These include students who are supported with tuition reductions or in the form of fellowships. In order to qualify for financial support, students should be from a low-income family and/or be protected by the law such as the “National Basic Living Security Act.” Furthermore, for each school, the Office of Education provided me with data on the share of students in each of the blocks mentioned in Section 3 for year 2009.

4.2. Empirical strategy

Let $T_i$ be a binary variable equal to one if student $i$ is assigned to a single-sex school and zero if assigned to a coeducational school. I denote an outcome $Y_{i}$ for student $i$ in single-sex schools as $Y_{i}(1)$, and $Y_{i}(0)$ for those in coeducational schools. In a canonical Rubin causal model framework under the “Fundamental Problem of Causal Inference” (Holland, 1986; Rubin, 1974), the average treatment effect on the treated, $\hat{\tau}$, is

$$\hat{\tau} = E[Y_i|T_i = 1] - E[Y_i|T_i = 0] = E[Y_i(1)|T_i = 1] - E[Y_i(0)|T_i = 1] - E[Y_i(0)|T_i = 0]$$ (1)

where the last two terms in Eq. (1) constitute a selection bias. The bias term is the difference in the expected values of $Y_i(0)$ between those in single-sex schools and those in coeducational schools. To give an example of how this bias may affect the treatment effect, suppose an educationally motivated parent chose to have their child attend a single-sex school believing that a coeducational school setting would negatively affect their child’s academic achievement. Then it would be likely that students who attend single-sex schools have a higher value of $Y_i(0)$, generating a positive bias; the resulting estimate would overstate the treatment effect.

$T_i$ in this study, is independent of potential outcomes because of the random assignment of students to either a single-sex or coeducational high school. Consequently, the bias terms in Eq. (1) get eliminated and we can estimate the treatment effect using a standard regression framework. Thus, for each outcome variable, I estimate the

\[\tau = \frac{1}{n} \sum_{i=1}^{n} (Y_{i}(1) - E[Y_{i}(0)] | T_{i} = 1) \]

22 For example, note that students who took CSAT 2002 entered high school in 1999. Accordingly, I have to drop schools which changed their school types within the three-year period (i.e., from 1999 to 2001) because students who took CSAT 2002 are affected by the modification of school types implemented during this three-year period.

23 It is important to use the number of first-year students because students are randomly assigned to high schools during the first year.

24 I would like to express my gratitude to the public officials at the Office of Education for providing me with this confidential data. Note, also, that the Office of Education no longer holds data for other years.
Table 1
Sample restrictions.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Number of students and schools</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Initial</td>
</tr>
<tr>
<td>CSAT 2002</td>
<td>191,518 (295)</td>
</tr>
<tr>
<td>CSAT 2003</td>
<td>176,335 (296)</td>
</tr>
<tr>
<td>CSAT 2004</td>
<td>174,176 (297)</td>
</tr>
<tr>
<td>CSAT 2009</td>
<td>146,046 (322)</td>
</tr>
<tr>
<td>CSAT 2010</td>
<td>159,774 (325)</td>
</tr>
<tr>
<td>CSAT 2011</td>
<td>163,603 (328)</td>
</tr>
<tr>
<td>CSAT 2012</td>
<td>164,777 (330)</td>
</tr>
</tbody>
</table>

Notes: Initial sample consists of students in Seoul. Step 1 excludes special purpose and vocational high schools. Step 2 excludes schools that do not admit students via the lottery system. Step 3 excludes schools that operate both during the day and at night. Step 4 excludes schools that modified their school type during the three-year period prior to CSAT year. For Final 1 sample, I exclude students who had already graduated from high school at the time of the test. Final 2 sample consists of public high school students only. Note that for CSAT 2002 to 2004, I exclude one district, where only one girls-only school is present. The number of high schools is in parentheses.

The following two specifications:

\[
\begin{align*}
y_{ids} = \begin{cases} 
\alpha_1 + \alpha_2 G_{isd} + \Gamma_{dc} + \epsilon_{isd}, & \text{if student } i \text{ is a female student} \\
\beta_1 + \beta_2 B_{isd} + \Gamma_{dc} + \eta_{isd}, & \text{if student } i \text{ is a male student.}
\end{cases}
\end{align*}
\]

where \(y_{ids}\) in Eq. (2) is either a continuous or discrete outcome (e.g., test scores, major choice) for student \(i\) in school \(s\) and in administrative district \(d\), where \(d \in \{1, 2, \ldots, 21\}\). Subscript \(c\) indicates the CSAT year \((c \in \{2002, 2003, 2004, 2009, 2010, 2011, 2012\})\). I include the district-by-CSAT year fixed effects \((\Gamma_{dc})\) so that the comparison of a student outcome is conducted within the CSAT year and district. Note that I am comparing girls in girls-only schools with girls in coeducational schools and boys in boys-only schools with boys in coeducational schools. Thus, the treatment indicator \(G_{isd}\) is equal to one if female students attend girls-only schools and zero if female students attend coeducational schools. On the other hand, \(B_{isd}\) is equal to one if male students attend boys-only schools and zero if male students attend coeducational schools. Lastly, \(\epsilon_{isd}\) and \(\eta_{isd}\) are the error terms.

The specifications, however, do not allow the statistical examination of the gender difference in impacts, if there is any. To examine whether there is a difference in the effect estimate between female and male students, therefore, I run a regression by pooling the female and male samples:

\[
y_{ids} = \gamma_0 + \gamma_1 F_{isd} + \gamma_2 S_{isd} + \gamma_3 (F_{isd} \times S_{isd}) + (F_{isd} \times \Gamma_{dc}) + \phi_{isd},
\]

where an indicator for female students \(F_{isd}\) is interacted with an indicator for single-sex school \(S_{isd}\) as well as the district-by-CSAT year fixed effects \(\Gamma_{dc}\). \(\phi_{isd}\) is an error term. In this setting, an estimate for coefficient \(\gamma_3\) provides the magnitude of the gender difference in the effect of single-sex schooling.

Note that the standard linear regression summarizes the average relationship between the outcome variable and the treatment based on the conditional mean function. This provides only a partial view of the relationship between the outcome variable \(y\) and the treatment variable \(T\) (i.e., mean effects of \(T\) on \(y\)). Estimating the mean impact, however, may miss the heterogeneous effect that the treatment has on students (Bitler, Gelbach, & Hoynes, 2006; Imai & Ratkovic, 2013). In the context of this paper, the effect of gender segregation may be different for high-, middle-, or low-performing students. I therefore estimate heterogeneous treatment effects by estimating the effect of treatment across the distributions of students' percentile ranks. To estimate the effect, I implement the quantile regression method pioneered by Koenker and Bassett (1978) and run the following using the pooled data:

\[
y_{ids}^q = \begin{cases} 
\alpha_1^q + \alpha_2^q G_{isd} + \Gamma_{dc} + \epsilon_{isd}^q, & \text{if student } i \text{ is a female student} \\
\beta_1^q + \beta_2^q B_{isd} + \Gamma_{dc} + \eta_{isd}^q, & \text{if student } i \text{ is a male student.}
\end{cases}
\]

where \(q \in (0, 1)\) denotes a quantile, and \(y_{ids}^q\) refers to the outcome in \(q\)th quantile. Hence, by running the quantile regression, we can retrieve different values of the treatment effect by choosing different values of quantile \(q\). When implementing the quantile regression, one needs to make a rank preservation assumption. The assumption, however, is quite a strong assumption in the current setting. A student who would have been at the \(p\)th percentile of the coeducational school distribution won’t necessarily be at the \(p\)th percentile of the single-sex school distribution. Nevertheless, one can still estimate meaningful effect estimates for policy purposes by taking the simple difference between quantiles of the marginal distribution of outcome under treatment and no treatment (i.e., unconditional quantile treatment effect). The method developed by Koenker and Bassett (1978) is used for estimating the conditional quantile treatment effect, and often times, conditional quantile treatment effects are hard to interpret and may not provide meaningful policy implications (Firpo, 2007). Hence, in order to estimate the unconditional quantile treatment effects, I use an estimator developed by Firpo, Fortin, and Lemieux (2009). Firpo et al. (2009)’s proposed method uses the so-called “influence function” that is used in robust estimation of statistical methods. Their method consists of conducting a regression of the influence function of the unconditional quantile of the
outcome variable on the explanatory variables.\textsuperscript{25} And the method allows for retrieving different marginal effects at various quantiles of the outcome variable.

For the analysis of mean effects, I provide results for the comprehensive school pool that includes students in private and public high schools as well as for the restricted school pool that consists of students in public high schools only. Two separate analyses allow us to examine whether the effect of single-sex schooling changes by controlling for endogenous sorting of teachers. For the unconditional quantile regression, I show the results for the restricted sample that consists of public schools only, because this paper improves upon the previous work by controlling teacher quality.

5. Validity of research design

In this section, I test for the validity of the research design by examining the following conditions: balance in predetermined covariates, non-compliance behaviors, and attrition.

5.1. Balance in predetermined covariates

Suppose students in single-sex schools performed better than those in coeducational schools. Although students are randomly assigned to high schools within their school district, those with a background favorable for one’s academic achievement may have been assigned to, say, a single-sex school by chance. If this is the case, one cannot conclude that single-sex schooling is beneficial for a student’s academic achievement even though students are randomly assigned. Hence, in this subsection, I test for balance in predetermined covariates that are purported to positively affect one’s academic achievement.

In Table 2, Panel A, I present tests of within-district balance in the share of students in each of the blocks presented in Fig. 2.\textsuperscript{26} For testing the share, it is necessary to use the ratio of which the numerator is the number of boys in each block and the denominator is the number of total boys in each school, because this study compares boys (girls) in boys-only (girls-only) schools with boys (girls) in coeducational schools. I received data on the number of students in each block—by gender—and calculated the shares of high-, middle-, and low-ranked students by dividing the number of boys (girls) in each category by the total number of boys (girls) in each school. As can be expected from the randomized block design-type assignment procedure described in Fig. 2, the shares of upper-, middle-, and lower-ranked students are almost identical between the two types of schools.\textsuperscript{27} Therefore, I conclude that differences do not exist in students’ previous academic achievements across school types.

I underscore the importance of checking the baseline covariates in a randomized control trial by presenting the shares of students in each block for high schools in which students are not randomly assigned (i.e., schools located in the circled area in Fig. 1). In Panel B of Table 2, I present the analysis that compares the share of boys-only (girls-only) schools with that of coeducational schools. Because students “apply” for the schools in this circled area, the share of students in each block differs significantly (especially for the female sample).

In Panel A of Table 3, I test balance in baseline student- and school-level covariates.\textsuperscript{28} Note that the tests of balance in Table 3 are conducted using school-level data. Admittedly, one needs to use data at the individual level, in order to conduct ideal tests of fidelity to the experimental design. Unfortunately, however, none of the variables in Table 3 are available at the individual level. But because the variables such as the share of students receiving lunch support are derived from student-level data, the results are—although not conclusive—indicative of the balance in baseline student characteristics. Note, further, that the shares of the variables in Table 3 are not separately available by gender. Accordingly, I cannot test the balance in these shares as in Table 2. Although there is no particular reason to believe that these shares differ by gender between single-sex and coeducational schools, the interpretation should reflect this limitation.

For student-related covariates, I test for balance in the following two variables: shares of government- and lunch-supported students. The share of government-supported students corresponds to the share of students financially supported by the government in the form of tuition waivers or fellowships. I use first-year high school students for calculating the shares because students are randomly assigned when entering high school. As can be seen from the table, none of the coefficients are significant, and the differences in the effect estimates are small. In order to gauge the extent to which there are differences in the distribution of student family income, I also test the difference in the annual per capita expenditure spent on lunch support.\textsuperscript{29} The estimated difference is about four dollars that is not statistically significant, suggesting that there are few differences in the expenditure. Next, I find almost no difference in either average class size or pupil-teacher ratio. The difference between the two types of schools is less than one student. These results regarding school-related characteristics are expected because the Office of Education is strongly committed to homogenizing these characteristics to abide by the “School Equalization Policy” and to sustain the random assignment policy. In sum, the results from the balancing test show that baseline student- and school-related characteristics are balanced between the two schools.

5.2. Non-compliance behavior

In a randomized control trial, if subjects do not comply with their initially assigned treatment, this will create

\textsuperscript{25} For the estimation, I benefited by the STATA command uploaded in the author’s website.

\textsuperscript{26} The Office of Education no longer holds the data prior to 2009, and consequently, I was only able to obtain data for the 2009 academic year.

\textsuperscript{27} All regressions are conditional on administrative district fixed effects.

\textsuperscript{28} All regressions are conditional on year-by-administrative fixed effects.

\textsuperscript{29} The dollar amount that students receive for their lunch support varies by student family income.
Table 2
Tests of within-district balance in baseline achievement using school-level data (2009).

<table>
<thead>
<tr>
<th>Outcome variable</th>
<th>Boys</th>
<th>Single-sex school (1=yes)</th>
<th>Number of observations</th>
<th>Girls</th>
<th>Single-sex School (1=yes)</th>
<th>Number of Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Randomly-assigned high school</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper-ranked students ($\psi &lt; 10$)</td>
<td>-0.001 (0.002)</td>
<td>66</td>
<td>-0.001 (0.003)</td>
<td>67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle-ranked students ($10 \leq \psi &lt; 50$)</td>
<td>0.001 (0.004)</td>
<td>66</td>
<td>-0.002 (0.003)</td>
<td>67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower-ranked students ($\psi \geq 50$)</td>
<td>0.000 (0.005)</td>
<td>66</td>
<td>0.002 (0.004)</td>
<td>67</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B. Non-randomly-assigned high school</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper-ranked students ($\psi &lt; 10$)</td>
<td>0.006 (0.013)</td>
<td>25</td>
<td>-0.059** (0.023)</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle-ranked students ($10 \leq \psi &lt; 50$)</td>
<td>0.020 (0.038)</td>
<td>25</td>
<td>-0.073** (0.033)</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower-ranked students ($\psi \geq 50$)</td>
<td>-0.026 (0.048)</td>
<td>25</td>
<td>0.332** (0.050)</td>
<td>19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Regressions in Panel A are conducted using the public school sample. Results in Panel B are obtained based on students attending non-randomly assigned high schools. I estimate all coefficients in Panel A by running a regression of each outcome variable on a dummy variable indicating the single-sex school conditional on administrative districts—twenty-one administrative districts. For Panel B, all coefficients are derived from running a regression of each outcome variable on the single-sex school indicator. For school type, coeducational school is the reference category. Standard errors in parentheses. $\psi$ denotes middle school graduate standing percentile rank—the smaller the number, the higher the ranking. * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Table 3

<table>
<thead>
<tr>
<th>Outcome variable</th>
<th>School type</th>
<th></th>
<th>School type</th>
<th></th>
<th></th>
<th>Number of school</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Student and school characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of government-supported students</td>
<td>0.034 (0.011)</td>
<td>0.100 (0.008)</td>
<td></td>
<td></td>
<td>265</td>
<td></td>
</tr>
<tr>
<td>Share of students receiving lunch support</td>
<td>-0.026 (0.022)</td>
<td>0.083 (0.013)</td>
<td></td>
<td></td>
<td>209</td>
<td></td>
</tr>
<tr>
<td>Expenditure spent on lunch support (W)</td>
<td>-4.533 (2.454)</td>
<td>31.082 (1.619)</td>
<td></td>
<td></td>
<td>208</td>
<td></td>
</tr>
<tr>
<td>Pupil–teacher ratio</td>
<td>0.277 (0.150)</td>
<td>15.440 (0.114)</td>
<td></td>
<td></td>
<td>275</td>
<td></td>
</tr>
<tr>
<td>Average class size</td>
<td>0.220 (0.238)</td>
<td>34.515 (0.181)</td>
<td></td>
<td></td>
<td>275</td>
<td></td>
</tr>
<tr>
<td><strong>Panel B. Share of moved-out students</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moved-out during the first year</td>
<td>-0.007 (0.003)</td>
<td>0.041 (0.003)</td>
<td></td>
<td></td>
<td>204</td>
<td></td>
</tr>
<tr>
<td>Moved-out during the second year</td>
<td>-0.005 (0.001)</td>
<td>0.012 (0.000)</td>
<td></td>
<td></td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>Moved-out during the third year</td>
<td>-0.001 (0.001)</td>
<td>0.002 (0.000)</td>
<td></td>
<td></td>
<td>196</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Regressions are conducted using the public school sample. I estimate all coefficients from running a regression of each outcome variable on a dummy variable indicating the school type conditional on year-by-administrative district fixed effects—three or four years and twenty-one administrative districts depending on a outcome variable. For “Share of students receiving lunch support,” “Expenditure spent on lunch support,” and Panel B, the data for 2006 is missing. There are some missing values for some of the variables. Standard errors in parentheses. In the ”Expenditure spent on lunch support” column, “W” denotes Korean Won (in thousands, local currency unit), and the number is presented in the local currency unit (approximately, 1W=1 dollar).

a threat to the validity of the analysis. In this case, validity of the analysis can be revitalized by using the initial assignment as an instrumental variable (e.g., Howell, Wolf, Campbell, & Peterson, 2002). Unfortunately, I lack data on students’ initial assignments. Based on several “facts” below together with some statistical tests, however, I argue that this study does not suffer from non-compliance behavior.

There are several reasons parents in Korea have very few incentives not to comply with their child’s initially assigned high school. First, the Office of Education refuses to accept requests for reassignment. The assignment is a once-for-all procedure in Korea. Second, Article 84(8) of the Education Act stipulates that if a student does not enroll in the high school that he or she has been assigned to, the student will not be assigned to another school during the same academic year. Third, Article 86 of the Education Act lists the cases in which reassignment procedures can be conducted, and according to this article, parents cannot request a reassignment just because they do not like the school their child has been initially assigned to. Fourth, Article 89(2) of the Education Act states that when all members of a student’s family have relocated to another school district after the initial assignment has taken place, then the parents may request a reassignment. Even if this is the case, however, parents cannot choose their child’s school within the district to which they have moved. That is, the student has to go through another computer-assisted lottery assignment procedure and is randomly assigned to a high school in the new school district. Accordingly, educationally motivated parents have few incentives to move to another school district just to gain a random reassignment opportunity. Lastly, one might argue that some parents may engage in lobbying the officials at the Office of Education before the lottery assignment takes place. In order to prevent corrupt
behaviors during this procedure, the lottery assignment is conducted by each school district’s High School Admis-
sion Lottery Management Committee which consists of parents, principals, assistant principals of high schools, and commissions of local education offices.

If the aforementioned facts are pervasive among par-
ents, then we should not observe any significant patterns in student move-out across school types. Panel B of Table 3 shows tests of within-district differences in the share of transferred students between single-sex and coeducational schools.\(^8\) During the school years 2007–2009, the share of students who moved out from coeducational schools during the first, second, and third year of high school is 4, 1, and 0%, respectively. The estimated differences are less than one percentage point with no statistical significance. These numbers do not necessarily “prove” the fact that parents are complying with their child’s initially assigned high school. I argue, however, that parents do comply with their child’s initially assigned high school, given the small shares of moved-out students across school types coupled with the reasons mentioned above.

5.3. Attrition

In Korea, students spend three years in high school. In Fig. 4, I demonstrate a time frame in high school. Once randomly assigned to a high school, students begin their first year in March. The following March, students start their second year of high school. The final year also begins the following March. Then during November of the third year, students take their college entrance exams; 30 months after entering high school. As a consequence, there is a probability that some students drop out during this time frame. And if dropped out students—who were originally assigned to coeducational high schools—had lower exam scores, on average, than those who were originally assigned to single-sex high schools, then the estimated treatment effect will be biased downwards. In Appendix A, I provide a mathematical illustration of attrition bias and the conditions in which one can bypass the problem of attrition.

A simple exercise (presented in Appendix A) shows that there are three cases under which a researcher can bypass attrition bias. The first case is when the attrition rate is close to zero. In Panel A of Table A.1, I provide, by school type, an average share of attrition by cohort and year of high school. The total attrition rates during the high school period range from 4 to 5%. Although the shares are quite small, we cannot ignore the possibility of attrition bias in the current setting. The second scenario is when the observation is missing at random. To test the second condition, one can check the baseline covariates of students who are missing and examine whether there are statistically signif-
ificant differences in these covariates between the two types of schools. Note, however, that I cannot test whether this holds because I observe only those students who took the exam. The last case is when there is a pattern in attrition but similar in expectation between the two types of schools. Although I cannot test this condition directly, I provide some tests and demonstrate—in Appendix A—that the pattern of attrition is likely to be similar, on average, between single-sex and coeducational schools. All in all, I argue that attrition is not a large source of bias in this study.

6. Estimation results

6.1. Mean effects on test scores

I first estimate the average treatment effects using a comprehensive school pool that contains both public and private schools. The effect estimates based on this sample correspond to the effects of attending single-sex schools. Next, I analyze the effect by restricting the sample to students in public schools. For the outcome variable, I use students’ test scores for reading and English exams. I do not use students’ test scores for math exams because students choose which math tests to take “after” they are assigned to their high school, and there may be non-random differential test-taking behavior as a result of students being assigned to single-sex schools (e.g., stereotype threat).\(^3\) Note, further, that I do not analyze the effect of attend-
ing single-sex schools on high school graduation because high school graduation rates are remarkably high in Seoul (about 95%), and few variation in this share exists across schools. Each student’s test score is converted to a per-
centile rank. I use two types of percentile ranks for the main analysis: one is calculated based on the national level; and the other based on the sample level. The reason for presenting the results for both types of percentile ranks is that using national norms would give some indication of where in the national distribution the students under the study were drawn from, even though the number of students in Seoul is large enough to exert a lot of influence on national test norms. Table 4 presents the es-
timated mean effects. Each cell in the table represents a separate regression.

Panel A provides effect estimates that are drawn from students in a comprehensive school pool. All the esti-
mates in Panel A are derived from running a regression of each outcome on dummies for single-sex schools and year-by-school district fixed effects, separately by gender. This specification basically estimates the effect that the previous literature derives. Interestingly, I find different results compared with the findings presented in Park et al. (2013). First, estimated single-sex effects on reading and English for female sample presented in Park et al. (2013) are about 0.065 of a standard deviation for both subjects. The corresponding effect estimates I find are 2.237 and 2.734 percentile points, which can be converted into z-scores of 0.074 and 0.090 using an inverse normal function. For female sample, therefore, I find effect estimates that are about 14 and 38% higher than those reported in the previous study. Second, effect estimates for male

\(^{10}\) Regressions are conditional on year-by-administrative district fixed ef-

certs.
**Fig. 4.** Time frame during high school.

Table 4
Mean effects of attending single-sex school.

<table>
<thead>
<tr>
<th>Outcome variable</th>
<th>Single-sex schooling effects</th>
<th>Tests of gender difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Female sample</td>
<td>Male sample</td>
</tr>
<tr>
<td>Panel A. Comprehensive school pool</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading percentile rank I</td>
<td>2.237***</td>
<td>1.972***</td>
</tr>
<tr>
<td></td>
<td>(0.341)</td>
<td>(0.319)</td>
</tr>
<tr>
<td></td>
<td>[237,548]</td>
<td>[272,387]</td>
</tr>
<tr>
<td>English percentile rank I</td>
<td>2.734***</td>
<td>2.752***</td>
</tr>
<tr>
<td></td>
<td>(0.472)</td>
<td>(0.424)</td>
</tr>
<tr>
<td></td>
<td>[237,013]</td>
<td>[271,318]</td>
</tr>
<tr>
<td>Panel B. Restricted school pool (administrative district fixed effects)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading percentile rank I</td>
<td>1.113</td>
<td>0.281</td>
</tr>
<tr>
<td></td>
<td>(0.814)</td>
<td>(0.565)</td>
</tr>
<tr>
<td></td>
<td>[88,561]</td>
<td>[95,218]</td>
</tr>
<tr>
<td>Reading percentile rank II</td>
<td>1.55</td>
<td>0.286</td>
</tr>
<tr>
<td></td>
<td>(0.846)</td>
<td>(0.590)</td>
</tr>
<tr>
<td></td>
<td>[88,561]</td>
<td>[95,218]</td>
</tr>
<tr>
<td>English percentile rank I</td>
<td>1.168</td>
<td>0.625</td>
</tr>
<tr>
<td></td>
<td>(1.063)</td>
<td>(0.750)</td>
</tr>
<tr>
<td></td>
<td>[88,324]</td>
<td>[94,795]</td>
</tr>
<tr>
<td>English percentile rank II</td>
<td>1.208</td>
<td>0.637</td>
</tr>
<tr>
<td></td>
<td>(1.097)</td>
<td>(0.775)</td>
</tr>
<tr>
<td></td>
<td>[88,324]</td>
<td>[94,795]</td>
</tr>
<tr>
<td>Panel C. Restricted school pool (school district fixed effects)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading percentile rank I</td>
<td>1.020</td>
<td>0.922</td>
</tr>
<tr>
<td></td>
<td>(0.728)</td>
<td>(0.569)</td>
</tr>
<tr>
<td></td>
<td>[88,561]</td>
<td>[95,218]</td>
</tr>
<tr>
<td>English percentile rank I</td>
<td>1.481</td>
<td>1.524*</td>
</tr>
<tr>
<td></td>
<td>(0.956)</td>
<td>(0.792)</td>
</tr>
<tr>
<td></td>
<td>[88,324]</td>
<td>[94,795]</td>
</tr>
</tbody>
</table>

Notes: A coefficient estimate in each cell represents a separate regression. Standard errors clustered at the school level are reported in parentheses. The number of observations are presented in brackets. For Panel A, the number of clusters for the female sample is 734; 769 for the male sample. For Panels B and C, the number of clusters for the female sample is 366; 361 for the male sample. Comprehensive school pool (Panel A) includes private and public schools. Percentile ranks I and II are based on the rank at the national and Seoul level, respectively. The estimates reported in the Tests of gender difference column are derived from running a regression of each outcome variable on a female indicator, a dummy for single-sex schools, and the female indicator interacted with the dummy for single-sex schools and year-by-district dummy vectors. For Panel B, all regressions are conditional on year-by-administrative district fixed effects. For Panels A and C, all regressions are conditional on year-by-school district fixed effects. * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.
sample provided in the previous study are 0.109 (reading) and 0.152 (English) of a standard deviation. The effect estimates derived in this study, on the other hand, are 1.972 (equivalent to 0.066) and 2.752 (equivalent to 0.092) percentile points; approximately a 39% reduction in the magnitude of the effect estimates. Third, the differences in the coefficient estimates between female and male sample are 0.044 and 0.087 for reading and English, respectively. Although a statistical significance of gender differences in the estimated single-sex effects cannot be determined, for the prior study does not test for such differences, the differences may imply that single-sex schooling is favorable for male students. This study, however, finds few gender differences in the estimated effects both statistically and practically; 0.008 for reading and 0.002 for English.

In Panel B, I present results obtained from running a regression of each outcome on an indicator for single-sex schools, and year-by-administrative district fixed effects. The sample is restricted to students in public school to account for the degree of endogenous sorting of teachers. For the outcome, I use two types of percentile ranks. Notably, the effects reduce to a great extent. The magnitude of the effect estimate on reading for the female sample is about one-half the effect derived in Panel A; 57% for the English exam. The corresponding reductions are 86 and 77% for the male sample. Moreover, while all the estimates in Panel A are statistically significant at the 1% level, the estimates in Panel B are no longer statistically significant. The results, in Panel B, therefore imply that when teacher quality is controlled for, to some extent, the difference in student achievement between single-sex and coeducational school is small.

As mentioned before, previous literature compares student achievement within school districts. Such comparison, however, cannot control for the degree of parental sorting within school districts. I examine to what extent effect estimates change when the comparison of students is conducted within school districts, rather than administrative districts. Panel C provides estimated effects obtained from running regressions that are conditional on year-by-school district fixed effects. Interestingly, I find larger estimates under the within-school district comparison (except for reading in the female sample). Although the differences between the estimates in Panel C and Panel B are not substantial for the female sample, it is for the male sample; about 1.5 times higher.

All in all, results in Table 4 highlight that although students in single-sex schools perform better, on average, than those in coeducational schools, the case no longer holds when the degree of teacher and parental sorting are controlled for to some extent. Although a strong conclusion on gender peer effects cannot be drawn from Table 4 because unobservable school- and teacher-level characteristics between the two types of school may still exist, the results may suggest that the presence of the opposite sex in school per se, may not have a negative influence on students' academic achievement.

The last column in Table 4 tests for differences in the coefficient estimates for the female and male sample. To test the statistical significance of these differences, I estimate standard errors by pooling the female and male samples and running a regression in which all the explanatory variables are interacted with a female dummy (see Eq. (3)). The effects are, in general, larger for females. We cannot, however, reject the hypothesis that there is a difference in the effect estimates between the female and male sample, regardless of the specifications.

6.2. Distributional impact on test scores

Estimating the mean effects, however, may not provide a full view of the relationship between the outcome and the treatment. In Fig. 5 and 6, therefore, I plot distributional effects of single-sex schooling on student achievement that are derived from analyzing public high school student scores, because the extent of the effects may vary depending on their academic performance. Quantile regressions are conditional on the CSAT year-by-administrative district fixed effects, and I use the percentile rank calculated at the national level. In the figure, I plot the quantile treatment effects for quantile values $q = \{0.05, 0.10, \ldots, 0.95\}$, and the upper horizontal lines correspond to the estimated mean effect. I also juxtapose 95% confidence intervals to gauge the statistical significance of the estimated distributional effects.

Panel A of Fig. 5 presents results for the effects on reading performance using the female sample. Although there seems to be some variation in the distribution of treatment effects, the estimated effects are, in general, concentrated around the mean effect. Thus, the effect of single-sex education on reading exams does not depend on student performance. I plot the distributional effects on English exams in Panel B. Contrary to reading, I find considerable heterogeneity in the estimated effects by student performance. The graph is roughly an inverted U-shape implying that the benefit of attending a single-sex school is mostly driven by the above-average students but below the top-performing students (i.e., 0.9). For example, the estimated effect for students in quantile 0.5–0.7 is approximately 1.8 percentile points, which is about three times larger than those observed for students in quantile 0.1 and 0.2. The effects for students located at the very bottom and the very top quantiles of the distribution of test scores, on the other hand, are about 0.4 and 0.8. For the statistical inference, I estimate bootstrapped standard errors. Most of the confidence intervals are above the zero horizontal line indicating that the estimated effects are statistically significant at the 5% level.

Quantile regression estimates based on the male sample are presented in Fig. 6. Patterns of the estimated effects are similar to those observed for the female sample. Panel A shows the results for reading. For the male sample, the mean impact of attending a single-sex school is 0.281. Although there are some variations, most of the quantile treatment effects are mostly around the mean impact suggesting that the single-sex schooling effect is not heterogeneous. Besides, we cannot reject the null hypothesis of the single-sex schooling effect, especially given the

---

32 Estimates of bootstrapped standard errors are based on 50 bootstrap samples. For unconditional quantile regression, a theoretical calculation of robust standard errors such as clustered standard errors is not available.
small impact observed for reading; the largest effect observed is 0.6.

Results for English, on the other hand, show marked heterogeneity in the effect estimates. The graph, again, is an inverted U-shape, but skewed more toward the above-average students. The test scores of male students at quantile 0.6–0.8—who attend male-only schools—are approximately 1.5 percentile points higher than those in coeducational schools. Yet, the estimated effects for students located at the very bottom (0.15 or below) and at the very top quantiles (0.95) of the distribution of test scores are small. Due to the small impact observed for the below-average and top-performing students, the estimated effects are statistically insignificant at the 5% level. Effect estimates for students near or above the average are all statistically significant at the 5% level.

In Fig. 7, I present tests of gender difference in the estimated effects for each quantile. As was the case in the mean analysis, I find a larger effect for female students for every quantile. And because the patterns of the estimated effects are similar between the female and male sample, gender differences in the effect estimates do not vary by the distribution of student achievement. Although the benefit of single-sex schooling appears to be larger for female students in every quantile and for both subjects, I cannot reject the null hypothesis of no gender differences because the estimated confidence intervals all encompass the zero horizontal line. Hence, as in the mean analysis, I caution against drawing a conclusion that female students benefit more by attending single-sex schools.

For both the female and male sample, I find a lot of heterogeneity in the effect estimates for English; few variations for reading. Moreover, the mean effects and distributional effects for English exams are larger than that of reading exams (especially for male students). This is probably because a large proportion of English education is conducted within schools and the curriculum and content of the subject matter is relatively well-defined and sequenced compared to reading education. Hence, it may suggest that single-sex education has a differential impact depending on the nature of subjects being tested. It would be interesting to analyze the effect of single-sex schooling
on students’ math performance, for math education is also well-defined and sequenced, and is mostly fulfilled within schools. But I cannot examine this hypothesis because the students took different math exam in this study.

6.3. Effects on major choice and test-taking behavior

One of the arguments often raised against coeducational settings is that female students, when surrounded by male students, suffer from a so-called “stereotype threat.” Literature on cognitive psychology, for example, argues that female students avoid taking math-intensive courses. Math-related subjects are oftentimes conceived as a “masculine” subject. Such stereotype undermines female students’ self-concept of ability with respect to these subjects (Leslie, Cimpian, Meyer, & Freeland, 2015; Nosek, Banaji, & Greenwald, 2002; Spencer, Steele, & Quinn, 1999). And because gender-related issues are more salient in coeducational school, the setting is more likely to intensify such perception and dissuade female students from taking such classes. Indeed, a report by the National Science Foundation (1996) finds that only 35% of undergraduates and 10% of graduate students are enrolled in math-related courses. Some literature in psychology have conducted experiments using a small number of observation, and find evidence that is supportive of the prevalence of stereotype threat (e.g., Inzlicht & Ben-Zeev, 2000; Kessels & Hannover, 2008).

There are two high school majors from which high school students in Korea are able to choose during their second year; natural or social sciences. This institutional setting provides an opportunity to analyze the effect of single-sex schooling on students’ major choice. This choice is very important in Korea. When students apply for college admissions, those who choose a natural science major often apply to science-related departments. Students in social sciences, on the other hand, apply to humanities or social sciences departments. The choice of high school major, therefore, has a tremendous impact on a student’s future career.

The CSAT data from 2001 to 2004 contains information on students’ major choice (i.e., natural or social sciences major), and I run a linear probability model that regresses a dummy for major choice on an indicator for single-sex school conditional on the year-by-district fixed effects. In Panel A of Table 5, I present results for the comprehensive school pool that does not account for endogenous parental and teacher sorting. Results show that female and male students in single-sex schools are 1.2 and 2.9 percentage points more likely to choose a natural science major. The tests of gender difference show that male students are, in fact, more likely to choose a natural science major compared with female students.

As previously mentioned, the analysis of a comprehensive school pool does not allow us to draw meaningful policy implications because we cannot determine whether teachers or gender-related factors have more impact on a student’s choice of major. If the estimated effects are driven mainly by teacher-related factors, the probability can be raised even at the coeducational setting. To control for teacher-related factors, I run the model again by focusing on students in public schools. Notably, all the positive and statistically significant effects of single-sex schooling disappear; there are no differences in the likelihood between single-sex and coeducational schools.

In addition to major choice, I examine whether single-sex schooling affects students’ test taking behavior. The CSAT data from 2009 to 2012 contain information on whether students took the advanced version of the math and science college entrance exams. In general, students should take the advanced version of these exams if they wish to apply for science-related majors in college. Therefore, analyzing such behavior is meaningful because of career-related consequences that come along with a student’s choice. Using a comprehensive school pool, I find that male students are 2.1 and 3.2 percentage points more likely to take advanced math and science exams. For the female sample, I find no practically significant effects. As was the case in the analysis of major choice, I find that the effects are larger for male students. Note, however, that the differences in the probability of taking the advanced version of the math and science exams reduce to zero, again, when the analysis is conducted using a restricted school pool. Moreover, a gender difference does not exist in the effect estimates due to the null effects.
Table 5
Effects of single-sex schooling on students’ major, math, and science selections.

<table>
<thead>
<tr>
<th>Outcome variable</th>
<th>Single-sex schooling effects</th>
<th>Panel A: Comprehensive school pool</th>
<th>Panel B: Restricted school pool</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Female sample</td>
<td>Male sample</td>
<td>Tests of gender difference</td>
</tr>
<tr>
<td>Natural science major (1 = yes)</td>
<td>0.012***</td>
<td>0.029***</td>
<td>−0.016***</td>
</tr>
<tr>
<td></td>
<td>[98,431]</td>
<td>[120,372]</td>
<td>[218,803]</td>
</tr>
<tr>
<td>Take advanced math exams (1 = yes)</td>
<td>0.005</td>
<td>0.021***</td>
<td>−0.017***</td>
</tr>
<tr>
<td></td>
<td>[119,141]</td>
<td>[152,058]</td>
<td>[291,199]</td>
</tr>
<tr>
<td>Take science exams (1 = yes)</td>
<td>0.009***</td>
<td>0.032***</td>
<td>−0.023***</td>
</tr>
<tr>
<td></td>
<td>[119,141]</td>
<td>[152,058]</td>
<td>[291,199]</td>
</tr>
</tbody>
</table>

Notes: A coefficient estimate in each cell represents a separate regression. Standard errors clustered at the school level are reported in parentheses. The number of observations are presented in brackets. Restricted school pool includes public schools only. Comprehensive school pool includes private and public schools. All regressions in Panel A are conditional on year-by-administrative district fixed effects, whereas for Panel A, all regressions are conditional on year-by-school district fixed effects. For the “Natural science major” variable, analyses are based on CSAT 2002 to 2004; for other variables, analyses are based on CSAT 2009 to 2012. For Panel A, the number of clusters for the female sample is 262 and 472; 287 and 482 for the female sample. For Panel B, the number of clusters for the female sample is 119 and 247; 118 and 243 for the male sample. The estimates reported in the Tests of gender difference column are derived from running a regression of each outcome variable on a female indicator, a dummy for single-sex schools, and the female indicator interacted with the dummy for single-sex schools and year-by-district dummy vectors. * significant at the 10% level; ‡ significant at the 5% level; *** significant at the 1% level.

The implication from the results in Table 5 is that, although it appears that students (especially male students) in single-sex schools are more likely to choose a natural science major, and the advanced version of the math and science exams, the differences in the likelihood decrease to zero when endogenous sorting of parents and teachers are accounted for. Contrary to the existing view, therefore, it seems that a single-sex setting makes little difference, at least in the Korean context, as to students’ major choice and test-taking behavior when the differences in student-, school-, and teacher-related characteristics are minimal between the two types of schools.

7. Discussion

As mentioned in Section 1, estimating the effects of attending single-sex schools does not, in general, provide meaningful policy implications unless one explores the underlying mechanisms. In Korea, teachers are able to select the private school they wish to work for. And because most of the private high schools are single-sexed, it is likely that differences in teacher-related characteristics between single-sex and coeducational schools may have given rise to the estimated effects. In Table 6, I present tests of within-district balance in teacher-related variables that are available. For a comprehensive school pool, I run a regression of each outcome variable on dummy variables indicating school type, private/public schools, and school district fixed effects. As can be seen from the results for a comprehensive school pool, many teacher characteristics are significantly different across school types. The F-tests that test the hypothesis that the school-type dummies jointly had no effect and the corresponding p-values show that most of the teacher characteristics are highly statistically different across school types. For example, an average share of female teachers in a coeducational school is 53%. The share, however, is significantly different in a boys-only school (i.e., only 23%). Not only do I find differences in the shares of female teachers, I also find significant differences in the share of senior, temporary, and part-time teachers, as well as teachers with a master’s degree. Moreover, because private schools are relatively flexible with respect to making decisions about inputs or instructions compared to public schools, I see some differences in principal characteristics (Panel B). For example, single-sex schools are 11.9 and 3.5% more likely to track
students (reading). Also, single-sex schools are 25.3 and 6.4% more likely to operate summer school programs.

Given the differences in teacher- and principal-related characteristics between the two kinds of schools, the relative magnitude of many possible factors—that may give rise to the estimated positive effect of single-sex education—is not clear from the effect estimates derived from the comprehensive school pool.

Focusing on students in public high schools, however, significantly reduces these differences. With the exception of the share of female teachers, I do not find any statistically significant differences in the observable teacher- and principal-related characteristics between the two types of schools. Although we do see a difference in the share of female teachers between single-sex schools and coeducational schools, the magnitude of the difference is much smaller than the one we observed for a comprehensive pool (especially between boys-only and coeducational schools).

Note that the possibility of unobserved differences in teacher- and school-related characteristics between single-sex and coeducational schools may still exist, even if the analysis is restricted to students in public schools. I argue, however, that such restriction reduces the differences in these characteristics between the two types of schools to a great extent given the rigid nature of the teacher assignment policy adopted for public high schools together with the Office of Education’s strong commitment to equalize school characteristics. Consequently, I believe the effect estimates, derived in this study under a restricted school pool, are not driven mainly by teacher- and school-related characteristics.

From a policy perspective, it would be interesting to isolate the single-sex schooling effect that is driven primarily by gender peer effects. If gender peer effects are substantial, student achievement can be improved by merely transitioning coeducational schools to single-sex schools. And although some short-run expenditure will be incurred by such transition, it is probable that such transition would not cost extra in the long run. For this reason, parents and education policy makers are keen to know whether such effects are significant, if any. This paper, however, cannot separate the effect of “gender peer effects” from other factors with which it is correlated, and the results of this study should be interpreted with this in mind.

Also worthy of note is the fact that Seoul is unique in many respects. Or at least, the city is quite different from the United States. For instance, Seoul is distinct in terms of gender roles, the degree of population homogeneity, and student attitudes and parental involvement with respect to learning. I therefore caution against drawing a strong conclusion on the effect of single-sex education in other contexts from this study.
Table A.1.  
Average share of attrition by cohort, and tests of within-district balance in attrition rate.

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Public school type</th>
<th>Year in high school</th>
<th>Panel A. Average share of attrition, by cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>First year</td>
<td>Second year</td>
</tr>
<tr>
<td>AY 2007</td>
<td>Single-sex school</td>
<td>0.032</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>Coeducational school</td>
<td>[0.019]</td>
<td>[0.010]</td>
</tr>
<tr>
<td>AY 2008</td>
<td>Single-sex school</td>
<td>0.023</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>Coeducational school</td>
<td>[0.011]</td>
<td>[0.005]</td>
</tr>
<tr>
<td>AY 2009</td>
<td>Single-sex school</td>
<td>0.029</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>Coeducational school</td>
<td>[0.012]</td>
<td>[0.009]</td>
</tr>
</tbody>
</table>

Notes: AY 2007 in the first column corresponds to cohorts that entered high school in academic year 2007, and the corresponding numbers in the First year, Second year, and Third year columns report the average share of dropouts for the first, second, and third grade students who entered high school in 2007. The number in the Total column is the sum of dropouts over the three-year period; the number reported in the Total column is slightly different from the sum of three prior columns because of the rounding. For Panel B, I run a regression of each outcome variable (i.e., the share of dropouts in the first, second, or third year of high school) on a dummy variable indicating the school type conditional on year-by-administrative district fixed effects—three years and twenty-one administrative districts. All regressions are conducted using data from 2007 to 2009. Sample size for the regression that uses the share in the first, second, and third year of high school is 204, 208, and 209, respectively. Standard deviations are in brackets. Standard errors are in parentheses.

8. Conclusions

Consider male or female students who move from coeducational schools to single-sex schools. The findings from this study imply that, on average, female students' performance increase around 2 and 2.5 percentile points in the reading and English sections of the college entrance exam; 1.5 and 2 percentiles points for male students. Moreover, male students are about three percentage points more likely to choose a natural science major and take the advanced version of the math and science tests if they attend a boys-only school; however, small effects on female students. The magnitude of the estimated effects, however, significantly decreases once the degree of parental and teacher sorting are controlled for by restricting the sample to students in public schools and comparing students within administrative districts. For female students, the estimated mean effect is one percentile point for reading and English; for male students, 0.3 and 0.6 percentile points. Furthermore, for either gender, I do not find any differences, practically or statistically, in students' major and test-taking behavior between the two types of schools; single-sex schooling may not make any differences in academic achievement of poor- or high-performing students, if teacher and school characteristics are similar between the two schools.

Although single-sex schooling appears to have little impact on students, in expectation—when there are few differences between single-sex and coeducational schools—unconditional quantile regression reveals that students above the middle quantiles but below the top quantiles (i.e., 0.9) benefit the most by being in a single-sex school; 1.8 and 1.2 percentile points. For reading, on the other hand, I do not find notable heterogeneity in the estimated quantile treatment effects. Furthermore, single-sex schooling has little impact on students located at the very bottom and the very top quantiles of the distribution of percentile ranks.

Lastly, the estimates derived in this study apply specifically to the Seoul metropolitan area. The exact magnitude of the estimated effects is therefore unlikely to generalize to other countries. Nevertheless, there is good reason to believe that once the differences in student-, teacher-, and school-related characteristics are minimized between single-sex and coeducational schools, the benefits of single-sex schooling may not be substantial. At a minimum, the results in this study highlight that the effect of single-sex schooling is heterogeneous.

Acknowledgments

I thank Steven Raphael, Jesse Rothstein, Michael Anderson, Eugene Smolensky, Aaron Chalfin, Candace Hamilton, Sarah Tahamont, Natalie Ahn, Roberto Hernandez, Layda Negrete for their comments and suggestions. I also thank editors and anonymous reviewers for providing me with invaluable advice on an earlier version of this paper. Moreover, I appreciate the Ministry of Education, Science, and Technology of South Korea and EduData Service System for generously providing me with the administrative records of students' test scores and school-level data. Finally, I am grateful to Hongyu Hwang and Kyungyeon Kim for
allowing me to access all the necessary administrative and confidential school-level data, and to Yoonchie Kim for the English editing. All errors are mine.

**Appendix A**

Let $S$ and $C$ denote an assignment to single-sex and coeducational schools, respectively. Note that student $i$ is assigned to a treatment $T_i \in \{S, C\}$, and let $Y_i(S)$ and $Y_i(C)$ indicate test scores of student $i$ assigned to single-sex and coeducational schools, respectively. The treatment effect we want to estimate is $\tau = E[Y_i(S)] - E[Y_i(C)]$. Now, some students might drop out of high school before taking exams. As a consequence, for student $i \in \{S, C\}$, we observe

$$D_i = \begin{cases} 1, & \text{if observed (i.e., not attrit)} \\ 0, & \text{if missing (i.e., attrit)}. \end{cases}$$

Furthermore, for student $i \in \{S, C\}$, we potentially have information on either of the following:

$$Y_i(S) = \begin{cases} Y_{i\text{obs}}(S), & \text{if observed} \\ Y_{i\text{miss}}(S), & \text{if missing} \end{cases} \quad \text{or} \quad Y_i(C) = \begin{cases} Y_{i\text{obs}}(C), & \text{if observed} \\ Y_{i\text{miss}}(C), & \text{if missing.} \end{cases}$$

In the presence of sample attrition, we end up estimating $\tau^* = E[Y_{i\text{obs}}(S)] - E[Y_{i\text{obs}}(C)]$ rather than estimating the $\tau = E[Y_i(S)] - E[Y_i(C)]$. Observe that for student $i \in S$,

$$cE[Y_i(S)] = E[D_i|T_i = S] \times E[Y_{i\text{obs}}(S)] + (1 - E[D_i|T_i = S]) \times E[Y_{i\text{miss}}(S)] = E[Y_i(S)].$$

In the same manner, for student $i \in C$,

$$E[Y_{i\text{obs}}(C)] = E[Y_i(C)] + \frac{1 - E[D_i|T_i = C]}{E[D_i|T_i = C]} (E[Y_i(C)] - E[Y_{i\text{miss}}(C)]).$$

Then by Eqs. (A.1) and (A.2), we have the following:

$$\tau^* = E[Y_{i\text{obs}}(S)] - E[Y_{i\text{obs}}(C)]$$

$$= E[Y_i(S)] - E[Y_i(C)]$$

$$+ \left( \frac{1 - E[D_i|T_i = S]}{E[D_i|T_i = S]} \right) \left( E[Y_i(S)] - E[Y_{i\text{miss}}(S)] \right)$$

$$\rho_S$$

$$- \left( \frac{1 - E[D_i|T_i = C]}{E[D_i|T_i = C]} \right) \left( E[Y_i(C)] - E[Y_{i\text{miss}}(C)] \right) \rho_C. \quad (A.3)$$

Thus, the last two terms ($\rho_S$ and $\rho_C$) in Eq. (A.3) correspond to the bias created by attrition in single-sex and coeducational schools, respectively. From the equation, there are three cases in which the bias terms disappear. First case is when $E[D_i|T_i = S]$ and $E[D_i|T_i = C]$ are close to one (i.e., there are few attrition). In this case, the two fraction terms approach zero and the bias terms vanish. Second, when the observation is missing at random (i.e., $E[Y_i(S)] - E[Y_{i\text{miss}}(S)] = E[Y_i(C)] - E[Y_{i\text{miss}}(C)] = 0$), then again both biases collapse to zero. Finally, when there is a pattern in attrition but similar in expectation between single-sex and coeducational schools, then $\rho_S$ and $\rho_C$ that constitute the bias will jointly cancel out, leaving no bias.

As the first case implies, we can ignore attrition problems when the attrition rate is close to zero. In the current setting, however, the attrition rate is not close to zero (about 5%). To test the second case, one can check the baseline covariates of students who are assigned to single-sex and coeducational schools and determine whether $E[Y_i(S)] - E[Y_{i\text{miss}}(S)] = 0$ and $E[Y_i(C)] - E[Y_{i\text{miss}}(C)] = 0$. I cannot, however, test whether this holds because I only have data for those who take the exam.

On the other hand, there is school-level data on a share of students who dropped out of high school, and by calculating the mean of the shares by school type, I can use these values as proxies for the expectation terms in $\rho_S$ and $\rho_C$. In Table A.1, I show an average share of high school dropouts by school type and by school grades. In the table, AY 2007 indicates students who entered high school in 2007, and on average, 2.5% of the students in coeducational schools quit school during their first year. When these students went on to second grade, another 1.5% of students in coeducational schools dropped out, on average. Lastly, 0.4% of students quit during their third year of school. Hence, the average share of dropouts during the high school period is 4.4% for students in coeducational schools who entered high school in 2007.

Therefore, based on the second to last column, an average share of high school dropouts in coeducational schools for three waves is 4.6%. Likewise, a corresponding dropout rate for single-sex schools is 4.6%. Using these two values, I can estimate two fraction terms in $\rho_S$ and $\rho_C$; $E[D_i|T_i = S] = 0.954$, and accordingly, the estimate for $\frac{1 - E[D_i|T_i = C]}{E[D_i|T_i = C]}$ in $\rho_S$ is 0.048. For coeducational schools, the estimate for $\frac{1 - E[D_i|T_i = C]}{E[D_i|T_i = C]}$ in $\rho_C$ is also 0.048.

As a consequence, the magnitude of the attrition bias is increased by the factor of 0.048 for both types of schools. According to Eq. (4), then, the size of the bias depends on the two terms, $E[Y_i(S)] - E[Y_{i\text{miss}}(S)]$ and $E[Y_i(C)] - E[Y_{i\text{miss}}(C)]$, multiplied by the above factor. Note that if these two values are similar, then the size of the attrition bias is minimal because the difference in the attrition multiplier between the treatment and the control groups is the same. In this study, I contend that the two values are similar in expectation because there are few reasons to believe that students who quit from coeducational schools are academically different, on average, from those who quit from single-sex schools. In Panel B of Table A.1, I present results from the tests of within-district differences in the dropout rates. The estimated differences in the dropout rates are almost zero. Although the results do not prove the “similar-in- expectation argument” above, I argue it is quite likely. Furthermore, even if the academic achievement is quite different, the study is less likely to suffer from the attrition bias as the attrition rate is small.

On the other hand, even if we assume that attrition is not similar in expectation, it is less likely that this study suffers from the attrition bias. For illustrative purposes, I analyze a hypothetical scenario in which attrition
is not similar in expectation. Suppose the estimated single-sex school effect is two percentile points. Also, assume that students who dropped out of boys-only schools are academically superior, on average, than those in coeducational schools. To derive the magnitude of the bias in this instance, I assume $E[Y_{j}^{miss}(S)] = 10$ and $E[Y_{j}^{miss}(C)] = 0$ (i.e., students who dropped out of boys-only schools may have earned 10 percentile points more than those who dropped out of coeducational schools. This is quite a worst-case scenario given the assumed treatment effect is two percentile points). Also, let $E[Y_{i}(S)] = 50$ and $E[Y_{i}(C)] = 48$ so that students in single-sex schools earn two percentile points more. Given the aforementioned settings, the magnitude of the attrition bias in Eq. (A.3) is 0.048(50 − 10) − 0.048(48 − 0) = −0.384. Thus, even if we assume quite a worst-case scenario, the magnitude of the attrition bias is quite small. Therefore, I conclude that attrition is not a large source of bias in this study.

References


