Economics of Education Review 91 (2022) 102330

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Health benefits of having more female classmates: Quasi-experimental evidence from China

Yuhe Guo [a], Shaoping Li [b], Siwei Chen [a], Yalin Tang [a], Chengfang Liu [a,*]

[a] China Center for Agricultural Policy, School of Advanced Agricultural Sciences, Peking University, China
[b] School of Economics and Management, Huzhou University

ARTICLE INFO

Keywords:
female classmates
physical health
random class assignment
junior high schools
China

ABSTRACT

We investigate the impact of having more female classmates on students' physical health outcomes. To do so, we draw on panel data from the first two waves of the China Education Panel Survey (CEPS) and take advantage of exogenous variation in classroom gender composition brought by the random assignment of students to classes when they entered junior high school. Results from the value-added model show that having more female classmates helps to improve the physical health outcomes of junior high school students in China. Specifically, a one-percentage-point (pp) increase in the proportion of female classmates in Grade 7 explains a 0.003-0.004 pp increase in one's probability of staying normal BmiAZ or going from abnormal to normal BmiAZ, a 0.029-0.031 increase in the BmiAZ score of those with low starting BmiAZ, a 0.003 increase in one's HAZ score, and a 0.157-0.165% or 0.166-0.177% decrease in the diopter of one's left or right lens of glasses from Grade 7 to Grade 8, respectively. One possible mechanism underlying these relationships is that more female classmates bring healthier behaviors. Heterogeneity analyses indicate that the beneficial effects of having more female classmates on health outcomes are more prominent among boys, students with less-educated parents, and those attending rural schools or schools with bigger class sizes.

I. Introduction

The latest statistics show that adolescents represent over 16 percent of the world's population, with almost 90 percent (or 1.11 hundred million) of them concentrating in low- and middle-income countries.[1] Although significant progress has been made in improving adolescents' health in developing countries in the 21st century, the pace has been slowing down in recent years (UN IGME, 2020; Ward et al., 2021). As a critical and unique stage of human development, investments in adolescence can bring benefits today, for decades to come, and for the next generations (Patton et al., 2016). Hence, it is of particular importance for developing countries to explore possible measures to promote adolescent health.

How to promote adolescent health? There has been growing evidence highlighting the strong impacts of social environment factors at various levels on adolescents' health. Specifically, at the national level, war and conflict (Klasen et al., 2010; Kohrt et al., 2010), national wealth (Torsheim et al., 2004), and income inequality (Due et al., 2009) are found to strongly affect adolescents' health in various aspects. At the community level, exposure to pollution (Gauderman et al., 2004), educational pressure (Morgan et al., 2012), and poor public infrastructure (Boehmer et al., 2007) have been blamed for adverse effects on adolescents' health. At the school level, adolescents' health is found strongly affected by school sanitation facilities (Freeman et al., 2012), classroom peer environment (Alexander et al., 2001), and student-teacher relationships (Kim, 2021). At the family level, some studies suggest consistent inequalities in children's health are closely related to family socioeconomic status (Currie et al., 2008; Elgar et al., 2015; Reiss, 2013), family structure (Blum et al., 2000), parenting style (Rhee et al., 2006), as well as family norms and attitudes (Malcolm et al., 2013).

Among those social environment factors, peers begin to play increasingly important roles in children's physical health as they reach adolescence (Viner et al., 2012). Classmates represent one crucial group of peers with whom students interact the most over the school day. Especially in countries where students usually stay with the same

* Corresponding author at: Boya Chair Professor and Deputy Director, School of Advanced Agricultural Sciences, China Center for Agricultural Policy, Peking University, China
E-mail address: cfliu.ccap@pku.edu.cn (C. Liu).

[1] World bank, 2021, World Development Indicators. Available at https://data.worldbank.org/?name_desc=true.

https://doi.org/10.1016/j.econedurev.2022.102330
Received 11 January 2022; Received in revised form 25 October 2022; Accepted 26 October 2022
Available online 3 November 2022

classmates for all subjects in the same class throughout the academic year, China is a good case in point. Previous studies have shown certain groups of classroom peers can help reduce multiple health-risk behaviors and enhance one's physical health (Anteghini et al., 2001; Campbell et al., 2008; Viner et al., 2012).

Females are one of such groups. Compared with their male counterparts, females have been widely observed to be less antisocial, violent, aggressive, and less likely to engage in smoking, alcohol using and drug addiction (Bertrand and Pan, 2013; Cornwell et al., 2013; Duckworth and Seligman, 2006; Jacob, 2002; Kritsotakis et al., 2016; Yang et al., 2014). Along this line, it seems that females may be more likely to help create a protective environment against a broad range of health-risk factors. Thus, an interesting question arises: is it beneficial to have more female classmates for one's physical health?

In this paper, we investigate whether having more female classmates positively affects junior high school students' physical health. This paper relates to the gender peer effects literature. The identification of gender peer effects may face two empirical challenges, including endogenous sorting and contextual confounding (Epple and Romano, 2011; Manski, 1993). The present study addresses these challenges by applying a quasi-experimental approach. To do so, we draw on data from the China Education Panel Survey (CEPS), a nationally representative longitudinal survey of junior high school (Grades 7-9) students in China. The CEPS offers two helpful features that facilitate our identification strategy. First, over 40 percent of the CEPS project schools randomly assigned incoming students to different Grade 7 classes, which helps to eliminate the endogenous sorting in these schools. Second, the rich information in the CEPS data enables us to address contextual confounding by controlling for a large set of covariates at the student, parent, household, and class levels, as well as a full set of school fixed effects in the estimation.

There has been substantial evidence across cultures that adolescents benefit from having more female classmates in terms of academic performance (Gong et al., 2021; Hill, 2015; Hoxby, 2000; Hu, 2015; Lu and Anderson, 2015; Briole, 2021), cognitive skills (Black et al., 2013; Lavy and Schlosser, 2011), non-cognitive skills (Gong et al., 2021), and mental health (Guo et al., 2021). However, to the best of our knowledge, little is known about whether having more female classmates would benefit adolescents' physical health. Hence, this study seeks to fill the knowledge gap by identifying the causal relationship between the proportion of female classmates and one's physical health in junior high schools in China. We believe understanding this causal relationship is particularly relevant for the optimal grouping of students in classrooms and schools in the broad context of mixed-gender education.

Our results show that being assigned to a class with more female classmates in Grade 7 has a statistically significant and positive effect on one's multiple physical health indicators in China. Specifically, a one-percentage-point (pp) increase in the proportion of female classmates in Grade 7 explains a 0.003-0.004 pp increase in one's probability of staying normal BmiAZ or going from abnormal to normal BmiAZ, a 0.029-0.031 increase in the BmiAZ score of those with low starting BmiAZ, a 0.003 increase in one's HAZ score, and a 0.157-0.165% or 0.166-0.177% decrease in the diopter of one's left or right lens of glasses from Grade 7 to Grade 8, respectively.

Our results also show that the mechanism underlying the observed gender peer effects works this way: more female classmates, more (less) likely to engage in healthy (unhealthy) behaviors. Specifically, a one pp increase in the proportion of female classmates in Grade 7 is associated with 0.973 minutes more on physical exercise per week, a 0.004 pp increase in one's probability of decreasing the amount of time spent on watching TV on both a weekday and a weekend day, as well as a 0.003 pp increase in one's probabilities of decreasing the amount of time spent on playing online games on a weekday. We further provide suggestive evidence that the observed positive gender peer effects on certain health measures (including the BmiAZ score for the subsample of BmiAZ<-2 in wave one, the HAZ score, as well as the diopters of left and right lens of glasses) were mainly driven by female classmates within homogeneous sub-groups (say, in terms of the *hukou* and left-behind status).

Finally, our results also reveal heterogeneity in the effects of having more female classmates. Specifically, the impacts of having more female peers on health outcomes are larger among boys, students with less-educated parents, and those attending rural schools or schools with bigger class sizes, which are consistent with the observed heterogeneous gender peer effects on health behaviors in this paper.

This study contributes to the literature in at least four ways. First, we add to a growing economic literature on the effects of external factors (including family, school, and other social environment factors) on adolescents' physical health. Second, we extend the gender peer effects literature by explicitly considering physical health as an output of the human capital production process. To the best of our knowledge, this study is the first to provide quasi-experimental evidence on the impacts of having more female classmates on adolescents' physical health. Third, beyond the gender peer effect analyses at the classroom level, we further extend it to the sub-classroom level by providing suggestive evidence that the observed effects were mainly driven by female peers within homogeneous sub-groups. Finally, we explore the potential working channels of gender peer effects, which helps improve our understanding of why classmates' gender composition matters during adolescence.

The rest of this paper is structured as follows. Section II introduces the data, followed by an empirical framework in Section III. Section IV presents our empirical findings. The final section concludes.

II. Data

A. Survey

Our analysis draws on data from the China Education Panel Survey (CEPS), a nationally representative survey of China's junior high school (Grades 7-9) students, designed and conducted by the Renmin University of China. In the academic year of 2013-2014, the CEPS adopted a multi-stage Probability-Proportional-to-Size sampling strategy to select the study sample. Administrative units and socioeconomic status (i.e., the average years of schooling and the proportion of migrants among the local population) were first used as the stratification variables to select 28 sample counties. Within each sample county, the enrollment sizes and school types were further used as the stratification variables to select 4 junior high schools.[2] In total, there were 112 junior high schools from 28 counties. In each of the 112 selected schools, two Grade 7 classes and two Grade 9 classes were randomly chosen. As it turned out, 10 sample schools had only one Grade 7 class and one Grade 9 class, so the total number of sample classes was 438 rather than 448. All students (a total of 19,487) in the sample classes participated in the baseline survey in the academic year 2013-2014. To date, four follow-up surveys have been conducted on a yearly academic basis, and the first two rounds of panel data (2013-2014 and 2014-2015) were currently publicly available.

This study draws on the first two waves of survey data from CEPS to examine the class gender composition at the start of junior high school on students' health outcomes one year after. In fact, CEPS conducted the 1st wave among 7th and 9th graders in two time periods in the 2013-2014 academic year. Specifically, 68 out of the 112 sample schools (or 61%) were surveyed in October 2013 (or the second month of the first semester) whereas the rest of 44 schools (or 39%) in March 2014 (or the first month of the second semester). All schools received their second wave of survey one year after the time when they got their baseline survey. In each wave of surveys, the survey team collected rich information, including students' characteristics and their health outcomes, which will be used in our study.

According to the combination of times when students started Grade 7

[2] There are three school types, including public, private, and migrant schools.

Table 1
Survey timing by cohorts of students

Cohorts	F1	F2	S1	S2
When got the 1st survey wave?	Oct., 2013	Mar., 2014	Oct., 2013	Mar., 2014
Grade	7	7	9	9
No. of students covered	6262	4017	5765	3443
When got the 2nd survey wave?	Oct., 2014	Mar., 2015	Not surveyed	Not surveyed
Grade	8	8	Out of the sample	Out of the sample
No. of students covered	5733	3716	0	0

Notes: The students in this table involved all the sample schools and students in CEPS.

and when they got their first survey wave, we grouped students into four cohorts (Table 1). Specifically, cohorts F1 and F2 were freshmen who started Grade 7 in September 2013 and were in their first year at junior high school when they got their first survey wave. The difference between these two cohorts is the time when they got their first survey wave. While F1 got their first survey wave in October 2013, F2 got theirs in March 2014. In the second survey wave, both F1 and F2 cohorts were in Grade 8. As to cohorts S1 and S2, both cohorts were seniors who started Grade 7 in September 2011 and were in their last year at junior high schools at the time of the first survey wave. Similar to the difference between cohorts F1 and F2, S1 got their first survey wave in October 2013 whereas S2 got theirs in March 2014. It should be noted that neither S1 nor S2 was covered in the second survey wave in CEPS as they had graduated from high school by the time of the second survey wave.

In this paper, we are interested in examining the gender peer effects of junior high school students on their health in terms of change in health outcomes from Grade 7 to Grade 8. To do so, we take advantage of the fact that a subsample of students from CEPS were randomly assigned into classes at the start of Grade 7 and remained in the same class until they entered Grade 8. In other words, the gender composition in their classes changed little during the one-year period following the beginning of Grade 7, assuming that transfer-in/out is not common in the first year of junior high school.

B. Class assignment and estimation sample

To implement our quasi-experimental design, we focus on cohorts F1 and F2 as they were covered in both survey waves. Following a four-step procedure proposed by Gong et al. (2021), we identified the subsample from CEPS for the purpose of our study. As depicted in Fig. 1, in the first step, among the 112 CEPS project schools, we excluded 3 schools with only one Grade 7 class to ensure there is within-school variation in class gender composition. This left us with 109 schools, each having two Grade 7 classes. Secondly, based on responses by school principals in the

Fig. 1. Study sample construction procedure
Note: G7 stands for Grade 7, RA stands for random assignment.

first survey wave to the question "Were newly enrolled students randomly assigned to Grade 7 classes?" we kept the 90 schools whose principals responded "yes." In the third step, based on responses by the head teachers of the two Grade 7 classes in the first wave to the question "Were Grade 7 students assigned to classes based on their entry academic performance or not?" we kept the 56 schools where both head teachers answered "no." These 56 schools are those that practiced random class assignment judging by responses from both school principals and head teachers.

In the final step, we need to confirm whether students in these 56 schools got reassigned to different classes from Grade 7 to Grade 8 or not. To do so, we compare the class IDs of students in both survey waves, and excluded 9 schools where students' class IDs are not the same in the two waves, which means they got reassigned to different classes from Grade 7 to Grade 8. After this process, we are left with 47 schools that include 4260 students from 94 classes. The data show that for these 4260 students, their class gender compositions in the beginnings of Grade 7 and Grade 8 are significantly correlated with a correlation coefficient at 0.96 ($p<0.001$). This further confirms that this subsample of students was not only randomly assigned into classes when they started Grade 7 but also remained with the same classmates over this one-year study period. Therefore, they become the final sample for our analyses in the rest of this study.

To understand how generalizable the effects presented in this paper are, we compared 47 sample schools that allocate students randomly to classes and do not reassign students in Grade 8 to those that have two Grade 7 classes but have been excluded from our study sample to see if they are statistically similar to each other. In total, we checked 26 observable characteristics at the school, principal, class, and head teacher levels.[3] Results show that only two out of the 26 characteristics (7.7 percent) come out statistically significant (Table A1), suggesting that the way we constructed our study sample schools will not severely affect the external validity of our research findings.

C. Variables

During each of the two waves of surveys, a set of questionnaires were administered to the sampled students themselves, their parents, subject teachers, head teachers, and school principals. For the purpose of the study, we draw on information from three modules in each round of the survey.

Classmate gender composition module. We measure the classmate gender composition of a student by the proportion of female classmates in the class excluding the student herself/himself in Grade 7. Therefore, students of the same gender in the same classroom share the same proportion of female classmates. Specifically, this proportion was based on class composition in the second month of the first semester (October 2013) for cohort F1 and the first month of the second semester (March 2014) for cohort F2, respectively.

Student physical health module. We focus on four student health indicators as follows. The first two indicators are calculated by using the weight and height information reported by students and referring to the WHO Growth reference data for children aged 5-19 years (WHO, 2006).[4] One is Height-for-age Z-score (HAZ), with higher HAZ indicating better health. The other is BMI-for-age z-score (BmiAZ). Following the WHO Child Growth Standards (2006), a child with a BmiAZ less than minus two is defined as "wasting" whereas a BmiAZ greater than one as "overweight". For those wasted children, higher BmiAZ indicates better health. In contrast, for those over-weighted children, lower BmiAZ indicates better health. Taking into account the starting BmiAZ of students, we will group the sample students into three categories, namely low starting BmiAZ (where BmiAZ<-2 in wave one), normal starting BmiAZ (where 2-≤BmiAZ≤1 in wave one) and high starting BmiAZ (where BmiAZ>1 in wave one) when examining their change in BmiAZ between the two survey waves. In addition, we could also examine whether a student transits from one BmiAZ category to another from wave one to wave two, say, staying in the normal category, or going from abnormal (either low or high starting BmiAZ category) to normal.

The third and fourth health indicators are both related to students' eyesight.[5] Studies from ophthalmology have shown that eyesight is closely related to factors including outdoor activity (Lee et al., 2013; Rose et al., 2008) and screen-based activity (Pan et al., 2012; Goldschmidt and Jacobsen, 2014; Holden et al., 2016).[6] Following the literature (Braun et al., 1996; Zhao et al., 2002; Bullimore and Brennan, 2019), for those students who self-reported as being myopic (or nearsighted), we measure their eyesight by the logarithms of self-reported diopter of their left and right lens of glasses. For those students who self-reported as not being nearsighted, we assume their diopters of both lenses of glasses as zero. The greater the diopter, the poorer the eyesight of the corresponding eye. Similarly, we will examine not only their change in the logarithm of diopters between the two survey waves, but also whether a student transits from one eyesight category to another from wave one to wave two, say, staying non-myopic or going from non-myopic to myopic.

Covariates. Following the literature, we control for characteristics at the student, parent, household, and class levels that might affect students' physical health. Specifically, we control for seven covariates at the student level (including age, gender, ethnicity, left-behind status, whether being the only child in the family, birth weight, and boarding status), five covariates at the parent level (including education of both parents, marital status, and both parents' age when the child under discussion was born), six covariates at the household level (including *hukou*, family social economics status, water facilities, toilet facilities, whether any family member drinks alcohol, whether any family member smokes), and four covariates at the classroom level (including class size, head teacher's gender, educational background, and teaching experience). We also include school fixed effects to control for factors at the school or the broader regional level that might affect student physical health. All standard errors are clustered at the grade level. Table 2 reports summary statistics of the key variables.

[3] Specifically, we checked nine characteristics at the school level (enrollment, the share of local students, the share of rural students, the share of left-behind students, public school or not, school's ranking within the county, school's location, scores of students' misbehavior in school, student-teacher ratio); five at the principal level (age, gender, education, and working experience, and whether graduated from normal university or majored in teaching); five at the class level (class gender composition, class size, the share of rural students in class, the share of local students in class, the share of left-behind students in class); and seven at the head teacher level (their average age, education, and teaching experience in years, the share of male head teachers, the share of head teachers graduated from normal university or majored in teaching, the average professional title of head teachers, the share of head teachers awarded in the past three years).

[4] In CEPS, both students and their parents were asked to report students' weight and height. The reports from both sources are significantly highly correlated ($p<0.001$). We used the report by students themselves in our study. But results remain similar when we use report by parents. Self-reported weight and height have been used as accurate indicators of actual weight and height (Stewart, 1982). Stommel and Schoenborn (2009) show that in health risk estimates associated with variations in BMI values, results based on the self-reported BMI or BMI measured on site are virtually the same.

[5] According to the National Health Commission (2021), the prevalence of myopia (or nearsightedness) among junior high school students in 2020 in China was 71.1%.

[6] The medical and ophthalmology literature also documents other factors closely related to eyesight, such as genetics and parental history of myopia (Jones et al., 2007; Lim et al., 2014), optical and environmental influences (Smith et al., 2014), urbanization (Zhang et al., 2010), educational pressure (Morgan et al., 2012), and near-work activities (Pan et al., 2012).

Table 2
Summary statistics

	(1) Mean	(2) SD	(3) Min	(4) Max
Panel A: Outcome variables Grade 7				
BmiAZ	-0.018	[1.307]	-5.16	4.58
Wasting (1=yes)	0.062	[0.242]	0	1
Overweight (1=yes)	0.213	[0.410]	0	1
HAZ	0.645	[1.142]	-5.04	5.14
Stunting (1=yes)	0.016	[0.124]	0	1
Myopia (1=yes)	0.476	[0.499]	0	1
Diopter of left lens of glasses	1.129	[1.547]	0	9.50
Diopter of right lens of glasses	1.182	[1.605]	0	9.99
Grade 8				
BmiAZ	-0.218	[1.267]	-5.19	4.75
Wasting (1=yes)	0.074	[0.262]	0	1
Overweight (1=yes)	0.169	[0.375]	0	1
HAZ	0.503	[1.044]	-4.45	4.30
Stunting (1=yes)	0.017	[0.128]	0	1
Myopia (1=yes)	0.522	[0.500]	0	1
Diopter of left lens of glasses	1.270	[1.644]	0	9.50
Diopter of right lens of glasses	1.323	[1.697]	0	9.99
Change in outcome variables (Grade 8-Grade 7)				
BmiAZ stays normal or goes from abnormal to normal (1=yes, full sample)	0.757	[0.429]	0	1
ΔBmiAZ (subsample of BmiAZ<-2 in wave one)	0.732	[1.157]	-2.01	4.03
ΔBmiAZ (subsample of BmiAZ>1 in wave one)	-0.458	[0.604]	-4.4	1.06
ΔHAZ (full sample)	-0.142	[0.485]	-1.14	2.52
Stays non-Myopic or goes from Myopic to non-Myopic (1=yes, full sample)	0.478	[0.499]	0	1
ΔDiopter of left lens of glasses (full sample)	0.141	[0.535]	-1.25	4.80
ΔDiopter of right lens of glasses (full sample)	0.141	[0.540]	-1.25	4.80
Panel B: Change in mechanism variables (Grade 8-Grade 7)				
ΔTime spent on physical exercises weekly (min)	45.938	[183.919]	-784	1040
Decreasing time spent on watching TV on a weekday (1=yes)	0.287	[0.452]	0	1
Decreasing time spent on watching TV on a weekend day (1=yes)	0.494	[0.500]	0	1
Decreasing time spent on playing online games on a weekday (1=yes)	0.195	[0.396]	0	1
Decreasing time spent on playing online games on a weekend day (1=yes)	0.366	[0.482]	0	1
Panel C: The proportion of female classmates in the following sub-groups (Grade 7)				
Same *hukou* and same left-behind status	0.501	[0.255]	0	1
Same *hukou* but different left-behind status	0.471	[0.240]	0	1
Different *hukou* but same left-behind status	0.482	[0.238]	0	1
Different *hukou* and different left-behind status	0.484	[0.147]	0	1
Panel D: Key explanatory variable (Grade 7)				
Proportion of female classmates (*FP*)	0.484	[0.076]	0.13	0.66
Panel E: Student characteristics (Grade 7)				
Age (months)	155.168	[8.249]	129	204
Girl (1=yes)	0.484	[0.500]	0	1
Ethnic minority (1=yes)	0.089	[0.285]	0	1
Left-behind children (1=yes)	0.183	[0.387]	0	1
Only child (1=yes)	0.552	[0.497]	0	1
Birth weight (kg)	3.534	[0.565]	1	5.90
Boarding (1=yes)	0.171	[0.377]	0	1
Panel F: Parent characteristics (Grade 7)				
Father's education (years)	11.081	[3.347]	0	19
Mother's education (years)	10.501	[3.568]	0	19
Parents married (1=yes)	0.897	[0.303]	0	1
Father's age at birth of this child (years)	28.791	[4.754]	14	60
Mother's age at birth of this child (years)	26.907	[4.437]	14	52
Panel G: Household characteristics (Grade 7)				
Rural *hukou* (1=yes)	0.428	[0.495]	0	1
Family social economics status (dummies):				
Very poor	0.018	[0.134]	0	1
Poor	0.109	[0.312]	0	1
Average	0.793	[0.405]	0	1
Rich	0.074	[0.262]	0	1
Very rich	0.005	[0.073]	0	1
Tap water (1=yes)	0.925	[0.263]	0	1
Flush toilet (1=yes)	0.858	[0.349]	0	1
Family members drink alcohol (1=yes)	0.288	[0.453]	0	1
Family members smoke (1=yes)	0.545	[0.498]	0	1
Panel H: Class characteristics (Grade 7)				
Class size	45.319	[12.258]	22	77
Headteacher is male (1=yes)	0.266	[0.442]	0	1
Headteacher's education (years)	15.862	[1.006]	15	19
Headteacher's teaching experience (years)	14.234	[9.148]	1	45
Number of students	**4260**			
Number of classes	**94**			
Number of schools	**47**			

Notes: The sample involved 47 schools with random class assignments in Grade 7 but did not reassign students in Grade 8.

D. Descriptive Statistics

The data show that in the first survey wave, the prevalence of wasting, overweight, and stunting among sample students were 6.2%, 21.3%, and 1.6%, respectively (Table 2). The second wave did see some improvement in the prevalence of overweight (16.9%), but not in wasting (7.4%) or stunting (1.7%). As for eyesight, the prevalence of myopia increased from 47.6% in the first survey wave to 52.2% in the second one. In the meantime, the diopters of sample students' left (right) lens of glasses increased from 1.1 (1.2) D to 1.3 (1.3) D. Moreover, almost half (48%) of the sample students were girls. An average student was about 13 years old in Grade 7, the years of schooling of her/his father and mother were 11.1 and 10.5, respectively. The average class size was 45, which is close to the national average of 48 reported for junior high schools in China (Ministry of Education of China, 2014). The head teachers had 15.9 years of education on average, almost the same as the average years of schooling of junior high school teachers in China (Ministry of Education of China, 2014). Their average years of teaching experience was 14.2. It is fair to say that our sample is quite representative.

III. Identification Strategy

A. Empirical Specifications

In this paper, we seek to examine the gender peer effects on the change in health indicators between the two surveys. In other words, we assume it takes about one year for gender composition to manifest its impact on adolescents' physical health outcomes, which is consistent with previous studies (e.g., Clark and Lohéac, 2007). It has been shown in the literature that not only do peer effects on health-risk behaviors

Table 3
Random assignment test

Outcomes (7^{th}):	(1) BmiAZ (full sample)	(2)	(3) BmiAZ (Subsample of BmiAZ<-2 in wave one)	(4)	(5) BmiAZ (Subsample of BmiAZ>1 in wave one)	(6)	(7) HAZ (full sample)	(8)	(9) Myopia (1=yes, full sample)	(10)	(11) Diopter of left lens of glasses (logarithm, full sample)	(12)	(13) Diopter of right lens of glasses (logarithm, full sample)	(14)
Panel A: 47 sample schools with random class assignments in Grade 7, but did not reassign students in Grade 8														
FP (7^{th})	1.149	-0.504	-1.297	0.304	0.662	0.929	1.966***	0.049	0.086	0.099	0.255	0.062	0.268	0.180
	(0.772)	(0.615)	(1.191)	(0.579)	(0.646)	(0.698)	(0.644)	(0.558)	(0.355)	(0.258)	(0.496)	(0.344)	(0.513)	(0.351)
N	4242	4242	260	260	884	884	4158	4158	4257	4257	4012	4012	4013	4013
R^2	0.076	0.065	0.162	0.114	0.006	0.116	0.205	0.218	0.000	0.124	0.001	0.135	0.001	0.140
Panel B: 62 schools with two Grade 7 classes but were excluded from our study sample														
FP (7^{th})	0.917**	-0.131	0.286	0.287	-0.472	0.007	1.390***	0.253	0.461**	0.466***	0.611**	0.464**	0.645**	0.488***
	(0.389)	(0.433)	(0.624)	(0.397)	(0.322)	(0.468)	(0.333)	(0.460)	(0.223)	(0.169)	(0.275)	(0.175)	(0.285)	(0.181)
N	5717	5717	493	493	840	840	5751	5751	5904	5904	5529	5529	5531	5531
R^2	0.067	0.057	0.127	0.105	0.004	0.156	0.215	0.194	0.005	0.129	0.007	0.146	0.007	0.151
Covariates:														
Student CHs	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Parent CHs	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Household CHs	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Class CHs	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
School FEs	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: *FP* stands for the proportion of female classmates, CHs stands for characteristics, FEs stands for fixed effects. The sample used in Panel A involved 47 schools with random class assignments in Grade 7 but did not reassign students in Grade 8. The sample used in Panel B involved 62 schools with two Grade 7 classes but were excluded from our study sample. Student characteristics include age, gender, ethnicity, left-behind status, whether being the only child in the family, birth weight, and boarding status. Parent characteristics include education of both parents, marital status, and both parents' age when the child under discussion was born. Household characteristics include *hukou*, family social economics status, water facilities, toilet facilities, whether any family member drinks alcohol, whether any family member smokes. Class characteristics include class size, head teacher's gender, educational background, and teaching experience. Standard errors reported in parentheses, clustered at the grade level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

take time to manifest (e.g., Getik and Meier, 2022), but also risky behaviors take time to generate consequences on human-capital accumulation (e.g., Kremer and Levy, 2008). For example, Clark and Lohéac (2007) found a significant correlation between the one-year lagged peer drinking measure and one's drinking behavior. Similarly, Card and Giuliano (2013) estimated the correlation between the risky behavior of one's friends and her/his risky behavior one year later. Eisenberg et al. (2014) reported a significant correlation between the 7-8 months lagged peers' risky behaviors and one's risky behaviors. Duncan et al. (2005) found a significant deviant peer effect of college roommates' drinking behavior back in high school on one's drinking behavior in college. Kremer and Levy (2008) found significant one- and two-year lagged effects of drinking peers on college students' academic performance.

To examine the causal effects of having more female classmates on adolescents' physical health, two identification concerns need to be addressed. One concern is endogenous sorting. For example, if schools sort students into different classes based on certain unobserved factors that simultaneously determine the classroom gender composition and students' physical health, the estimated correlation between them may come out significant even if no causal relationship exists. We circumvent the potential non-random sorting problem by exploiting random class assignments of newly-enrolled students in the 47 sample schools (as noted before). Since Grade 7 is the first grade in China's junior high school education system, this assignment rule helps create a quasi-experiment in which newly-enrolled students are randomly mixed with classmates with different individual characteristics and family backgrounds. Hence, conditional on school fixed effects, the random classroom assignment helps overcome the influences of unobserved factors and thus address the concern of endogenous sorting.

The other concern in identifying the causal effect of having more female classmates on adolescents' physical health is contextual confounding. Contextual factors simultaneously affect all members within certain groups. For example, if a school assigned teachers who attach more importance to students' physical health to classes with higher proportions of female students by chance, we might end up observing a positive relationship between having more female classmates and physical health. We deal with this concern by controlling for a set of class characteristics along with school fixed effects in the estimation.[7]

Specifically, following Koedel et al. (2015), we employ a linear value-added model to identify the gender peer effects on students' health as follows,

$$\Delta Health_{ics} = \beta_0 + \beta_1 \overline{FP}^{G7}_{-i,cs} + \beta_2 Health^{G7}_{ics} + \mathbf{Z}'_{ics}\boldsymbol{\gamma} + \alpha_s + u_{ics} \quad (1)$$

where $\Delta Health_{ics}$ denotes the change in health indicators of student i in class c of school s from Grade 7 to Grade 8. $\overline{FP}^{G7}_{-i,cs}$ is the proportion of female classmates of student i in Grade 7 excluding student i herself/himself (as indexed by "-i"). $Health^{G7}_{ics}$ denotes the level of health of the student in Grade 7. $\mathbf{Z}$ is a set of covariates composed of four vectors. The first vector is student characteristics, including age, gender, ethnicity, left-behind status, whether being the only child in the family, birth weight, and boarding status. The second vector is parent characteristics, including years of schooling of both parents, marital status, and their ages at the birth of this child. The third vector is household characteristics, including *hukou*, family social economic status, water and toilet

[7] Although classrooms are randomly assigned, the choice of schools by households may be not random (e.g. Black, 1999; Burgess et al., 2015; Abdulkadiroğlu et al., 2020), which may bias our estimates when not controlling for school fixed effects. For example, if the proportion of girls in schools with better sanitation conditions happens to be low due to the nonrandom choice of schools by households, the gender peer effects may not come out significant without school fixed effects. Hence, to address the possible endogeneity of school choice, we take advantage of the random within-school across-classes variation in gender composition.

Table 4
Effects of the proportion of female classmates in Grade 7 on the change in adolescents' physical health indicators from Grade 7 to Grade 8

	(1)	(2)	(3)	(4)
Panel A: Full sample (Dependent variable: BmiAZ stays normal or goes from abnormal to normal, 1= yes)				
FP (7^{th})	0.337*	0.340*	0.348*	0.377**
	(0.182)	(0.181)	(0.178)	(0.169)
N	4242	4242	4242	4242
R^2	0.090	0.092	0.094	0.094
Panel B: Subsample of BmiAZ<-2 in wave one (Dependent variable: ΔBmiAZ)				
FP (7^{th})	3.045***	3.052***	2.862***	3.023***
	(0.776)	(0.802)	(0.804)	(0.834)
N	260	260	260	260
R^2	0.301	0.325	0.383	0.399
Panel C: Subsample of BmiAZ>1 in wave one (Dependent variable: ΔBmiAZ)				
FP (7^{th})	-0.021	-0.031	-0.211	-0.476
	(0.486)	(0.472)	(0.440)	(0.478)
N	884	884	884	884
R^2	0.322	0.328	0.351	0.362
Panel D: Full sample (Dependent variable: ΔHAZ)				
FP (7^{th})	0.255*	0.254*	0.253*	0.250**
	(0.145)	(0.137)	(0.130)	(0.112)
N	4158	4158	4158	4158
R^2	0.188	0.194	0.200	0.201
Panel E: Full sample (Dependent variable: staying non-Myopic or goes from Myopic to non-Myopic, 1= yes)				
FP (7^{th})	0.044	0.038	0.044	0.056
	(0.100)	(0.099)	(0.101)	(0.099)
N	4257	4257	4257	4257
R^2	0.803	0.804	0.804	0.804
Panel F: Full sample (Dependent variable: Δlogarithm of diopter of left lens of glasses)				
FP (7^{th})	-0.163**	-0.159**	-0.165**	-0.157**
	(0.073)	(0.073)	(0.073)	(0.060)
N	4012	4012	4012	4012
R^2	0.060	0.063	0.066	0.068
Panel G: Full sample (Dependent variable: Δlogarithm of diopter of right lens of glasses)				
FP (7^{th})	-0.168*	-0.166*	-0.177*	-0.169**
	(0.091)	(0.090)	(0.092)	(0.083)
N	4013	4013	4013	4013
R^2	0.065	0.067	0.070	0.071
Covariates:				
Student CHs	Yes	Yes	Yes	Yes
Parent CHs	No	Yes	Yes	Yes
Household CHs	No	No	Yes	Yes
Class CHs	No	No	No	Yes
School FEs	Yes	Yes	Yes	Yes

Notes: Student characteristics include age, gender, ethnicity, left-behind status, whether being the only child in the family, birth weight, and boarding status. Parent characteristics include education of both parents, marital status, and both parents' age when the child under discussion was born. Household characteristics include *hukou*, family social economics status, water facilities, toilet facilities, whether any family member drinks alcohol, whether any family member smokes. Class characteristics include class size, head teacher's gender, educational background, and teaching experience. Standard errors reported in parentheses, clustered at the grade level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

facilities, family members' health behaviors. The final vector is class characteristics, including class size, head teacher's gender, years of schooling, and teaching experience. α_s denotes school fixed effects. Standard errors are clustered at the grade level.

B. Evidence of random class assignments in Grade 7 in sample schools

The identification strategy of this paper depends on the assumption that the class assignment in Grade 7 in the 47 sample schools was random. Before presenting the results about the impact of having more female classmates on the physical health outcomes of junior high school students, we have to verify that the class assignment in Grade 7 was random in terms of their entry physical health status. To do so, we will take three approaches to verify simultaneously.

Our first approach is straightforward by examining the correlations between one's physical health outcomes and her/his classmates' gender composition in Grade 7. As shown in Panel A of Table 3, with the only exception of HAZ score, we did not find any significant correlation between physical health outcomes in Grade 7 and the proportion of female classmates (presented in odd columns of Panel A). When we further controlled for observable characteristics at the student, parental, household and class levels, as well as school fixed effects, none of the Grade 7 health outcome measures that we are interested in had any significant correlation with the proportion of female classmates (presented in even columns of Panel A). This lack of correlation provides evidence in support of the random assignment in Grade 7 in the 47 sample schools.

Our second approach is similar to a "falsification" test. Specifically, we estimated the correlations between one's physical health outcomes and the proportion of female classmates in Grade 7 in the 62 schools that have two Grade 7 classes but were not included in our study sample as they did not assign newly-enrolled students randomly in Grade 7 or reassigned students into different classes during the study period. Results based on data from these schools show that after controlling for observable characteristics at the student, parental, household and class levels, as well as school fixed effects, four out of the seven physical health measures are not correlated with the proportion of female classmates, whereas the rest three physical health measures are still significantly correlated with the proportion of female classmates (Table 3, Panel B, Columns 10, 12, and 14). This result lends further evidence in support of the random assignment in Grade 7 in the 47 sample schools.

The last approach is to conduct a simulation test by randomly reassigning sampled students in the 47 sample schools to two classes

Table 5
Mechanism analysis

	(1)	(2)	(3)	(4)	(5)
Panel A: Gender difference in health behaviors in Grade 7					
Outcomes (7^{th}):	Time spent on physical exercises weekly (min)	Spent more than one hour watching TV on a weekday (1=yes)	Spent more than one hour watching TV on a weekend day (1=yes)	Spent more than one hour playing online games on a weekday (1=yes)	Spent more than one hour playing online games on a weekend day (1=yes)
Coefficient on the female student dummy of this health behavior in Grade 7	-39.150***	-0.048***	-0.027*	-0.066***	-0.124***
	(4.78)	(0.01)	(0.01)	(0.01)	(0.02)
School FEs	Yes	Yes	Yes	Yes	Yes
N	3693	3864	3870	3865	3865
R^2	0.369	0.429	0.429	0.344	0.344
Panel B: Effects of the proportion of female classmates in Grade 7 on the change in adolescents' health behaviors from Grade 7 to Grade 8					
Outcomes:	ΔTime spent on physical exercises weekly (min)	Decreasing time spent on watching TV on a weekday (1=yes)	Decreasing time spent on watching TV on a weekend day (1=yes)	Decreasing time spent on playing online games on a weekday (1=yes)	Decreasing time spent on playing online games on a weekend day (1=yes)
FP (7^{th})	97.305**	0.379**	0.410***	0.339**	0.171
	(48.083)	(0.146)	(0.087)	(0.145)	(0.168)
Student CHs	Yes	Yes	Yes	Yes	Yes
Parent CHs	Yes	Yes	Yes	Yes	Yes
Household CHs	Yes	Yes	Yes	Yes	Yes
Class CHs	Yes	Yes	Yes	Yes	Yes
School FEs	Yes	Yes	Yes	Yes	Yes
N	3693	3864	3870	3865	3865
R^2	0.369	0.429	0.429	0.344	0.344

Notes: The estimation sample involved 47 schools with random class assignments in Grade 7 but did not reassign students during the study period. Student characteristics include age, gender, ethnicity, left-behind status, whether being the only child in the family, birth weight, and boarding status. Parent characteristics include education of both parents, marital status, and both parents' age when the child under discussion was born. Household characteristics include *hukou*, family social economics status, water facilities, toilet facilities, whether any family member drinks alcohol, whether any family member smokes. Class characteristics include class size, head teacher's gender, educational background, and teaching experience. Standard errors clustered at the grade level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

within the same grade for 1000 times. Then we calculate the standard deviation of the proportion of female classmates in a class based on the simulation results, and draw the kernel density of the standard deviation. As can be seen in Fig. A1, it seems the standard deviation fraction of females basically follows a normal distribution with a mean of 0.081, which is close to the actual standard deviation in our study sample (0.076). This simulation provides further evidence for our assumption of random assignment of students to classrooms.

Even if the class assignment was verified to be random in Grade 7, there is still a concern for the potential non-balanced allocation of educational resources within schools. Specifically, if different sample classes within the same sample school have different resources, say, class size or teacher qualifications, the estimated impacts of class gender composition on the physical health outcomes of students would still be biased. Therefore, it is necessary to check if the allocation of resources was "balanced" across classes within a school. To do so, we follow Wang et al. (2018) and Wang and Zhu (2021) to test whether the class means of observable characteristics of students and their households[8] are correlated with class characteristics. We focus on five class characteristics, namely class size, age, gender, years of schooling and years of teaching experiences of head teachers. The results show that out of the 90 correlation coefficients, the majority (82 or 91 percent) came out statistically insignificant after we controlled for school fixed effects (Table A2). The eight correlation coefficients (9 percent) might just come out significant by chance. Therefore, it seems reasonable to say that the allocation of resources was "balanced" across classes within sample schools.

The essence of our empirical strategy is to compare the change in physical health outcomes of students from two classes within the same school from Grade 7 to Grade 8, with the fact that one class has a relatively higher proportion of female students than the other purely by chance. Before conducting the estimations, another potential concern is that there is a deficiency in the within-school across-classes variation in gender composition. To address this concern, we follow Wang and Zhu (2021) to perform a regression of the class-level proportion of female classmates in Grade 7 on school fixed effects and then collect the residuals. Fig. A2, plotting both the original distribution of classmates' gender composition and its residuals obtained from the above regression, showed that there was still a reasonable amount of variation in the proportion of female peers across classes after controlling for school fixed effects.

Taken together, results from the above verifications/tests provide strong evidence that both students and educational resources were randomly assigned to Grade 7 classes within sample schools, and our female classmates measure has sufficient variations for identification purposes. This provides us with a unique opportunity to identify the impact of having more female classmates on the physical health outcomes of junior high school students.

Ⅳ. Results

A. Main results

Our regression results show that having more female classmates has a statistically significant and positive impact on adolescents' multiple physical health measures (Table 4). When we look at the BmiAZ, a one-percentage-point (pp) increase in the proportion of female classmates in Grade 7 is associated with 0.003-0.004 pp increase in one's probability

[8] Specifically, we checked seven characteristics at the student level (i.e., age, gender, ethnicity, left-behind status, whether being the only child in the family, birth weight, and boarding status), five characteristics at the parent level (i.e., education of both parents, marital status, and both parents' age when the child under discussion was born), and six characteristics at the household level (i.e., hukou status, family social economics status, water facilities, toilet facilities, whether any family member drinks alcohol, whether any family member smokes).

Table 6
Gender peer effects within homogeneous sub-groups

Outcomes:	(1) BmiAZ stays normal or goes from abnormal to normal (1=yes, full sample)	(2) ΔBmiAZ (subsample of BmiAZ<-2 in wave one)	(3) ΔBmiAZ (subsample of BmiAZ>1 in wave one)	(4) ΔHAZ (full sample)	(5) Staying non-Myopic or goes from Myopic to non-Myopic (1=yes, full sample)	(6) ΔLogarithm of diopter of left lens of glasses (full sample)	(7) ΔLogarithm of diopter of right lens of glasses (full sample)
FP in the following sub-groups (7^{th}):							
Same *hukou* and same left-behind status (β1)	0.125***	1.016***	-0.076	0.201**	-0.013	-0.047***	-0.039**
	(0.039)	(0.287)	(0.051)	(0.098)	(0.016)	(0.017)	(0.019)
Same *hukou* and different left-behind status (β2)	0.066***	-0.164	-0.061	-0.008	-0.007	-0.001	-0.019
	(0.024)	(0.222)	(0.088)	(0.027)	(0.013)	(0.012)	(0.015)
Different *hukou* and same left-behind status (β3)	0.059***	0.434	-0.050	0.013	0.020	-0.017	-0.018
	(0.018)	(0.362)	(0.083)	(0.039)	(0.015)	(0.017)	(0.019)
Different *hukou* and different left-behind status (β4)	0.043	0.281	-0.082	-0.054	-0.002	0.025	0.012
	(0.040)	(0.239)	(0.069)	(0.041)	(0.013)	(0.025)	(0.024)
Student CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parent CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4242	260	884	4158	4257	4012	4013
R^2	0.138	0.257	0.250	0.061	0.884	0.039	0.041
H0: β1=β2=β3=β4 (*p* value)	0.1928	**0.0149**	0.2187	**0.0593**	0.1047	**0.0003**	**0.0108**

Notes: The estimation sample involved 47 schools with random class assignments in Grade 7 but did not reassign students during the study period. Student characteristics include age, gender, ethnicity, left-behind status, whether being the only child in the family, birth weight, and boarding status. Parent characteristics include education of both parents, marital status, and both parents' age when the child under discussion was born. Household characteristics include *hukou*, family social economics status, water facilities, toilet facilities, whether any family member drinks alcohol, whether any family member smokes. Class characteristics include class size, head teacher's gender, educational background, and teaching experience. Standard errors clustered at the grade level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

of staying normal BmiAZ or going from abnormal BmiAZ to normal BmiAZ from Grade 7 to Grade 8 (Panel A).[9] This result may be driven by those students who were wasted in the first wave, as a one pp increase in their proportion of female classmates in Grade 7 increases their BmiAZ score by 0.029-0.031 (Panel B) whereas the estimates are not significant for those students who were over-weighted in the first wave (Panel C).

As for HAZ, our results also show that a one pp increase in the proportion of female classmates in Grade 7 is associated with a 0.003 increase in one's HAZ score from Grade 7 to Grade 8. The point estimates are quite robust to different model specifications conditional on school fixed effects (Panel D).

With regard to eyesight, although the estimates of one's change in myopic status from Grade 7 to Grade 8 are not significant, having more female classmates is associated with a reduction in diopters. Specifically, a one pp increase in the proportion of female classmates in Grade 7 is associated with a 0.157-0.165% (Panel F) or 0.166-0.177% (Panel G) decrease in the diopter of her/his left or right lens of glasses, respectively. Our findings are consistent with previous estimates, especially those that found positive effects of having more female classmates on adolescents' human-capital outcomes (e.g., Black et al., 2013; Gong et al., 2021; Hill, 2015; Hoxby, 2000; Hu, 2015; Lavy and Schlosser, 2011). The estimated coefficients on control variables are also quite informative and consistent with previous findings.[10]

B. Mechanism analysis

The results reported above show that adolescents in classes with a higher proportion of female classmates tend to be in better physical health. Why is it like this? A close examination of the literature reveals that gender peer effects may work through various channels (Hu, 2015; Briole, 2021; Gong et al., 2021).[11] One important channel that has been verified in several studies is the improved peer environment brought by

[9] In this study, replacing one boy with one girl is equivalent to a two pp increase in the proportion of girls in Grade 7, according to the means of class sizes and the proportion of female students in wave one.

[10] For examples, having an in-door flush toilet is found to be positively associated with the BmiAZ score of those with low starting BmiAZ, which have been widely observed in developing countries (Esrey, 1996; Pickering, 2015). The eyesight of left-behind children tends to drop relatively more than their non-left-behind peers. While few papers in the literature have examined the correlation between adolescents' left-behind status and their eyesight, several studies have documented the negative impacts of left-behind status on various health measures (e.g., Zhao and Yu, 2016; Jin et al., 2020). Detailed results are available upon request.

[11] For example, Hu (2015) found that the benefits of female peers on student academic performance may be related to a better learning environment brought by a higher proportion of girl peers. Briole (2021) found suggestive evidence that gender peer effects partially operated through adjustments in student and teacher behaviors based on the gender composition of the classroom. Gong et al. (2021) further found that strengthened teacher behaviors, greater student efforts, and an improved classroom environment were the primary channels through which peers' gender influences student outcomes.

table 7
Heterogeneity in gender peer effects on adolescents' physical health outcomes

Outcomes:	(1) BmiAZ stays normal or goes from abnormal to normal (1=yes, full sample)	(2) ΔBmiAZ (subsample of BmiAZ<-2 in wave one)	(3) ΔBmiAZ (subsample of BmiAZ>1 in wave one)	(4) ΔHAZ (full sample)	(5) Staying non-Myopic or goes from Myopic to non-Myopic (1=yes, full sample)	(6) ΔLogarithm of diopter of left lens of glasses (full sample)	(7) ΔLogarithm of diopter of right lens of glasses (full sample)	(8) BmiAZ stays normal or goes from abnormal to normal (1=yes, full sample)	(9) ΔBmiAZ (subsample of BmiAZ<-2 in wave one)	(10) ΔBmiAZ (subsample of BmiAZ>1 in wave one)	(11) ΔHAZ (full sample)	(12) Staying non-Myopic or goes from Myopic to non-Myopic (1=yes, full sample)	(13) ΔLogarithm of diopter of left lens of glasses (full sample)	(14) ΔLogarithm of diopter of right lens of glasses (full sample)
	Boy							Girl						
FP (7^{th})	0.607***	3.101***	0.197	0.384*	-0.021	-0.190*	-0.229**	0.219	2.152	0.870	0.050	0.099	-0.122	-0.184*
	(0.217)	(0.968)	(0.610)	(0.224)	(0.164)	(0.107)	(0.107)	(0.232)	(1.396)	(1.071)	(0.188)	(0.109)	(0.100)	(0.102)
N	2186	135	618	2130	2195	2065	2070	2056	125	262	2028	2062	1947	1943
R^2	0.091	0.459	0.385	0.207	0.793	0.080	0.091	0.179	0.487	0.604	0.278	0.823	0.099	0.095
	Father's education < Median							Father's education ≥ Median						
FP (7^{th})	0.497***	3.993***	-1.481*	0.436***	0.106	-0.176**	-0.242*	0.313	2.438**	0.503	0.061	-0.016	-0.175*	-0.138
	(0.176)	(1.227)	(0.735)	(0.136)	(0.083)	(0.071)	(0.130)	(0.330)	(1.089)	(0.695)	(0.159)	(0.188)	(0.099)	(0.094)
N	2097	150	384	2052	2107	1972	1972	2144	110	498	2106	2149	2039	2040
R^2	0.119	0.423	0.509	0.209	0.811	0.087	0.092	0.121	0.553	0.428	0.240	0.794	0.090	0.094
	Mather's education < Median							Mather's education ≥ Median						
FP (7^{th})	0.305*	5.817***	-1.495***	0.274*	-0.041	-0.221*	-0.262**	0.341	0.827	-1.057	0.238*	0.198	-0.167*	-0.179*
	(0.165)	(1.649)	(0.517)	(0.162)	(0.092)	(0.113)	(0.107)	(0.283)	(1.001)	(0.770)	(0.123)	(0.181)	(0.091)	(0.101)
N	2092	149	375	2056	2104	1965	1963	2149	111	502	2102	2152	2046	2049
R^2	0.151	0.447	0.467	0.191	0.818	0.090	0.102	0.100	0.543	0.458	0.246	0.789	0.092	0.094
	Rural schools							Urban schools						
FP (7^{th})	0.144	2.407**	0.572	0.224	0.066	-0.237**	-0.314*	0.540	3.355	0.181	0.238	-0.123	-0.082	-0.108
	(0.186)	(1.090)	(0.720)	(0.141)	(0.107)	(0.106)	(0.170)	(0.332)	(2.254)	(0.687)	(0.232)	(0.207)	(0.070)	(0.090)
N	1491	91	290	1454	1499	1391	1387	2751	169	594	2704	2758	2621	2626
R^2	0.110	0.369	0.367	0.213	0.823	0.070	0.079	0.107	0.618	0.421	0.244	0.776	0.082	0.082
	Class size < Median							Class size ≥ Median						
FP (7^{th})	-0.029	4.531***	0.552	0.140	0.107	-0.190	-0.215	0.532***	4.984***	0.046	0.217	-0.064	-0.097**	-0.095*
	(0.313)	(1.274)	(0.736)	(0.224)	(0.167)	(0.114)	(0.150)	(0.136)	(1.543)	(0.337)	(0.166)	(0.083)	(0.043)	(0.049)
N	1990	123	407	1941	1998	1852	1852	2251	137	476	2217	2258	2159	2160
R^2	0.079	0.449	0.383	0.213	0.799	0.080	0.093	0.151	0.581	0.465	0.220	0.814	0.082	0.080
Student CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parent CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The estimation sample involved 47 schools with random class assignments in Grade 7 but did not reassign students during the study period. Student characteristics include age, gender, ethnicity, left-behind status, whether being the only child in the family, birth weight, and boarding status. Parent characteristics include education of both parents, marital status, and both parents' age when the child under discussion was born. Household characteristics include *hukou*, family social economics status, water facilities, toilet facilities, whether any family member drinks alcohol, whether any family member smokes. Class characteristics include class size, head teacher's gender, educational background, and teaching experience. Standard errors clustered at the grade level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8
Heterogeneity in gender peer effects on adolescents' health behaviors

Outcomes:	(1) ΔTime spent on physical exercises weekly (min)	(2) Decreasing time spent on watching TV on a weekday (1=yes)	(3) Decreasing time spent on watching TV on a weekend day (1=yes)	(4) Decreasing time spent on playing online games on a weekday (1=yes)	(5) Decreasing time spent on playing online games on a weekend day (1=yes)	(6) ΔTime spent on physical exercises weekly (min)	(7) Decreasing time spent on watching TV on a weekday (1=yes)	(8) Decreasing time spent on watching TV on a weekend day (1=yes)	(9) Decreasing time spent on playing online games on a weekday (1=yes)	(10) Decreasing time spent on playing online games on a weekend day (1=yes)
	Boy					Girl				
FP (7^{th})	121.436	0.402*	0.456**	0.494***	0.458***	-6.391	0.430**	0.269	0.253	0.162
	(129.379)	(0.214)	(0.184)	(0.105)	(0.158)	(52.587)	(0.199)	(0.190)	(0.183)	(0.236)
N	1878	1967	1972	1965	1971	1815	1897	1898	1900	1894
R^2	0.301	0.166	0.154	0.274	0.251	0.596	0.122	0.131	0.306	0.225
	Father's education < Median					Father's education ≥ Median				
FP (7^{th})	44.206	0.422**	0.244	0.438***	0.354*	105.512	0.150	0.506**	0.118	0.221
	(81.820)	(0.203)	(0.226)	(0.116)	(0.179)	(86.515)	(0.181)	(0.204)	(0.236)	(0.171)
N	1784	1869	1871	1863	1861	1908	1995	1998	2002	2003
R^2	0.420	0.195	0.148	0.287	0.244	0.347	0.120	0.126	0.247	0.280
	Mather's education < Median					Mather's education ≥ Median				
FP (7^{th})	-0.881	0.440**	0.332	0.450***	0.436**	151.729	0.158	0.447**	0.189	0.129
	(64.338)	(0.193)	(0.215)	(0.121)	(0.170)	(93.522)	(0.186)	(0.195)	(0.194)	(0.201)
N	1759	1864	1860	1856	1853	1933	2000	2010	2009	2012
R^2	0.412	0.182	0.152	0.269	0.259	0.355	0.120	0.121	0.245	0.279
	Rural schools					Urban schools				
FP (7^{th})	10.914	0.536**	0.190	0.556***	0.255*	73.747	0.536***	-0.291	0.341*	-0.152
	(69.238)	(0.201)	(0.207)	(0.123)	(0.147)	(131.035)	(0.161)	(0.285)	(0.194)	(0.236)
N	1260	1343	1342	1334	1334	2433	2521	2528	2531	2531
R^2	0.410	0.196	0.132	0.294	0.226	0.350	0.122	0.125	0.232	0.287
	Class size < Median					Class size ≥ Median				
FP (7^{th})	181.360	0.382**	0.303	-0.047	0.232	10.494	0.568***	0.052	0.354**	0.449***
	(110.979)	(0.175)	(0.181)	(0.127)	(0.260)	(69.898)	(0.133)	(0.263)	(0.169)	(0.088)
N	1672	1775	1781	1774	1774	2021	2089	2089	2091	2091
R^2	0.318	0.164	0.157	0.259	0.254	0.437	0.143	0.096	0.286	0.237
Student CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parent CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The estimation sample involved 47 schools with random class assignments in Grade 7 but did not reassign students during the study period. Student characteristics include age, gender, ethnicity, left-behind status, whether being the only child in the family, birth weight, and boarding status. Parent characteristics include education of both parents, marital status, and both parents' age when the child under discussion was born. Household characteristics include *hukou*, family social economics status, water facilities, toilet facilities, whether any family member drinks alcohol, whether any family member smokes. Class characteristics include class size, head teacher's gender, educational background, and teaching experience. Standard errors clustered at the grade level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

a higher proportion of female classmates (Hu, 2015; Gong et al., 2021). As far as physical health is concerned, an improved peer environment can be less exposure to health-risk behaviors as females are less likely to involve in such behaviors than their male peers during adolescence (Byrnes et al., 1999; Shakya et al., 2019; Myers, 2010; Kritsotakis et al., 2016), which in turn, yields beneficial spillovers (Eisenberg et al., 2014; Fletcher, 2012). Therefore, we propose a hypothesis that having more female classmates enhances the physical health outcomes of adolescents by helping create a protective environment against a broad range of risk factors and positively shaping their health behaviors.

To test this hypothesis, we conducted a three-step exercise. In the first step, we took advantage of the rich information collected in CEPS and created five variables to measure student's health behaviors that have been identified by the literature on adolescents' physical health: time spent on physical exercises weekly, time spent on watching TV on a weekday, time spent on watching TV on a weekend day, time spent on playing online games on a weekday, time spent on playing online games on a weekend day.

In the second step, we compare whether there was any significant difference in these five health behavior variables by gender when the students were in Grade 7 using school fixed effects model. Results show that female students are less likely to engage in four out of the five health-risk behaviors than their male peers in Grade 7 (Panel A of Table 5). Specifically, girls are less likely than boys to spend more than one hour watching TV or playing online games both on a weekday and on a weekend day. In contrast, girls tend to spend less time than boys on physical exercises, which is partly due to the gender-related differences in physical activity patterns in adolescence and is consistent with a large body of previous studies (French et al., 1994; Troiano et al., 2008; Kritsotakis et al., 2016).

In the final step, we further explore whether the improved peer environment brought by a higher proportion of female classmates indeed results in any increase (decrease) in adolescents' healthy (unhealthy) behaviors. Results from school fixed effects models show that having more female classmates in Grade 7 significantly improves four out of the five adolescents' health behaviors from Grade 7 to Grade 8 (Panel B of Table 5), which echo the results from step two above (Panel A of Table 5). Specifically, a one pp increase in the proportion of female classmates in Grade 7 is associated with 0.973 minutes more on physical exercises per week, a 0.004 pp increase in one's probability of decreasing time spent watching TV on both a weekday and a weekend day, as well as a 0.003 pp increase in one's probabilities of decreasing time spent playing online games on a weekday. In contrast, having more female classmates has no impact on one's probability of decreasing time spent on playing online games on a weekend day.

Taken together, these results provide evidence supporting the

Table 9
Robustness check of controlling for the classroom averages of observed covariates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcomes:	BmiAZ stays normal or goes from abnormal to normal (1=yes, full sample)	ΔBmiAZ (subsample of BmiAZ<-2 in wave one)	ΔBmiAZ (subsample of BmiAZ>1 in wave one)	ΔHAZ (full sample)	Staying non-Myopic or goes from Myopic to non-Myopic (1=yes, full sample)	ΔLogarithm of diopter of left lens of glasses (full sample)	ΔLogarithm of diopter of right lens of glasses (full sample)
FP in Grade 7	0.381***	3.518***	-0.895	0.335**	0.020	-0.179*	-0.224**
	(0.118)	(1.051)	(0.815)	(0.152)	(0.145)	(0.099)	(0.109)
Pre-determined classroom peers' characteristics in Grade 7 (on average or proportion):							
Personality CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parent CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other covariates:							
Student CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parent CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class CHs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4242	260	884	4158	4257	4012	4013
R^2	0.110	0.510	0.436	0.225	0.806	0.081	0.088

Notes: The estimation sample involved 47 schools with random class assignments in Grade 7 but did not reassign students during the study period. Besides controlling for covariates at the student, parent, household, and class levels, as well as a full set of (47) school fixed effects in the estimation, all models reported in this table further control for 21 pre-determined peer characteristics including four personality characteristics of peers (including conscientiousness, social interaction, emotional stability, and self-efficacy) and seventeen classmates' averages of observed covariates at the student (including age, ethnicity, left-behind status, whether being the only child in the family, birth weight, and boarding status), parent (including education of both parents, marital status, and both parents' age when the child under discussion was born) and household (including hukou, family social economics status, water facilities, toilet facilities, whether any family member drinks alcohol, whether any family member smokes) level. Standard errors clustered at the grade level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

hypothesis that a higher proportion of female classmates does help to improve the peer environment and healthier behaviors.

C. Gender peer effects within homogeneous sub-groups

Is it possible that the gender peer effects observed above might have been driven by female peers within homogeneous sub-groups? This needs to be considered given the possible subgroup formation within classrooms (Carrell et al., 2013), especially a positive selection into subgroups with similar classmates (Currarini, et al., 2009; McPherson et al., 2001). For example, previous studies have found larger spillovers among homogeneous than among heterogeneous classroom peer groups (Opper, 2019; Lu and Anderson, 2015).

Following Opper (2019), we take a three-step approach to explore this possibility. The first step is to classify one's classmates into sub-groups by their homogeneity. Specifically, based on two dummies of student characteristics (namely, hukou and left-behind status) that have been frequently used in conducting sub-group analyses as well as in the targeting of policies or intervention programs in China (Liu, 2005; Huang et al., 2015), four mutually exclusive sub-groups of classmates are constructed: with the same hukou and the same left-behind status as the student under discussion, with the same hukou but different left-behind status as the student under discussion, with different hukou but the same left-behind status as the student under discussion, with different hukou and different left-behind status as the student under discussion. With these classifications, the second step is to calculate the proportion of female classmates within each sub-group. The last step is to replace the proportion of female classmates variable in Equation (1) with the four proportions that we just constructed and reran the regressions.

Results in Table 6 show that for four out of the seven physical health outcomes that we examined, the observed gender peer effects are indeed driven by female peers within homogeneous sub-groups. Specifically, for BmiAZ score for the subsample of BmiAZ<-2 in wave one, the HAZ score, as well as the logarithm of diopter of left and right lens of glasses, it is only in the sub-group where classmates share the same hukou and left-behind status as the student under discussion where the proportion of female classmates come out statistically significant. In contrast, the effect of having more female classmates on the other three health outcomes (including the transition in BmiAZ or myopia category from the first to the second wave, the BmiAZ score for the subsample with BmiAZ >1 in wave one) is not statistically different between sub-groups. However, it should be kept in mind that this result does not say anything about how classmates interact within classrooms. Future studies examining students' subgroup formation in the classroom are needed.

D. Heterogeneity in gender peer effects on adolescents' health outcomes and health behaviors

There is increasing evidence that peer effects vary by student, parent, household, and school characteristics (Black et al., 2013; Hu, 2015; Lu and Anderson, 2015; Gong et al., 2021; Wang and Zhu, 2021). In this sub-section, we further investigate the potential heterogeneous effects of having more female classmates on adolescents' physical health outcomes (behaviors) by repeating analyses reported in Table 4 (Panel B of Table 5) using sub-samples, results are presented in Table 7 (Table 8). We focus on five dimensions of subgroups: student gender (Panels A-B), maternal and paternal education (Panels C-F), school location (Panels G-H), and class size (Panels I-J).

Results from heterogeneous effects on health outcomes reveal some informative patterns (Table 7). First, we find that boys benefit more from having more female classmates than girls, which is consistent with previous studies that also show apparent gender differences in gender peer effects (e.g., Black et al., 2013; Hu, 2015; Lu and Anderson, 2015). One possible explanation for this gender heterogeneity might be that boys are more sensitive than girls to the improved classroom environment accompanied by a higher share of female classmates (Diette and Ruth, 2014; Legewie and DiPrete, 2012). Second, students with less-educated mothers or fathers tend to benefit more from having more female classmates in terms of most physical health outcomes. For these

students, female peers may serve as a potential remedy for their relatively disadvantaged family backgrounds in human capital production (Hu et al., 2021). Finally, students attending schools with unfavorable conditions (e.g., those located in rural areas or with larger classes) also benefit more from having more female classmates, which further provides suggestive evidence for the potential substitutability between school characteristics and peer characteristics in the role of shaping adolescents' physical health (Viner et al., 2012).

In addition, results from heterogeneous effects on adolescents' health behaviors also show some sub-groups of students benefit more from having more female classmates in certain health outcomes (Table 8). Specifically, boys benefit more in terms of the probabilities of decreasing time spent watching TV on a weekend day and the time playing video games on both a weekday and a weekend day from Grade 7 to Grade 8. In contrast, girls benefit more in terms of the probability of decreasing time spent on watching TV on a weekday from Grade 7 to Grade 8. Both students with less-educated parents and those attending schools with unfavorable conditions (e.g., those located in rural areas or with larger classes) benefit more in terms of the probabilities of decreasing time spent watching TV on a weekday and the time playing video games on both a weekday and a weekend day from Grade 7 to Grade 8. However, both students with better-educated mothers and those with better-educated fathers benefit more in terms of the probability of decreasing time spent on watching TV on a weekend day. Overall, these heterogeneous effects on health behaviors are consistent with those that we have found on health outcomes.[12] Such consistency further supports the hypothesis that the observed gender peer effects on adolescents' physical health result from the healthier behaviors brought by a higher proportion of female classmates.[13]

E. Robustness check

Were the observed gender peer effects driven by peer characteristics other than gender? Some recent studies have found that peers' personalities (Golsteyn, Non, & Zölitz, 2021) and household characteristics (Wang and Zhu, 2019; Yin et al., 2020) exert a significant impact on adolescents' academic performance and non-cognitive skills. If these pre-determined peer characteristics are correlated with the gender composition of classmates, the positive impact of having more female classmates on one's physical health outcomes that we observed above might have picked up certain spillovers of these peer characteristics.

To deal with this concern and also as a robustness check, we reran Equation (1) in Table 9 by including pre-determined peer characteristics and tested whether the estimated coefficient associated with the proportion of female classmates (our key explanatory variable) would change. Following the literature (Poropat, 2009; Rahmani and Lavasani, 2012; Wang and Zhu, 2019; Yin et al., 2020), we include 21 pre-determined peer characteristics, including four personality characteristics of peers (conscientiousness, social interaction, emotional stability, and self-efficacy)[14] and 17 characteristics that measure the classmates' averages of covariates at the student (exclude gender), parent and household levels, which we have described in the "Variables" sub-section under the "Data" section above.

Results from the robustness check show that the estimated coefficient associated with the proportion of female classmates remained substantially the same even after we included these characteristics of peers. This result provides evidence that the effect of having more female classmates on adolescents' physical health outcomes was not driven by other characteristics of their peers.

V. Conclusions and Discussions

In this paper, we have tried to understand whether having more female classmates would benefit adolescents' physical health outcomes. By exploiting the random within-school across-classes variation in gender composition brought by the random class assignments of 4260 newly-enrolled students in 47 junior high schools in China, we took a quasi-experimental approach and identified significant positive impacts of having more female classmates on adolescents' physical health. We also explored the mechanism underlying these research findings as the higher the proportion of female classmates, the more (less) likely to engage in healthy (unhealthy) behaviors during adolescence. Beyond these, we further provide suggestive evidence that the observed effects on certain health measures were mainly driven by female classmates within homogeneous sub-groups (specifically in terms of the *hukou* and left-behind status). Results from heterogeneity analyses show that the beneficial effects of having more female classmates on health outcomes are more pronounced among boys, students with less-educated parents, and those attending schools with unfavorable conditions (i.e., rural schools or schools with bigger class sizes), which are consistent with the observed heterogeneous gender peer effects on health behaviors.

We acknowledge at least three limitations of our study. First, due to data constraints, we were only able to examine the impact of having more female classmates in Grade 7 on adolescents' change in physical health outcomes one year later (from Grade 7 to Grade 8). Classmates may play different roles in different grades as the stock of health capital would evolve as adolescent advances to higher grades. Future studies based on data from longer panels may be able to detect interesting dynamic patterns of classroom gender peer effects.

Second, we defined the proportion of female classmates at the class level. Yet, as we found, the observed effects may be driven by female classmates within homogeneous sub-groups. Hence, such a definition of the gender composition of classmates may be susceptible to measurement errors that mask peer interactions within sub-groups of classmates. If data permit, better measurements of the gender composition of classmates should be constructed to identify students' actual peer groups, such as social network analysis. Despite that, our estimates indeed provide evidence in support of the health benefits of having more female classmates, which can be interpreted as a lower bound of the actual benefits of having more female classmates.

Finally, an inherent concern in the literature on gender peer effect is that the observed effect might be driven by characteristics of the classmates that are correlated with the gender of classmates per se. While we did show the estimated effects of having more female classmates on adolescents' physical health outcomes remain substantially the same even when we included a set of peer characteristics, we still cannot rule out this possibility. However, from a policy perspective, we believe it is still important to know that exposure to female peers can have positive

[12] The three exceptions are girls benefit more in terms of the probability of decreasing time spent on watching TV on a weekday from Grade 7 to Grade 8, students with better-educated mothers and those with better-educated fathers benefit more in terms of the probability of decreasing time spent on watching TV on a weekend day. Nonetheless, we do not think these three exceptions violate our interpretation, considering they echo the positive gender peer effects on their eyesight in Table 7 (column 14, Panel B; column 13, Panel D; columns 13-14, Panel F).

[13] We also investigated whether there is nonlinearity in gender peer effects by including a squared term of classmates' gender composition in Equation (1) and reran the regressions. However, we did not find any evidence in favor of the nonlinearity hypothesis in our study context. The results are available upon request.

[14] These four personality measures are adapted from the "Big Five" personality traits that have been widely used in psychological studies (Poropat, 2009; Rahmani and Lavasani, 2012). The score of each measure was by standardizing the summed raw score that a student got for each self-valued statement under each dimension (Please refer to Table A3 for a copy of the instrument).

effects on adolescents' physical health, at least in the context of China.

Nevertheless, despite these limitations, we believe that the present study provides useful information on the existence, magnitude, and underlying mechanisms of having more female classmates on adolescents' physical health outcomes. These findings shed light on educational policies that seek to improve adolescents' physical health in at least three aspects. We found that having more female classmates helps improve adolescents' physical health by improving their health behaviors. This finding has implications for practices that aim to promote health by controlling risk behaviors. For example, school principals might consider assigning teachers with more positive health attitudes to classes with fewer female students. And teachers are suggested to pay more attention to creating a protective environment against adolescents' health-risk behaviors (such as too much exposure to TV or online games, etc.) to mitigate the adverse impacts of having fewer female students on adolescents' health outcomes. Moreover, we found that boys, students with less-educated parents, and those attending rural schools or schools with bigger class sizes tend to benefit more from having more female classmates. This implies that rearranging these students with a gender perspective might serve as a low-cost opportunity for improving aggregate health outcomes. Finally, our results suggest that the observed gender peer effects are partly driven by female peers within homogeneous sub-groups. This finding has implications for the group organization within sub-classroom microenvironments in order to improve adolescents' physical health.

Author Contributions

Conceptualization, Y.G., S.C., C.L.; methodology, Y.G., S.L., S.C., Y.T., and C.L.; software, Y.G.; validation, Y.G., S.C., and C.L.; formal analysis, Y.G., S.C., and S.L.; investigation, Y.G., C.L. and S.L.; resources, C.L.; data curation, Y.G.; writing—original draft preparation, Y.G., S.L., and, C.L.; writing—review and editing, Y.G., S.C., Y.T., and C.L.; visualization, Y.G.; supervision, C.L.; project administration, C.L.; funding acquisition, C.L. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data that support the findings of this study are available in Chinese National Survey Data Archive at http://www.cnsda.org/index.php?r=projects/view&id=72810330 and http://www.cnsda.org/index.php?r=projects/view&id=61662993.

Declaration of Competing Interest

The authors declare no conflict of interest.

Data Availability

Data will be made available on request.

Acknowledgments

The authors acknowledge the financial support of the National Natural Science Foundation of China (grant numbers 71925009, 72141014 and 71861147003).

Appendix A

Tables A1,A2

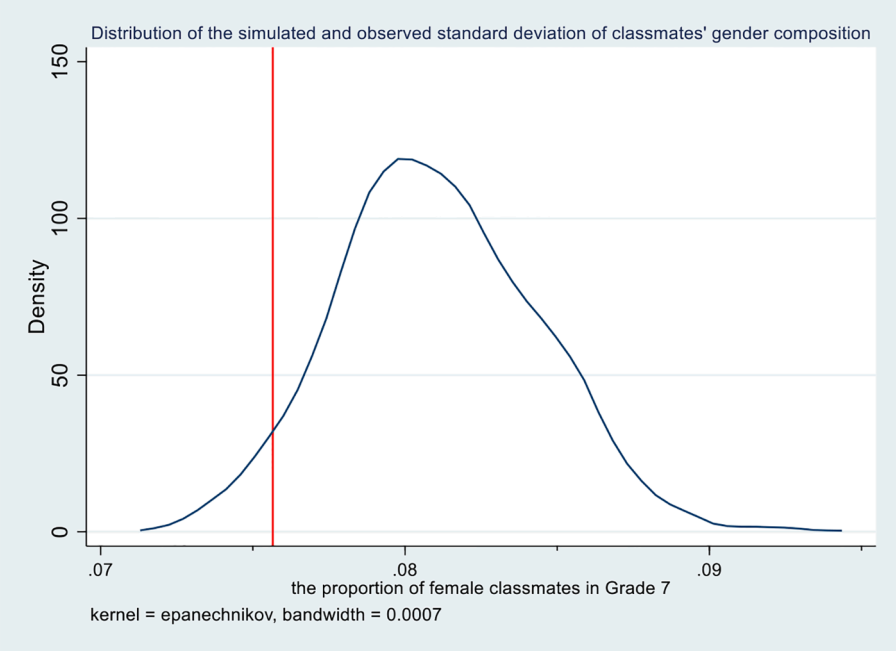

Fig. A1. Distribution of the simulated and observed standard deviation of classmates' gender composition in Grade 7
Notes: The blue line depicted the distribution of the simulated standard deviation of the proportion of female classmates which was calculated by randomly re-assigning sampled students to two classes within grade for 1000 times. The red line depicted the observed actual standard deviation of the proportion of female classmates in our sample.

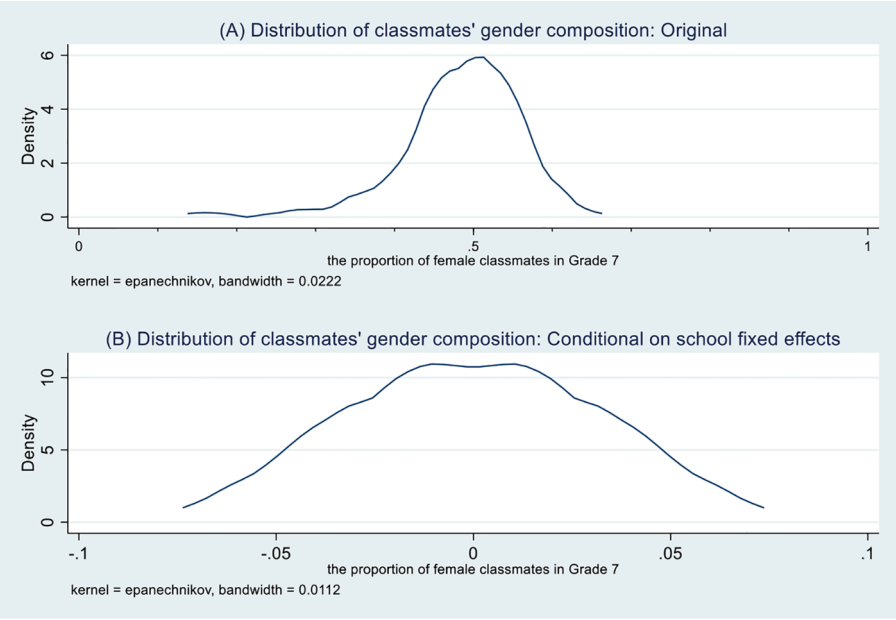

Fig. A2. Distributions of classmates' gender composition in Grade 7
Notes: The analysis reported in this figure is done at the class level. (A) Original distribution of classmates' gender composition in Grade 7. (B) Conditional distribution of classmates' gender composition in Grade 7, which is the distribution of residuals obtained from regressing the proportion of female classmates in Grade 7 on school fixed effects.

Table A1
External validity test

	(1)	(2)	(3)	(4)	(5)
	Schools with two Grade 7 classes but were excluded from our study sample		Schools with random class assignments in Grade 7, but did not reassign students in Grade 8		Diff
	Mean	SD	Mean	SD	(3)-(1)
Panel A: School level					
Enrollment	363.134	[257.141]	359.420	[228.090]	-3.714
The share of local students	0.766	[0.203]	0.837	[0.176]	0.070*
The share of rural students	0.545	[0.275]	0.534	[0.291]	-0.011
The share of left-behind students	0.218	[0.151]	0.233	[0.162]	0.014
Public school or not	1.067	[0.252]	1.061	[0.242]	-0.005
School's ranking within the county	3.817	[0.911]	4.020	[0.661]	0.204
School's location	2.533	[1.501]	2.837	[1.663]	0.303
Scores of students' misbehavior in school	9.850	[2.767]	9.673	[1.886]	-0.177
Student-teacher ratio	13.073	[4.227]	12.783	[4.795]	-0.290
Panel B: Principle level					
Age	44.857	[4.875]	44.191	[6.749]	-0.666
Gender	0.845	[0.365]	0.813	[0.394]	-0.032
Education	15.786	[1.217]	15.667	[1.117]	-0.119
Whether graduated from normal university or majored in teaching	0.983	[0.131]	1	[0]	0.017
Working experience	11.732	[6.011]	10.542	[7.593]	-1.190
Panel C: Class level					
Classroom gender composition	0.470	[0.064]	0.484	[0.052]	0.014
Class size	45.867	[11.075]	45.837	[11.068]	-0.030
The share of local students in class	0.739	[0.227]	0.814	[0.193]	0.075*
The share of rural students in class	0.435	[0.260]	0.423	[0.273]	-0.012
The share of left-behind students in class	0.158	[0.148]	0.185	[0.172]	0.027
Panel D: Headteacher level					
Average age	35.958	[5.407]	35.044	[5.504]	-0.914
The share of male head teachers	0.226	[0.336]	0.267	[0.331]	0.041
Average education	15.599	[0.585]	15.755	[0.846]	0.156
Average teaching experience	15.655	[6.246]	13.602	[6.824]	-2.053
The share of head teachers graduated from normal university or majored in teaching	0.934	[0.120]	0.893	[0.162]	-0.041
The average professional title of head teachers	2.633	[0.666]	2.715	[0.604]	0.082
The share of head teachers awarded in the past three years	0.584	[0.283]	0.598	[0.296]	0.014
Number of schools	**62**		**47**		

Notes: The sample involved 109 schools with two Grade 7 classes. Comparing 26 observable characteristics at the school, principal, class, and headteacher levels in Grade 7. * $p < 0.1$
** $p < 0.05$
*** $p < 0.01$.

Table A2
Balancing test for the allocation of educational resources across classes within schools in Grade 7

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	Average student age	Proportion of girls	Prop. of ethnic minority	Prop. of left-behind children	Prop. of only child	Average student Birth weight	Proportion of boarding students	Average years of paternal education	Average years of maternal education	Prop. of students parents married	Average mother's age at birth of this child	Average father's age at birth of this child	Prop. of students with rural " Hukou"	Prop. of students with "not poor" family SES	Prop. of students households have tap water	Prop. of students households have flush toilet	Prop. of students family members drink alcohol	Prop. of students family members smoke
Class size	0.009	0.003	-0.002*	0.003	-0.001	0.005	0.013	-0.001	0.009	0.001	-0.006	0.004	0.006	-0.000	0.003	0.001	-0.001	-0.003
	(0.053)	(0.002)	(0.001)	(0.003)	(0.004)	(0.013)	(0.013)	(0.021)	(0.030)	(0.003)	(0.033)	(0.041)	(0.005)	(0.003)	(0.002)	(0.003)	(0.003)	(0.005)
Teacher's:																		
Age	-0.044	-0.001	-0.001	0.001	0.002	-0.002	-0.005	0.016	0.008	0.000	-0.020	-0.006	0.001	0.005	0.002	-0.001	0.002	0.002
	(0.053)	(0.002)	(0.001)	(0.003)	(0.004)	(0.008)	(0.004)	(0.021)	(0.023)	(0.003)	(0.027)	(0.034)	(0.004)	(0.003)	(0.002)	(0.001)	(0.003)	(0.004)
Gender	0.490*	-0.009	-0.004	0.017	0.008	-0.004	-0.005	-0.231	-0.036	-0.015	0.149	0.082	0.019	-0.036	0.018*	-0.003	0.012	0.017
	(0.267)	(0.014)	(0.010)	(0.018)	(0.022)	(0.075)	(0.023)	(0.141)	(0.160)	(0.016)	(0.214)	(0.208)	(0.027)	(0.032)	(0.010)	(0.018)	(0.024)	(0.029)
Education	-0.065	-0.001	-0.001	-0.006	-0.006	0.035	-0.011	0.009	0.132	-0.007	-0.167*	-0.092	-0.015	0.018	0.009*	-0.008*	0.013	0.002
	(0.226)	(0.007)	(0.004)	(0.011)	(0.016)	(0.024)	(0.011)	(0.088)	(0.124)	(0.006)	(0.096)	(0.108)	(0.012)	(0.014)	(0.005)	(0.005)	(0.014)	(0.014)
Teaching experience	0.041	0.000	0.001	-0.002	-0.001	-0.001	0.002	-0.010	-0.001	0.000	0.052***	0.040**	-0.000	-0.001	-0.000	0.001	0.001	-0.001
	(0.026)	(0.001)	(0.001)	(0.002)	(0.003)	(0.003)	(0.002)	(0.012)	(0.012)	(0.001)	(0.012)	(0.017)	(0.003)	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)
School FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	156.702***	0.407***	0.209**	0.141	0.606*	3.018***	-0.077	10.546***	7.704***	0.961***	31.795***	27.876***	0.365	2.519***	0.564***	0.943***	0.055	0.617**
	(4.815)	(0.139)	(0.087)	(0.270)	(0.307)	(0.699)	(0.447)	(1.699)	(2.407)	(0.187)	(2.157)	(2.601)	(0.256)	(0.244)	(0.140)	(0.128)	(0.286)	(0.292)
N	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94
R^2	0.978	0.837	0.984	0.913	0.944	0.757	0.953	0.964	0.958	0.940	0.918	0.868	0.955	0.918	0.984	0.982	0.852	0.855

Notes: The analysis reported in this table is done at the class level. Standard errors are reported in parentheses. * $p < 0.1$
** $p < 0.05$
*** $p < 0.01$.

Table A3
Personalities measurement in the CEPS in Grade 7

Dimensions	Items	Scores	Mean	SD	N
Conscientiousness	Even if I feel a little uncomfortable or have other reasons to stay home, I would still try my best to go to school.	Strongly disagree = 1; Somewhat disagree = 2;	3.371	0.851	4260
	Even if I don't like my homework, I would try my best to finish it.	Somewhat agree = 3;	3.407	0.788	4260
	Even if my homework takes a long time to finish, I would try my best to do it.	Strongly agree = 4.	3.520	0.751	4260
Social interaction	Most of the classmates are nice to me.	Strongly disagree = 1;	3.338	0.817	4260
	My class is in good atmosphere.	Somewhat disagree = 2;	3.280	0.859	4260
	I often participate in school/class activities.	Somewhat agree = 3;	2.963	0.998	4260
	I feel very close to people at school.	Strongly agree = 4.	3.103	0.917	4260
Emotional stability	Frequency of feeling blue in the past seven days	Always = 1;	2.142	0.985	4260
	Frequency of feeling depressed in the past seven days	Often = 2;	1.851	1.026	4260
	Frequency of feeling unhappy in the past seven days	Sometimes = 3;	2.146	1.035	4260
	Frequency of feeling life has no meaning in the past seven days	seldom = 4;	1.662	1.034	4260
	Frequency of feeling pessimistic in the past seven days	Never = 5.	1.909	1.001	4260
Self-efficacy	I was able to express myself clearly.	Strongly disagree = 1;	3.208	0.789	4260
	I was able to give quick responses.	Somewhat disagree = 2;	3.122	0.768	4260
	I could learn new things very quickly.	Somewhat agree = 3;	3.110	0.795	4260
	I was curious about new things.	Strongly agree = 4.	3.526	0.773	4260

Notes: The sample involved 47 schools with random class assignments in Grade 7 but did not reassign students in Grade 8.

References

Abdulkadiroğlu, A., Pathak, P., Schellenberg, J., & Walters, C. (2020). Do parents value school effectiveness? *American Economic Review, 110*(5), 1502–1539.

Alexander, C., Piazza, M., Mekos, D., & Valente, T. (2001). Peers, schools, and adolescent cigarette smoking. *Journal of Adolescent Health, 29*(1), 22–30.

Anteghini, M., Fonseca, H., Ireland, M., & Blum, R. (2001). Health risk behaviors and associated risk and protective factors among Brazilian adolescents in Santos, Brazil. *Journal of Adolescent Health, 28*, 295–302.

Bertrand, M., & Pan, J. (2013). The trouble with boys: Social influences and the gender gap in disruptive behavior. *American Economic Journal: Applied Economics, 5*(1), 32–64.

Black, S., Devereux, P., & Salvanes, K. (2013). Under pressure? The effect of peers on outcomes of young adults. *Journal of Labor Economics, 31*(1), 119–153.

Black, S. (1999). Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics, 114*(2), 577–599.

Blum, R., Beuhring, T., Shew, M., Bearinger, L., Sieving, R., & Resnick, M. (2000). The effects of race/ethnicity, income, and family structure on adolescent risk behaviors. *American Journal of Public Health, 90*(12), 1879–1884.

Boehmer, T., Hoehner, C., Deshpande, A., Ramirez, L., & Brownson, R. (2007). Perceived and observed neighborhood indicators of obesity among urban adults. *International Journal of Obesity, 31*(6), 968–977.

Braun, C., Freidlin, V., Sperduto, R., Milton, R., & Strahlman, E. (1996). The progression of myopia in school age children: data from the Columbia Medical Plan. *Ophthalmic epidemiology, 3*(1), 13–21.

Briole, S. (2021). Are girls always good for boys? Short and long term effects of school peers' gender. *Economics of Education Review, 84*, Article 102150.

Bullimore, M., & Brennan, N. (2019). Myopia control: Why each diopter matters. *Optometry and Vision Science, 96*(6), 463–465.

Burgess, S., Greaves, E., Vignoles, A., & Wilson, D. (2015). What parents want: school preferences and school choice. *Economic Journal, 125*(587), 1262–1289.

Byrnes, J., Miller, D., & Schafer, W. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin, 125*(3), 367–383.

Campbell, R., Starkey, F., Holliday, J., Audrey, S., Bloor, M., Parry-Langdon, N., & Moore, L. (2008). An informal school-based peer-led intervention for smoking prevention in adolescence (ASSIST): a cluster randomised trial. *Lancet, 371*(9624), 1595–1602.

Card, D., & Giuliano, L. (2013). Peer effects and multiple equilibria in the risky behavior of friends. *The Review of Economics and Statistics, 95*(4), 1130–1149.

Carrell, S., Sacerdote, B., & West, J. (2013). From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica, 81*(3), 855–882.

Clark, A., & Lohéac, Y. (2007). It wasn't me, it was them!" Social influence in risky behavior by adolescents. *Journal of Health Economics, 26*(4), 763–784.

Cornwell, C., Mustard, D., & Van Parys, J. (2013). Noncognitive skills and the gender disparities in test scores and teacher assessments: Evidence from primary school. *Journal of Human Resources, 48*(1), 236–264.

Currarini, S., Jackson, M., & Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica, 77*(4), 1003–1045.

Currie, C., Molcho, M., Boyce, W., Holstein, B., Torsheim, T., & Richter, M. (2008). Researching health inequalities in adolescents: The development of the Health Behaviour in School-Aged Children (HBSC) Family Affluence Scale. *Social Science & Medicine, 66*(6), 1429–1436.

Diette, T., & Ruth, U. (2014). Gender and race heterogeneity: The impact of students with limited English on native Students' performance. *American Economic Review, 104*(5), 412–417.

Duckworth, A., & Seligman, M. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of educational psychology, 98*(1), 198.

Due, P., Damsgaard, M., Rasmussen, M., Holstein, B., Wardle, J., Merlo, J., & Plus, H. (2009). Socioeconomic position, macroeconomic environment and overweight among adolescents in 35 countries. *International Journal of Obesity, 33*(10), 1084–1093.

Duncan, G., Boisjoly, J., Kremer, M., Levy, D., & Eccles, J. (2005). Peer effects in drug use and sex among college students. *Journal of Abnormal Child Psychology, 33*(3), 375–385.

Eisenberg, D., Golberstein, E., & Whitlock, J. (2014). Peer effects on risky behaviors: New evidence from college roommate assignments. *Journal of Health Economics, 33*, 126–138.

Elgar, F., Pförtner, T., Moor, I., De Clercq, B., Stevens, G., & Currie, C. (2015). Socioeconomic inequalities in adolescent health 2002–2010: A time-series analysis of 34 countries participating in the Health Behaviour in School-aged Children study. *Lancet, 385*(9982), 2088–2095.

Epple, D., Romano, R., Benhabib, J., Bisin, A., & Jackson, M. O. (2011). *Peer effects in education: A survey of the theory and evidence*. Chapter 20. In *Handbook of Social Economics, 1* pp. 1053–1163). North-Holland.

Esrey, S. (1996). Water, waste, and well-being: A multicountry study. *American Journal of Epidemiology, 143*(6), 608–623.

Fletcher, J. (2012). Peer influences on adolescent alcohol consumption: evidence using an instrumental variables/fixed effect approach. *Journal of Population Economics, 25*, 201–218.

Freeman, M., Greene, L., Dreibelbis, R., Saboori, S., Muga, R., Brumback, B., & Rheingans, R. (2012). Assessing the impact of a school-based water treatment, hygiene and sanitation programme on pupil absence in Nyanza Province, Kenya: a cluster-randomized trial. *Tropical Medicine & International Health, 17*(3), 380–391.

French, S., Perry, C., Leon, G., & Fulkerson, J. (1994). Food preferences, eating patterns, and physical-activity among adolescents - correlates of eating disorders symptoms. *Journal of Adolescent Health, 15*(4), 286–294.

Gauderman, W., Avol, E., Gilliland, F., Vora, H., Thomas, D., Berhane, K., & Peters, J. (2004). The effect of air pollution on lung development from 10 to 18 years of age. *New England Journal of Medicine, 351*(11), 1057–1067.

Getik, D., & Meier, A. (2022). Peer gender and mental health. *Journal of Economic Behavior & Organization, 197*, 643–659.

Goldschmidt, E., & Jacobsen, N. (2014). Genetic and environmental effects on myopia development and progression. *Eye, 28*(2), 126–133.

Golsteyn, B., Non, A., & Zölitz, U. (2021). The impact of peer personality on academic achievement. *Journal of Political Economy, 129*(4), 1052–1099.

Gong, J., Lu, Y., & Song, H. (2021). Gender peer effects on students' academic and noncognitive outcomes evidence and mechanisms. *Journal of Human Resources, 56*(3), 686–710.

Guo, Y., Li, S., Yang, W., & Liu, C. (2021). The peer effects of gender on student mental health: Quasi-experimental evidence from the China education panel survey. *Peking University Education Review, 19*(02), 19–40. +188. (in Chinese).

Hill, A. (2015). The girl next door: the effect of opposite gender friends on high school achievement. *American Economic Journal: Applied Economics, 7*(3), 147–177.

Holden, B., Fricke, T., Wilson, D., Jong, M., Naidoo, K., Sankaridurg, P., & Resnikoff, S. (2016). Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology, 123*(5), 1036–1042.

Hoxby, C. (2000). Peer effects in the classroom: learning from gender and race variation. *National Bureau of Economic Research (NBER) Working Paper 7867*.

Hu, F. (2015). Do girl peers improve your academic performance? *Economics Letters, 137*, 54–58.

Hu, N., Yuan, M., Liu, J., Coplan, R., & Zhou, Y. (2021). Examining reciprocal links between parental autonomy-support and Children's peer preference in Mainland China. *Children-Basel, 8*(6), 508.

Huang, Y., Zhong, X., Li, Q., Xu, D., Zhang, X., Feng, C., & Deng, B. (2015). Health-related quality of life of the rural-China left-behind children or adolescents and

influential factors: a cross-sectional study. *Health and Quality of Life Outcomes, 13*(1), 29.
Jacob, B. (2002). Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education Review, 21*(6), 589–598.
Jin, X., Chen, W., Sun, I., & Liu, L. (2020). Physical health, school performance and delinquency: A comparative study of left-behind and non-left-behind children in rural China. *Child Abuse & Neglect, 109*, Article 104707.
Jones, L., Sinnott, L., Mutti, D., Mitchell, G., Moeschberger, M., & Zadnik, K. (2007). Parental history of myopia, sports and outdoor activities, and future myopia. *Investigative Ophthalmology & Visual Science, 48*(8), 3524–3532.
Kim, J. (2021). The quality of social relationships in schools and adult health: differential effects of student-student versus student-teacher relationships. *School Psychology, 36*(1), 6–16.
Klasen, F., Oettingen, G., Daniels, J., Post, M., Hoyer, C., & Adam, H. (2010). Posttraumatic resilience in former Ugandan child soldiers. *Child Development, 81*(4), 1096–1113.
Koedel, C., Mihaly, K., & Rockoff, J. (2015). Value-added modeling: A review. *Economics of Education Review, 47*, 180–195.
Kohrt, B., Jordans, M., Tol, W., Perera, E., Karki, R., Koirala, S., & Upadhaya, N. (2010). Social ecology of child soldiers: child, family, and community determinants of mental health, psychosocial well-being, and reintegration in Nepal. *Transcultural Psychiatry, 47*(5), 727–753.
Kremer, M., & Levy, D. (2008). Peer effects and alcohol use among college students. *Journal of Economic Perspectives, 22*(3), 189–206.
Kritsotakis, G., Psarrou, M., Vassilaki, M., Androulaki, Z., & Philalithis, A. (2016). Gender differences in the prevalence and clustering of multiple health risk behaviours in young adults. *Journal of advanced nursing, 72*(9), 2098–2113.
Lavy, V., & Schlosser, A. (2011). Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics, 3*(4), 1–33.
Lee, Y., Lo, C., Sheu, S., & Lin, J. (2013). What factors are associated with myopia in young adults? A survey study in Taiwan military conscripts. *Investigative Ophthalmology & Visual Science, 54*(2), 1026–1033.
Legewie, J., & DiPrete, T. (2012). School context and the gender gap in educational achievement. *American Sociological Review, 77*(3), 463–485.
Lim, L., Gong, Y., Ah-Kee, E., Xiao, G., Zhang, X., & Yu, S. (2014). Impact of parental history of myopia on the development of myopia in mainland china school-aged children. *Ophthalmology and eye diseases, 6*, 31–35.
Liu, Z. (2005). Institution and inequality: the Hukou system in China. *Journal of Comparative Economics, 33*(1), 133–157.
Lu, F., & Anderson, M. (2015). Peer effects in microenvironments: The benefits of homogeneous classroom groups. *Journal of Labor Economics, 33*(1), 91–122.
Malcolm, S., Huang, S., Cordova, D., Freitas, D., Arzon, M., Jimenez, G., & Prado, G. (2013). Predicting condom use attitudes, norms, and control beliefs in Hispanic problem behavior youth: The effects of family functioning and parent-adolescent communication about sex on condom use. *Health Education & Behavior, 40*(4), 384–391.
Manski, C. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies, 60*(3), 531–542.
McPherson, M., Smith-Lovin, L., & Cook, J. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27*, 415–444.
Ministry of Education of China. (2014). *Educational Statistics Yearbook of China 2014*. Beijing: People's Education Press.
Morgan, I., Ohno-Matsui, K., & Saw, S. (2012). Ophthalmology 2 myopia. *Lancet, 379*(9827), 1739–1748.
Myers, L. (2010). Health risk behaviors among adolescents in the rural south: A comparison of race, gender, and age. *Journal of Human Behavior in the Social Environment, 20*, 1024–1037.
National Health Commission of the People's Republic of China. (2021). In 2020, the overall myopia rate of children and adolescents was 52.7%; and the problem of myopia aging is still prominent [Online]. Available: http://health.people.com.cn/n1/2021/0713/c14739-32156389.html [Accessed 2022-09-02 2022].
Opper, I. (2019). Does helping john help sue? Evidence of spillovers in education. *American Economic Review, 109*(3), 1080–1115.
Pan, C., Ramamurthy, D., & Saw, S. (2012). Worldwide prevalence and risk factors for myopia. *Ophthalmic and Physiological Optics, 32*(1), 3–16.
Patton, G., Sawyer, S., Santelli, J., Ross, D., Afifi, R., Allen, N., & Viner, R. (2016). Our future: a Lancet commission on adolescent health and wellbeing. *Lancet, 387*(10036), 2423–2478.
Pickering, A., Djebbari, H., Lopez, C., Coulibaly, M., & Laura Alzua, M. (2015). Effect of a community-led sanitation intervention on child diarrhoea and child growth in rural Mali: a cluster-randomised controlled trial. *Lancet Global Health, 3*(11), E701–E711.
Poropat, A. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*, 322–338.
Rahmani, S., & Lavasani, M. (2012). Gender differences in five factor model of personality and sensation seeking. *Procedia-Social and Behavioral Sciences, 46*, 2906–2911.
Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: a systematic review. *Social Science & Medicine, 90*, 24–31.
Rhee, K., Lumeng, J., Appugliese, D., Kaciroti, N., & Bradley, R. (2006). Parenting styles and overweight status in first grade. *Pediatrics, 117*(6), 2047–2054.
Rose, K., Morgan, I., Ip, J., Kifley, A., Huynh, S., Smith, W., & Mitchell, P. (2008). Outdoor activity reduces the prevalence of myopia in children. *Ophthalmology, 115*(8), 1279–1285.
Shakya, H., Domingue, B., Nagata, J., Cislaghi, B., Weber, A., & Darmstadt, G. (2019). Adolescent gender norms and adult health outcomes in the USA: a prospective cohort study. *Lancet Child Adolescent Health, 3*(8), 529–538.
Smith, E., Hung, L., & Arumugam, B. (2014). Visual regulation of refractive development: insights from animal studies. *Eye, 28*(2), 180–188.
Stewart, A. (1982). The reliability and validity of self-reported weight and height. *Journal of Chronic Diseases, 35*(4), 295–309.
Stommel, M., & Schoenborn, C. (2009). Accuracy and usefulness of BMI measures based on self-reported weight and height: findings from the NHANES & NHIS 2001-2006. *BMC Public Health, 9*(1), 421.
Torsheim, T., Currie, C., Boyce, W., Kalnins, I., Overpeck, M., & Haugland, S. (2004). Material deprivation and self-rated health: a multilevel study of adolescents from 22 European and North American countries. *Social Science & Medicine, 59*(1), 1–12.
Troiano, R., Berrigan, D., Dodd, K., Masse, L., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and Science in Sports and Exercise, 40*(1), 181–188.
United Nations Inter-Agency Group for Child Mortality Estimation (UN IGME). (2020). *Levels and trends in child mortality: Report*. https://www.unicef.org/reports/levels-and-trends-child-mortality-report-2020.
Viner, R., Ozer, E., Denny, S., Marmot, M., Resnick, M., Fatusi, A., & Currie, C. (2012). Adolescence and the social determinants of health. *Lancet, 379*(9826), 1641–1652.
Wang, H., & Zhu, R. (2019). China's only children and their spillover effects on academic performance in the classroom. *The B.E. Journal of Economic Analysis & Policy, 19*(4), 1–10.
Wang, H., & Zhu, R. (2021). Social spillovers of China's left-behind children in the classroom. *Labour Economics, 69*, Article 101958.
Wang, H., Cheng, Z., & Smyth, R. (2018). Do migrant students affect local students' academic achievements in urban China? *Economics of Education Review, 63*, 64–77.
Ward, J., Azzopardi, P., Francis, K., Santelli, J., Skirbekk, V., Sawyer, S., & Viner, R. (2021). Global, regional, and national mortality among young people aged 10-24 years, 1950-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet, 398*(10311), 1593–1618.
World Health Organization. (2006). *WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development*. Geneva, Switzerland: World Health Organization.
Yang, M., Mamy, J., Zhou, L., Liao, Y., Wang, Q., Seewoobudul, V., & Hao, W. (2014). Gender differences in prevalence and correlates of antisocial personality disorder among heroin dependent users in compulsory isolation treatment in China. *Addictive Behaviors, 39*(3), 573–579.
Yin, G., Huang, H., & Huang, W. (2020). The intergenerational spillover of human capital-evidence from "Parents of Others. *China Economic Quarterly, 19*(04), 1491–1514 (in Chinese).
Zhang, M., Li, L., Chen, L., Lee, J., Wu, J., Yang, A., & Congdon, N. (2010). Population density and refractive error among Chinese children. *Investigative Ophthalmology & Visual Science, 51*(10), 4969–4976.
Zhao, F., & Yu, G. (2016). Parental migration and rural left-behind Children's mental health in China: A meta-analysis based on mental health test. *Journal of Child and Family Studies, 25*(12), 3462–3472.
Zhao, J., Mao, J., Luo, R., Li, F., Munoz, S., & Ellwein, L. (2002). The progression of refractive error in school-age children: Shunyi District, China. *American Journal of Ophthalmology, 134*(5), 735–743.