The Impact of Low-Ability Peers on Cognitive and Noncognitive Outcomes

Random Assignment Evidence on the Effects and Operating Channels

Di Xu
Qing Zhang
Xuehan Zhou

ABSTRACT

This article presents new experimental estimates of the impact of low-ability peers on own outcomes using nationally representative data from China. We exploit the random assignment of students to junior high school classrooms and find that the proportion of low-ability peers, defined as having been retained during primary school ("repeaters"), has negative effects on nonrepeaters' cognitive and noncognitive outcomes. An exploration of the mechanisms shows that a larger proportion of repeater peers is associated with reduced after-school study time. The negative effects are driven by male repeaters and are more pronounced among students with less strict parental monitoring at home.

Di Xu is an associate professor of Education Policy and Social Context at the School of Education, University of California Irvine (dix3@uci.edu). Qing Zhang is a postdoctoral fellow at the Instituite for Policy Research at Northwestern University. Xuehan Zhou is a doctoral student at the School of Education, University of California, Irvine. *The authors of this article declare no conflicts of interest with respect to the research, authorship, and publication of this article. The authors are extremely grateful to Rachel Baker, Damon Clark, Greg Duncan, Paco Martorell, Jonah Rockoff, and two anonymous reviewers for their valuable comments and suggestions. The data used in this article can be obtained upon request from the website of Chinese National Survey Data Archive (http://www.cnsda.org/index.php?r=projects/view&id=72810330).*

[Submitted July 2018; accepted December 2019]; doi:10.3368/jhr.57.2.0718-9637R2
JEL Classification: I21, I28, and D62
ISSN 0022-166X E-ISSN 1548-8004
Supplementary materials are freely available online at: http://uwpress.wisc.edu/journals/journals/jhr-supplementary.html

Di Xu https://orcid.org/0000-0002-8160-4783
Qing Zhang https://orcid.org/0000-0002-2967-6069
Xuehan Zhou https://orcid.org/0000-0002-1750-3494

I. Introduction

Peer effects are central to understanding the education production function because, if they exist, the composition and characteristics of peers could potentially affect own behaviors, preferences, and performance. Yet, the econometric difficulties of estimating peer effects have been well documented in the literature. Any study that attempts to provide a causal estimate of peer effects on own outcomes is subject to several methodological challenges, including self-selection into peer groups, the simultaneous influence of peers' and own outcomes—the so-called "reflection problem" (Manski 1993), common shocks that make it difficult to separate the peer effects from other shared treatment effects, and measurement error that can lead to overestimation of peer effects in settings without random assignment (Feld and Zölitz 2017). Ideal data for providing a clean estimate of peer effects therefore would need to contain orthogonal-to-baseline peer group variation, preexisting measures that precisely capture peers' ability and are unlikely to have been affected by own ability, and clear distinction between the subjects of a peer effects investigation and the peers who provide the mechanism for causal effects (Angrist 2014).

While the field has started to gather experimental evidence on peer effects in education settings where students are assigned to peer groups exogenously, the majority of these studies occurred at the postsecondary education level (for example, Booij, Leuven, and Oosterbeek 2017; Carrell, Fullerton, and West 2009; Carrell, Sacerdote, and West 2013; Duflo, Dupas, and Kremer 2011; Lyle 2007; Sacerdote 2001; Zimmerman 2003). In contrast, existing studies that examine peer effects at the K–12 level typically exploit exogenous between-cohort variations in fixed student characteristics (such as gender) or prior achievement to identify peer effects (for example, Ammermueller and Pischke 2009; Bifulco, Fletcher, and Ross 2011; Burke and Sass 2013; Carrell and Hoekstra 2010; Carrell, Hoekstra, and Kuka 2018; Gould, Lavy, and Paserman 2009; Hoxby 2000; Lavy and Schlosser 2011; Lefgren, 2004). One caveat with this approach, however, is that peer characteristics at the grade or cohort level may only serve as a rough approximation of the peer interactions at primary and middle schools, since students typically spend more hours with their classmates and limited hours with their other schoolmates. Additionally, due to data limitations, there is far less empirical evidence or consensus on the potential mechanisms through which peer effects operate. Despite all of this, understanding the operating channels of peer effects is important because it would inform policies or interventions to optimize the education production process and outcomes.

This work provides experimental evidence on peer effects and possible mechanisms in middle schools by examining whether having low-ability peers, defined as ever being retained during primary school (referred to as "repeaters" hereafter), has any effect on the cognitive and noncognitive outcomes of nonrepeater classroom peers. We exploit a unique setting where junior high students in China are randomly assigned to classes upon initial school enrollment. It is important to note that if having low-ability peers indeed impacts a junior high student's academic performance and motivation, then these influences occur at a critical juncture in the life-cycle. An extensive literature indicates that academic choices and career aspirations based on individual aptitudes, self-concept, and values are formulated during early adolescence (Eccles, Vida, and Barber 2004; Wang

2013). More importantly, the nine-year compulsory education ends at Grade 9 in China, and junior high graduates are then required to choose between different academic paths that may lead to distinct educational attainment and labor market outcomes.[1]

We begin by documenting the differences between repeaters and nonrepeaters in observed characteristics. Descriptive statistics show that repeaters are not only consistently associated with lower academic performance relative to nonrepeaters, but also are more likely to experience negative emotions, show lower levels of school engagement, and have lower educational expectations. These strong correlations motivate our main question: Do higher proportions of repeaters affect nonrepeaters' cognitive and noncognitive outcomes? Our subsequent analyses based on the random assignment design indicate that having greater proportions of repeaters in the classroom significantly affects nonrepeaters' academic performance, cognitive assessment score, and school engagement.

Drawing on the rich information included in the survey, we also examine three possible channels through which peer effects may operate: (i) student perceived student–teacher interaction at school, (ii) peer relationship and classroom atmosphere, and (iii) daily study hours after school. Our findings show that reduced study time after school is the most robust channel among the three, although the size of the effect is fairly small. We also find that being exposed to a larger proportion of repeater peers increases a nonrepeater's probability of having a "best friend" who regularly plays at internet cafés, which has been seen as one of the main causes for school absenteeism and neglect of daily routine among Chinese teenagers (Reuters 2007). These results provide suggestive evidence that one of the main channels through which troubled kids negatively affect their peers is through social networks and joint activities after school.

Taken together, the results from this study make two distinct contributions to the existing literature on peer effects. First, we provide clean estimates of having repeater peers on own outcomes in a unique setting, based on a natural experimental design where middle school students are randomly assigned to classes, and therefore class peers, within a school. A flurry of studies that are closely related to ours have examined how peers with particularly low academic ability or high potential of being disruptive may influence own outcomes (for example, Aizer 2008; Carrell and Hoekstra 2010; Carrell, Hoekstra, and Kuka 2018; Figlio 2007; Lavy, Paserman, and Schlosser 2012). This strand of research typically exploits exogenous variations in peer composition across cohorts and concludes that exposure to low-ability or disruptive peers not only has negative impact on short-term academic outcomes, but also has long-run educational and labor market consequences.

We build on this literature but extend it through a research design with several advantages that make our estimates less susceptible to bias. First, the random assignment of students to classes purges our estimates of bias from potential confounders associated

1. Junior high graduates in China are required to choose between a high school that may eventually lead to postsecondary education or a vocational school that is oriented towards obtaining occupation-specific skills. Students who choose to attend a high school are also required to choose whether to enter the STEM track or the non-STEM track. Students cannot easily switch tracks after they make their initial choice because each track prepares students for a content-specific college entrance exam. Accordingly, students' academic performance, educational aspirations, and attitudes toward different subjects formulated in junior high are likely to influence their decisions of academic paths.

with peer quality. This setting also allows us to examine peer effects at the classroom level, which is arguably a better approximation of peer interactions than at the grade or cohort level. In addition, the status of being a repeater was determined during primary school, which means that repeaters likely had limited opportunities to interact with their junior high peers before being labeled repeaters, since they started primary school in different cohorts. Most importantly, Angrist (2014) points out that peer effects may be biased, even in a randomly assigned setting, due to a mechanical relationship between the measures of own and peer ability. Following his recommendation, our research design is able to address this mechanical relationship by making a clear distinction between the subjects of a peer effects investigation (that is, nonrepeaters) and the peers who provide the mechanism for causal effects (that is, repeaters).

Second, our data include new noncognitive measures that not only enable us to include important social-emotional and behavioral measures as key outcomes in addition to academic performance, but also provide insight into the mechanisms driving the peer effects. Extensive evidence suggests that noncognitive measures, such as school engagement and educational aspirations, serve as strong predictors of lifelong success, even conditional on academic ability (Borghans et al. 2008; Fredricks, Blumenfeld, and Paris 2004; Moffitt et al. 2011). The data set used in our study, the China Education Panel Survey (CEPS), is the first nationally representative survey for junior high students in China. The CEPS not only includes student test scores in each subject, but also directly collected information on noncognitive measures, such as mental stress, school absenteeism, educational expectations, and social and emotional engagement with the school. Findings regarding these outcomes hence assist us to achieve a more comprehensive understanding of peer effects.

Additionally, the CEPS data also include important measures of the educational process, such as student's perceived interactions with teachers, student's perceived classroom environment and peer relationships, as well as after-school study time. Understanding the impact of repeater peers on these measures therefore would contribute to the small but emerging literature on the operating channels of peer effects, which have yielded rather mixed findings (Booij, Leuven, and Oosterbeek, 2017; Duflo, Dupas, and Kremer 2011; Feld and Zölitz 2017; Gong, Lu, and Song, 2021; Lavy, Paserman, and Schlosser 2012). For example, based on middle school students' answers to a "school environment survey" in Israel, Lavy, Paserman, and Schlosser (2012) find evidence for peer effects on both group functioning and teacher functioning, where a high proportion of low-ability students negatively influences teachers' pedagogical practices, raises the level of disruption and violence within the class, and worsens both student–teacher relationships and interstudent relationships. In the context of Kenyan primary schools, Duflo, Dupas, and Kremer (2011) also find evidence for classroom ability composition on teacher functioning, where teachers assigned to a class of high-achieving students display more effort. In contrast, based on data from university contexts, both Booij, Leuven, and Oosterbeek (2017) and Feld and Zölitz (2017) find evidence for peer effects on group functioning but no evidence on teacher functioning.

Taken together, these findings indicate that the channels through which peers influence own outcomes are highly dependent on the specific contexts. Our study builds on the existing literature and examines peer effects in a new setting. Moreover, in addition to perceived activities and interpersonal relationships at school, we are able to include

after-school study time as a possible mechanism, which to our knowledge, has never been studied in the previous literature. Understanding the impact of peers on own behaviors and activities outside of the classroom is particularly important, as evidence that supports this operating channel would imply that adolescents may be negatively influenced by troubled peers on a more proactive and voluntary basis, rather than simply through class time spent together at school.

The rest of the paper proceeds as follows. Section II introduces the specific context of junior high and grade retention in China. Section III describes the CEPS data and provides descriptive information about the characteristics of repeaters relative to non-repeaters. Section IV reviews the methodological challenge in estimating peer effects and presents our methodological approach in addressing these challenges. Section V presents and discusses our results and underlying mechanisms of the estimated peer effects. Section VI concludes the paper with a brief discussion of the interpretation of the findings.

II. Background

A. Compulsory Education in China and Class Assignment

Education in China is a state-administered system of public education, where the Ministry of Education standardizes textbooks and curriculum and enforces national education standards. In an effort to attain a universal education for all school-aged children, the government enacted the Law on Nine-Year Compulsory Education in 1986, which requires nine years of free and universal education in the country (six years of primary education and three years of junior high school education). The local authorities are responsible for following the guidelines formulated by the central authorities and implementing nine-year compulsory education tailored to local conditions.

Before the 1990s, all students in primary schools were required to take the junior high school entrance exam administered either by a district or by individual schools. After that, students were placed into different junior high schools on the basis of their exam scores. Starting in the mid-1990s, the Ministry of Education reformed the compulsory education system to promote equal and fair opportunities for all students and canceled the junior high entrance exam. After graduating from elementary school (at the end of the sixth grade), students would directly enter a state-run junior high school, typically based on the location of their hukou (an official household registration record that identifies a person's residency status in an area). According to the Ministry of Education (Ministry of Education of the People's Republic of China 2015a,b), in 2014 there were 201,000 primary schools and 52,000 junior high schools in China. On average four primary schools feed into one junior high school. Students can only attend schools assigned by their hukou location, which are the schools closest to their neighborhoods. Since the assigned schools are usually in the vicinity of the students' community, it is common for students to walk or bike to school both at the primary and junior high level (Li and Liu 2014). To enforce the policy implementation, the local governments closely monitor the school assignment process and prevent schools from charging extra fees for admitting students from other districts (Organization for Economic Co-operation and Development 2016).

Partly due to the lack of clear indicators of prior academic achievement and partly following the Ministry of Education's promotion of "equal and fair opportunity for all students," the majority of junior high schools in China now follow random assignment of students to classes upon initial school enrollment, where students are expected to remain within the initial homeroom class assignment throughout their junior high years (Grade 7–9).[2] However, as students move beyond the seventh grade, teachers and school administrators increasingly gather knowledge about each student's behavior and academic performance at school, and hence they are more likely to take actions to improve student outcomes, such as communicating with the parents, transferring students between classes, or allocating more resources and time to target lower performing students or students showing problematic behaviors at school. To minimize possible bias due to teachers' and schools' responses to students' academic performance and behaviors, we therefore restrict our analyses to schools that randomly assign students to classes upon initial school enrollment and focus on students in the seventh grade only, as this is the time when schools have the least knowledge of students' academic ability and are least likely to take additional actions to respond to students' academic behaviors and performance.

B. Grade Retention in China

Grade retention has been widely used in China as a solution to address insufficient grade-level achievements (for example Chen et al. 2009; Zhang 2014). The formulation and implementation of grade retention policies have been mostly decided by local school administrators following provincial policy guidelines. Following the Law on Nine-Year Compulsory Education in 1986, many provinces and municipalities had specific regulations on when a student should be retained in primary and junior high schools. The criteria to retain students vary across grades, schools, areas, and provinces. In general, while schools are required to follow provincial policies regarding grade retention, the specific decision is made primarily at the school level mainly on the basis of academic performance (Chen et al. 2009). Students are typically retained when they receive an F in one or more core courses (such as Chinese and math) or in multiple supplemental subjects (such as PE) (Chen 2013). In Beijing, for example, "students in primary schools who fail the annual examinations in either Chinese or math should take make-up tests; those who fail the make-up tests should be retained. In junior high schools, students who fail five or more subjects in the annual examinations should be retained" (Chen 2013).

In 1994, in line with the objectives to promote equal and fair opportunity for all students regardless of academic merit, the Ministry of Education enacted a policy encouraging school districts to experiment with the abolishment of retention policy. In

2. Junior high schools in China use a homeroom teacher system where students remain in the homeroom class throughout the school day and subject teachers rotate classes. The homeroom teacher typically teaches one main subject (English, math, or Chinese); in addition, they also assume the responsibility of looking after the students' general academic performance in all subjects, personal development, and social activities at school (Chen et al. 2009; Liu and Barnhart 1999; Zhao 2014). Students are usually expected to stay within the same homeroom class throughout their junior high years, though schools may make adjustments from time to time (Chen et al. 2009; Liu and Barnhart 1999).

the following two decades, there was a nationwide effort to reduce retention rates (Chen 2013). Yet, the process of rescinding retention was lengthy and slow. In a 2009 study based on survey and transcript data drawn from 1,653 students from 36 primary schools in Shannxi province (Chen et al. 2009), one of the less economically developed provinces in northwest China, more than one-third of the students repeated at least one grade before they entered Grade 6.

The cohort of students covered in this study entered junior high school during the fall of 2013. Therefore, the majority of students were in primary school between 2007 and 2013, when the average retention rate was still fairly high in primary schools. Our data set does not include information on the specific province the student is in; therefore, we were not able to examine the extent of variation across provinces in the proportion of newly admitted junior high students who had been retained in primary school. Yet, among all the 112 schools in the national survey data used in the current study, 92 percent had at least one grade repeater, with an average of 17 percent repeaters across all schools, indicating that grade retention in primary schools was still a fairly common phenomenon for the cohort of students examined.

III. Data

A. China Education Panel Survey

The China Education Panel Survey (CEPS) is the first nationally representative survey for junior high students in China. The primary analyses of this study are based on data from the baseline wave that was collected during the spring of the 2013–2014 academic year. The CEPS employs a stratified, multistage sampling scheme. In the first stage, 28 counties/districts were chosen from 2,870 counties/districts. Four schools from each county/district were selected, and two classrooms were then randomly selected from seventh grade and another two from ninth grade in each sample school.[3] Finally, all students from these selected classrooms were surveyed, resulting in a sample of approximately 20,000 students in 438 classrooms of 112 schools.

The CEPS administered five separate questionnaires to the (i) sample students, (ii) homeroom teachers, (iii) subject teachers for the three main subjects (Chinese, English, and math), (iv) students' parents, and (v) school administrators. Hao and Yu (2015) provide a comprehensive summary of CEPS in their recent report; here we briefly review the features of the baseline survey most relevant for our analysis. The school administrator questionnaire solicited information about school resources, school management, and other school-level statistics about teachers and students. One important question asked in the school administrator questionnaire is whether the school randomly assigns students to different classrooms. Among all the 112 schools sampled, 83 percent ($N = 93$) reported that students were randomly assigned to classrooms upon entry into the junior high school. The homeroom teacher and main subject teachers were then randomly assigned to each class. In the methodology section, we present statistical evidence to show that students in the seventh grade in these schools indeed seem to be randomly assigned to classes, as indicated by preexisting student demographic and

3. If the sample schools had only one or two classes in each grade, all classes were included in the sample.

family characteristics. The teacher questionnaire collected information from each homeroom and main subject teacher on teachers' demographic characteristics, subject taught, and teaching experience. Finally, the student questionnaire collected information about student demographic and family background characteristics, cognitive measures such as academic performance in each of the three main subjects (Chinese, English, and math), and noncognitive measures such as students' perceived interaction with teachers of the three main subjects, perceived classroom climate, experience of negative emotions, school disengagement, and educational expectations. In the following section, we explain in more detail the outcome and control variables used in this study and how they are constructed.

B. Key Measures and Variable Definitions

We focus on five domains of student outcomes: academic performance, cognitive assessment, mental stress, school disengagement, and educational expectations.[4] Following Kling, Liebman, and Katz (2007) and Deming (2009), we create summary indexes for domains that contain multiple survey items.[5] The aggregation can partly address the problem of multiple hypothesis testing by reducing the number of outcome measures; it also has the potential to increase the statistical power to detect effects that go in the same direction within each domain (Kling, Liebman, and Katz 2007). Specifically, we first normalize each component of a domain to have a mean of zero and a standard deviation of one for all nonrepeaters at the school level. We then take the equally weighted average of the z-scores of the components with equalized signs so that larger values of the index represent higher levels in that domain. We describe each summary index and the specific survey items used to construct them below. It is worth noting that since we focus on the seventh grade (or the first grade of junior high school), students in our analytical sample would have spent almost one year with their classmates at the time of the survey.

1. Academic performance

Students were required to report their most recent midterm test score for each of the three main subjects: Chinese, English, and math. Within a grade at a school, teachers who teach the same subject use the same syllabus; the mid-term and final exams are administered by school and are common across all students in a grade. Student test scores are therefore comparable within a particular grade at each school. The raw score is on a 150-point scale.

2. Cognitive assessment

The CEPS also required all the participating schools to administer a 15-minute standardized cognitive ability test. The cognitive ability test assesses a student's aptitude on

4. We use principal factor analysis to group our noncognitive measures into three domains (that is, mental stress, school disengagement, and educational expectations) following Gong, Lu, and Song (2018). A detailed description of the procedure can be found in Online Appendix 1.

5. For all of our main findings, we also present the estimates for specific survey items that measure student outcomes in each domain (rather than summary indices) in Online Appendix Table 2.

reasoning and problem-solving in three dimensions: (i) language, (ii) vision and space, and (iii) arithmetic and logic.[6] The test follows similar practices used in a number of cognitive tests conducted in other countries, such as the Taiwan Education Panel Survey and the National Education Panel Study in Germany (Zhao et al. 2017).

In addition to the two cognitive outcome measures, we also examine three domains of noncognitive outcomes. Specifically, mental stress was measured by five survey items, asking whether the respondent has ever experienced feelings such as feeling down, depressed, unhappy, not enjoying life, or sad in the past seven days, which was based on a well-established clinical depression assessment using a five-point Likert scale (Löwe, Kroenke, and Gräfe 2005) ranging from 1 (never) to 5 (always). Higher values thus indicate that students experienced more negative feelings and therefore higher levels of mental stress.

3. School disengagement

School disengagement was measured by six survey questions: "I am often late for school," "I am often absent from school," "I seldom participate in school or class activities," "I do not feel close to people at this school," "I feel bored at school," and "I want to attend another school." All the questions about school disengagement were measured on a four-point scale ranging from 1 (strongly disagree) to 4 (strongly agree), with higher scores indicating higher levels of school disengagement.

4. Educational expectations

Student educational expectations were measured through questions about the highest level of education the student expected to receive and the confidence level a student has about their future. For the highest level of education, it was measured by ten options ranging from "Drop out now" to "Get a doctoral degree." We code this variable as expected years of education for the purpose of our analysis.[7] For a student's confidence level about their future, the response format consisted of a four-point scale, ranging from 1 (not confident at all) to 4 (very confident). We first standardize each item and then create an index variable of "educational expectations" by taking the average score of the two items.

5. Student, teacher, and classroom control variables

Control variables include information collected at the student level, teacher level, and classroom level. Specifically, student background characteristics include gender, whether

6. Specifically, the language dimension assesses verbal analogy and verbal reasoning. The vision and space dimension examines geometric pattern recognition, a paper-folding test, and application of geometric graphs. The arithmetic and logic dimension assesses applied problem solving, replacement of expressions by self-defined symbols, number sequence completion, numerical pattern recognition, probability, and quantitative comparison and reverse thinking.

7. Specifically, we recode this variable to the student's expected years of schooling based on the normal time to acquire a degree in China: seven years of education for students who chose "dropping out now" or "does not matter," nine years of education for "junior high school," 12 years of education for "technical secondary school," "vocational high school," or "senior high school degree," 15 years for "junior college degree," 16 years for "bachelor's degree," 18 years for "master's degree," and 21 years for "doctoral degree."

the student lives in a rural or urban hukou, whether the student is the only child in the household, age, parents' education level, family income, and potential family risk factors (that is, parental absence, whether the father drinks regularly, and whether parents always quarrel with each other). Teacher background information was collected through individual teacher surveys, which include information on their demographic characteristics (that is, gender and age), whether they graduated from a normal institution versus a comprehensive university, whether they have a teaching certificate, highest level of education, any prior teaching awards, teaching title, and years of teaching experience. Summary statistics of teacher background characteristics are presented in Online Appendix Table 3. Finally, since all the students in a sample class were surveyed, we are also able to calculate the average classroom characteristics and include them as control variables in the analysis, including class size, percentage of boys, percentage of low-income families, and percentage of students with at least one family risk factor.

C. Summary Statistics

For the purpose of our analysis, we restrict the analytical sample to the seventh-graders in 93 schools (out of the 112 schools in the full sample) that use a random algorithm to assign students and teachers to classes, resulting in a sample of 8,520 students.[8] Online Appendix Table 4 compares the characteristics of schools that claimed to use random assignment and those that did not. It seems that schools that did not claim to use random assignment are more likely to be located in rural regions and have a teacher force that is older than the schools that claimed to use random assignment. We further drop 66 students who have missing information on grade retention status.[9] The final sample consists of 1,392 repeaters and 7,062 nonrepeaters.

The missing rates for the outcome and mechanism measures among nonrepeaters are generally low, ranging between 1 percent and 4 percent. To examine whether students with certain characteristics are more likely to answer the survey, we run ordinary least squares regressions that correlate student characteristics and whether a student has a missing value for a specific outcome measure, controlling for school fixed effects. Online Appendix Table 5 shows the missing patterns for various domains of outcome variables.[10] Overall, except for a handful of cases, student characteristics do not seem to be consistently correlated with a student's probability of answering a survey question. In addition, results in Online Appendix Table 5 also show small and insignificant correlations between our key treatment variable (proportion of repeater peers) and nonrepeaters' probability of responding to each outcome measure. Finally, for students who

8. As mentioned, we focus on seventh-graders only due to the concern that there might be nonrandom between-class mobility in later years of junior high school.

9. Among the 66 students who have missing information on grade retention status, the majority (87 percent) are male students, about half are from urban families, and 14 percent are from low-income families. The average age of the 66 students is 14 years old.

10. To save space, we only present the results on the variable that is subject to the largest missing rate from each domain of the outcome and mechanism measures. For example, the variable "not enjoying life" has the largest missing rate among the outcome measures under the domain of "mental stress." The missing rate across all variables is low nevertheless: less than a 4 percent missing rate for covariates with missing values.

are missing information on other covariates, we retain them in the analytical sample and include indicators for missing data on those variables.

On average, approximately 16 percent of the seventh-graders were ever retained in primary school. Among these students, the majority (80 percent) were retained only once. The first column in Table 1 shows the demographic and family characteristics for all students, while Columns 2 and 3 present descriptive information for repeaters and nonrepeaters, respectively. Overall, the comparison between the two groups indicates that the repeaters are more likely to be older in age, have siblings, live in a rural hukou, have parents with lower levels of education, and be from families with lower income. It also appears that repeaters are more likely to be subject to potential family problems, such as having fathers who drink regularly and parents who fight with each other more often. Yet, these descriptive patterns might be partly due to between-school differences in the share of repeaters. For example, if repeaters are more likely to concentrate in certain types of schools, the raw difference between repeaters and nonrepeaters could thus partly reflect between-school distinctions in student characteristics.[11] To address between-school variations in the share of repeaters and student characteristics, Column 4 of Table 1 further presents the average difference between repeaters and nonrepeaters within the same school. The results generally echo the patterns shown in Columns 2 and 3, where repeaters are more likely to live in a rural hukou, have siblings, be older, and have fathers with drinking problems.

Table 2 presents descriptive information on each outcome measure. The first two columns summarize student mean scores for repeaters and nonrepeaters, respectively. Column 3 presents the gaps between the two groups, adjusting for school fixed effects to take into account overall variations in student outcomes across schools. Finally, Column 4 shows the number of nonrepeaters with nonmissing values for each outcome measure examined. Unsurprisingly, repeaters are associated with consistently lower test scores relative to nonrepeaters across all subject areas. The average raw test score among repeaters is 69 on a 150-point scale, versus 80 among nonrepeaters. When compared within schools, repeaters still score 5.5 points lower on average than nonrepeaters. The cognitive assessment score for repeaters is also significantly lower than that of nonrepeaters even after we compare the students within the same school.

Results from Table 1 seem to suggest that repeaters are more likely to be from higher risk families, which may not only result in lower academic ability, but may also induce psychological and behavioral problems that could influence class peers. To shed light on this possibility, Table 2 further presents students' noncognitive outcomes in junior high, including their mental stress, school disengagement, expected years of education, and confidence about the future.[12] Indeed, it is immediately apparent that repeaters are

11. We directly examine whether the share of repeaters at a school is associated with available school characteristics; the results are presented in Online Appendix Table 6. Most of the coefficients are small and nonsignificant, with two exceptions: repeaters seem to be more concentrated at schools that enroll a higher proportion of low-income families and also at schools with smaller enrollment size.

12. It is important to note that since "expected years of education" and "confidence about future" are measured in different scales, we show the descriptive statistics for the raw score of each variable separately. In all of the subsequent regression analyses, we combine them into a summary index by first standardizing each item and then taking the average score of the two. We also present the estimates on the two measures separately instead of the summary index in Online Appendix Table 2.

Table 1
Descriptive Statistics of Student Demographics and Family Characteristics

	All (1)	Repeaters (2)	Nonrepeaters (3)	Raw Gap (4)
Female	0.476	0.417	0.488	–0.007
	(0.499)	(0.493)	(0.500)	(0.007)
Urban hukou	0.483	0.252	0.528	–0.024*
	(0.500)	(0.434)	(0.499)	(0.011)
Only child	0.423	0.145	0.478	–0.022***
	(0.494)	(0.352)	(0.500)	(0.007)
Student age in years	13.585	14.306	13.443	0.179***
	(0.746)	(0.806)	(0.645)	(0.014)
Mother education in years	9.618	7.509	10.033	–0.002
	(3.573)	(3.206)	(3.494)	(0.002)
Father education in years	10.393	8.716	10.725	–0.002
	(3.141)	(2.433)	(3.160)	(0.002)
Low income	0.212	0.370	0.180	0.006
	(0.408)	(0.483)	(0.385)	(0.011)
Mother occupation				
Civil servants and government officials	0.034	0.012	0.038	
	(0.182)	(0.109)	(0.192)	
Executives and managers	0.052	0.012	0.060	0.012
	(0.223)	(0.109)	(0.238)	(0.019)
Teachers, engineers, doctors, and lawyers	0.060	0.019	0.068	0.013
	(0.237)	(0.136)	(0.252)	(0.018)
Technicians (including drivers)	0.043	0.049	0.042	0.026
	(0.204)	(0.216)	(0.201)	(0.023)
Manufacturing workers	0.117	0.137	0.113	–0.003
	(0.322)	(0.344)	(0.317)	(0.020)
Marketing, sales, and service	0.150	0.104	0.159	0.008
	(0.357)	(0.306)	(0.366)	(0.018)
Self-employed	0.162	0.116	0.170	0.005
	(0.368)	(0.321)	(0.376)	(0.018)
Farmers	0.219	0.394	0.185	0.010
	(0.414)	(0.489)	(0.389)	(0.028)
Unemployed	0.092	0.075	0.095	0.010
	(0.289)	(0.264)	(0.293)	(0.020)
Other	0.070	0.080	0.068	0.046*
	(0.255)	(0.272)	(0.252)	(0.021)

(*continued*)

Table 1 (*continued*)

	All (1)	Repeaters (2)	Nonrepeaters (3)	Raw Gap (4)
Father occupation				
Civil servants and government officials	0.052	0.025	0.058	Omitted
	(0.223)	(0.155)	(0.233)	
Executives and managers	0.085	0.018	0.098	–0.031+
	(0.279)	(0.133)	(0.297)	(0.016)
Teachers, engineers, doctors, and lawyers	0.071	0.052	0.075	0.009
	(0.257)	(0.221)	(0.263)	(0.016)
Technicians (including drivers)	0.176	0.179	0.176	–0.016
	(0.381)	(0.383)	(0.381)	(0.018)
Manufacturing workers	0.093	0.106	0.090	–0.028
	(0.290)	(0.308)	(0.287)	(0.019)
Marketing, sales, and service	0.166	0.117	0.176	–0.019
	(0.273)	(0.253)	(0.276)	(0.019)
Self-employed	0.166	0.117	0.176	–0.028
	(0.372)	(0.321)	(0.381)	(0.018)
Farmers	0.189	0.340	0.160	0.003
	(0.392)	(0.474)	(0.367)	(0.028)
Unemployed	0.025	0.022	0.026	–0.033
	(0.157)	(0.146)	(0.159)	(0.024)
Other	0.061	0.073	0.058	–0.023
	(0.239)	(0.261)	(0.234)	(0.020)
Family risk factors				
At least one parent is absent from home	0.246	0.399	0.220	0.001
	–0.431	–0.450	–0.412	(0.009)
Father gets drunk regularly	1.082	1.110	1.076	0.031*
	(0.274)	(0.313)	(0.266)	(0.013)
Parents always quarrel	1.096	1.126	1.090	0.004
	(0.295)	(0.332)	(0.287)	(0.014)
Observations	8,454	1,392	7,062	8,454

Notes: All variables are binary unless otherwise noted. While some variables have missing values, the missing rates for student characteristics variables are all below 4 percent. Standard deviations in parentheses from Columns 1–3. In Column 4, we regress the indicator of whether a student is a repeater against student and family characteristics while controlling for school fixed effects to examine whether the differences between repeaters and nonrepeaters reach statistical significance. Standard errors are in parentheses. $+p<0.10$, $*p<0.05$, $**p<0.01$, $***p<0.001$.

Table 2
Descriptive Statistics of Student Cognitive and Noncognitive Outcomes

	Repeaters (1)	Nonrepeaters (2)	Raw Gap (3)	*N* (4)
Academic performance	69.260	80.480	–5.507***	8,210
	(21.93)	(21.37)	(0.908)	
Cognitive assessment	9.040	11.060	–0.733***	8,408
	(3.243)	(3.433)	(0.129)	
Mental stress	2.180	1.973	0.102**	8,028
	(0.789)	(0.806)	(0.032)	
School disengagement	1.796	1.555	0.071***	8,113
	(0.492)	(0.460)	(0.015)	
Expected years of education	15.230	16.710	–0.839***	8,078
	(3.443)	(2.975)	(0.152)	
Confidence about future	3.125	3.329	–0.101***	8,328
	(0.756)	(0.684)	(0.022)	
Observations	1346	7062		

Notes: We present the descriptive statistics of "expected years of education" and "confidence about future" separately because they are in different scales. Columns 1 and 2 present the means of outcomes with standard deviations in parentheses. In Column 3, we regress each outcome variable on the repeater dummy to examine the raw gap between repeaters and nonrepeaters while controlling for school fixed effects; standard errors are in parentheses. Column 4 shows the number of nonrepeaters in the regression analysis sample. The analysis sample is restricted to students with nonmissing values on the outcome variables. The missing rates of outcome variables are all below 4 percent. $+p<0.10$, $*p<0.05$, $**p<0.01$, $***p<0.001$.

substantially more likely to experience negative emotions, show higher levels of mental stress, be less engaged behaviorally and emotionally in school, and have fewer expected years of education and lower confidence about the future. Such differences remain within schools.[13] Taken together, the descriptive statistics seems to suggest that repeaters are not only low academic achievers but may also have other social-emotional or behavioral problems that may be associated with negative externalities for their peer classmates.

13. It is worth noting that these outcome measures may also partly reflect the impact of the "repeater" status on repeaters' academic and emotional outcomes during the first year in junior high. For example, some researchers suggest that retained students, being older than the majority of their classmates, may feel less attached to school (Jimmerson 2001). However, studies examining the impacts of retention on students' academic performance and motivation have led to inconsistent results, and among those that report negative effects of retention on student outcomes, the effect sizes are generally small (for meta-analytic reviews, see Allen et al. 2009 and Jimmerson 2001). Therefore, the large gap between repeaters and nonrepeaters in their cognitive and noncognitive outcomes should at least partially capture the pre-enrollment differences between these two groups.

IV. Methodology

A. Validity of the Random Assignment

The validity of the causal inferences that follow rests on the successful randomization of students to classes. While the institutional setting we study makes it clear whether or not administrators in the schools follow random assignment upon students' initial school enrollment, we cannot rule out the possibility that in some cases students were lobbied to be placed in a class with a better teacher or better peers. For example, influential parents might pull their children out of a class with a particularly high proportion of repeaters. This problem is less of a concern in this particular context, considering that four primary schools feed into one junior high school on average, and therefore it is unlikely that parents would know all the repeaters in the feeding primary schools and manage to avoid them. Additionally, restricting the sample to seventh-graders only also helps reduce the extent of the problem, since it is at the beginning of the junior high years when parents and teachers have the least knowledge about each student's academic ability. However, we cannot rule out the possibility that teachers and parents may still gather such information through various channels. Below we evaluate the validity of the randomization formally to provide empirical evidence for our identification.[14]

To assess whether the student randomization protocol was implemented as designed, we test for balance in predetermined student family background characteristics, examining whether the background characteristics of a nonrepeater, such as parental education and occupation, are correlated with the proportion of repeater peers they are assigned to after conditioning on school fixed effects. It is worth noting that although the CEPS survey was conducted after the students in our sample had enrolled in a junior high school, and thus the family and parental background characteristics are not measured prior to random assignment, the background characteristics (such as parents' educational level) are unlikely to be impacted by the child's class assignment in junior high.

We first establish that these family and parental characteristics are in fact strong predictors of the key outcome measures. In Columns 1–5 of Table 3, we regress nonrepeaters' outcomes on all available student family background characteristics, controlling for school fixed effects. The results indicate that several demographic and family characteristics, such as gender, age, parental education level, family income, and parental occupation, are highly significant predictors of student cognitive and noncognitive outcomes.

Having identified a set of baseline characteristics that predict key outcome measures, we evaluate randomization of students into the treatment by regressing the proportion of repeater peers in a class to which a nonrepeater is assigned on that nonrepeater's

14. Additionally, in one of our robustness checks, we further use one question in the school administrator questionnaire to identify schools in our sample that might violate the random assignment due to pressure from parents. Specifically, this question asked the school administrators whether parents made special requests to assign students to certain classrooms, on a scale of 1 (not true at all) to 4 (true). Among the 93 schools that indicated using random assignment, only 16 school administrators answered true (4) or somewhat true (3) to this question. We thus conducted a robustness check of all of our analyses (including main outcomes and mechanisms) excluding the 16 schools. Results are presented in Online Appendix Table 7 and are fairly consistent with our main findings.

Table 3
Randomization Check of Student Assignment

	Academic Performance (1)	Cognitive Assessment (2)	Mental Stress (3)	School Disengagement (4)	Educational Expectations (5)	Treatment: Proportion of Peer Repeaters (%) (6)
Female	0.403***	–0.009	0.087**	–0.040*	0.017	–0.003
	(0.026)	(0.032)	(0.029)	(0.019)	(0.029)	(0.119)
Urban hukou	0.028	0.040	–0.020	–0.012	–0.010	–0.563
	(0.031)	(0.048)	(0.031)	(0.018)	(0.022)	(0.391)
Only child	0.119***	0.041	–0.034	–0.014	0.032	–0.209
	(0.035)	(0.040)	(0.026)	(0.022)	(0.027)	(0.232)
Age	–0.162***	–0.185***	0.016	0.018	–0.102***	0.101
	(0.024)	(0.026)	(0.021)	(0.017)	(0.019)	(0.134)
Mother education in years	0.009	0.006	–0.004	0.000	0.011*	–0.038
	(0.006)	(0.007)	(0.006)	(0.004)	(0.005)	(0.036)
Father education in years	0.025***	0.012	–0.005	–0.011**	0.032***	–0.016
	(0.006)	(0.008)	(0.006)	(0.004)	(0.005)	(0.038)
Low income	–0.086*	–0.126**	0.136***	0.069*	0.018	0.219
	(0.038)	(0.043)	(0.030)	(0.029)	(0.037)	(0.444)

(*continued*)

Table 3 (*continued*)

	Academic Performance (1)	Cognitive Assessment (2)	Mental Stress (3)	School Disengagement (4)	Educational Expectations (5)	Treatment: Proportion of Peer Repeaters (%) (6)
Mother occupation (Reference group: Civil servants and government officials)						
Executive and managerial	–0.064	–0.094	–0.101	–0.065	0.095	–0.082
	(0.084)	(0.132)	(0.091)	(0.063)	(0.060)	(0.488)
Teachers, engineers, doctors, and lawyers	0.147	0.298*	–0.006	–0.168***	0.057	–0.513
	(0.097)	(0.120)	(0.076)	(0.042)	(0.067)	(0.444)
Technicians (including drivers)	0.007	0.143	0.016	–0.116+	–0.118	0.409
	(0.086)	(0.106)	(0.121)	(0.060)	(0.080)	(0.710)
Manufacturing workers	0.055	0.181	0.072	–0.085	–0.146*	–0.037
	(0.078)	(0.112)	(0.090)	(0.054)	(0.069)	(0.562)
Marketing, sales, and service	0.029	0.188+	0.010	–0.081	–0.136+	–0.157
	(0.088)	(0.104)	(0.083)	(0.050)	(0.070)	(0.450)
Self-employed	0.106	0.181*	0.026	–0.117*	0.030	–0.710
	(0.069)	(0.082)	(0.075)	(0.046)	(0.070)	(0.543)
Farmers	0.069	0.091	–0.040	–0.100+	–0.025	–0.139
	(0.086)	(0.112)	(0.084)	(0.057)	(0.066)	(0.450)

(*continued*)

Table 3 (*continued*)

	Academic Performance (1)	Cognitive Assessment (2)	Mental Stress (3)	School Disengagement (4)	Educational Expectations (5)	Treatment: Proportion of Peer Repeaters (%) (6)
Unemployed	0.046	0.157	0.060	–0.016	–0.121	–0.004
	(0.086)	(0.115)	(0.083)	(0.057)	(0.074)	(0.610)
Other	–0.014	0.026	0.013	–0.108*	–0.055	–0.214
	(0.077)	(0.100)	(0.079)	(0.047)	(0.061)	(0.357)
Observations	6,939	7,062	6,904	7,029	7,020	7,062
Father occupation (Reference group: Civil servants and government officials)						
Executive and managerial	–0.076	–0.110	0.087	–0.035	–0.080	0.118
	(0.059)	(0.084)	(0.068)	(0.056)	(0.077)	(0.380)
Teachers, engineers, doctors, and lawyers	–0.009	–0.044	–0.018	–0.034	–0.054	–0.331
	(0.075)	(0.091)	(0.069)	(0.052)	(0.104)	(0.694)
Technicians (including drivers)	–0.072	–0.025	0.024	–0.025	–0.022	0.045
	(0.068)	(0.084)	(0.063)	(0.052)	(0.083)	(0.424)
Manufacturing workers	–0.120+	–0.070	–0.003	–0.035	–0.123	0.512*
	(0.063)	(0.107)	(0.069)	(0.053)	(0.097)	(0.254)
Marketing, sales, and service	–0.142*	–0.176+	0.019	–0.061	–0.053	0.765*
	(0.070)	(0.102)	(0.074)	(0.051)	(0.093)	(0.383)

(*continued*)

Table 3 (*continued*)

	Academic Performance (1)	Cognitive Assessment (2)	Mental Stress (3)	School Disengagement (4)	Educational Expectations (5)	Treatment: Proportion of Peer Repeaters (%) (6)
Self-employed	–0.071	–0.017	0.012	–0.051	–0.052	0.086
	(0.059)	(0.084)	(0.060)	(0.040)	(0.097)	(0.304)
Farmers	–0.075	–0.053	–0.043	–0.070	–0.112	0.460
	(0.073)	(0.097)	(0.068)	(0.059)	(0.095)	(0.537)
Unemployed	–0.100	–0.104	0.017	–0.058	0.086	1.018
	(0.110)	(0.124)	(0.102)	(0.085)	(0.143)	(0.726)
Other	–0.021	–0.074	0.036	0.001	–0.061	0.374
	(0.082)	(0.111)	(0.071)	(0.058)	(0.099)	(0.591)
Family risk factors						
At least one parent is absent from home	–0.091**	–0.033	0.063*	0.031	–0.038	0.191
	(0.034)	(0.032)	(0.025)	(0.027)	(0.024)	(0.286)
Father gets drunk regularly	–0.093*	–0.090+	0.125*	0.017	–0.156**	0.063
	(0.044)	(0.054)	(0.051)	(0.039)	(0.048)	(0.306)
Parents always quarrel	–0.027	0.010	0.359***	0.189***	–0.183***	–0.038
	(0.040)	(0.056)	(0.052)	(0.035)	(0.048)	(0.283)
Observations	6939	7062	6904	7029	7020	7062

Notes: All models include school fixed effects. Standard errors in parentheses. $+p<0.10$, $*p<0.05$, $**p<0.01$, $***p<0.001$.

demographic and family characteristics. In calculating the proportion of repeater peers, we divide the total number of repeater students in a class by the total number of students in that class minus one. If students are indeed randomly assigned to classes within a school, then, once conditional on school fixed effects, there should be no systematic association between student characteristics and assignment to proportion of repeaters as peers. The results shown in Column 6 of Table 3 indicate that overall these predetermined variables do not seem to predict systematically the likelihood that a student is assigned to a class with higher proportions of repeaters. An F-test for the joint significance of all the predetermined demographic and family characteristics is also insignificant, providing support for the validity of the randomization.

Another way to examine the validity of the randomization is to correlate the proportion of repeater peers a nonrepeater is assigned to with the nonrepeater's predicted outcomes using available demographic characteristics. If the randomization is indeed successfully implemented, we should expect no correlation between the predicted outcomes and proportions of repeaters assigned. We use cognitive outcomes for this additional check because they are less noisy than noncognitive measures and are more likely to detect possible correlation, if there is any. Specifically, we first predict the four cognitive outcomes (that is, Chinese, math, and English midterm test scores and cognitive assessment scores) of nonrepeaters by regressing these variables on all the available observables, including student, homeroom teacher, and classroom characteristics. We then regress the fitted value of cognitive outcomes on the proportion of repeater peers. Results presented in Online Appendix Table 8 and Online Appendix Figure 1 indicate that there is no significant correlation between the predicted outcomes of nonrepeaters and the proportion of repeater peers they are assigned to, except for English scores, where we identify a small correlation that is marginally significant at the 0.1 level.

Even though students are randomly assigned to classes, one potential threat to our identification strategy is that schools might assign teachers in a systematic way. For example, a school might assign more experienced homeroom teachers to classes that have more students with behavioral problems. Although this is less likely to happen given that we are focusing on the first year of junior high, when schools have minimal information on students' academic ability and behaviors, we conduct two balance tests at the class level to explore the extent of the problem more formally. The first one examines possible correlation between the proportions of repeaters and teacher assignment. Specifically, we regress the proportion of repeaters in a class on the characteristics of the homeroom teacher assigned to that class. Results in Table 4 indicate that, once controlling for school fixed effects, there is no systematic correlation between the characteristics of the homeroom teacher and the proportions of repeaters in that class. In addition, we also examine the correlation between the average characteristics of students in a class and key characteristics of the homeroom teacher assigned to that class, where we use students' background characteristics aggregated at the class level to predict homeroom teachers' gender, age, whether they have a college degree or higher, and teaching experience. The results are presented in Online Appendix Table 9 and further support the validity of randomization, where none of the teacher characteristics are systematically correlated with the average characteristics of the students in a class.

Table 4
Randomization Check of Homeroom Teacher Assignment

	Proportion of Repeaters in Class
Homeroom teacher female	–2.145
	(2.083)
Homeroom teacher age in years	–0.437
	(0.292)
Education level	
(Reference group: Professional college education)	
Adult higher education	0.936
	(3.768)
College education or higher	–0.185
	(4.167)
Graduated from a normal university	–3.038
	(2.525)
Have a teaching certificate	2.494
	(5.195)
Teaching experience in years	0.428*
	(0.190)
Teaching title	
(Reference group: No title)	
Level 2 or 3	8.130
	(5.704)
Level 1	6.219
	(6.635)
Senior teacher or higher	6.475
	(8.990)
Teaching award	
(Reference group: School level or no award)	
County or district level	–0.928
	(2.483)
City level	–0.888
	(2.678)
Provincial or national level	–3.044
	(2.833)
Class size	–0.203
	(0.385)
Observations	183

Notes: Data are aggregated to the class level. Accordingly, each observation represents one class. The model controls for school fixed effects. Standard errors in parentheses. $+p<0.10$, $*p<0.05$, $**p<0.01$, $***p<0.001$.

B. Econometric Specification for Student-Level Analysis

Having validated the randomization fidelity, our empirical model writes as follows:

$$(1)\quad Y_{igs} = a + \beta \%repeater_{gs} + student\_control_i + teacher\_control_{gs} + classroom\_control_{gs} + \delta_s + \mu_{igs}$$

where Y_{igs} is an outcome such as cognitive assessment score for nonrepeater student i randomly assigned to class g at school s. The variable $\%repeater_{gs}$ is the proportion of repeater peers for a nonrepeater in that particular class, which is calculated by dividing the number of repeaters by the total number of students in that class minus one. For easier interpretation, we standardize our outcome measures and multiply the proportion of repeater peers by 100 (such as converting 2 percent to 2). β hence measures the estimated effect on the outcome measure in standard deviations given a one percentage point increase in the proportion of classroom repeater peers. $student\_control_i$ is a vector of individual background characteristics listed in Table 1. $teacher\_control_{gs}$ further controls for the characteristics of the homeroom teacher assigned to this class listed in Table 4. $classroom\_control_{gs}$ is a vector of classroom average peer characteristics such as percentage of boys and percentage of low-income students. δ_g represents the school fixed effects.

One concern with the school fixed effects model is that only schools with between-class variations in the proportion of repeaters would contribute to the estimate of β. Figure 1 shows the distribution of the proportion of repeaters across all classes in our analytical sample. Among all 93 schools in our analytical sample, there is substantial variation in the proportion of repeaters, ranging from 0 percent to 65.5 percent. Since our identification draws on the between-class variations in proportion of repeaters within a school, Figure 2A further pairs the two classrooms in each school and creates a scatter plot, where each dot represents a school, with the share of repeaters in Classroom 1 on the *x*-axis and the share of repeaters in Classroom 2 (of the same school) on the *y*-axis. The results indicate that while there are strong correlations in the proportion of repeaters between each pair of classes within a school, very few of the schools fall exactly on the 45° diagonal line, indicating that most of the schools have variations in the share of repeaters between classrooms.[15]

Figure 2B further shows the distribution of within-school between-classroom differences in the share of repeaters. Among the 93 schools in our analytical sample that specifically indicated using a random algorithm to assign students to classes, three schools only had one class. Among the remaining 90 schools, 81 (or 87 percent) have between-class variations within school in the proportion of repeaters, with an average within-school variation of 5.51 percentage points and a median of 3.65 percentage points, therefore providing sufficient within-school variation to support our analyses. Additionally, while the within-school between-class difference in proportions

15. While schools without variation in the proportion of repeaters will not contribute to the estimate of the peer effects, these observations can still contribute to the estimation of other coefficients and are thus retained in our analysis. Standard errors are clustered at the school level to accommodate correlations among students as well as between classes within the same school. In our robustness checks, we also cluster the standard errors at the class level; the results indicate that the school-level clustering generates the largest standard errors and therefore represent more conservative estimates.

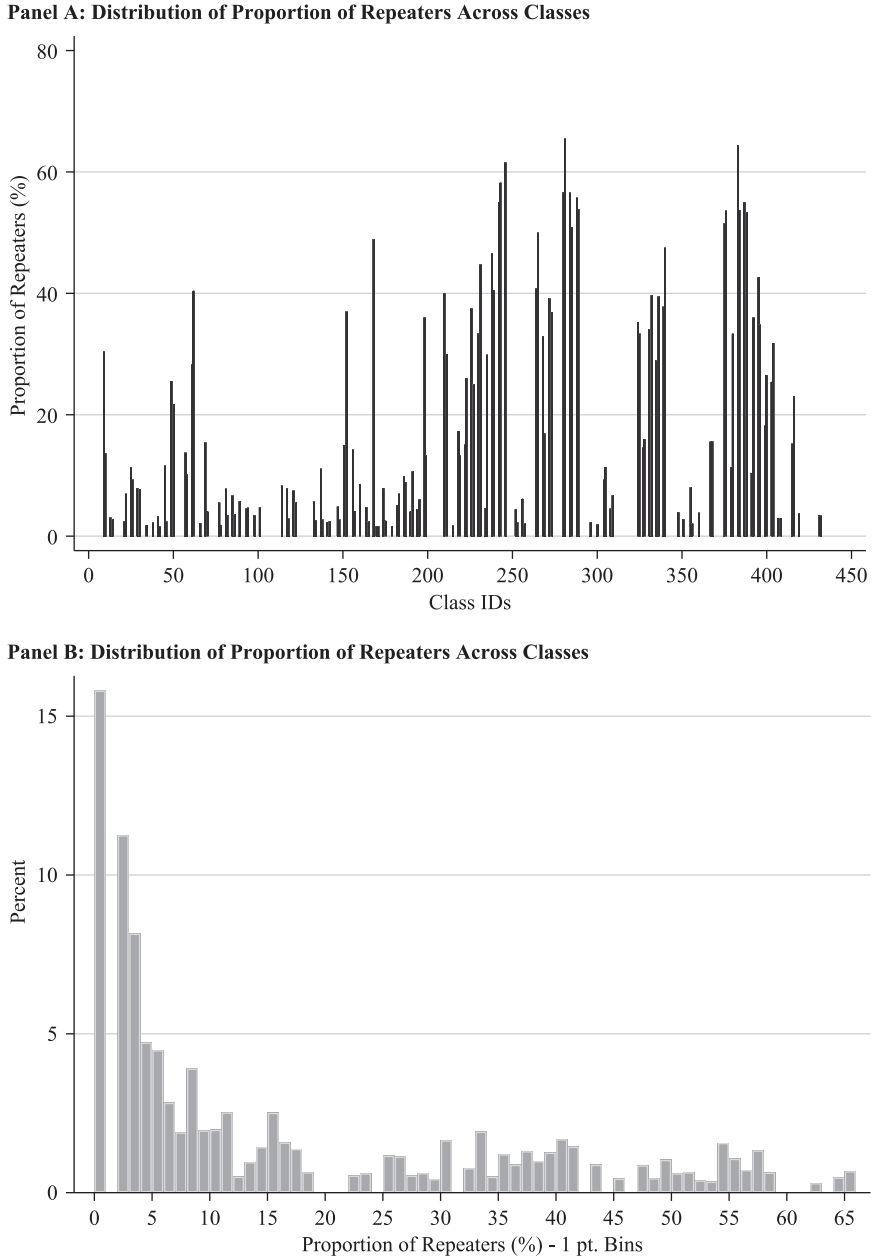
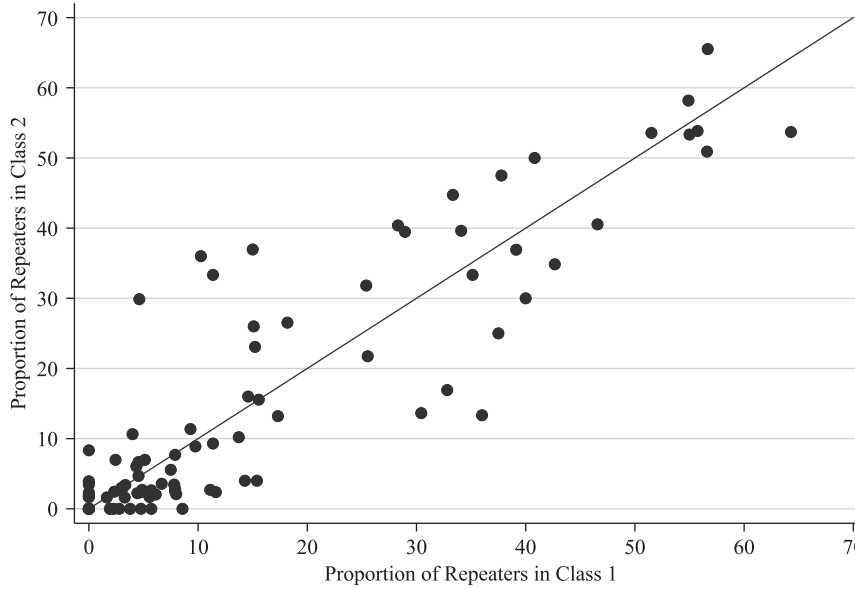


Figure 1

Distribution of the Proportion of Repeaters across Classes

Notes: The proportion of repeaters ranges from 0 to 66 percent. The numbers are all rounded to the nearest integer and shown in equally sized bins of one percentage point.

Panel A: Distribution of the Proportion of Repeaters in the Two Classes Sampled within Each School (Scatterplot)



Panel B: Distribution of the Difference in the Proportion of Repeaters between the Two Classes Sampled within Each School (Histogram)

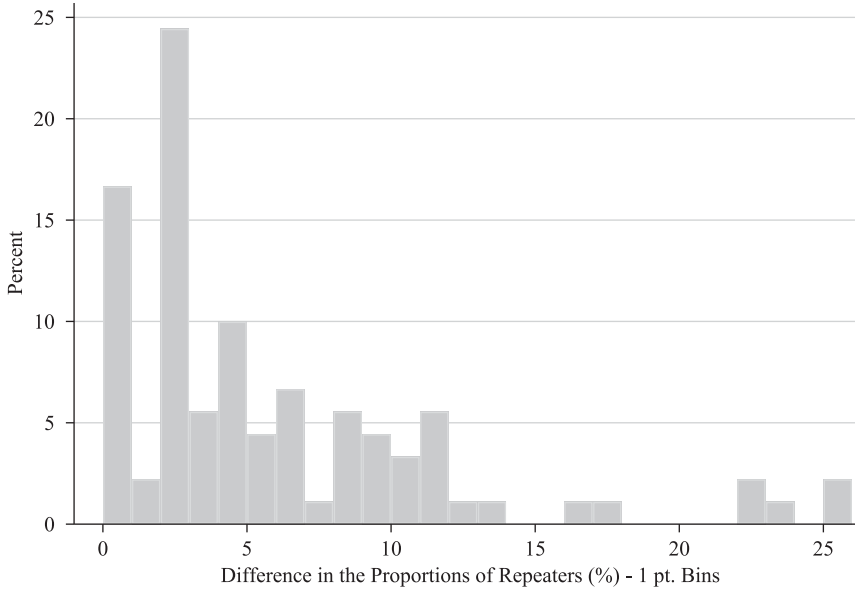


Figure 2

Distribution of the Proportion of Repeaters in Two Classrooms within Each School

Notes: The within-school between-class difference in the proportion of repeaters ranges from 0 to 26 percent. The numbers are all rounded to the nearest integer and shown in equally sized bins of one percentage point.

of repeaters ranges from 0 percent to 26 percent, the vast majority (83 percent) have a within-school variation less than 10 percent, indicating that the estimated effects are unlikely to be driven primarily by a small set of schools that have a larger variation in peer composition.[16]

C. First Difference at the School Level

In addition to the student-level analysis, we also aggregate the data at the class level and run all the analyses first differenced at the school level. Since each school includes two classes and since randomization occurred at the school level, these first-differenced estimates are a straightforward variation of the matched-pairs design that provides a conservative robustness check for the analyses conducted at the student level. More specifically, we relate first differences in average outcome of nonrepeaters against first differences in the proportions of repeaters:

$$(2)\quad (\bar{Y}_{1s} - \bar{Y}_{2s}) = \beta(\%repeater_{1s} - \%repeater_{2s}) + (\bar{X}_{1s} - \bar{X}_{2s}) + (T_{1s} - T_{2s}) + \varepsilon_s$$

where $\bar{Y}_{1s}$ and $\bar{Y}_{2s}$ represent the average outcomes of nonrepeaters in Class 1 and Class 2 at school s, respectively; $\%repeater_{1s}$ and $\%repeater_{2s}$ are the proportion of repeaters in the corresponding class; $\bar{X}$ is the average background characteristics of the nonrepeaters in a particular class; and T is a vector of the characteristics of the homeroom teacher assigned to a class. Among the 93 schools that used random assignment of students to classes, three schools only had one class, therefore providing us with 90 observations for the first-difference estimation. All the first-difference analyses using Equation 2 are weighted by class size.

V. Results

A. Main Effects

We begin by plotting the aggregate distributions of the average outcomes of nonrepeaters against proportions of repeaters, first differenced at the school level. Figure 3 visually shows the correlations in terms of all the five outcome indexes. First of all, there are noticeable negative correlations between proportions of repeaters in a class and both cognitive outcome measures of nonrepeater classmates, academic performance and cognitive assessment score. There also seems to be a negative but less pronounced correlation between proportions of repeater peers and educational expectations, as well

16. To further rule out the possibility that schools with particular characteristics may be more likely to have a larger variation in peer composition and therefore drive the estimated effects, we directly examine the correlation between within-school variation in the share of repeaters and a set of school characteristics, including school size, average class size, proportion of students with rural hukou, average parental education, whether the school is located in rural areas, school funding per student in the current year, and proportion of students from low-income families. Results from this correlation analysis are presented in Online Appendix Table 10 and indicate that within-school variation in the share of repeaters is not associated with any of the school characteristics mentioned above, therefore providing additional support that the estimated effects are unlikely to be driven by particular types of schools.

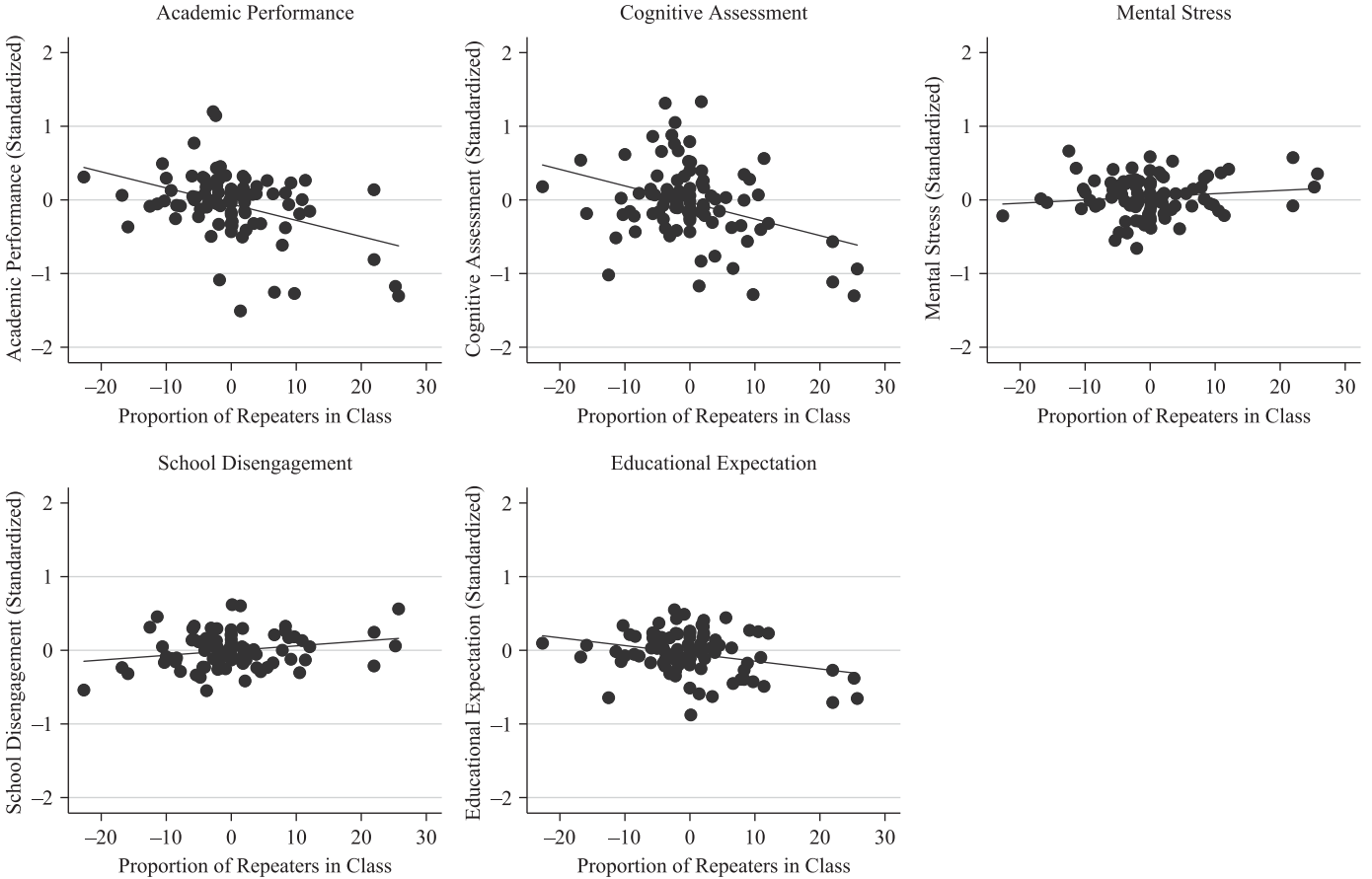


Figure 3

Relationship between First Differences in Average Outcomes of Nonrepeaters and First Differences in the Proportion of Repeaters

Notes: Data used in these figures are first aggregated to the class level and then first differenced between classes within each school. Accordingly, these figures show the relationships between the first differences in the average outcomes of nonrepeaters (*y*-axis) and the first differences in the proportion of repeaters (*x*-axis).

Table 5
Impact of Repeaters on Nonrepeaters' Cognitive and Noncognitive Outcomes

	(1)	(2)	(3)	(4)	(5)
Academic performance	−0.029***	−0.026***	−0.024***	−0.021***	−0.018*
	(0.008)	(0.007)	(0.007)	(0.006)	(0.007)
Observations	6,939	6,939	6,939	6,939	90
Cognitive assessment	−0.025**	−0.023**	−0.026***	−0.023***	−0.023**
	(0.008)	(0.008)	(0.007)	(0.006)	(0.008)
Observations	7,062	7,062	7,062	7,062	90
Mental stress	0.006+	0.006	0.003	0.003	0.004
	(0.004)	(0.003)	(0.004)	(0.003)	(0.005)
Observations	6,904	6,904	6,904	6,904	90
School disengagement	0.010*	0.009**	0.007*	0.007*	0.006+
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Observations	7,029	7,029	7,029	7,029	90
Educational expectations	−0.013**	−0.011**	−0.005	−0.003	−0.003
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	7,020	7,020	7,020	7,020	90
School fixed effects	Yes	Yes	Yes	Yes	
Student characteristics	No	Yes	Yes	Yes	Yes
Homeroom teacher characteristics	No	No	Yes	Yes	Yes
Classroom characteristics	No	No	No	Yes	Yes
First difference number of repeaters					Yes

Notes: All outcome variables are standardized among nonrepeaters within each school. Each coefficient represents a separate regression using the proportion of repeater peers to predict each outcome variable. Student characteristics include gender, hukou status, only child, age, parental educational attainment, parental occupations, family income, and family risk factors. Homeroom teacher characteristics include homeroom teacher gender, age, educational level, graduated from a normal university (a university specializing in teaching), have a teaching certificate, teaching experience in years, teaching title, and teaching award. Classroom characteristics include class size, percentage of boys, percentage of low-income families, and percentage of students who have at least one family risk factor. Column 5 shows the first-difference estimates at the school level. Standard errors are in parentheses. $+p<0.10$, $*p<0.05$, $**p<0.01$, $***p<0.001$.

as slightly positive correlations between proportions of repeater peers and nonrepeaters' level of mental stress and school disengagement.

Table 5 quantifies these correlations on the basis of five different model specifications. We start with a model that controls only for school fixed effects (Column 1) and then progressively add the controls for individual (that is, nonrepeaters') (Column 2), homeroom teacher (Column 3), and classroom average peer characteristics (Column 4). Finally, Column 5 presents the first-difference estimates at the school level based on Equation 2.

The estimates echo the patterns shown in Figure 3. Even in the most highly specified model that controls for school fixed effects with the full set of controls (Table 5, Column 4), exposure to higher proportions of repeater peers is associated with a significant decrease in nonrepeaters' academic performance and cognitive assessment score. Specifically, a one percentage point increase in the proportion of repeater peers decreases academic performance by almost 2.1 percent of a standard deviation among nonrepeaters. In other words, adding one more repeater to a class of 46 (which is the average class size in our analytical sample—roughly a 10 percent standard deviation increase, or an increase in the proportion of repeater peers by two percentage points) is associated with a decline in average academic performance for nonrepeaters in that class by 4.2 percent of a standard deviation toward the end of the seventh grade. Similar magnitude of the negative effects is also observed for cognitive assessment score (by 2.3 percent of one standard deviation). Column 5 presents the first-difference estimates aggregated at the school level. Even with this relatively more conservative approach with substantially larger standard errors, the negative correlations between the proportion of repeater peers and nonrepeaters' academic performance and cognitive assessment scores remain significant. Online Appendix Table 2 also shows the estimated impacts of proportions of repeaters on the values measured by each single survey item instead of the summary indexes, and the sizes of the effects on the test scores of the three subject areas are strikingly consistent.

In terms of noncognitive outcomes, student-level analyses controlling for school fixed effects (Column 1) indicate that exposure to a higher proportion of repeaters negatively influences nonrepeaters' mental health, school engagement, and educational expectations. However, only the impact on school disengagement remains significant when we further control for individual, homeroom teacher, and classroom average peer characteristics (Column 4) or in the first-difference analysis at the school level (Column 5).

One potential concern regarding our analyses is that as we test more and more outcomes, the problem of false positives could arise from multiple hypothesis testing, where even a randomized experiment could yield some *p*-values that appear to be statistically significant purely by chance if a sufficient number of hypotheses are tested. We have partly addressed this concern by aggregating outcome measures within the same domain and creating summary indexes following previous studies (for example, Anderson 2008; Deming 2009; Kling, Liebman, and Katz 2007). Another approach that has been commonly used in the existing literature to address the multiple hypothesis testing problem is to adjust the *p*-values controlling for the familywise error rate—the probability of rejecting at least one true null hypothesis—using the stepwise resampling method (for example, Anderson 2008; Kling, Liebman, and Katz 2007; Romano and Wolf 2016). We therefore follow the procedures described in Romano and Wolf (2016) to jointly test the null hypothesis that there is no treatment effect on any of the outcomes or mechanisms. The adjusted *p*-values presented in Online Appendix Table 11 indicate that the significant effects of having repeater peers on academic performance, cognitive assessment, and school disengagement are unlikely to be an artifact of multiple hypothesis testing.

B. Possible Mechanisms

Having found that repeater classmates impose significant externalities on classroom peers, which are particularly robust in terms of academic performance, we further explore possible channels driving these effects. Specifically, we draw on students' responses to

several survey items to shed light on three possible mechanisms in this particular research context: student–teacher interaction, student–student classroom interaction, and study hours after school.

Student–teacher interaction is measured by eight survey items on nonrepeaters' perceived interaction with their homeroom and subject teachers teaching any of the three main subject areas—Chinese, English, and math. Students were first asked about their interactions with each of the three subject teachers, including whether the subject teacher asked the students to answer questions in class frequently and whether the student felt that the teacher praised them frequently. Two additional questions asked about students' interaction with their homeroom teacher, including whether the student felt criticized by the homeroom teacher frequently and whether the student felt praised by the homeroom teacher frequently. All the questions are based on a four-point scale ranging from 1 (strongly disagree) to 4 (strongly agree).[17] If teachers indeed adjust their expectations of students and behaviors based on classroom peer composition, we would expect that having a greater proportion of repeaters in a class influences how teachers interact with nonrepeaters. For student–student interaction, students were asked to respond to two statements, "most of my classmates are nice to me" and "my class has a good atmosphere." Both questions were answered in a four-point scale ranging from 1 (strongly disagree) to 4 (strongly agree). Finally, students were asked to report their time spent on study after school (in hours), which includes study time on school work, tutoring, and assignments from tutoring. We create a summary index for both student–teacher interaction and student–student interaction following the same procedures for creating the summary indexes for our outcome measures. We also standardize students' self-reported study hours for easier interpretation.

The first two columns in Table 6 show the mean and standard deviation of repeaters and nonrepeaters separately for each index. On average, nonrepeaters report higher values for student–teacher interaction, student–student interaction, and more time spent on study after school. Columns 3–5 present the estimated effects of repeater peers on these measures, starting with a school fixed effect regression and progressively adding individual, homeroom teacher, and classroom average peer controls. The results suggest that proportion of repeater peers in a class is not significantly associated with nonrepeaters' perceived interaction with their teachers.[18] However, a greater proportion of repeater peers is negatively associated with nonrepeaters' perceived peer relationships. Specifically, a one percentage point increase in the proportion of repeater peers leads to approximately 1 percent of a one standard deviation decrease in students' evaluation of classroom peer interaction. In other words, adding one more repeater to a class of 46 is associated with a decline in nonrepeaters' perceived classroom peer interaction by 2 percent of one standard deviation ($2 \times 0.010 = 0.020$). Yet, the estimated effect becomes insignificant once we further control for available characteristics of the homeroom teacher and classroom average peer characteristics in Columns 4 and 5.

The third row of Table 6 presents the impact of having a larger proportion of repeater peers on nonrepeaters' daily hours spent on study after school. The results indicate that

17. We reverse code the item that asked about students' perception of criticism by the homeroom teacher when we aggregate the eight items into a single measure, so that higher values indicate stronger and more positive interactions between the student and the teacher.

18. We also separately examine students' perceived interaction with subject teachers and homeroom teachers; none of the analyses yield significant estimates.

Table 6
Mechanisms through Which Repeater Peers Influence Nonrepeaters

	Descriptive Statistics		Regression Coefficients			
	Repeaters (1)	Nonrepeaters (2)	School FE with Individual Controls (3)	School FE with Individual and Homeroom Teacher Controls (4)	School FE with Individual, Homeroom Teacher, and Classroom Controls (5)	*N* (6)
Positive student–teacher interaction	2.560 (0.633)	2.745 (0.621)	–0.002 (0.003)	–0.000 (0.003)	–0.001 (0.002)	7,038
Positive student–student interaction	2.996 (0.775)	3.246 (0.734)	–0.010* (0.005)	–0.004 (0.004)	–0.002 (0.004)	7,016
After-school study time (hours)	1.238 (1.200)	1.334 (1.016)	–0.010*** (0.003)	–0.008* (0.003)	–0.010* (0.004)	6,498
Having friends with general disciplinary problems	0.334 (0.472)	0.193 (0.395)	0.001 (0.002)	0.002 (0.002)	0.002 (0.002)	6,943
Having friends who go to internet cafés regularly	0.212 (0.409)	0.087 (0.282)	0.001 (0.002)	0.003+ (0.001)	0.002 (0.001)	6,929
Observations	1,391	7,061				
School fixed effects			Yes	Yes	Yes	
Student characteristics			Yes	Yes	Yes	
Homeroom teacher characteristics			No	Yes	Yes	
Classroom characteristics			No	No	Yes	

Notes: Columns 1 and 2 present the raw mean values for repeaters and nonrepeaters with standard deviations in parentheses. Columns 3–5 present estimates based on regressions with different specifications, in which we progressively add student characteristics, homeroom teacher characteristics, and classroom characteristics to the school fixed effects models. Outcome variables in Column 3–5 are standardized among nonrepeaters within each school. Standard errors in parentheses. $+p<0.10$, $*p<0.05$, $**p<0.01$, $***p<0.001$.

having a greater proportion of repeater classmates significantly reduces nonrepeaters' after-school study time. A one percentage point increase in the proportion of repeater peers is associated with a significant reduction in after-school study hours by approximately 1 percent of a standard deviation. In other words, adding one more grade repeater to a class of 46 is associated with an average decline in nonrepeaters' self-reported after-school study time by approximately 1.2 minutes daily ($2 \times 0.01 \times 60 = 1.2$). The association between repeater peers and self-study hours remains significant in models that further control for homeroom teacher and classroom average peer characteristics. Although the effect size seems small for each student on a daily basis, it adds up to a loss of study time of more than 200 minutes for each nonrepeater during the academic year, considering that middle school students usually spend almost 180 days at school in a typical academic year.

To summarize, we find some suggestive evidence for peer effects on interstudent relationships but do not find evidence for student–teacher interactions. Interestingly, after-school study hours seem to be the most robust of the three discussed channels in our setting. This finding suggests that peers may influence own academic efforts by influencing the student's time investments after school. One possibility is that having low-ability peers may induce nonrepeaters to be more relaxed and thus exert less effort after school. However, perhaps a more compelling channel is that repeaters may spend time with nonrepeater classmates, and thus repeaters, who tend to spend less time on study on average as shown in Table 6, could influence their nonrepeater friends' after-school activities and time use through group activities and peer pressure.

To further shed light on the specific channel of after-school activities, we examine the association between proportion of repeater classmates and a nonrepeater's probability of having friends who play at internet cafés regularly. According to a recent study on internet addiction using a nationally representative sample of Chinese primary and middle school students, surfing and playing video games in internet cafés has become the most important risk factor leading to internet addiction among teenagers (Li et al. 2014). By 2016, there were more than 140,000 internet cafés in China (Zhiyan Consulting Group 2016), which have been seen as one of the main reasons for school absenteeism, neglect of studies, and "hotbeds of juvenile crime" among Chinese teenagers (Reuters 2007). Although CEPS does not include information on individuals' time spent in internet cafés, students were asked whether any of their top five best friends play at internet cafés regularly. Since peer influence is an important factor in internet and digital game addictions (for example, Gunuc 2016), understanding the impact of repeater peers on one's probability of having friends with risky behaviors could shed light on why exposure to greater proportions of repeater classmates may negatively influence one's study time after school. In addition to having friends who play at internet cafés regularly, students were also asked whether any of their top five best friends ever had other behavioral problems, including skipping classes, violating school rules, fighting, drinking, and smoking. We aggregated the information and created a variable to indicate whether any of a student's top five best friends has general disciplinary problems.[19]

19. Among students who had friends with disciplinary problems, the majority had only one such friend. In a separate robustness check, we also code the variable as the number of best friends regularly going to internet cafés or showing general disciplinary misbehaviors instead. The results are almost identical.

Results presented in the last two rows of Table 6 indicate that having repeater classmates is not significantly associated with nonrepeaters' probability of having friends with general disciplinary problems. Yet, having a greater proportion of repeater peers increases a nonrepeater's probability of having a friend who regularly plays at internet cafés after school. Based on the model specification with school fixed effects, individual characteristics, and homeroom teacher characteristics (Column 4), a one percentage point increase in repeater peers is associated with an increased probability of having friends who regularly go to internet cafés by 0.3 percentage points. These results provide suggestive evidence that the negative impacts of repeaters on their nonrepeater classmates may operate through social networks and joint activities after school. Yet, the estimated effect becomes insignificant once we further control for classroom average peer characteristics in Column 5.

C. Heterogeneous Effects by Parental Monitoring and Mother Education

Having found that repeater classmates impose significant externalities on nonrepeaters on average, we further explore whether there is any evidence that such spillovers are heterogeneous by the characteristics of nonrepeaters. In particular, given the extensive evidence on the impact of parental involvement on students' academic outcomes, the negative externalities of repeater peers might be mitigated if a student has parents who regularly check homework and monitor after-school behaviors, whereas students with less involved parents and less strict discipline at home might be more vulnerable to having repeater friends, particularly given that one of the most important mechanisms seems to be reduced study time after school.

To explore this possibility, we construct two measures as proxies of parental involvement in monitoring their children's academic and after-school activities based on survey questions from the parents' questionnaire: (i) a composite score of child disciplinary practices at home that consists of eight specific questions regarding various activities[20] and (ii) a survey question that asked whether the parents check their child's homework regularly at home. Panel A of Table 7 presents the heterogeneous effects of repeater peers on nonrepeaters' cognitive and noncognitive outcomes. Results show that repeaters' negative spillovers on academic performance, cognitive assessment score, and school disengagement seem to be stronger among students who are from families without strict child discipline at home (Column 1), although none of these differences reach statistical significance (Column 3). We observe similar patterns of results when we divide the sample by whether the parents check their child's homework regularly. Results presented in Columns 4 and 5 show that having a larger proportion of repeater peers affects students from families without regular homework checking more severely on cognitive assessment score and school engagement compared to students whose parents check their homework regularly. Finally, given the ample evidence that establishes the connection between mother's education and children's cognitive and social development (for example, Menaghan and Parcel 1991; Parcel and Menaghan 1994)

20. The eight questions asked parents whether they have strict rules regarding their child's (i) academic test scores, (i) behaviors and activities at school, (iii) going to school on time every day, (iv) going back home on time every day, (v) rules about choosing the right friends, (vi) dress code, (vii) time spent on internet, and (viii) time spent on watching TV.

Table 7
Heterogeneous Effects of Repeaters on Nonrepeaters: Outcomes and Mechanisms

	Child Discipline at Home			Regular Homework Check			Mother's Education		
	Nonstrict (1)	Strict (2)	Difference (3)	No (4)	Yes (5)	Difference (6)	College or Higher (7)	High School or Less (8)	Difference (9)
Panel A: Cognitive and Noncognitive Outcomes									
Academic performance	–0.023*	–0.019***	0.004	–0.019*	–0.021**	–0.002	–0.002	–0.022***	–0.020
	(0.009)	(0.006)	(0.007)	(0.009)	(0.007)	(0.011)	(0.019)	(0.006)	(0.018)
Observations	1,356	5,157	6,513	1,564	5,301	6,865	1,074	5,712	6786
Cognitive assessment	–0.030**	–0.019**	0.011	–0.043***	–0.015*	0.028***	–0.014	–0.024***	–0.011
	(0.010)	(0.006)	(0.009)	(0.007)	(0.006)	(0.007)	(0.018)	(0.006)	(0.018)
Observations	1,385	5,237	6,622	1,603	5,377	6,980	1,084	5,821	6,905
Mental stress	0.007	0.003	–0.003	0.004	0.003	–0.001	–0.006	0.002	0.008
	(0.006)	(0.004)	(0.006)	(0.006)	(0.003)	(0.005)	(0.010)	(0.004)	(0.010)
Observations	1,359	5,126	6,485	1,585	5,252	6,837	1,063	5,697	6,760
School disengagement	0.011*	0.004	–0.007	0.011*	0.005	–0.007+	–0.012	0.007*	0.021+
	(0.005)	(0.003)	(0.006)	(0.004)	(0.003)	(0.004)	(0.012)	(0.003)	(0.011)
Observations	1,379	5,219	6,598	1,597	5,363	6,960	1,080	5,795	6,875
Educational expectations	–0.002	–0.003	–0.002	–0.006	–0.002	0.005	0.023+	–0.004	–0.027*
	(0.005)	(0.004)	(0.005)	(0.007)	(0.004)	(0.006)	(0.012)	(0.004)	(0.012)
Observations	1,373	5,212	6,585	1,594	5,348	6,942	1,079	5,785	6,864

(*continued*)

Table 7 (*continued*)

	Child Discipline at Home			Regular Homework Check			Mother's Education		
	Nonstrict (1)	Strict (2)	Difference (3)	No (4)	Yes (5)	Difference (6)	College or Higher (7)	High School or Less (8)	Difference (9)
Panel B: Mechanisms									
Positive student–teacher interaction	−0.001	−0.002	−0.001	−0.001	−0.002	−0.001	−0.011	−0.001	0.009
	(0.006)	(0.003)	(0.008)	(0.005)	(0.003)	(0.007)	(0.012)	(0.002)	(0.011)
Observations	1,379	5,229	6,608	1,598	5,372	6,970	1,084	5,800	6,884
Positive student–student interaction	−0.009+	−0.001	0.008	−0.002	−0.002	−0.000	0.003	−0.002	−0.003
	(0.005)	(0.004)	(0.007)	(0.005)	(0.004)	(0.006)	(0.023)	(0.004)	(0.022)
Observations	1,378	5,209	6,587	1,595	5,355	6,950	1,078	5,786	6,864
After-school study time (hours)	−0.025***	−0.007+	0.019**	−0.016**	−0.010+	0.005	−0.003	−0.010*	−0.006
	(0.007)	(0.004)	(0.006)	(0.005)	(0.005)	(0.007)	(0.015)	(0.004)	(0.015)
Observations	1,280	4,829	6,109	1,479	4,964	6,443	987	5,380	6,367
Having friends with general disciplinary problems	0.004	0.001	−0.003	0.003	0.000	−0.003	−0.012	0.003	0.014*
	(0.005)	(0.002)	(0.004)	(0.003)	(0.002)	(0.003)	(0.007)	(0.002)	(0.007)
Observations	1,356	5,164	6,520	1575	5305	6880	1067	5726	6793
Having friends who go to internet cafés regularly	0.006*	0.001	−0.004+	0.006*	−0.000	−0.006**	−0.005	0.003*	0.008
	(0.003)	(0.001)	(0.003)	(0.003)	(0.001)	(0.002)	(0.005)	(0.002)	(0.005)
Observations	1,355	5,153	6,508	1,571	5,297	6,868	1,067	5,714	6,781

Notes: Columns 3, 6, and 9 are results of joint *F*-tests examining whether the two groups of coefficients are significantly different from each other. All models include student, homeroom teacher, and classroom control variables, as well as school fixed effects. Standard errors in parentheses. $+p<0.10$, $*p<0.05$, $**p<0.01$, $***p<0.001$.

and delinquency (Hillman, Sawilowsky, and Becker 1993; McCord 1991), we further examine whether the negative effects of repeaters are moderated by mothers' education level, where we divide the sample in half by mothers with "a college degree or higher" versus mothers with "a high school degree or less." Results show that the negative spillovers of repeaters on academic performance and cognitive assessment score are primarily driven by students whose mothers have below-college education (Panel A, Column 8). For these students, a one percentage increase in the proportion of repeaters in their class leads to a 2.2 percent of a standard deviation decrease in their academic performance and a 2.4 percent of a standard deviation decrease in their cognitive assessment score toward the end of the seventh grade. In contrast, the size of the coefficients on students whose mothers have a college education is substantially smaller and no longer significant.

Taken together, the results from the heterogeneity analyses presented in Panel A of Table 7 indicate that the negative effects of having repeater peers seem to be more pronounced among nonrepeaters from families with less strict parental monitoring. Among students from these families, academic performance and cognitive assessment score are the areas that are most severely and persistently affected by having a larger proportion of repeater peers. One possible explanation for the heterogeneous impact is that students from families without strict parental monitoring are more likely to interact with disruptive peers and therefore reduce their daily study time after school.

Panel B of Table 7 empirically explores this possibility by examining the heterogeneous effects of repeater peers on available mechanism measures. Indeed, while we do not find any heterogeneous effects on nonrepeaters' in-school activities (such as student–teacher interaction and student–student interaction), we find fairly consistent patterns that repeater peers have a noticeably larger impact on daily after-school study time and the probability of having friends who go to internet cafés among nonrepeaters from families that lack strict child discipline at home, do not regularly check homework, or have less educated mothers. For example, among students from families without strict child discipline at home (Column 1), a one percentage point increase in the proportion of repeater peers is associated with a significant reduction in daily after-school study hours by approximately 2.5 percent of a standard deviation.[21]

D. Male Repeaters vs. Female Repeaters

Finally, we explore whether the spillover effects depend on the gender of the repeaters, based on two considerations. First, ample research has shown that male teenagers are associated with higher levels of disciplinary and misbehavior problems than girls (Mendez and Knoff 2003). For example, boys are more likely to show personal and physical aggression than girls (McGee et al. 1992; Zoccolillo 1993). Indeed, descriptive information shown in Online Appendix Table 13 indicates that male repeaters are associated with higher levels of school disengagement at school than female repeaters. Male repeaters also spend the least amount of hours on study among all students. Second, our results in the previous section indicate that important mechanisms for the spillovers

21. Given the rural–urban differences in school quality in China, we have conducted additional heterogeneity analysis based on school location. The results are presented in Online Appendix Table 12 and follow similar patterns.

Table 8
Impacts of Male and Female Repeaters on Nonrepeaters: Outcomes and Mechanisms

	Proportion of Male Repeaters (1)	Proportion of Female Repeaters (2)
Panel A: Cognitive and noncognitive outcomes		
Academic performance	–0.028***	–0.012
	(0.007)	(0.010)
Observations	6,939	6,939
Cognitive assessment	–0.021**	–0.028*
	(0.007)	(0.011)
Observations	7,062	7,062
Mental stress	0.001	0.004
	(0.005)	(0.005)
Observations	6,904	6,904
School disengagement	0.007+	0.006
	(0.004)	(0.005)
Observations	7,029	7,029
Educational expectations	–0.008+	0.006
	(0.005)	(0.007)
Observations	7,020	7,020
Panel B: Mechanism outcomes		
Positive student–teacher interaction	0.001	–0.006
	(0.003)	(0.005)
Observations	7,038	7,038
Positive student–student interaction	–0.009+	0.011
	(0.005)	(0.008)
Observations	7,016	7,016
After-school study time (hours)	–0.011*	–0.011
	(0.005)	(0.008)
Observations	6,498	6,498
Having friends with general disciplinary problems	0.006**	–0.006*
	(0.002)	(0.002)
Observations	6,943	6,943
Having friends who go to internet cafés regularly	0.004**	–0.001
	(0.002)	(0.002)
Observations	6,929	6,929

Notes: All models include student, homeroom teacher, and classroom control variables, as well as school fixed effects. Standard errors in parentheses. $+p<0.10$, $*p<0.05$, $**p<0.01$, $***p<0.001$.

might include social networks and joint activities together after school. Existing literature consistently reveals gender differences in teenagers' patterns of intimacy, where girls are more likely to establish intimacy through discussion and self-disclosure, whereas boys tend to establish intimacy through shared activities (Bauminger et al. 2008; McNelles and Connolly 1999).

Table 8 presents the effects of having male and female repeater peers on each outcome and mechanism measure of nonrepeaters based on the model specification that controls for school fixed effects, individual characteristics, homeroom teacher characteristics, and classroom average peer characteristics.[22] The results indicate that male repeaters primarily cause the negative externalities on nonrepeater students' academic performance, cognitive assessment score, school engagement, and educational expectations. For example, the coefficient for male repeaters on nonrepeaters' academic performance (–0.028) implies that adding one male repeater peer to a classroom of 46 students decreases nonrepeater students' test scores by nearly 5.6 percent of one standard deviation. Estimates from other outcome variables predict that adding one more male repeater peer to a classroom of 46 students decreases nonrepeater students' cognitive assessment score by nearly 4.2 percent of one standard deviation, increases nonrepeater students' school disengagement by nearly 1.4 percent of one standard deviation, and decreases nonrepeaters' level of educational expectations by 1.6 percent of one standard deviation.[23] Interestingly, when focusing on after-school study hours, which is the most important mechanism identified in the previous section, we also find that having a greater proportion of male repeaters in class reduces nonrepeaters' study hours, but the same is not true for having greater proportions of female repeaters. In addition, male repeaters also increase nonrepeater peers' probability of having friends who play at internet cafés regularly, as well as having friends with general disciplinary problems.[24]

VI. Discussion and Conclusion

Understanding peer effects of low-ability children in the classroom is important, but estimating peer effects credibly has been difficult due to data and methodological limitations. In this study, we overcome these identification problems by utilizing nationally representative data from a unique setting where students are randomly assigned to classes and are therefore randomly assigned to classroom peers. By identifying low-ability students by their grade retention history in primary school, we

22. The proportion of male and female repeaters are calculated by dividing the number of male/female repeaters by the total number of students in the class minus one; 46 repeaters have missing information on gender, and are thus dropped from this analysis. Among repeaters who have nonmissing information on gender ($N = 1{,}346$), 785 (58 percent) are male and 561 (42 percent) are female.

23. We also examine whether male repeaters have larger impact on male nonrepeaters than female nonrepeaters and do not find any significant gender-matching effects.

24. Similar to our main analyses, we also conduct multiple hypothesis testing adjustments for our analyses that differentiate between male and female repeaters. The adjusted *p*-values are presented in Columns 4 and 6 in Online Appendix Table 11. Although the adjustments lead to higher *p*-values in general, the estimated effects of male repeaters remain significant on nonrepeaters' academic performance, cognitive assessment score, after-school study time, having friends who have general disciplinary problems, and having friends who go to internet cafés regularly.

are able to address "the reflection problem" (Manski 1993) and make a clear distinction between the subjects of a peer effects investigation (that is, nonrepeaters) and the peers who provide the mechanism for causal effects (that is, repeaters).

Our results indicate that having a greater proportion of repeater classmates leads to a statistically significant reduction in nonrepeaters' academic performance and cognitive ability. Specifically, our findings indicate that a one percentage point increase in the proportion of repeater peers decreases nonrepeaters' academic performance by 2.1 percent of a standard deviation and cognitive ability by 2.3 percent of a standard deviation. This finding is consistent with the existing literature on the negative impact of low-ability or disruptive peers on cognitive outcomes but has a notably larger effect size (Carrell and Hoekstra 2010; Carrell, Hoekstra, and Kuka 2018; Lavy, Paserman, and Schlosser 2012). For example, Lavy, Paserman, and Schlosser (2012) find that a one percentage point increase in the proportion of repeaters leads to a decrease of between 0.47 percent and 1.13 percent of a standard deviation in the average scores of nonrepeater students. Similarly, Carrell, Hoekstra, and Kuka (2018) report that a one percentage point increase in the proportion of disruptive peers reduces achievement by 0.35 percent of one standard deviation. One possible explanation for the larger effect size identified in our study is that we measure peer composition at the classroom level instead of at the grade or cohort level, where the former is likely to capture peer interactions in a more precise way.

We also identify negative impacts of low-ability peers on nonrepeaters' noncognitive outcomes, such as school engagement and educational expectations, though these estimates are less robust to different model specifications. Our subsequent heterogeneity analyses indicate that male repeaters primarily drive the negative effects of repeaters, which is consistent with the findings in Carrell and Hoekstra (2010) and Carrell, Hoekstra, and Kuka (2018). Additionally, the negative externalities of repeater peers seem to be more pronounced among nonrepeater students from families with less strict parental monitoring at home.

In exploring possible mechanisms driving these effects, we find that reduced after-school study hours is the most robust of the three channels explored in our study, although the size of the effect is fairly small. We further find that a greater proportion of repeater classmates increases nonrepeaters' probability of having friends who play at internet cafés regularly. These results therefore provide suggestive evidence in support of friendship and social networks as an operating channel through which repeaters may influence nonrepeaters' academic performance.

A large volume of literature from psychology indicates that unsupervised after-school time, especially when spent with peers with pre-existing disciplinary misbehaviors or poor relations to the school, is likely to negatively influence one's own behaviors (for example, Mahoney and Stattin 2000; Mahoney, Stattin, and Lord 2004). Results from our study strengthen this argument by providing quasi-experimental evidence that having peers who have a greater level of disciplinary problems and invest less time in after-school study is associated with negative externalities on students. The possibility that peer effects operate beyond school time therefore highlights the importance of schools and parents monitoring after-school activities as a way to mitigate possible negative influence of disruptive peers.

This study is conducted in the specific context of the Chinese education system, but the results might have policy implications for other systems, especially for settings with large proportions of low-ability students or students with disruptive behaviors. Existing

studies on peer effects at the primary and secondary education levels emphasize the negative influence that low-ability and poorly behaved students can have on teacher pedagogy and student group functioning in school. Yet, our results suggest that low-ability or troubled kids may also bring about negative externalities on their classmates through social networks and joint activities after school. This finding highlights the importance for policymakers to pay close attention to after-school supervision in settings where students are likely to be subject to negative peer influences.

References

Aizer, Anna. 2008. "Peer Effects and Human Capital Accumulation: The Externalities of ADD." NBER Working Paper 14354. Cambridge, MA: NBER.

Allen, Chiharu S., Qi Chen, Victor L. Willson, and Jan N. Hughes. 2009. "Quality of Design Moderates Effects of Grade Retention on Achievement: A Meta-analytic, Multilevel Analysis." *Educational Evaluation and Policy Analysis* 31(4):480–99.

Ammermueller, Andreas, and Jörn-Steffen Pischke. 2009. "Peer Effects in European Primary Schools: Evidence from The Progress in International Reading Literacy Study." *Journal of Labor Economics* 27(3):315–48.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103(484):1481–95.

Angrist, Joshua D. 2014. "The Perils of Peer Effects." *Labour Economics* 30:98–108.

Bauminger, Nirit, Ricky Finzi-Dottan, Sagit Chason, and Dov Har-Even. 2008. "Intimacy in Adolescent Friendship: The Roles of Attachment, Coherence, and Self-Disclosure." *Journal of Social and Personal Relationships* 25(3):409–28.

Bifulco, Robert, Jason M. Fletcher, and Stephen L. Ross. 2011. "The Effect of Classmate Characteristics on Post-Secondary Outcomes: Evidence From the Add Health." *American Economic Journal: Economic Policy* 3(1):25–53.

Booij, Adam S., Edwin Leuven, and Hessel Oosterbeek. 2017. "Ability Peer Effects in University: Evidence from a Randomized Experiment." *Review of Economic Studies* 84(2):547–78.

Borghans, Lex, Angela. L. Duckworth, James. J. Heckman, and Bas Ter Weel. 2008. "The Economics and Psychology of Personality Traits." *Journal of Human Resources* 43:972–1059.

Burke, Mary A., and Tim R. Sass. 2013. "Classroom Peer Effects and Student Achievement." *Journal of Labor Economics* 31(1):51–82.

Carrell, Scott E., Richard L. Fullerton, and James E. West. 2009. "Does Your Cohort Matter? Measuring Peer Effects in College Achievement." *Journal of Labor Economics* 27(3):439–64.

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

Carrell, Scott E., Mark Hoekstra, and Elira Kuka. 2018. "The Long-Run Effects of Disruptive Peers." *American Economic Review* 108(11):3377–415.

Carrell, Scott E., Bruce I. Sacerdote, and James E. West. 2013. "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation." *Econometrica* 81(3):855–82.

Chen, Huiying. 2013. "Yi Wu Jiao Yu Jie Duan Liu Ji Zhi Du Zhi Shen Si." ("Examining China's Grade Retention Policy in Compulsory Education.") *Educational Measurement and Evaluation* 2013(1):23–7.

Chen, Xinxin, Chengfang Liu, Linxiu Zhang, Yaojiang Shi, and Scott Rozelle. 2009. "Does Taking One Step Back Get You Two Steps Forward? Grade Retention and School Performance in Poor Areas in Rural China." *International Journal of Educational Development* 30(6):544–59.

Deming, David. 2009. "Early Childhood Intervention and Life-cycle Skill Development: Evidence from Head Start." *American Economic Journal: Applied Economics* 1(3):111–34.
Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review* 101(5):1739–74.
Eccles, Jacquelynne S., Mina N. Vida, and Bonnie. Barber. 2004. "The Relation of Early Adolescents' College Plans and Both Academic Ability and Task-Value Beliefs to Subsequent College Enrollment." *Journal of Early Adolescence* 24(1):63–77.
Feld, Jan, and Ulf Zölitz. 2017. "Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects." *Journal of Labor Economics* 35(2):387–428.
Figlio, David N. 2007. "Boys Named Sue: Disruptive Children and Their Peers." *Education Finance and Policy* 2(4):376–394.
Fredericks, Jennifer. A., Phyllis C. Blumenfeld, and Alison H. Paris. 2004. "School Engagement: Potential of the Concept, State of the Evidence." *Review of Educational Research* 74:59–109.
Gong, Jie, Yi Lu, and Hong Song. 2018. "The Effect of Teacher Gender on Students' Academic and Noncognitive Outcomes." *Journal of Labor Economics* 36(3):743–78.
Gong, Jie, Yi Lu, and Hong Song. 2021. "Gender Peer Effects on Students' Academic and Noncognitive Outcomes: Evidence and Mechanisms." *Journal of Human Resources* 56(3):686–710.
Gould, Eric D., Victor Lavy, and M. Daniele Paserman. 2009. "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence." *Economic Journal* 119(540):1243–69.
Gunuc, Selim. 2016. "Peer Influence in Internet and Digital Game Addicted Adolescents: Is Internet/Digital Game Addiction Contagious?" *International Journal of High Risk Behaviors and Addiction* 6(2):1–8.
Hao, Lingxin, and Xiao Yu. 2015. "Rural–Urban Migration and Children's Access to Education: China in Comparative Perspective." Paper Commissioned for the Education for All Global Monitoring Report 2015, Education for All 2000–2015: Achievements and Challenges.
Hillman, Stephen B., Shlomo S. Sawilowsky, and Marilyn J. Becker. 1993. "Effects of Maternal Employment Patterns on Adolescents' Substance Use and Other Risk-Taking Behaviors." *Journal of Child and Family Studies* 2(3):203–19.
Hoxby, Caroline M. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." NBER Working Paper 7867. Cambridge, MA: NBER.
Jimmerson, Shane. 2001. "Meta-analysis of Grade Retention Research: Implications for Practice in the 21st Century." *School Psychology Review* 30(3):420–37.
Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75(1):83–119.
Kaiser, Henry F. 1974. "An Index of Factorial Simplicity." *Psychometrika* 39(1):31–6.
Lavy, Victor, M. Daniele Paserman, and Analia Schlosser. 2012. "Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom." *Economic Journal* 122(559):208–37.
Lavy, Victor, and Analia Schlosser. 2011. "Mechanisms and Impacts of Gender Peer Effects at School." *American Economic Journal: Applied Economics* 3(2):1–33.
Lefgren, Lars. 2004. "Educational Peer Effects and the Chicago Public Schools." *Journal of Urban Economics* 56(2):169–91.
Li, Li, and Haoming Liu. 2014. "Primary School Availability and Middle School Education in Rural China." *Labour Economics* 28:24–40.
Li, Yajun, Xinghui Zhang, Furong Lu, Qin Zhang, and Yun Wang. 2014. "Internet Addiction among Elementary and Middle School Students in China: A Nationally Representative Sample Study." *Cyberpsychology, Behavior, and Social Networking* 17(2):111–16.
Liu, Jingqiu, and Ruth Barnhart. 1999. "Homeroom Teacher and Homeroom Class: The Key to Classroom Management in China's Schools." *Educational Forum* 63(4):380–84.

Löwe, Bernd, Kurt Kroenke, and Kerstin Gräfe. 2005. "Detecting and Monitoring Depression with a Two-Item Questionnaire (PHQ-2)." *Journal of Psychosomatic Research* 58(2):163–71.
Lyle, David S. 2007. "Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point." *Review of Economics and Statistics* 89(2):289–99.
Mahoney, Joseph L., and Håkan Stattin. 2000. "Leisure Activities and Adolescent Antisocial Behavior: The Role of Structure and Social Context." *Journal of Adolescence* 23(2):113–27.
Mahoney, Joseph L., Håkan Stattin, and Heather Lord. 2004. "Unstructured Youth Recreation Centre Participation and Antisocial Behavior Development: Selection Influences and the Moderating Role of Antisocial Peers." *International Journal of Behavioral Development* 28(6):553–60.
Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60(3):531–42.
McCord, Joan. 1991. "Family Relationships, Juvenile Delinquency, and Adult Criminality." *Criminology* 29(3):397–417.
McGee, R.O.B., Michael Feehan, Sheila Williams, and Jessie Anderson. 1992. "DSM-III Disorders from Age 11 to Age 15 Years." *Journal of the American Academy of Child & Adolescent Psychiatry* 31(1):50–59.
McNelles, Laurie R., and Jennifer A. Connolly. 1999. "Intimacy between Adolescent Friends: Age and Gender Differences in Intimate Affect and Intimate Behaviors." *Journal of Research on Adolescence* 9(2):143–59.
Menaghan, Elizabeth G., and Toby L. Parcel. 1991. "Determining Children's Home Environments: The Impact of Maternal Characteristics and Current Occupational and Family Conditions." *Journal of Marriage and the Family* 53(2):417–31.
Mendez, Linda M. Raffaele, and Howard M. Knoff. 2003. "Who Gets Suspended from School and Why: A Demographic Analysis of Schools and Disciplinary Infractions in a Large School District." *Education and Treatment of Children* 26(1):30–51.
Ministry of Education of the People's Republic of China. 2015a. "Number of Schools, Classes of Junior Secondary Education." http://www.moe.gov.cn/s78/A03/moe_560/jytjsj_2014/2014_qg/201509/t20150906_205423.html (accessed September 2, 2021).
Ministry of Education of the People's Republic of China. 2015b. "Number of Schools, External Teaching Sites, and Classes in Primary Schools." http://www.moe.gov.cn/s78/A03/moe_560/jytjsj_2014/2014_qg/201509/t20150901_204658.html (accessed September 2, 2021).
Moffitt, Terrie E., Louise Arseneault, Daniel Belsky, Nigel Dickson, Robert J. Hancox, HonaLee Harrington, Renate Houts, Richie Poulton, Brent W. Roberts, Stephen Ross, Malcolm R. Sears, W. Murray Thomson, and Avshalom Caspi. 2011. "A Gradient of Childhood Self-Control Predicts Health, Wealth, and Public Safety." *Proceedings of the National Academy of Sciences* 108(7):2693–98.
Organization for Economic Co-operation and Development. 2016. "Education in China: A Snapshot." https://www.oecd.org/china/Education-in-China-a-snapshot.pdf (accessed September 9, 2021).
Parcel, Toby L., and Elizabeth G. Menaghan. 1994. "Early Parental Work, Family Social Capital, and Early Childhood Outcomes." *American Journal of Sociology* 99(4):972–1009.
Reuters. 2007. "Worried China Bans New Internet Cafes for a Year." https://www.reuters.com/article/us-china-internet-ban/worried-china-bans-new-internet-cafes-for-a-year-idUSPEK24407720070307 (accessed September 9, 2021).
Romano, Joseph P., and Michael Wolf. 2016. "Efficient Computation of Adjusted *P*-values for Resampling-Based Stepdown Multiple Testing." *Statistics & Probability Letters* 113:38–40.
Sacerdote, Bruce. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics* 116(2):681–704.
Wang, Xueli. 2013. "Why Students Choose STEM Majors: Motivation, High School Learning, and Postsecondary Context of Support." *American Educational Research Journal* 50(5):1080–121.

Zhang, Chi. 2014. "Dui Wo Guo Yi Wu Jiao Yu Jie Duan Liu Ji Zhi Du De Si Kao." ("China's Grade Retention Policy in Compulsory Education.") *Journal of Educational Development* 2014(1):25–7.
Zhao, Guochang, Jingjing Ye, Zhengyang Li, and Sen Xue. 2017. "How and Why Do Chinese Urban Students Outperform Their Rural Counterparts?" *China Economic Review* 45:103–23.
Zhao, Zhenzhou. 2014. "The Teacher–State Relationship in China: An Exploration of Homeroom Teachers' Experiences." *International Studies in Sociology of Education* 24(2):148–64.
Zhiyan Consulting Group. 2016. "Statistics and Prospects of Internet Café in China: 2016." http://www.chyxx.com/research/201609/445408.html#catalogue (accessed September 9, 2021).
Zimmerman, David J. 2003. "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment." *Review of Economics and Statistics* 85(1):9–23.
Zoccolillo, Mark. 1993. "Gender and the Development of Conduct Disorder." *Development and Psychopathology* 5(1–2):65–78.