Peer Effects in Classrooms – Evidence from Random Assignment

Lei Xu[a, c, d *] and Yu Zhu[b, c, d]

[a] Department of Accounting, Finance and Economics, Bournemouth University, BH12 5BB

[b] University of Dundee School of Business, 3 Perth Road, Dundee, DD1 4HN

[c] IZA Institute of Labor Economics, Schaumburg-Lippe-Strasse 5-9, 53113 Bonn, Germany

[d] Global Labor Organization

Version 26-Feb-23

Abstract

Identifying peer effect is a daunting empirical challenge. By exploiting a random assignment design, this paper estimates distinctive peer effects resulting from the interaction of students with peers in the same classroom who have different attributes. To identify the direct impacts from peers, we use the term "peer" to denote the direct impact of the closest friends, who generate causal peer effect and use the term "classmates" to denote the impact of other students in a class. We find that being exposed to more advantaged peers increases test scores and the estimated peer effects vary by peers' attributes. Motivated by the recent development on measurement error in the peer effect literature, we provide evidence on the violation of non-randomisation in previous research and the potential weakness of the balancing test (Guryan et al., 2009). We show that estimates based on the commonly used *leave-own-out* measures are highly sensitive to non-random tracking in the sample.

JEL Classification: I20, I24

Keywords: peers' parents, random assignment, China

*: Corresponding Author: Business School, Department for Accounting, Finance and Economics, Bournemouth University. Email: lxu1@bournemouth.ac.uk.

1. Introduction

The persistent, or even widening, gap in academic achievements between advantaged and disadvantaged students, typically measured by parental qualifications, occupations, and ethnic groups, have been well documented by economists (Fryer and Levitt, 2004). Among sociologists and ethnographers, the cultural reproduction theory, which argues that the lack of equal access to social capital underpins the widening gap in academic achievements, has been particularly influential (Bourdieu, 1977). To the extent that students with better family background may possess different cultural dispositions (Giroux, 1983), and tend to have more social capital and beneficial habitus (Gaddis, 2013), schools are not institutions promoting equality of opportunity but enhancing the gap between social classes (Collins, 2009).

Contrary to the cultural reproduction theory, there is growing evidence on the role of schools in providing an upward mobility (Gaddis, 2013). DiMaggio (1982) argue that social status acts as a cultural process rather than as an attribute of individuals and find that cultural capital is less strongly tied to parental background. Those students with poorer backgrounds may gain cultural capital or have higher academic achievements by actively participating in prestigious status cultures. Gaddis (2013) find that the habitus as measured by high-arts participation and reading habits has positive impacts on academic achievements for disadvantaged students. The theory and the finding lead to the discussion of the role of schools and skills acquisition in schools. Although advantaged students from better socio-economic backgrounds may have stronger cognitive or socio-economic skills, disadvantaged students may benefit from the interaction with advantaged students.

The spillover effect from advantaged students has attracted great attention by economists. The research on peer effects is effectively based on an assumption that ‘advantaged’ students have higher academic achievements and there is spillover effect from them. The literature has grown exponentially recently (see Sacerdote 2011). Identifying peer effects in empirical research is challenging due to the selection and the reflection problems (Manski 1993). In addition, the very existence of peer effect has also been questioned in the presence of measurement error and potential misspecification of identification strategies (Angrist 2014; Feld and Zolitz 2017).

To address self-selection into groups, previous research has exploited cross-cohort variations within a school and exogenous policy changes (Figlio and Ozek 2019). While most peer effect studies rely on cross-cohort variation within a group, these estimates might be biased due to the prevalence of self-sorting and streaming. This has motivated an emerging literature using random class assignment design, quasi-experiment design, or random control trials (Kang 2007; Duflo et al 2011; Wang et al 2018; Mendolia et al 2018; Huang and Zhu 2020). Most of the research using a random assignment has employed linear regression to examine peer effects and the mechanisms.

However, the estimates using group mean variables or *leave-own-out* measures may suffer from a number of limitations.[1]

The most important limitation perhaps, is whether group mean variables correctly measure peer effects.[2] The research has examined peer effects after being exposed to higher proportion of students with different backgrounds, measured by *leave-own-out* variables, including single child (Cai, Fan, and Yuan, 2022), child with educated parents (Chung and Zou, 2020), child with alcoholic parents (Zhao and Zhao, 2021), repeaters (Huang and Zhu, 2020; Xu et al., 2020), higher same-gender ratio (Luo and Yang, 2022), migrant children (Fent 2018). They interpret the variation in the backgrounds as altering the environment which influences others in the same class, known as the contextual peer effect (Bramoulle, Djebbari, and Fortin, 2020).

The fundamental ambiguity of peer effect using *leave-own-out* variables to measure peers' quality results from omitted variables and the endogenous peer group formation.[3] The peer effect, measured by *leave-own-out* variables, represents the impact of peers-of-peers and the impact of endogenous group interaction. Following the summary of Bramoulle, Djebbari, and Fortin (2020), we view the peer effects in classrooms, estimated by *leave-own-out* variables, as the causal impact of studying in classrooms with different students' backgrounds rather than the direct impact of interacting with classmates having a certain background in the presence of omitted variables.[4] For instance, they argue that the causal impact of the proportion of smoking peers does not estimate

[1] A group mean variable refers to the mean value of a variable in a group and the leave-one-out measure refers to the mean value of a variable after excluding the subject concerned.

[2] Hanushek et al (2003) argue that peer effect is sensitive to measurement and specification of peers' attributes. The peer effect might also be biased because of ambiguous measurement errors resulting from group mean variables or leave-one-out measures in a group (Feld and Zolitz 2017).

[3] Peers provide mechanisms for causal effects but not the research subject (Angrist 2014).

[4] The presence of omitted variable connects with the so-called 'cross effect' in the literature of peer effect. An individual could be affected by different behaviours measured by a background variable.

the impact of smoking but the impact of being connected with students who are smoking. The research based on *leave-own-out* measures is thus ambiguous in the interpretation of the peer effect, give audience the impression that the peer effect has been driven by classmates' backgrounds.

Recent studies have used self-collected data to examine peer effect in the context of self-selection into peer groups (Carrell et al 2013). Calvano, Immordino, and Scognamiglio (2022) find that high-ability students interact more with other high-ability students. The recent peer effect literature discussing mechanisms of peer effect mainly focuses on the changes in individuals' inputs and only a few research has focused on the interaction and peers' attributes. In an experimental study, Li et al (2014) suggest that group incentive rather than individual incentive drive significant peer effect. Babcock et al (2020) shed light on how cooperation and contagion within peer groups generate peer effect. The most recent research has established a close connection between peers' personality and peer effect (Glosteyn et al 2020). Various studies have discovered non-linear peer effect and complex mechanisms driving peer effect (Carrell et al 2009; Booji et al 2017; Garlick 2018). The complexity largely results from complex interactions between peers and the associations between attributes and peer effects. The growing evidence on endogenous peer group formation may challenge the seemingly large peer effects found in research based on a random assignment design, as significant peer effects might be generated through the peers of peers and model misspecification such as violation of randomisation assumption. A research line has tried to model the complex social network and estimate the average effect from an endogenous peer group (Bramoulle et al., 2009; Lee et al., Jochman, 2022).[5]

Second, most of the recent research on peer effect using *leave-own-out* variables may suffer from Weak Instruments in a random assignment and the mechanical relationship between measures for peers' quality and outcomes (Angrist 2014). Intuitively, *leave-own-out* variables should not be statistically significant between randomly assigned groups, resulting in no treatment effect especially in groups with large number of observations. Otherwise, the random assignment assumption might be violated. The small differences between groups may result in a smaller treatment effect. On the other hand, there might be a mechanical relationship between outcomes and *leave-own-out* variables in the OLS regression when it somehow involves the dependent

[5] Bramoulle et al. (2009) have used a *leave-own-out* model to identify the interaction through a social network.

variable (Angrist 2014). Recent studies have suggested restricting the targeted group to mitigate this mechanical relationship (Angrist 2014; Carrell et al 2018).

Third, peer effect may have an impact on teacher's pedagogy, resulting in indirect impact on students. Disadvantaged peers may not have direct contact with advantaged peers and significant peer effects may be caused by changes in other inputs, such as teaching practices (Lavy et al 2012; Feld and Zolitz 2017; Zhao and Zhao 2021). Lavy et al (2012) find a significant negative effect of repeaters on classmates as teachers divert attention from regular students to repeaters. The indirect impact may be misinterpreted as peer effect in an econometric framework in which the responses of teachers and schools have not been accounted for.

In this paper, we examine the peer effects of being exposed to students whose parents are degree-educated and contribute to the literature by examining the relationship between peer's quality and peer effects. Taking advantage of random class assignment and direct responses on peers' quality in the China Education Panel Survey (CEPS), we employ a 2SLS strategy to measure the influence of peer's interaction resulting from distinctive peers' characteristics in a value-added framework by using the *leave-own-out* measure as proxies for peers' quality in the previous literature. We are one of the first papers to examine the distinctive impacts induced by peers' attributes using **naturally occurring peer groups.** In this paper, to identify the direct impacts from peers, we use the term "**peers**" to denote the direct source of impact from the closest friends in the endogenous "**peer group**", consisting of the self-reported up to 5 closest friends. In contrast, we use the term "**classmates**" to represent the source of impact from all other students in a class.

We regard measurement errors as the reflection of using *leave-own-out* variables to proxy for peers' quality. In this paper, we measure peers' quality directly using the responses on the characteristics of the closest friends in the class and rely on the randomly assigned classes to address the potential self-selection into classes. We also take advantage of the availability of information on all students in the class in the CEPS to derive a precise measure of the average quality of a class, using the pre-determined *leave-own-out* mean characteristics of all classmates in the classroom. Although we don't observe each peer in peer groups, we examine the extent to which students form new peer groups after entering a new class and the average peers' quality based on their backgrounds. Hence, peers' attributes are measured by the responses of individuals on their closest friends.

Instead of disentangling peer effects through changes in respondent's inputs, we focus on measuring peer's quality and behaviours, and estimate the impacts of peers' quality on academic and cognitive scores. Following the literature, we estimate the effect on academic tests of being exposed to advantaged classmates in the same classroom, defined as classmates with at least one college educated parent (Fruehwirth and Gagete-Miranda 2019; Chung and Zou 2020), and provide evidence on the violation of randomisation. We show that students with seemingly the same background may have different observed attributes, resulting in distinctive peer effects. Hence, we expect that the estimate of peer effect measured by students' backgrounds represents **weighted average peer effect** that is measured and estimated by directly observed peers' attributes. The distinctive peer effects using different group mean variables may be explained by different weights and estimates measured by observed peers' attributes.

We have three main observations.

First, we present novel evidence that the unbiasedness of *leave-own-out* measures strongly depend on the validity of random assignment, by exploiting the availability of information on classroom random assignment from alternative sources. While standard balance tests are not sensitive to minor violation of random assignment, estimates of peer effect are highly sensitive to the violation of the random assignment condition and the specifications of the reduced form equation (Hanushek et al 2003). Based on the responses on classroom randomisation from both principals and subject teachers, we show substantial inconsistency in their responses, with disagreements in about 60% of the cases. The conventional balancing tests also suggest that the sample only based on principals' responses have higher risks of non-randomisation. The bootstrapping results suggest that the samples based on two sample selection rules are likely to be drawn from different distributions after examining the distribution of balancing test used in the literature including other measures, such as classmates with alcoholic parents (Zhao and Zhao 2021), and classmates who have repeated grades in primary schools (Huang and Zhu 2020; Xu et al 2020). To examine the extent to which peer effects are driven by the non-randomisation in the data, we estimate peer effects using different sample selection rules and demonstrate that peer effects are highly sensitive to selection bias when using *leave-own-out* measures. The highly significant and strong peer effects diminish massively in the sub-sample using the strictest sample selection rule.

Second, we show that being exposed to advantaged peers in the OLS results, over an extended period, generally increases test scores based on stricter sample selection rule on randomised classes. The effect is more pronounced among students with advanced peers, suggested by the endogenous peer group formation (Carrell et al 2013) and the non-linear peer effect (Carrell et al 2009). Inspired by these heterogenous results, we show that the impacts of peers on test scores are different after being exposed to peers with distinctive observed behavioural attributes. Being exposed to hard-working peers could increase test scores more than being exposed to peers who have higher aspirations for college and have better academic achievements. The results imply that it is not the peers *per se* but peers' attributes that generate peer effects and the multiplier of the social returns. Being exposed to more diligent peers may encourage disadvantaged students to work harder on studies. On the other hand, college aspiration alone may not generate strong externalities on peers. This is related to the research examining the relationships between group incentive and peers' personality and peer effects (Li et al 2014).

Third, we observe complex social interaction between students. By examining the heterogeneous peer effects, we show that peer effects are more pronounced among rural students, or students without degree-educated parents. However, the peer effect is more robust among students with better academic background. The heterogenous OLS results imply the importance of interactions between students. One possible explanation is that students with poorer family backgrounds could benefit more from advantaged peers as peer group formation is endogenous. Students with stronger academic background might be more likely to interact with advantaged peers compared to their counterparts. [6] However, using the 2SLS strategy, contrary to the OLS results, we find urban students can benefit from advantaged peers, possibly due to the higher social capital possessed by urban students, resulting in higher probability of interacting with advantaged students.

The remainder of the paper is organised as follows. Section 2 reviews the relevant literature. Section 3 introduces the baseline methodology and the data used in this paper, and our *lax* and

[6] We also find heterogeneous peer effects that peers who have strong academic backgrounds and are more willing to interact with peers may benefit more from advantaged peers, consistent with the non-linear peer effect found in previous research. The evidence on mechanisms indicates that social interaction and its correlation with individuals' characteristics are complex and indeed under-studied.

strict selection rules with regard to random class assignment to define the analytical samples. Section 4 demonstrates the 2SLS design and presents the main results and potential bias from the violation of randomness. We also discuss the mechanisms through which peer effects could be generated. Lastly, section 5 concludes.

2. Literature review

The peer effect in schools has attracted great attention and indeed the knowledge has been rapidly advanced over the last decades. The empirical strategy of estimating peer effect has been constantly updated. Hoxby (2000) makes use of variation of demographic composition of a grade in a school to estimate the peer effect. Kang (2007) exploits a quasi-experiment in Korea and use science test score as an IV for the math score of peers. Fruehwirth and Gagete-Miranda (2019) include school fixed effect to address the selection across schools and assume the assignment of student within schools is random. Mendolia et al (2018) have proposed to use peers-of-peers' ability in primary school as an instrument for the mean high school peer ability.

Various research has employed the *leave-own-out* measure based on random assignment design to address the self-selection into classes or schools. The *leave-own-out* measure is sufficient to identify the endogenous peer effect if the exogeneity holds (Bramoulle, Djebbari, and Fortin, 2020). Making use of the random assignment element in China Education Panel Survey, various research has found sizable peer effect in classroom for student in secondary education based on different students' backgrounds. Chung and Zou (2020) identify the relationship between peers' maternal education and academic achievements, using *leave-own-out* maternal education in classes. Zhao and Zhao (2021) use the proportion of peers with alcoholic fathers to show the adverse effect on test scores for having peers with troubled families. Huang and Zhu (2020) look at the effect of more grade repeaters on non-repeaters' outcomes using the same data but focus more on the heterogeneity and dynamics of the peer effect. They show that the peer effects are strongest at the bottom end of the achievement distribution but insignificant at the top end. Xu *et al.* (2020) show that the proportion of grade repeaters during primary school has negative effects on non-repeaters' cognitive and non-cognitive outcomes. They further show that much of the negative repeater peers' effect is driven by reduced after-school study time, male repeaters and students with less strict parental monitoring at home. Wang et al (2018) have found educational spillover effects of migrant students on local students' academic achievement, using the proportion of migrant students to measure the treatment intensity. Balestra et al (2021) find exposure to gifted students could raise educational achievements and the likelihood of getting into academic track and STEM fields. They also find the positive externalities is higher for female students. Chung (2020) examines the effect of exposure to peers with educated parents by exploiting within-school cohort variations in parental compositions and argue that students are disproportionally affected by their peers and the peer effect is more pronounced among less advantaged students.

There is rapidly growing research in conducting experiments to understand the interaction between peers. Making use of a quasi-experiment through which students have been randomly reallocated across groups, Carrell et al (2009) have found heterogeneous peer effects across different subjects. The peers' quality is measured by pre-treatment characteristics. Carrell et al (2013) conduct an experiment to examine the peer effect and argue that the estimated peer effect could be affected by endogenous peer groups formation. Duflo et al (2011) has used a natural experiment to examine the effect of random allocation of peers. The large amount of empirical research has suggested that being exposed to peers with favourable backgrounds tend to have a positive social multiplier.

However, a growing literature has highlighted the importance of measurement errors and the factors driving the large peer effects. Most of the research is based on the *leave-own-out* measures of pre-determined variables. However, the balance tests might be insufficient to ensure the randomness of students' assignment, the crucial assumption to the estimation of peer effect using the *leave-own-out* measures. Although the empirical strategy based on *leave-own-out* measures and lagged dependent variable has various merits, it also raises questions regarding the inability to capture time-variant individual's behaviours induced by their peers, resulting in difficulty in interpreting the peer effect (Hanushek et al 2003). Hanushek et al (2003) has argued that the estimated peer effects are also sensitive to the measurement. Feld and Zolitz (2017) have examined the role of measurement errors on the estimation of peer effect when using group mean variables and argue that it introduces ambiguous bias into the results, resulting in either overestimated or underestimated estimates. Using the background variables may not only introduce measurement errors but also lead to ambiguous interpretations of peer effect. Angrist (2014) has argued that peer effects are the mechanisms of connecting peers in a group rather than the subject for study, suggesting that the advantaged students, measuring by pre-determined variables, provide the mechanisms of affecting disadvantaged peers.

Moreover, the recent literature has documented nonlinear and heterogeneous peer effect (Carrell 2009; Booji, et al 2017; Garlick 2018). Carrell et al (2013) document the evidence of self-selection into peer groups and argue that students tend to form more homogeneous peer groups. The positive peer effect of middle ability group is resulting from the less interaction students in the low ability group. They find a negative peer effect among the lowest ability students when being exposed to high ability peers. Consistent with Carrell et al (2013), Booji et al (2017) argue that peer effects are not linear and find negative peer effects among low ability students when being exposed to high ability students, resulting from the negative relationship between interaction and group diversity. They also find positive peer effect among middle ability students benefiting from ability tracking.

The complex results may result from the lack of understanding on mechanisms and its connection with attributes of peers. Li et al (2014) conduct an experiment to study peer effect in schools and argue that the peer effect results from group incentives rather than individual incentives or better peers. It reveals an important implication that the social interaction is complex but an important source of generating the social multiplier. Babcock et al (2020) also conduct an experiment to shed light on peer mechanisms and argue that the peer effect is likely the result from coordination between students. The two studies generate important insight on how peers interact each other and generate the peer effect. A more recent study, using random assignment design in a university, examines the relation between student's personality and peer effect and argue that different characteristics can generate distinctive peer effect (Golsteyn et al (2021)). Being exposed to conscientious peers have influence on students' performance, while the exposure to risk-tolerant peers does not. Peer anxiety and self-confidence do not generate significant peer effects.

Another line of research is to model the endogenous peer network and quantify the differential peer effects in the network matrix. In the presence of the correlation between peer group formation and outcomes, researchers have tried to capture the social network to estimate peer effects (Bramoulle, Djebbari, and Fortin, 2009; Hsieh, and Lee, 2016). Johnsson and Moon (2021) propose a semi-parametric estimation to implement a control function method to address the self-selection into peer groups in the presence of the correlation of unobserved factors between peer group formation and outcomes. Goldsmith-Pinkham and Imbens (2013) investigate the extent to which the peer effect is generated through indirectly channel based on the Manski's linear-in-means model. Based on the model estimated by a Bayesian approach, they argue that both the direct effect from the connection and the indirect effect from peers of peers contribute to the correlation between current friends and current grades. Jochmans (2022) argues that individual's outcome could be influenced by individual's own characteristics, peers' outcome, and peers' characteristics and the *leave-own-out* measures have predictive power in individual's link behaviours. He proposes to *leave-own-out* characteristics as an instrument to proxy for a network matrix.

3. Baseline empirical strategy and Data

To overcome the selection problem, we take advantage of **both** the principals' and the subject teachers' responses on random assignment to identify random assigned classes at Grade 7. Following the literature, we employ a value-added framework to estimate the peer effect (Hanushek et al, 2003), in which baseline scores are used to proxy cumulative past inputs. We first follow the literature to estimate the relationship between *leave-own-out* means and outcomes.

$$y_{igst} = \alpha y_{isg,t-1} + \beta Cla_{-ig} + \delta X_i + \theta T_i + \varepsilon_{it} \tag{1}$$

where y_{igst} represents the standardised cognitive test scores or exam scores for student *i* in class g at grade *t*, while s denotes the subject. Cla_{-ig} denotes the *leave-own-out* measure using proportion of classmates with parents having a degree, and X_i denotes time-invariant individual, including individuals' characteristics, family background and attrition rate between Grade 7 and 8. T_i denotes teachers' characteristics. The lagged dependent variable $y_{isg,t-1}$ denotes the corresponding scores in Grade 7. The parameter of interest is the coefficient on Cla_{-ig}, which denotes the influence of classmates on test scores. The variable measures the influence of studying in a classroom after being exposed to higher proportion of advantaged students, resulting from the impact of peers of peers and the direct interaction with advantaged peers in endogenous peer groups. The measure does not say anything with regard to what advantaged students could do to create the peer effect.

The China Education Panel Survey (CEPS) is a large-scale, nationally representative longitudinal survey by the National Survey Research Centre (NSRC) at Renmin University of China, starting with two cohorts – the 7th and 9th graders in the baseline survey conducted in in the academic year 2013-14 (https://ceps.ruc.edu.cn/index.php?r=index/index&hl=en). The baseline survey contains 5 different questionnaires for the sampled students, parents, class headteachers,[7] core subject teachers other than headteachers, and school principals respectively. Moreover, the survey includes a standardized cognitive ability test for students in each grade respectively and an internet-based personality test for all sample students and collects transcripts of important (mid-term) examinations. The CEPS follows a stratified, multistage sampling design with probability proportional to size (PPS), randomly selecting a school-based, nationally representative sample of approximately 20,000 students in 438 classrooms of 112 schools in 28 county-level administrative units in mainland China. In each relevant grade, all students from two randomly selected classes are included in the survey.

The student questionnaire covers students' demographic characteristics, mobility and migration status, childhood experience, health status, household structure, parent-child interactions, in-school performance, extracurricular activities, relationship with teachers and peers, social behaviour development, and expectations for the future. The parent questionnaire covers parents' demographic characteristics and lifestyles, parent-child interactions, educational environment and investment for child, community environment, parent-teacher interactions, and parents' perceptions of school education and expectations for the future of the child. The questionnaires for headteachers and core subject teachers cover teachers' demographic characteristics, teaching experience, comments on student behaviours, parent-teacher interactions, comparison between local and non-local students, perceptions of education, and degree of stress and job satisfaction. The questionnaire for school principals asks about their demographic characteristics, perceptions of education, school's educational facilities, daily management, enrolment of students, statistics of the student body and staff body, and other school characteristics. Importantly for our study, both the principal's questionnaire and the subject teachers' questionnaire ask whether students are randomly assigned to classes in the relevant grade.

Our main sample includes all Grade 7 students in the 2013-14 baseline survey, as well as the follow-up survey of the Grade 7 cohort in the following academic year.[8] The paper has examined various aspects of peer effect. The small differences in the number of observations are due to the

[7] A class headteacher under the Chinese education system is a designated teacher with overall responsibility for a particular class, and is responsible for establishing class rules, leading class actions and providing non-academic support to all students in the class.

[8] We exclude Grade 9, due to failure of randomisation arising from regrouping, attrition and entry.

small missing values of different variables when we explore the channels affecting the peer effect and the formation of peer groups.

We firstly show how we make use of the responses of both schools' principal and subject teachers on the randomisation of students into classes. Table 1 shows the roadmap of sample selection. There are 106 schools with full information on class assignment in Grade 7. Of these, 40 and 66 schools follow non-random and random assignment rule respectively, according to the principals' questionnaire only. The latter group forms our analytical sample, by what we term the *lax rule* which is in keeping with previous CEPS studies. In contrast, the *strict rule* is defined as random assignment cases with full agreement between principals and teachers, which is satisfied by only 27 (40%) schools in the analytical sample. Table A1 presents the summary statistics.

Table 1: Roadmap of sample selection

	Steps	**Sample composition**
Data preparation	Linked student, parent, teacher and principal's data in the CEPS 2013-2014 baseline survey	112 school, 438 classes
	Drop 8 classes due to missing information in teacher's data	112 schools, 430 classes
	Grade 7 only	112 schools, 218 classes
	Schools with class-pairs in Grade 7 (sample used for balancing tests in Table 1)	106 schools, 212 classes
Analytical sample	***Lax Rule*** (Schools with random class assignment in Grade 7 according to principals' questionnaires only)	66 schools, 132 classes
Of which	***Strict Rule*** (Subset of schools where there is full agreement between principals and all subject teachers)	27 schools, 54 classes

4. Results.

4.1. Average peer effect.

Following the literature having examined the impacts of advantaged peers in the research of peer effect, we start by estimating the peer effect using the linear-in-means model on test scores based on the value-added framework. We use the strict rule on sample selection to avoid the disruption from non-randomisation and has exclude the corresponding (dis)advantaged group to test the robustness due to the mechanical relationship between the group mean variables and the dependent variable.

We find significant effect of advantaged groups on test scores under three different specifications. Having educated parents peers will have positive impacts on test scores, consistent with the literature. It is worth noting that the empirical setting differs from previous research. First, we are employing a value-added framework using the test score in Grade 7 as the measure for cumulative academic performance. Second, all our results are based on the strict rule on sample selection. We will discuss the impact of sample selection and specifications in the next subsection of the paper.

Table 2. Peer effects measured by the *leave-own-out* measure using the strict sample

	(1)	(2)	(3)
		Dependent variable: Test scores	
Educated parents peers' share	0.441**	0.441**	0.628***
	(0.20)	(0.20)	(0.21)
Educated parents peers' share at school level		5.387	0.578
		(12.06)	(0.76)
Constant	-1.158***	-4.558	-1.659**
	(0.31)	(7.63)	(0.73)
Specification			Excluding students with degree-educated parents
N	5,856	5,856	4,314

Notes: Grade 7-8 panel data. The results are based on the strict sample selection rule. The estimates of having educated parents peers in the third column have excluded students with parents having a degree. Exam scores are normalised by school and standard errors (in parentheses) are clustered at the class-subject level. Controls variables include individual characteristics such as corresponding scores in Grade 7, attrition rate, class size, gender, hukou, age, single child, parent's highest education, school and subject fixed effects, and subject teachers' characteristics include gender, age, educational level; teacher characteristics include gender, teaching certificate, experience, and qualification. The following regressions have the same control variables unless specified otherwise.

The *leave-own-out* measures do not shed light on how the peer effect has been generated. Due to the fact that peer group formation is endogenous, peer effects is likely to vary with students' backgrounds, resulting in non-linear peer effects. The peer effect is affected by individuals' attributes, as well as peers' attributes. To examine the heterogenous effect resulting from different attributes, we first estimate the peer effect based on students' interaction with peers following the same specification in Table 2.

Table 3 reports the heterogenous peer effects based on friends' attributes, such as whether friends have good academic achievement, whether friends work hard, or whether friends are more willing to attend a university. We test whether students' attributes matter for estimating peer effect given the fact that peer groups are self-selected. The results suggest that the peer effect measured by the *leave-own-out* measure is much larger amongst students who do not have friends with helpful attributes for academic achievements. However, due to the endogenous peer group formation, advantaged students are more likely to form peer groups with advantaged students although it seems that disadvantaged students may benefit more from the interaction with other advantaged students, resulting from higher probability of interacting with advantaged students due to the higher density of advantaged students. The heterogenous peer effects motivate us that individuals' attributes matter when estimating peer effects due to the endogenous peer group formation.

Table 3. Heterogenous peer effects varied by peer groups

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Test scores					
	Having many friends who study well	Not having many friends who study well	Having many friends who work hard	Not having many friends who work hard	Having many friends who are degree-motivated	Not having many friends who degree-motivated
Educated parents peers' share	0.319	0.572*	0.373*	0.485*	0.252	1.183***
	(0.20)	(0.30)	(0.20)	(0.28)	(0.22)	(0.33)
Educated parents peers' share at school level	55.596***	-37.895**	35.338**	-22.533	32.634**	-71.520***
	(17.47)	(16.69)	(14.87)	(17.09)	(14.06)	(20.32)
N	2811	3045	2970	2886	4341	1515

Note: The results including estimates measured by the *leave-own-out* variable. The setting is the same as that in Table 2.

4.2. Peer' attributes and peer effect

Due to the presence of omitted variables, a *leave-own-out* variable measures the reflection of having higher proportion of students with certain backgrounds rather than how classmates' attributes generate peer effect.[9] Students without being exposed directly to advantaged classmates may be affected by the peers of peers. It also cast doubts on the issue of weak IV as the random assigned groups are supposed to be similar. The average treatment effect, estimated by instruments (*leave-own-out* variables), might be small due to the nature of a random assignment. Any significant differences in backgrounds between groups would cast doubt on the validity of random assignment.

Following the work constructing the peer group network (Jachmons 2022), we employ Two-Stage Least Squares (2SLS) to examine the distinctive peer effects resulting from the interaction with peers having different attributes in a value-added framework.[10] The *leave-own-out* variable across randomly assigned classes is used as instruments to estimate the network matrix. Instead of using *leave-own-out* variables as proxies for peers' quality, our data contains detailed responses on peers' behaviours used to measure peers' quality directly. Therefore, our empirical strategy estimates the Local Average Treatment Effect (LATE), induced by being exposed to the advantaged peers. To satisfy the exclusion restriction, we further employ a Difference-in-Difference design and make use of the distinctive impacts after being exposed to advantaged students between advantaged and disadvantaged students due to the endogenous peer group formation. In another word, we expect the probabilities of forming friendship with advantaged students differ by students' family background. Disadvantaged students may benefit more in a class with higher density of advantaged students compared to advantaged students. Our empirical specification is below:

$$Qua_{igt} = \gamma_1 Cla_{-ig} * Educated\ parents + \gamma_2 Cla_{-ig} + \gamma_3 Educated\ parents + \delta X_i + \theta T_i + \varepsilon_{it} \quad (2)$$

[9] Even after accounting for the self-selection into schools and classes, the peer effects estimated by average group measures have been questioned with regard to self-selection into peer groups, measurement errors, and the mechanical relationship (Lavy et al 2012; Angrist 2014; Feld and Zolitz 2017).

[10] Our 2SLS estimate is different from the notation referred by Angrist (2014) in which he refers to the results estimated by average group variables as 2SLS estimates.

$$y_{igt} = \beta_1 \widehat{Qua}_{igt} + \beta_2 Cla_{-ig} + \beta_3 Educated\ parents + \alpha y_{igt-1} + \delta X_i + \theta T_i + \varepsilon_{it} \tag{3}$$

Qua_{igt} represents the observed peers' quality based on the questionnaire, such as whether your friends are hard-working or whether your friends are going to a university. The variable $Educated\ parents$ equals to one if a students whose parents have at least a degree, regarded as advantaged classmates. γ_1 represents the causal impact of classmates' backgrounds on observed peers' quality and γ_2 is our coefficient of interest measuring the peer effect. The interaction between the *leave-own-out* measure and the parental educational background works as the instrument. To further control teachers' practices, we have included teacher's characteristics, including gender, qualification, years of experience, and teaching certificate. We control for class size to ensure the peer effect is not driven by variation in class sizes. A caveat of using panel data is that we also observe attrition in Grade 8.[11] We have included the attrition rate to control for the potential bias. If the attrition between random assigned classes is random, the estimated peer effect is valid. The value-added framework will further address possible non-randomisation and idiosyncratic factors.

The main merit of employing the 2SLS strategy is to construct the network of endogenous peers and to estimate the distinctive peer effect resulting from peers' different attributes. The *leave-own-out* variables have assumed that classmates with certain backgrounds can bring beneficial or detrimental effect on classmates. Peer group formation is endogenous, and students are most likely affected by their closest peers. It is likely that they have distinctive impacts on their friends based on a virtual distance measuring the closeness. Our empirical strategy measures the quality of students' network directly based on the responses on the behaviours of five closest friends. In addition, amongst strictly randomised groups, differences in students' background are expected to be small, resulting in smaller treatment effect. It may also address the potential weak IV problem. By measuring the responses of every student in a class, instrumented by the *leave-own-out* variable, peer effect is estimated as the Local Average Treatment Effect (LATE) of peer groups. Consequently, the peer effect, estimated by the 2SLS strategy, estimates the impacts of peers'

[11] It is likely that the significant peer effect is driven by the attrition if students' attrition is associated with the unobserved factors in the class and the change in class composition may bias the results. Due to the relatively small sample size, we could not restrict the analysis further to subgroups.

attributes from the network of peers rather than the peer effect stemming from the connection with students with certain backgrounds.

Our proposed empirical approach provides the opportunity to examine mechanisms of peer effect, through which attributes individuals are affected by their peers. Inspired by Angrist (2014), we argue that the peer effect is generated through peers' characteristics and interactions between them rather than students' background, implying that it is not the peers themselves but the attributes of peers and the interaction between peers that affect outcomes. For instance, being exposed to students with better academic achievements may not necessarily generate positive peer effect if peer groups are self-formed and peers are not motivated to share the experience on study. On the other hand, having helpful and cooperative peers may increase the interaction between peers and increase test scores by sharing information or motivating each other, implying that different attributes of advantaged classmates may have distinctive impacts.

Table 4 presents our main results of peer effects. Panel A and Panel B show the impacts of peers' attributes and the impacts of having advantaged peers on peers' quality, respectively. In Panel B, we find that compared to advantaged peers, disadvantaged peers could much more significantly benefit from higher proportion of advantaged students due to endogenous peer group formation. With more advantaged students in a class, disadvantaged students have higher probability of forming friendship with advantaged students. In Panel A, the results suggest that having advantaged peers (closest friends) has positive impacts on own test scores. More importantly, the results suggest that the peers may have impacts on test scores through different channels. Although these three attributes all have positive impacts on test scores, the magnitude differs by the quality of students. Although the impact of having working-hard peers is less significant, the magnitude is the largest among the three attributes, possibly resulting from the weaker first-stage estimation. We argue that being exposed to peers with similar background but having different revealed characteristics in randomly assigned classes will induce distinctive peer effect through the nature of characteristics and the way that peers interact.

Table 4. Main results (2SLS)

Panel A, peer effect

	(1)	(2)	(3)
		Dependent variable: Test scores	
Peers studying well	0.533**		
	(0.24)		
Peers studying hard		0.908*	
		(0.51)	
Peers degree-motivated			0.660**
			(0.33)
Educated parents	-0.054	-0.061	0.092
	(0.17)	(0.19)	(0.12)
Educated parents peers' share	0.229	-0.236	-0.003
	(0.25)	(0.48)	(0.32)
Educated parents peers' share at school level	-14.601	9.023	-1.326
	(15.73)	(22.97)	(13.01)
N	5856	5856	5856

Panel B, First stage

	(1) Peers studying well	(2) Peers studying hard	(3) Peers degree-motivated
Educated parents peers' share	0.575***	0.850***	0.817***
	(0.17)	(0.20)	(0.22)
Educated parents	0.678***	0.407**	0.327***
	(0.16)	(0.19)	(0.11)
Educated parents peers' share X Educated parents	**-0.527*****	**-0.309***	**-0.426*****
	(0.15)	**(0.18)**	**(0.14)**
Educated parents peers' share at school level	27.972	-9.595	2.486
	(17.32)	(25.59)	(11.51)
N	5,856	5,856	5,856

Note: The estimation has the same set of control variables in the estimation above.

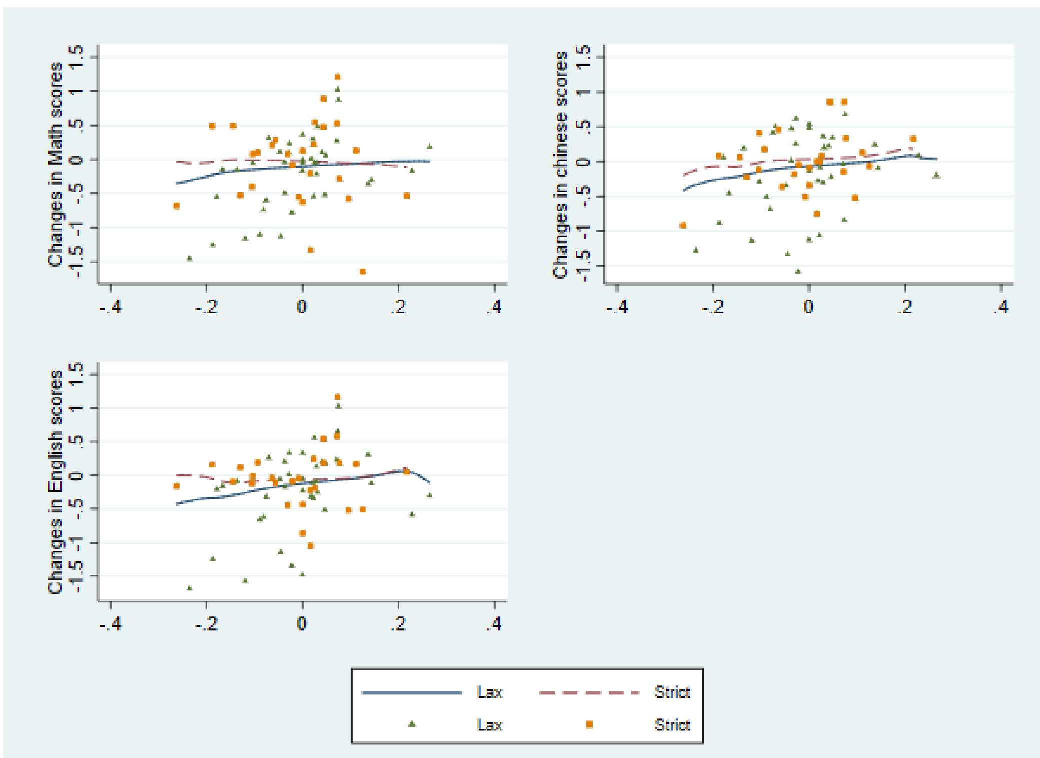
4.2. Bias from non-randomisation

Randomisation is the fundamental identifying assumption underpinning peer effect estimation. Previous research has heavily relied on random assignment classes in CEPS to remove self-selection into classes.

In this section, we highlight how the *leave-own-out* empirical design is highly sensitive to non-randomisation which results in biased estimates. Following the existing literature examining peer effect, we construct a new sample using students in Grade 7, and select the sample based on different rules according to the responses from principals and subject teachers on randomisation. We find that there is a strong inconsistency of the responses on randomisation between principals and subject teachers. We then show the statistically differences in students' backgrounds between classes based on the responses. We suspect that the responses of principals may not reflect the reality on the randomisation in schools. Based on this feature, we examine how the violation of randomisation affect the causal estimation of peer effects by affecting the extent to which selected sample violates the validity of randomisation.

Figure 1 presents the scatter plots of the class-pair differences in mean academic performances against differences in mean shares of college educated parents. The differences in the polynomial fitted lines between the *lax* and *strict rules* strongly indicate an upward bias in the estimated peer effects if the measurement errors in random assignment under the *lax rule* are overlooked.

Figure 1: Scatter plots of the class-pair differences in mean academic performances against mean differences in advantaged peers share, by random assignment type



Notes: Each dot plots the differences in mean academic performances against mean differences in mean shares of college educated parents for Math, English and Chinese. The fitted lines are based on local polynomial smoothing.

Table 5 presents the balancing test in means of key backgrounds between within-school class-pairs based on different sample selection rules. In cases of clear non-random assignment according to the response from schools' principal, between 10% and 30% of class-pairs report significant differences in the proportion of students with college educated parents, repeaters, and rural *hukou* students. What is really striking is that these patterns do not change significantly under the *lax rule*, especially for educated parents peers' share. In contrast, there are little differences between class-pairs when the *strict rule* is applied.

Table 5: Balancing test by class assignment type

	Non-random	Random assignment	
		Lax	Strict
Educated parents peers' share	10.0%	10.6%	7.4%
Repeated peers' share	10.0%	4.5%	0%
Alcoholic parents peers' share	30.0%	13.6%	7.4%
Obs (class-pairs)	40	66	27

Note: Proportion of within-school class-pairs with statistically significant different means in key characteristics at the 5% level.

The existing study of peer effect has largely followed Guryan et al (2009) to carry out the balancing test to testify the randomness of sample. The we have replicated the balancing tests following the previous research making use of CEPS that has examined the peer effects resulting from different students' backgrounds, such as educated parents (Chung and Zou, 2020), repeaters (Huang and Zhu, 2020; Xu et al., 2020), alcoholic parents (Zhao and Zhao, 2021). We firstly replicate the balancing tests based on different samples. The three panels in Table 6 shows the balancing tests of three different *leave-own-out* measures, including shares of educated parents, shares of repeaters, and shares of alcoholic parents based on three distinctive samples. Although the balancing tests in previous research have not suggested significant relationship between the *leave-own-out* measures and individual's backgrounds, our sample using the lax rule has shown that the shares of educated parents and the shares of repeaters have predictive power in students' backgrounds in Panel A. In Panel B, the *leave-own-out* measures shrink massively and become insignificant. We also show that the relationships are still significant after excluding schools following the strict rule, shown in Panel C. The sharp differences between two samples have suggested that the inconsistent responses on the randomisation have caused suspicion on the validity of randomisation.

Table 6. Balancing test

Panel A, Lax Rule

	(1) Educated parent	(2) Repeater	(3) Alcoholic parent
Educated parents peers' share	0.106**		
	(0.05)		
Educated parents peers' share at school level	-108.321***		
	(27.08)		
Repeated peers' share		0.149***	
		(0.06)	
Repeated peers' share at school level		-151.030***	
		(7.88)	
Alcoholic parents peers' share			0.013
			(0.02)
Alcoholic parents peers' share at school level			-121.102***
			(21.37)
_cons	22.887***	20.295***	59.519***
	(5.68)	(1.05)	(10.40)
N	5834	5834	5834

Panel B, Strict Rule

	(1) Educated parent	(2) Repeater	(3) Alcoholic parent
Educated parents peers' share	0.042		
	(0.09)		
Educated parents peers' share at school level	-82.150**		
	(34.58)		
Repeated peers' share		0.074	
		(0.06)	
Repeated peers' share at school level		-131.500***	
		(11.42)	
Alcoholic parents peers' share			0.002
			(0.03)
Alcoholic parents peers' share at school level			-95.731***
			(30.14)
_cons	20.608**	15.913***	46.715***
	(8.58)	(1.37)	(14.54)
N	2,331	2,331	2,331

Panel C, Lax Rule Excluding Classes using the Strict Rule

	(1) Educated parent	(2) Repeater	(3) Alcoholic parent
Educated parents peers' share	0.117**		
	(0.04)		
Educated parents peers' share at school level	-142.817***		
	(8.87)		
Repeated peers' share		0.098*	
		(0.06)	
Repeated peers' share at school level		-168.102***	
		(7.42)	
Alcoholic parents peers' share			0.032*

			(0.02)
Alcoholic parents peers' share at school level			-151.171***
			(6.76)
_cons	26.436***	24.098***	74.620***
	(1.63)	(1.06)	(3.32)
N	3,503	3,503	3,503

Moreover, the validity of the balancing test based on *leave-own-out* measures has been questioned by the recent research done by Jochmans (2020), in which he argues that the frequently used balancing test (Guryan et al., 2009) has failed to detect violations of the null of random assignment. He proposes a revised balancing test to present the correct size in large samples and to address the low power of the previous balancing test. Table 7 shows the results based on two samples, suggesting that the classes based on principal's response on randomisation may have higher risk of non-randomisation.

Table 7. Balancing test following Jochmans (2020)

	Lax rule excluding strict sample	Strict rule
Educated parent's peers	2.13	1.68
Repeater peers	1.89	-1.34
Alcoholic parent's peers	-0.20	0.77

Note: The results include T-test following the balancing test proposed by Jochmans (2020) after addressing the mechanical relationship between dependent variables and *leave-own-out* measures in Guryan et al (2009). The null hypothesis of the test is absence of correlation.

The previous exercises have suggested that the validity of randomisation is correlated with sample selection. However, the exams have only shed light on the differences between means. We also want to testify the differences between Grade 7 and Grade 9 classes. Previous research has made use of both grades based on the assumption that the responses on randomisation are correct. However, due to the pressure on examination in the Grade 9, more schools might be motivated to adopt ability streaming to boost the enrolment of the key high schools. Moreover, we do not have accurate information on the randomisation of Grade 8 and the classes in Grade 9 may have higher risks of reshuffle due to the speculation that students may transfer between schools, altering peer group formation.

As the data designers do not have any control over classes, we cannot rule out the risk of non-randomisation. We then test if the samples are drawn from the same distribution based on the two sample selection rules by Grade7 and Grade 9, respectively. We employ bootstrap to calculate the distribution of the balancing test (Guryan et al., 2009) originating from randomly selecting 80% of the classes based on two sample selection rules.

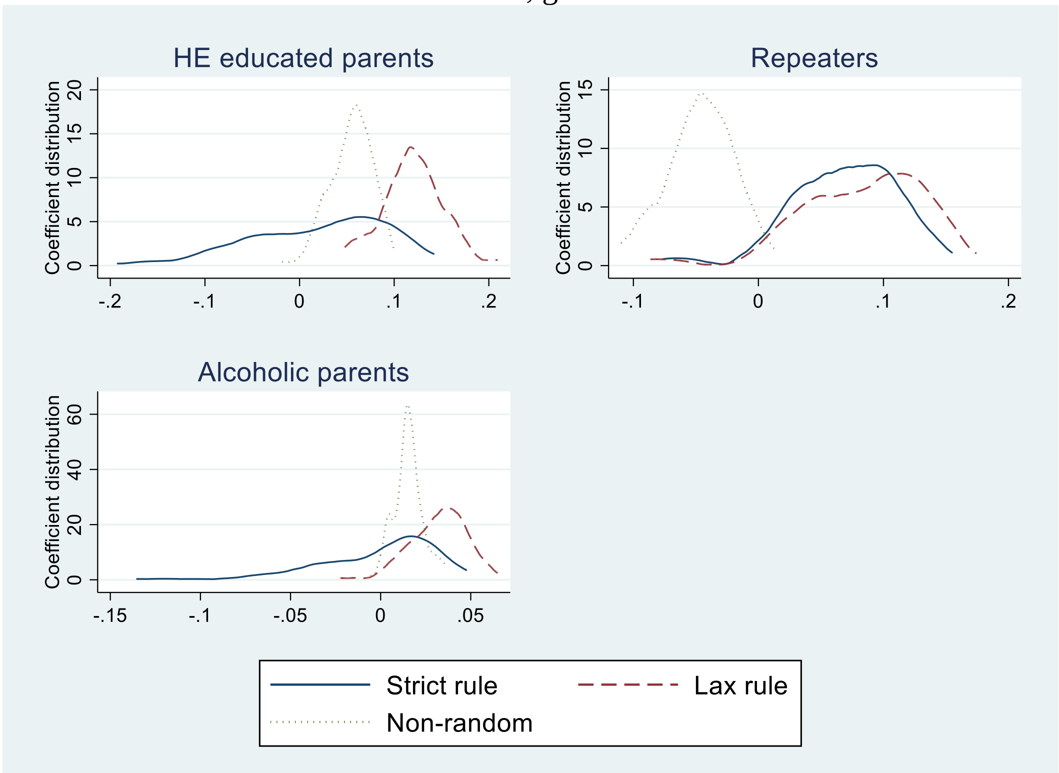
Figure 2 shows the distribution of the result from the balancing test between Grade 7 and Grade 9. Previous research does not reach a consensus on the randomisation between the two grades. While the majority of research has made use of two grades, few research has only relied on classes in Grade 7 (Gong, Lu, and Song, 2021). Due to the pressure of examination, classes in Grade 9 are more likely to be non-randomised. However, as the data is not collected in an experiment, the validity of randomisation remains unknown.

The sub-figures in Panel A show the distributions of balancing test estimated from three *leave-own-out* measures, educated parents, repeaters, and alcoholic parents. Each sub-figure includes three distributions, sample using the lax rule, sample using the strict rule, and non-randomised sample that has not been used in the peer effect research. It suggests that three sub-figures show that the distributions of balancing test using different sample selection rules are rather different. The distributions of non-randomised sample are even more centred around zero compared to the sample using the lax rule. This casts doubt on the principals' response on the randomisation. The distribution of balancing test using the strict rule have larger standard error and closer to zero.

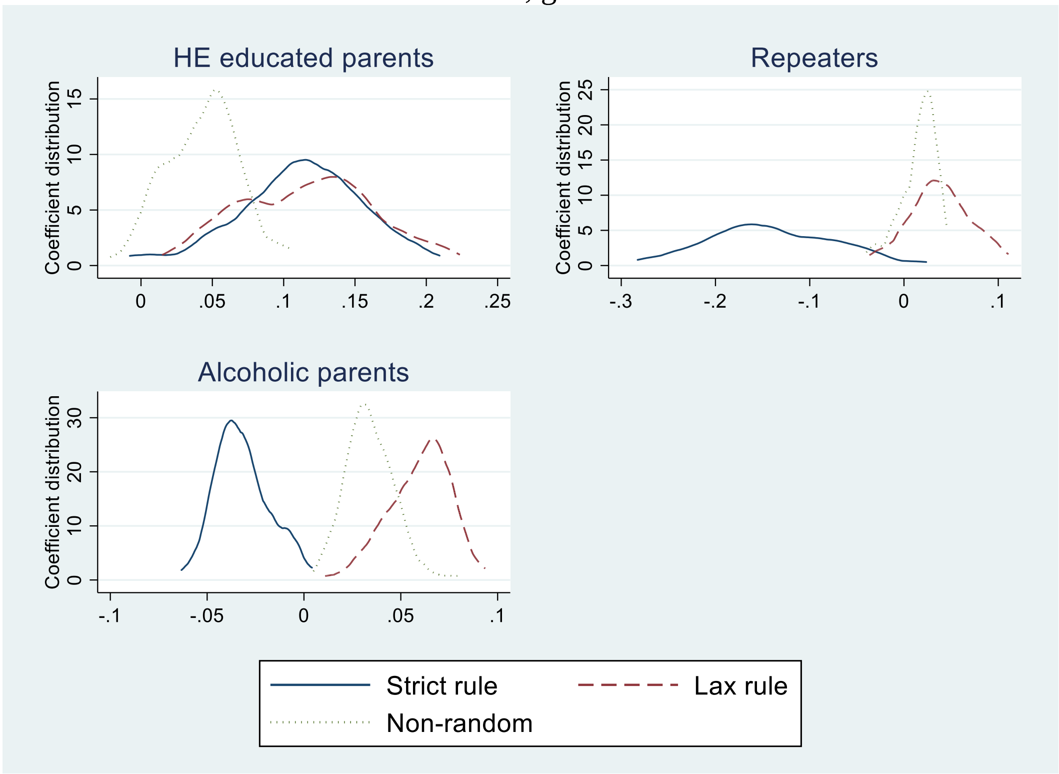
Panel B shows that the distributions based on Grade 9 classes. The distributions are less straightforward to understand. By eyeballing, the distribution of non-random sample is even closer to zero compared both the lax rule and the strict rule. This may suggest that classes in Grade 9 are not strictly randomised due to unobservables. It may result in higher sensitive results when we use the Grade 9 sample.

Figure 2. Distribution of balancing tests using bootstrap

Panel A, grade 7



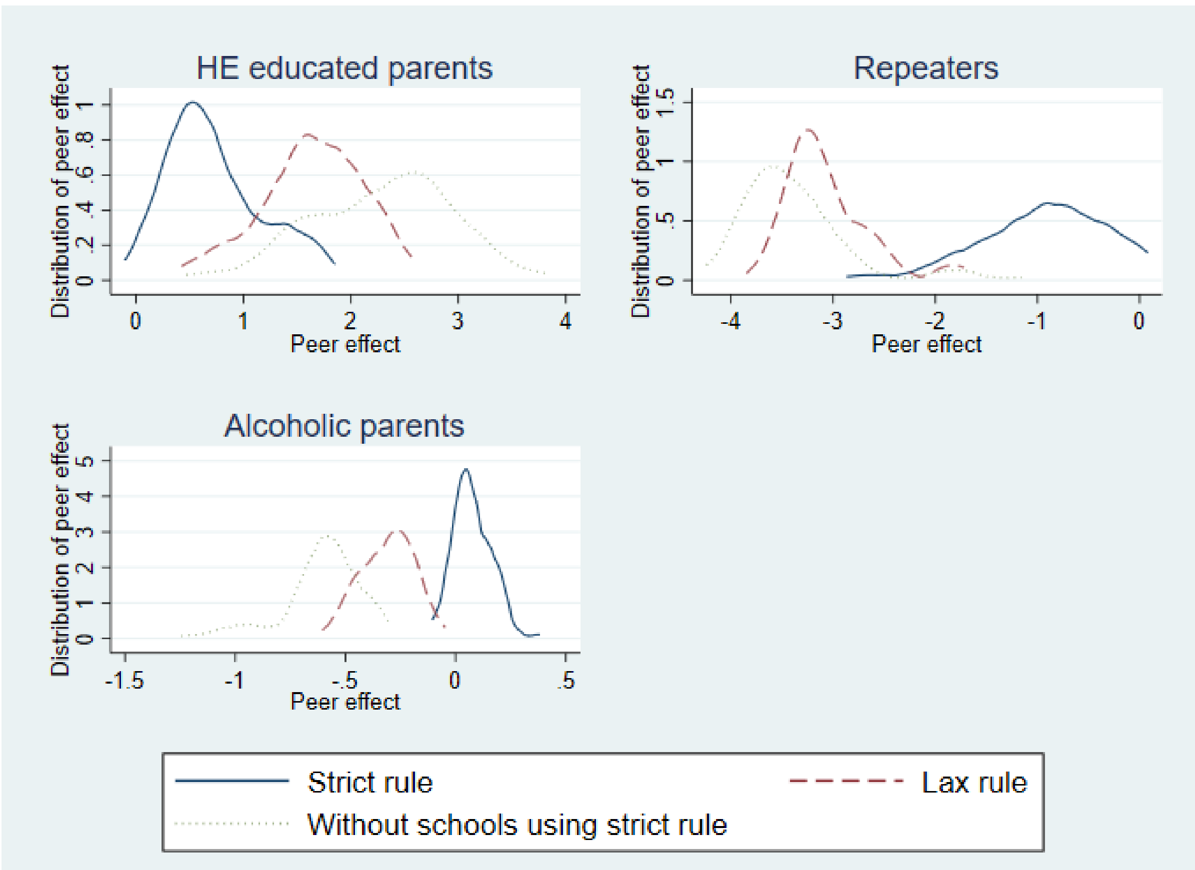
Panel B, grade 9



Note: The bootstrap includes randomly select 80% of the corresponding sample for 100 times.

To show the impacts of sample selection rule on the estimation of peer effect, we examine the peer effect following previous research. Figure 3 shows the distributions of peer effects in Grade 7, generated from bootstrapping. It clearly shows that the peer effects have largely shift to the origin when we use the strict sample rule while the results show the largest peer effect after excluding the schools following the strict rule. Taken together with the evidence above, we conclude that using the lax rule has higher risks of non-randomisation, resulting in large peer effects.

Figure 3. Distributions of peer effect using bootstrap



Note: Grade 7. Following the setting in previous research.

4.3. Robustness check and heterogeneity

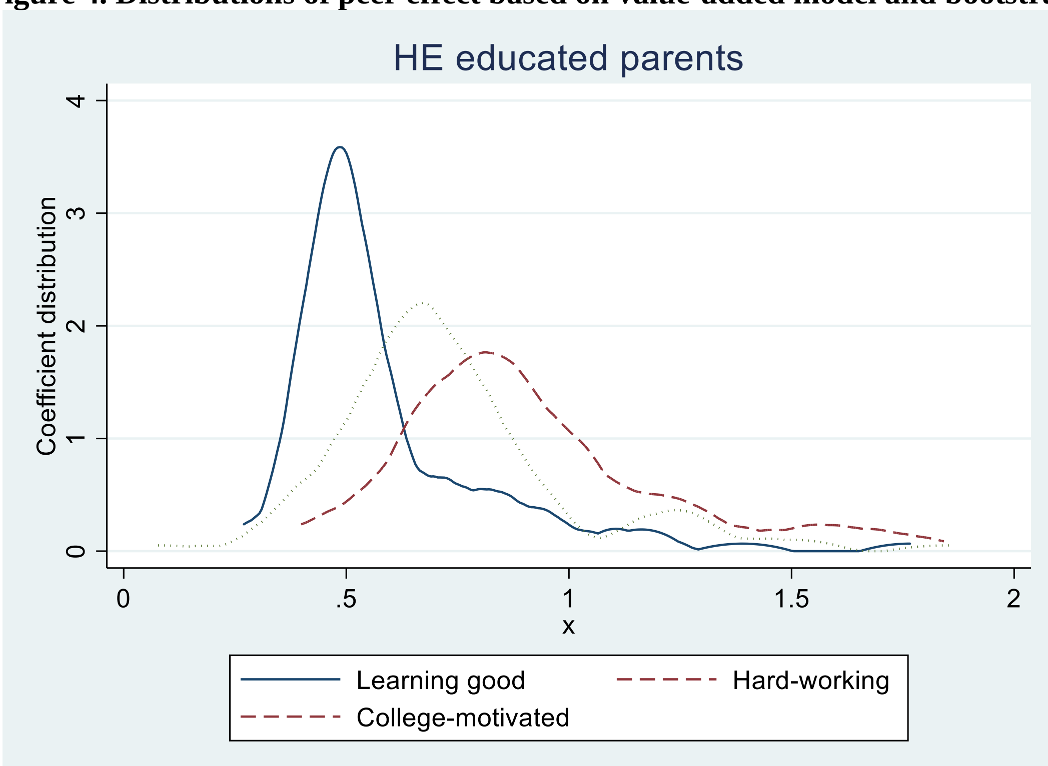
Identifying the peer effect making use of a random assignment design might be subject to the impact of size of sample. The small power of treatment due to the balanced observed characteristics across cells may result in sensitive results. Therefore, we provide the sensitivity test by bootstrapping results. The bootstrapping involves randomly selecting 90% of the sample and generate the distribution of the treatment effect shown in Table 7 and Figure 4.

Table 7 results show that attributes present different robustness of the result. Staying with peers studying well is more robust than having friends with the other two attributes. Figure 4 shows the distribution of the treatment effect. It clearly shows that the effect of 'Studying well' attribute is stronger and has smaller standard error.

Table 7. Treatment distribution based on value-added model and bootstrap

Outcomes	Non-significant	Significant	Total
Peers studying well	37	63	100
Peers studying hard	75	25	100
Peers degree-motivated	73	27	100
Total	185	115	300

Figure 4. Distributions of peer effect based on value-added model and bootstrap



Note: The distributions are drawn from 2SLS setting.

When discussing the mechanisms driving the peer effect, numerous research has examined the impacts on other inputs that may impact test scores, such as time use, teaching practices, behaviours, etc. While we think those changed inputs are important, the results do not shed light on how the peer effect is originated. Without an experiment, we also cannot observe interactions between peers in the data. Hence, we have examined the heterogenous peer effects to shed light on where the peer effect comes from.

Table 8 shows the heterogenous average treatment effect based on OLS. It suggests that some socio-economic disadvantaged students may benefit from more advantaged peers, such as rural students and students whose parents are not degree-educated. On the other hand, the peer effect of students who have a better academic background is strong.

Although the results look contradictory, we believe it may have implication on the importance of social interaction. Those socio-economic disadvantaged students may not interact with advantaged students more easily than advantaged students due to the lack of socio-economic skills. As peer groups formation is endogenous, advantaged students are more likely to form friendship with other advantaged students. On the other hand, students having better academic backgrounds might be more confident in classroom and can easily form study groups with other studying-well students.

We then examine the heterogenous peer effects in the 2SLS design. In Table 9, the students having stronger academic background benefit more from higher share of advantaged students due to stronger first stage results, suggesting the endogenous peer group formation.

However, contrary to the OLS estimation, peer effects only exist amongst urban students in Table 10. Being exposed to more advantaged students cannot increase the responses on the attributes of peers. These results highlight the importance of endogenous peer group formation.

Students with weaker socio-economic skills can benefit from more advantaged students with the increasing probability of being exposed to advantaged students. However, due to the endogenous peer group formation, disadvantaged students are less likely to interact with advantaged students compared to their counterparts, resulting less gain in the network.

Table 8. Heterogenous average peer effect

	(1) Rural	(2) Urban	(3) Feeling difficult	(4) Not feeling difficult	(5) Educated parents	(6) Not educated parents
Educated parent's peers	0.818***	0.215	0.013	0.598***	0.061	0.619***
	(0.30)	(0.24)	(0.49)	(0.19)	(0.33)	(0.21)
N	2541	3297	1646	4192	1538	4300

Table 9. Heterogeneity (2SLS)

Panel A, second stage

	(1)	(2)	(3)	(4)	(5)	(6)
	Feeling difficult			Not feeling difficult		
Peers studying well	0.733			0.500*		
	(0.69)			(0.26)		
Peers studying hard		6.271			0.766	
		(38.08)			(0.48)	
Peers degree-motivated			-2.659			0.504*
			(7.22)			(0.27)
Educated parents	-0.273	0.292	-0.212	0.035	0.019	0.215*
	(0.30)	(3.15)	(0.80)	(0.19)	(0.24)	(0.12)
Educated parents peers' share	-0.194	-8.078	1.996	0.403*	0.185	0.248
	(0.52)	(49.50)	(5.57)	(0.24)	(0.34)	(0.26)
Educated parents peers' share at school level	-49.764	261.963	-78.634	-0.185	21.574	15.747
	(33.81)	(1916.86)	(135.43)	(18.45)	(22.41)	(14.60)
N	1,646	1,646	1,646	4,192	4,192	4,192

Panel B, first stage

	(1)	(2)	(3)	(4)	(5)	(6)
	Feeling difficult			Not feeling difficult		
Educated parents peers' share	0.439	1.309***	0.702	0.569***	0.656***	0.872***
	(0.47)	(0.45)	(0.49)	(0.19)	(0.22)	(0.23)
Educated parents	0.450	-0.037	-0.101	0.710***	0.485**	0.349***
	(0.36)	(0.35)	(0.30)	(0.17)	(0.20)	(0.13)
Educated parents peers' share X Educated parents	-0.554	-0.065	0.153	-0.509***	-0.333	-0.506***
	(0.40)	(0.40)	(0.39)	(0.18)	(0.20)	(0.15)
Educated parents peers' share at school level	-1.369	-49.871	-10.479	32.990*	-6.850	1.149
	(35.62)	(44.23)	(36.61)	(18.20)	(27.31)	(15.45)
N	1,646	1,646	1,646	4,192	4,192	4,192

Table 10. Heterogeneity (2SLS)

Panel A, second stage

	(1)	(2)	(3)	(4)	(5)	(6)
	Rural students			Urban students		
Peers studying well	-3.276			0.427***		
	(8.51)			(0.15)		
Peers studying hard		0.811			0.744**	
		(0.90)			(0.29)	
Peers degree-motivated			-16.265			0.557***
			(220.60)			(0.20)
Educated parents	3.823	-0.437	1.762	0.009	0.009	0.235
	(11.24)	(0.54)	(27.30)	(0.19)	(0.19)	(0.16)
Educated parents peers' share	0.387	-0.117	31.727	-0.037	-0.255	0.215
	(1.59)	(1.07)	(419.01)	(0.28)	(0.34)	(0.25)
Educated parents peers' share at school level	584.643	-95.574	401.221	-1.143	20.316	9.208
	(1702.86)	(80.13)	(6164.03)	(15.68)	(19.13)	(12.88)
N	2,541	2,541	2,541	3,297	3,297	3,297

Panel B, first stage

	(1)	(2)	(3)	(4)	(5)	(6)
	Rural students			Urban students		
Educated parents peers' share	-0.155	1.248***	1.896***	0.974***	0.851***	0.294
	(0.36)	(0.21)	(0.32)	(0.28)	(0.30)	(0.26)
Educated parents	1.229**	0.287	0.121	1.075***	0.617***	0.418*
	(0.50)	(0.66)	(0.34)	(0.23)	(0.23)	(0.22)
Educated parents peers' share X Educated parents	0.161	-0.650	0.032	-0.874***	-0.502**	-0.670***
	(0.40)	(0.55)	(0.44)	(0.19)	(0.20)	(0.16)
Educated parents peers' share at school level	198.389***	37.357	28.681	36.220*	-8.062	9.179
	(61.08)	(75.83)	(59.19)	(18.41)	(26.86)	(13.07)
N	2,541	2,541	2,541	3,297	3,297	3,297

5. Conclusions.

Using random group assignment and natural experiments, the empirical research on peer effects has grown exponentially in recent years. This paper mainly contributes to the literature by uncovering the differential peer effects induced by randomly generated differences in characteristics of students and by examining the potential bias resulting from the violation of randomisation when using OLS and using the group mean variable to proxy for the quality of peers.

Although numerous empirical research has found strong evidence of peer effects on various outcomes, few research has explored the underlying mechanisms and the interaction between peers (Li et al 2014; Babcock et al 2020). The active interaction within peer groups generates complex peer effects and some studies have argued that peer effects are non-linear (Carrell et al 2009; Booji et al 2017; Garlick 2018). Our results suggest that the complexity may result from two sources, the complex social interaction, and measures of peer effect. Although the interaction between peers provides the opportunity to generate peer effects, students' characteristics will impact how peers affect each other. Students with similar background may have various characteristics, and few will dominate in the context of school environment. This implies that characteristics will generate differential peer effects and the observed peer effect measured by group mean variables is a combination of various peer effects induced by the interaction and characteristics.

Inspired by the recent development in the empirical strategies to address measurement errors in estimating peer effects, we propose to estimate the peer effect by measuring peer's quality directly using predetermined *leave-own-out* measure in random assigned classes as instruments to address the selection into peer groups. Our empirical framework estimates the causal peer effect of having advantaged classmates on disadvantaged students in the same classroom, through the attributes of the self-reported friendship-based peer group (i.e. 5 closest friends).

We have made a few contributions to the peer effect literature. First of all, we find significant effects of being exposed to advantaged classmates. However, the peer effects are heterogeneous with respect to the different characteristics, suggesting peers' attributes have distinctive mechanisms in the way they affect peers. We find that although being exposed to peers who have attributes of studying-well, hard-working, and degree-aspiring could significantly improve test scores, being exposed to hard-working peers has twice as large as the effects of being exposed to the other two attributes. The heterogenous results suggest that students with different backgrounds benefit from advantaged students distinctively, possibly due to different socio-economic skills possessed by students. Students with weaker socio-economic students might be less likely to form friendship with advantaged students due to endogenous peer group formation.

Second, our paper contributes to the discussion of the existence of peer effect in two ways. We start by examining the potential bias coming from non-randomised groups. On the basis of the different responses on randomisation between principals and subject teachers, we examine the extent to which the agreement between principals and subject teachers fails. Almost half of the class pairs classified as randomly assigned classes according to principal's response on randomisation alone have significant differences in students' backgrounds. More importantly, these classes have been considered as randomly assigned after passing the balance tests in previous research, implying that balance tests might not be informative when examining the unbalance between groups, especially in a case where the differences between groups are small. Given that randomisation is the fundamental identifying assumption underpinning peer effects estimation, we highlight how the *leave-own-out* empirical design is highly sensitive to the non-randomisation which results in biased estimates. The large and significant peer effects reported in the literature using the CEPS data may arise from the imperfect randomisation between groups. The results also question the validity of randomisation of Grade 9 classes. The results have important implication on the examination of randomisation in future research on peer effect.

The magnitude of the social multiplier may depend on correctly mixing students with different backgrounds over an extended period, as the results suggest the different peer effects are generated by different peers' attributes. In an education system with very intense competition and growing inequality, future designs of educational policies in China must take the impacts of peer effects seriously.

References

Angrist, J.D. (2014). The perils of peer effects. *Labour Economics* 30: 98-108.

Babcock, P., Bedard, K., Fischer, S., Hartman. 2020. Coordination and contagion: Individual connections and peer mechanisms in a ranodmized field experiment. *Journal of Public Economics*, Vol. 185, 104069.

Balestra, S., Sallin, A., and Wolter, S. C. 2021. High-ability influencers? The heterogeneous effects of gifted classmates. *Journal of Human Resources*, forthcoming.

Booji, A. S., Leuven, E., and Oosterbeek, H. 2017. Ability peer effects in university: Evidence from a randomized experiment. *Review of Economic Studies*, Vol. 84 (2), 547-579.

Bourdieu, P., 2018. Cultural reproduction and social reproduction. In *Knowledge, education, and cultural change* (pp. 71-112). Routledge.

Bramoullé, Y., Djebbari, H. and Fortin, B., 2009. Identification of peer effects through social networks. *Journal of econometrics*, *150*(1), pp.41-55.

Bramoullé, Y., Djebbari, H. and Fortin, B., 2020. Peer effects in networks: A survey. *Annual Review of Economics*, *12*, pp.603-629.

Cai, X., Fan, Q. and Yuan, C., 2022. The impact of only child peers on students' cognitive and non-cognitive outcomes. *Labour Economics*, *78*, p.102231.

Calvano, E., Immordino, G. and Scognamiglio, A., 2022. What drives segregation? Evidence from social interactions among students. *Economics of Education Review*, *90*, p.102290.

Carrell, S. E., Fullerton, R. L., and West, J. E. 2009. Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labor Economics*, Vol. 27 (3).

Carrell, S. E., Hoekstra, M. Kuka, E. (2018). The long-run effects of disruptive peers. *American Economic Review* 108(11): 3377-3415.

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

Chiu, M. M., Chow, B. W. Y., Joh, S. W. (2017). Streaming, tracking and reading achievement: A multilevel analysis of students in 40 countries. *Journal of Educational Psychology*, *109*(7), 915-934.

Chung, B.W. (2020). Peer's parents and educational attainment: The exposure effect. *Labour Economics*. Online First.

Chung, B. and Zou, J., 2020. *Understanding spillover of peer parental education: Randomization evidence and mechanisms* (No. 2020-045).

Chuang, B. W., and Zou, J. 2020. Understanding spillover of peer parental education: Randomization evidence and mechanisms. HCEO, Working Paper, No. 045.

Collins, J., 2009. Social reproduction in classrooms and schools. *Annual review of Anthropology*, *38*(1),pp.33-48.

DiMaggio, P., 1982. Cultural capital and school success: The impact of status culture participation on the grades of US high school students. *American sociological review*, pp.189-201.

Duflo, E., Dupas, P., and Kremer, M. 2011. Peer effects, teacher incentive, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, Vol. 101(5): 1739-1774.

Feld, J., and Zolitz, U. 2017. Understanding peer effects: On the nature, estimation, and channels of peer effects. *Journal of Labor Economics*, Vol. 35 (2): 387-428.

Figlio, D., and Ozek, U. 2019. Unwelcome guests? The effects of refugees on the educational outcomes of incumbent students. *Journal of Labor Economics*, Vol. 37 (4): 1061-1096.

Fitzgerald, J., Gottschalk, P., Moffitt, R. (1998). An analysis of sample attrition in panel data: The Michigan penal study of income dynamics. *Journal of Human Resources*, 33 (2), 251-299.

Fruehwirth, J.C., Gagete-Miranda, J. (2019). Your peers' parents: Spillovers from parental education. *Economics of Education Review* 73(C), 10910.

Fryer Jr, R.G. and Levitt, S.D., 2004. Understanding the black-white test score gap in the first two years of school. *Review of economics and statistics*, *86*(2), pp.447-464.

Gaddis, S.M., 2013. The influence of habitus in the relationship between cultural capital and academic achievement. *Social science research*, *42*(1), pp.1-13.

Garlick, R. 2018. Academic peer effects with different group assignment policies: Residential tracking versus random assignment, *American Economic Journal: Applied Economics*, Vol. 10 (3), 345-369.

Giroux, H., 1983. Theories of reproduction and resistance in the new sociology of education: A critical analysis. *Harvard educational review*, *53*(3), pp.257-293.

Goldsmith-Pinkham, P. and Imbens, G.W., 2013. Social networks and the identification of peer effects. *Journal of Business & Economic Statistics*, *31*(3), pp.253-264.

Golsteyn, B. H. H., Non, A., and Zolitz, U. 2021. The impact of peer personality on academic achievement. *Journal of Political Economy*, Vol. 129 (4), 1052-1099.

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

Hanushek, E. A., Kain, J. F., Markman, J. M., and Rivkin, S. G. 2003. Does peer ability affect student achievement? *Journal of Applied Econometrics*, Vol. 18: 527-544.

Hsieh, C.S. and Lee, L.F., 2016. A social interactions model with endogenous friendship formation and selectivity. *Journal of Applied Econometrics*, *31*(2), pp.301-319.

Hoxby, C. 2000. Peer effects in the classroom: Learning from gender and race variation. NBER Working Paper, No. 7867.

Hu, F., 2018. Migrant peers in the classroom: Is the academic performance of local students negatively affected?. *Journal of Comparative Economics*, *46*(2), pp.582-597.

Huang, B., Zhu, R. (2020). Peer effects of low-ability students in the classroom: evidence from China's middle schools. *Journal of Population Economics* 33**,** 1343–1380.

Jochmans, K., 2020. Testing random assignment to peer groups.

Jochmans, K., 2022. Peer effects and endogenous social interactions. *Journal of Econometrics*.

Kang, C. 2007. Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea. *Journal of Urban Economics*, Vol. 61 (3): 458-495.

Lavy, V., Paserman, M. D., and Schlosser, A. 2012. Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *Economic Journal*, Vol. 122 (559), 208-237.

Lavy, V., Silva, O., and Weinhardt, F. 2012. The good, the bad, and the average: Evidence on ability peer effects in schools. *Journal of Labor Economics*, Vol. 30 (2), 367-414.

Lee, L.F., Liu, X., Patacchini, E. and Zenou, Y., 2021. Who is the key player? A network analysis of juvenile delinquency. *Journal of Business & Economic Statistics*, *39*(3), pp.849-857.

Li, T., Han, L., Zhang, L., and Rozelle, S. 2014. Encouraging classroom peer interactions: Evidence from Chinese migrant schools. *Journal of Public Economics*, Vol. 111, 29-45.

Luo, Y. and Yang, S., 2022. Gender peer effects on students' educational and occupational expectations. *China Economic Review*, p.101898.

Manski, C.F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies* 60(3): 531-42.

Mendolia, S., Paloyo, A. R., and Walker, I. 2018. Heterogeneous effects of high school peers on educational outcomes. *Oxford Economic Papers*, Vol. 70 (3): 613-634.

Sacerdote, B. (2011). Peer effects in education: How might they work, how large are they and how much do we know thus far? *Handbook of the Economics of Education*, Volume 3, Chap 4.

Wang, H., Cheng, Z., and Smyth, R. 2018. Do migrant students affect local students' academic achievements in urban China? *Economics of Education Review*, Vol. 6, 64-77.

Xu, D., Zhang, Q., Zhou, X. (2020). The Impact of Low-Ability Peers on Cognitive and Non-Cognitive Outcomes: Random Assignment Evidence on the Effects and Operating Channels. *Journal of Human Resources*. Online First.

Zhao, L., and Zhao, Z. 2021. Disruptive peers in the classroom and students' academic outcomes: Evidence and mechanisms. *Labour Economics*, Vol. 68, 101954.

Appendix:

Table A1. Summary of variables

VARIABLES	(1) mean	(2) sd
Cognitive scores in grade 7 and 9	0.10	0.87
Cognitive scores in grade 8	0.37	0.81
Chinese scores	70.52	9.55
Math scores	70.48	9.66
English scores	70.57	9.59
Class size	49.04	12.33
Rural hukou	0.44	0.50
Male	0.51	0.50
Single	0.49	0.50
Parental degree	0.23	0.42
Repeaters	0.11	0.31
Alcoholic parents	0.46	0.50

Table A2. Impacts of peers on other attributes

	(1) **Study well** b/se	(2) **Hard-working** b/se	(3) **Degree expectation** b/se	(4) **Skip class** b/se
Educated parents peers' share	0.575***	0.850***	0.817***	0.162
	(0.17)	(0.20)	(0.22)	(0.23)
Educated parents	0.678***	0.407**	0.327***	-0.244**
	(0.16)	(0.19)	(0.11)	(0.11)
Educated parents peers' share X Educated parents	-0.527***	-0.309*	-0.426***	-0.089
	(0.15)	(0.18)	(0.14)	(0.14)
Educated parents peers' share at school level	27.972	-9.595	2.486	-39.930***
	(17.32)	(25.59)	(11.51)	(15.04)
N	5856	5856	5856	5679

	(5) **Having penalties** b/se	(6) **Fighting** b/se	(7) **Smoking and drinking alcohol** b/se	(8) **Game centre** b/se
Educated parents peers' share	0.018	-0.552***	0.215	0.534**
	(0.20)	(0.20)	(0.20)	(0.23)
Educated parents	-0.323**	-0.216*	-0.095	-0.186
	(0.13)	(0.12)	(0.10)	(0.12)
Educated parents peers' share X Educated parents	-0.234	0.127	-0.030	0.001
	(0.15)	(0.16)	(0.11)	(0.15)
Educated parents peers' share at school level	-64.968***	-38.018**	-42.434***	-47.023***
	(16.75)	(14.87)	(13.58)	(13.58)
N	5856	5679	5526	5679