Friendship formation and peer effect: Using seat distribution as an instrument

Abstract

Peers play a crucial role in the production of education. Effectively assigned peer groups can optimize academic performance in school and facilitate the accumulation of human capital. Identifying causal effects in peer group research is difficult, however, as changes in student academic achievement may result from endogenous or correlated effects, not from exogenous peer effects. This paper aims to control for the endogeneity of peer group selection and identify causal relationships between peer effects and academic performance inside classroom microenvironments. We collected data from 2,956 primary school students in rural China and use network theory to model the structure of study groups. Estimates based on instrumental variable approach indicate that study groups enhance student achievement by 0.11 standard deviations, and lower-ranked students benefit more from this effect. Self-concept and intrinsic motivation are channels behind these effects. Study groups, however, have no substantial effect on the learning outcomes or behavior of high-achieving students. Additionally, peer effects are more pronounced in study groups characterized by greater cohesiveness or among male students.

Keywords: peer effect; peer network; instrumental variable; rural education

JEL codes: A20; D85; I25

Friendship formation and peer effects: Using seat distribution as an instrument

1. Introduction

Education enhances both labor productivity and the innovative capacity of economies, leading to increased output and growth (Mankiw et al., 1992). Additionally, it facilitates the diffusion of knowledge essential for implementing new technologies, further fueling economic development (Aghion et al., 1998; Benhabib and Spiegel, 1994; Romer, 1990). A number of studies have shown that groups of students who sit together and/or often study together, can improve the academic achievement of all members of the group (Duflo et al., 2011; Marotta, 2017). However, identifying the causality behind peer effects among groups poses challenges due to the various ways peers influence each other, such as endogenous, exogenous, and correlated effects (Manski, 1993). These effects give rise to three fundamental challenges: reflection problem and self-selection bias (Brock and Durlauf, 2001; Manski, 1993). The reflection problem leads to simultaneous changes in individual behaviors among interacting agents, complicating the differentiation between endogenous, correlated, and exogenous effects (Bramoullé et al., 2020).

Our approach helps to address the identification issues caused by correlated effects, including the reflection problem and self-selection bias. We intend to explore how peer groups influence students in primary schools in rural China. The educational system in China requires that students spend most of their time with a specific group of peers within classrooms (usually those seated nearby). Specifically, in most primary schools in China, desks are arranged in a specific way and students are assigned to their seats by their teacher (Rivera et al., 2010). Importantly, student initial academic performance is not permitted to be disclosed and considered a significant factor when determining seating arrangements. Because of this, seat assignments are not correlated to academic performance. Students in these schools stay in a fixed seat in a single, fixed classroom for most of their classes throughout the day, while it is teachers that rotate through the classrooms. This system creates the conditions for students to have sustained physical proximity to the same group of peers, which may induce social interaction and enhance friendship formation (Back et al., 2008; Hare and Bales, 1963; McAndrew, 1993). Distance between students thus becomes an important and exogenous index on which to predicate study relationships between two students.

We conducted two survey waves in primary schools in rural China and collected seating distribution information in each classroom. Next, we utilize seat distribution data to compute the relative distance between students, serving as an instrumental variable. Our study aims to construct peer groups and identify the causal impact of peer networks on academic performance. Additionally, we seek to explore the mechanisms of peer effects in this context and examine whether peer effects vary across different types of peer groups.

Our findings indicate that peer groups have a significant effect on student academic performance, especially for students with lower initial test score. At the beginning of the semester, students showed more confidence and intrinsic motivation when they had high-ranking students as study peers. Intrinsically motivated activities are defined as those which a person does for no apparent reward except the activity itself or the feelings which result from the activity (Deci, 1973). These students then cooperated more with their study peers during the semester and showed greater improvement by the end of the semester. Specifically, we found that the intrinsic motivation generated at the beginning of semester had a lasting positive effect on student academic performance at the end of academic year by increasing time investment into their study of math.

This study contributes to the existing literature in the following ways. First, other studies rely on random or quasi-random assignment of students to classrooms or dormitories (Carrell et al., 2013; Foster, 2006; Lyle, 2007). These studies, however, often ignore the nature of personal relationships among group members. In addition, an instrumental variable (IV) approach based on the nature of personal relations between students can be employed to analyze the causality of peer effects on student academic performance (Hinke et al., 2019). Past studies that take this approach typically have created instrumental variables by using either the education levels of the parents of peers or the personal backgrounds and characteristics of the friends of a student's peers. These studies assume that an individual's academic outcomes are affected by the academic ability of their peers, but not either by the personal characteristics and backgrounds of their peers, or by the personal characteristics or background of the friends of their peers. However, this assumption does not always hold in practice. By developing an instrumental variable using naturally and exogenously assigned seats based on classroom geography, we identify the causal effects of peer group formation without manipulating seat distribution.

In addition, the current literature has somehow ignored the process of subgroup formation within the classroom. Most studies leverage variation in peer groups at the classroom or school level (Bramoullé et al., 2009; Burke & Sass, 2013; Hinke et al., 2019; Min et al., 2019). However, in larger peer groups, students are more likely to form subgroups with peers of the same gender or other social identity (e.g., by ethnicity, etc.) (McKeown et al., 2016). Thus, peer effect analyses at the classroom or school level may miss important interactions within sub-classroom groups (Anderson and Lu, 2017). Because of this, some studies have divided individuals into smaller subgroups in an attempt to better understand the importance of peer effects (Berthelon et al., 2019), but they do not take the intensity of study relationships into account. As an indication that the intensity of study relationships is important, research shows that the relationship between peer effects and academic outcomes tends to be stronger among peers with closer relationships (Mora and Oreopoulos, 2011). In this study, we established a study group within a classroom based on students' personal choices and predicted the intensity of their interactions by considering the distance between them.

2. Conceptual Framework

2.1 Peer effect in social interactions

Manski defines three ways in which an individual might be affected by their social interactions, all of which are potentially relevant to the study of peer effects in education (Manski, 1993). These three types of effects are: a.) endogenous effects; b.) exogenous effects; and c.) correlated effects. Under endogenous effects, the propensity of an individual to behave in a certain way varies directly with the prevalence of that behavior in the individual's group. For example, if a student belongs to a high-performing peer group, the student is more likely to have high academic performance as well. Under exogenous effects, the propensity of an individual to behave in a certain way varies with the characteristics of the individual's group. One example of an exogenous effect identified by Manksi is the education level of the parents of peers, which may have an indirect effect on student academic performance but must be carefully distinguished from the direct effect of peer academic performance on student performance. Finally, under correlated effects, individuals in a group behave similarly to one another because they have similar individual characteristics or face similar institutional environments.

Stemming from these three types of effects are three fundamental challenges to identifying the causal effects of the academic performance of peers on student scores (Brock and Durlauf, 2001; Manski, 1993). First, there is a reflection problem, as student academic outcomes may affect the outcomes of their peers and vice versa. Second, similar students tend to join or be assigned to the same group, which is referred to as self-selection bias. Correlation in outcomes and endogenous selection of peers are the primary issues with correlated effects. Third, as a consequence of the reflection problem, individual behaviors may change simultaneously among all interacting agents, making it difficult to separate endogenous effects from correlated effects and exogenous effects (Bramoullé et al., 2020).

2.2 Investigating the possible mechanisms

Many existing studies explore the mechanisms by which peer effects can operate within a study group. The first of these mechanisms is academic anxiety among peer members when they are studying on math, which can be considered a cost of study group membership. Academic anxiety is defined as the feelings of fear, tension, and apprehension that students may experience when engaging with study materials (Ashcraft, 2002), and specifically math in the context of this study. Academic anxiety is negatively related to math achievement because it disrupts the working memory resources with which students use advanced problem-solving strategies to solve difficult math problems (Ashcraft, 2002; Ramirez et al., 2016). One source of anxiety is social penalties within the group. When observed by peers, students within a group may try to avoid social penalties by conforming to the social norms in the group (Santor et al., 2000). For example, under the pressure of social penalties, students with high abilities in mathematics may decide to underachieve in order to avoid social exclusion (e.g. being called names associated with being too good in math) in their schools (Boehnke, 2008). At the same time, however, students with low math performance may also experience academic anxiety if they observe the high performance of other group members and feel excluded because of their low grades.

Students can also experience peer effects within their peer groups via intrinsic or instrumental motivation. Intrinsic motivation refers to engagement in an activity for one's own sake, such as to assimilate with one's social and physical surroundings (Ryan and Deci, 2017). Intrinsic motivation has been widely shown to have strong positive effects on academic achievement (Froiland and Oros, 2014). In contrast, instrumental motivation refers to motivation from future goals and activities that have utility value, such as motivations related to academic grades, career opportunities, financial gains, job promotion, etc. Simons et al. (2000). In China, the highly competitive learning environment and the importance of examinations may increase the influence of instrumental motivation, and students may experience high utility from ranking higher than their classmates (Li and Liu, 2020). Because of this phenomenon, it is particularly important to consider the role of instrumental motivation when studying academic achievement in the context of China.

Finally, peer effects can occur because peer groups often serve as a frame of reference or standard of comparison which can help students to form an academic self-concept (ASC). ASC is defined as the mental representation of one's own abilities (Brunner et al., 2010) and refers to how individuals view themselves in specific academic domains (e.g. Byrne, 1984). When studying within a group, students may compare self-beliefs of their own skills with the perceived skills of other students (Marsh et al., 2015). For example, in a situation in which students are working together on a math assignment, they may observe how many problems their peers can solve and compare such observations with their own achievements. When students perform better than their peers, they may form a higher ACS, which has been found to positively predict academic performance (Altermatt et al., 2002; Gest et al., 2008; Marsh and Hau, 2003).

3. Data

3.1 Sampling

The data for the present study were collected from Shaanxi Province. Among the 31 mainland provinces, Shaanxi ranked 14^{th} in terms of GDP per capita in 2019. The per capita disposable income of rural households in the province was 11,213 RMB in 2018, or approximately 1,495 USD (National Bureau of Statistics, 2019). There are 2.7 million students enrolled in primary school in the province, representing 2.6 percent of all primary school students in China (Ministry of Education, 2019).

The sampling strategy for our survey followed a three-step protocol. First, we restricted our sampling frame to two rural prefectures within Shaanxi province.

Second, nine counties in these two prefectures were randomly chosen to be included in our sampling frame, all nine of which are nationally designated poverty counties (People's Daily Online, 2014). Third, we obtained a list of all primary schools from the local bureaus of education in each sample county and selected all schools with more than 10 boarding students per classroom as participants in the survey. We set this condition to ensure that all schools in our sample would have a sufficient number of students studying together in groups after class. Out of the 90 schools that met this condition, the principals of four schools declined to participate, resulting in a final sample of 86 schools. Finally, we randomly selected one fourth, fifth, and sixth grade class in each school, enrolling all students in the sample classes into our study. In total, 3,025 students in 86 schools participated.

3.2 Data collection

In China, the academic year is typically divided into two semesters. The first semester starts in September and ends in January of the following year. The second semester begins in February and concludes in June. According to this academic calendar, the longitudinal data used for this study were aggregated from two survey waves conducted during the 2017-2018 school year by the enumeration team. The baseline survey was administered in October 2017, one month after September when no students have left or joined the class during that time and the academic term would be well underway with the same group of students who initially enrolled at the beginning of the semester. Follow-up survey was administered in June 2018.

Enumerators received two days of training before visiting sample schools. At each sample school, enumerators followed a strict protocol when administering each part of the survey.

The baseline survey consisted of a 35-minute standardized math exam with strict time limits and a standardized questionnaire consisting of four sections. The first section of the questionnaire measured the math learning attitudes of students. This section included questions about math-related anxiety, self-concept, and intrinsic and instrumental motivation, as well as questions about the amount of time they spent studying, reading, relaxing, and sleeping on a typical weekday.[1] Enumerators also asked teachers how often students studied mathematics with classmates each week and how often students were distracted during class each week, on average. Teachers could respond with: “never,” “less than once a week,” “twice or three times a week,” or “three or four times a week.” In the second section, students were asked to list the names of their classroom study partners. The third section collected information about student seating arrangements. The final section collected socioeconomic and demographic data from students and their teachers.

The follow-up survey included a 35-minute standardized math exam and the same questions as the first and second sections of the baseline survey. In addition, teachers were asked to describe the rules and process of seating assignments in their

[1] Specific questions can be found in Appendix II.

classrooms. The description of data collection protocols for each section can be found in more detail below.

3.3 Definitions and measurement of variables

3.3.1 Academic performance

In this study, we take math score as the key measure of a student's academic performance. The math exam administered in this study was grade-appropriate and tailored to the national and provincial-level mathematics curriculum and was constructed by trained psychometricians using a multi-stage process. Exam items were first selected from the standardized mathematics curriculum for each grade. The content validity of these items was checked by multiple experts, including local teachers and professors at Shaanxi Normal University. The psychometric properties of the exam were then validated through extensive pilot testing and data analysis. We standardized math score into z-scores using the mean and standard deviation of each grade for both baseline and follow-up surveys before excluding students who had no study partners in their class.

3.3.2 Math learning attitude

Student attitudes toward math learning were measured in the first section of both the baseline and follow-up surveys. Firstly, math anxiety, self-concept, intrinsic motivation, and instrumental motivation were measured using specially designed and validated items from the 2012 Programme for International Student Assessment (OECD, 2014), which have been widely used as measures of math learning attitudes in many countries (Lee, 2009; Pitsia et al., 2017; Thien et al., 2015). In each of the four scales, students responded to items with either “strongly agree,” “agree,” “disagree,” or “strongly disagree.” We condensed student responses to the items into four single measures using the GLS weighting procedure described in Anderson (2008). Positive values on the math self-concept index and motivation index corresponded to higher student-reported math self-concept or intrinsic and instrumental motivation than the average student. Meanwhile, positive math anxiety scores indicated that a student’s level of math anxiety was higher than that of the average student.

The survey also included three continuous variables about math learning attitude. The first variable was studying time ratio per day, or the percentage of time spent studying compared to the total time spent relaxing, reading, sleeping, and studying. The second and third variables were the average cooperative studying occurrences and the average distraction occurrences, which refer to how often students studied mathematics with classmates or were distracted during class on average each week. A summary of the above student and study group outcomes from the baseline and follow-up surveys is reported in Table 1.

<Place Table 1 about here>

3.3.3 Study groups and their structure

In recent years, network theory has provided a conceptual and empirical framework to analyze peer effects among microenvironments. In subgroups within a classroom, network theory can be utilized to define the characteristics of the group and the study relationships among group members. In the literature, a network is defined as a group or system of interconnected people or things, or more technically, a set of nodes and relations or interconnections between nodes, called links (Jackson, 2008). In the context of academic peer effects, students are the nodes; the relationship in the course of studying between two students is the link; and the complete set of students in a class and their relationships form a study network within the class. In this way, network theory can be used to describe variations in the make-up of each student's study group beyond those captured in non-network-based peer effect studies. Specifically, network theory makes it possible to investigate the variation and cohesiveness of study groups.

In the second section of the baseline and follow-up surveys, we collected a list of each student's study partners by asking each student to list up to 10 classmates with whom they most frequently studied or discussed math together. For each student surveyed, we generated a study group using all students included in their self-reported study partner list. Utilizing these study partner lists and the classroom maps (see below for a description of the maps), we were able to identify links between pairs of students, as well as networks among individual students and their study groups. Links in this study were undirected according to the definition of network theory (Jackson,

2008) since the students were not asked to provide a clear direction of the study relationship (e.g., who was asking for help and who was providing it). In this mutual academic support partnership, each study partner could be expected to play the same role, which means each student could either ask for help or provide help. We excluded 69 sample students who did not report any study partners, leaving 2,956 students in our final sample.

After creating the networks, the research team then estimated study group academic diversity and cohesiveness. Diversity is defined as the extent of the variation of academic performance among group members (Berthelon et al., 2019; Jackson, 2008). To measure variation, we used the standard deviation of math scores of the group members as the indicator of diversity in each study group, taking this as zero when a student had only one study partner. For each student, cohesiveness refers to the extent to which each student's study partners also studied with each other (Berthelon et al., 2019; Jackson, 2008). Additionally, we created a cohesiveness indicator utilizing the ratio of peers from the student study group who, aside from the student themselves, nominated their study companions as belonging to this group.

3.3.4 Distance between students

In this study, we used relative distance as an instrumental variable to the study relationship between classmates, as relative distance is assumed to be exogenous to academic performance (given that seating charts are assigned exogenously with respect to academic performance, as explained above). In the third section of the baseline survey, enumerators drew a chart that reproduced the distribution of the classroom seating assignments. Based on the distribution data (Figure A1) collected in this section, we calculated the relative distance between pairs of students using the following steps. First, we created a coordinate plane in the seat distribution table by using the first row of the classroom as the horizontal axis and the first column of the classroom as the vertical axis, taking the distance between two adjacent desks on either axis as one desk. Second, we assigned coordinates (a_i, b_j) to the students according to their seat order relative to the origin, defined as the first seat of the first row. Finally, we calculated two relative distances (Figure 1): the direct distance and the step distance. Direct distance refers to the distance calculated by the Pythagorean theorem, whereas the step distance combines horizontal and vertical distances between two students. When there are two students with coordinates (a_{i1}, b_{j1}) and (a_{i2}, b_{j2}), their direct distance is calculated by model 1, and their step distance is calculated by model 2.

<Place Figure 1 about here>

$$Direct\ distance = \sqrt{(a_{i1} - a_{i2})^2 + (b_{j1} - b_{j2})^{\,2}} \tag{1}$$

$$Step\ distance = abs(a_{i1} - a_{i2}) + abs\left(b_{j1} - b_{j2}\right) \tag{2}$$

3.3.5 Control variables

Identifying causality in peer effect studies requires overcoming both the reflection problem and the risk of self-selection bias. Most prior studies have sought to resolve the reflection problem by controlling for observable covariables in their analytical models. One common way to achieve this is to measure the academic ability of peers by prior test scores (Lavy et al., 2012). At the same time, to reduce selection bias, previous studies have taken several approaches. Some studies control for individual student and school characteristics, such as age, ethnicity, and grade level (Burke and Sass, 2013; Hanushek et al., 2004; Lavy et al., 2012). Therefore, in this study, besides assessing students' initial academic performance, we also gather data on their socioeconomic backgrounds during the data collection phase.

The socioeconomic and demographic data that were collected from students and teachers in the final section of the baseline survey were used as control variables. For students, these control variables included age, gender, grade, boarding status, parental education levels (i.e., whether their father or mother graduated from junior high school), number of siblings, and household asset value. Teacher control variables included age, gender, whether they graduated from at least a two-year college, and teaching experience (in years).

Individual and household characteristics are summarized in Table 2. The mean age for sample students was 10.9 years old. In our sample, 53% of the students were male, and 38% were boarding students. There were, on average, five family members in each student's household, and around half of mothers (44%) and fathers (51%) completed junior high school. In terms of teacher characteristics, the mean age of teachers was 36.9, and half (50%) were male. More than half of the teachers had graduated from junior college (58%), and the average teaching experience was 16 years. There were 32 students per class, on average, and students had an average of five study partners. The average direct distance between students and their peers was 2.7 desks, and the average step distance was 3.4 desks.

<Place Table 2 about here>

4. Econometric Approach

One primary goal of this study was to estimate the effect of study groups on student academic performance. To do so, we used a reduced form model:

$$Score_{i,endline} = \beta_0 + \gamma\overline{G_{-\iota}} + g_{i,baseline} + \beta_1 X_i + \varepsilon_i \qquad (3)$$

where $Score_{i,endline}$ is the standardized mathematics score of student *i* at the time of the follow-up survey; $\overline{G_{-\iota}}$ represents the average score of the study group, excluding the student's own math score; X_i is a vector of controls (individual, household, or teacher characteristics); and $g_{i,baseline}$ is the standardized mathematics score of student *i* at baseline. ε_i is a random error term.

To explore the mechanisms and analyze the effect of study groups on student math learning attitude, we substituted $Score_{i,endline}$ in model (3) with

$Attitude_{i,baseline}$ (math learning attitude variables at baseline survey) and $Attitude_{i,endline}$ (math learning attitude variables at follow-up survey), respectively. When $Attitude_{i,endline}$ was included in the outcome variables, $Attitude_{i,baseline}$ was controlled in the model.

Building on the idea that the intensity of the interaction among peers, the relative distance between students may be able to serve as an instrumental variable for student peer relationships at the subgroup level. Research has shown that classmates that sit closer together have a closer relationship . When individuals are repeatedly exposed to a stimulus, such as the frequent communication that may arise between students when they are seated near one another, the literature shows that such students tend to develop feelings of familiarity and positivity toward their nearby peers (Rhodes et al., 2001; Van Den Berg and Cillessen, 2015). In terms of peer effects, there is ample evidence that when students are in close physical proximity to one another and often meet each other in an academic context, it is possible that they can influence each other's academic performance. In contrast, a lack of such close contact may prevent this influence (Marmaros and Sacerdote, 2006; Rivera et al., 2010).

To control for endogeneity of network formation, we employed an IV strategy where we instrumented for observed study group average scores $\overline{G_{-\iota}}$ using predicted study group average scores $\widehat{G_{-\iota}}$. Referencing a social network model from Berthelon et al. (2019), observed group scores were constructed using the following equation:

$$\overline{G_{-\iota}} = \frac{\sum_{j\neq i} I_{ij,baseline}\, g_{j,baseline}}{\sum_{j\neq i} I_{ij}} \tag{4}$$

where $g_{j,baseline}$ is the mathematics score of student *j* at baseline, and $I_{ij,baseline}$ indicates whether student *i* and student *j* were study partners at baseline. To generate an instrument for $\overline{G_{-i}}$, we needed exogenous instruments or predetermined measures of $I_{ij,baseline}$. Distance between students served as an instrument of study relationships in this paper. Two-Stage least squares (2SLS) regression analysis (2SLS) with linear first and second stages have been used in this paper.

4.1 Relationship (network) formation

To study the probability of relationship formation, we first estimated a model of the determinants of forming a relationship (link) between any two given students, based on the relative distance between them and their predetermined characteristics. As an instrument, relative distance is exogenous to academic performance, as seat assignment generates exogenous variation in the probability of link formation between students that is not correlated with academic performance.

The link formation model in its general form can be represented as follows:

$$I_{ij,baseline} = f\left(\delta distance_{ij} + \beta M_{ij} + w_{ij}\right) \quad (5)$$

where $I_{ij,baseline}$ is the indicator variable of a relationship between students *i* and *j*, and is equal to one if i and j study together. $Distance_{ij}$ is the relative distance between two students. M_{ij} are variables measuring shared personal characteristics between students, including whether both students are the same gender (yes/no); whether both are boarding students (yes/no); whether both students have a father or mother who completed junior high school (yes/no); absolute difference of sibling numbers between the two students; absolute difference of family asset value between the two students; and absolute difference of baseline math scores. These variables cannot be affected by student decisions made at school. We took w_{ij} as an independently distributed random disturbance.

We estimated Equation (5)—the relationship model—with a Logit regression and estimated probabilities of relationship formation by predicting $\widehat{I_{ij,baseline}}$. Following this estimation strategy, Column 1 of Table 3 shows a significant and negative correlation between direct distance and study relationship: students with a smaller direct distance at baseline were more likely to form a study relationship (row 1). Similarly, Column 2 of Table 3 demonstrates that there is a significant and negative correlation between step distance and study relationship (row 2). Therefore, we conclude that distance is a realizable instrumental variable capable of predicting student relationship formation.

<Place Table 3 about here>

4.2 Instruments for study group average score

We constructed instruments for study group average score by using the estimated probability of relationship formation ($\widehat{I_{ij,baseline}}$) and the baseline math scores of group members. The model is as follows:

$$\widehat{G_{-i}} = \frac{\sum_{j \neq i} \widehat{I_{ij,baseline}} g_{j,baseline}}{\sum_{j \neq i} I_{ij}} \tag{6}$$

where $\widehat{G_{-i}}$ is the predicted study group average score for $student_i$, and $g_{j,baseline}$ is the standardized mathematics score of student j at baseline. Overall, this procedure allowed us to predict study group average score based on exogenous probabilities of relationships and predetermined math scores of peers.

5. Main results

5.1 Peer effect on student academic performance

To analyze the effect of study group average score on student academic achievement, we conducted both a simple linear regression (OLS) and an IV analysis. We first estimated Eq. (3) with a simple linear regression (OLS) without accounting for the endogeneity of group formation. Results of this regression are reported in column (1) of Table 4. We found that a 1 SD increase in study group's average score result in a 0.084 SDs increase in the student's score, which suggests that the academic ability level of a student's study group members was positively correlated with the student's own grades (row 1).

We then examined the effect of study groups on student academic performance while controlling for the endogeneity of group formation (columns 2 and 3). When using relative distance to predict the formation of study relationships, the instrument was significant in the first stage, as revealed by F-statistics of 350 and 363, and by the fact that the instrument explained about 59% of the variation in the first stage (rows 8 and 9). The second-stage results indicated that study groups had a significantly positive effect on student academic performance. Specifically, when the average score of a student's study group, as predicted by direct distance, increased by one standard deviation, the final score of the student increased by 0.114 standard deviations (column 2, row 2). Similarly, when the study group average score, as predicted by step distance, increased by one standard deviation, the final score of the student increased by 0.113 SDs (column 3, row 3).

<Place Table 4 about here>

5.2 Robustness analysis

Two potential sources of bias exist when we use relative distance as an instrumental variable to estimate the effects of study groups on academic performance, necessitating robustness checks. One potential concern is that relative distance between students may be endogenous to their math test score. This concern arises when teachers assign seats to students based on their academic performance. The second is that relative distance between students may not be constant throughout the academic year. While we collected seat distribution tables in the baseline survey, in some cases, student seating arrangement may be re-assigned between the baseline and follow-up surveys. To address these two concerns, we included questions about seating arrangements in the follow-up survey addressed to both students and teachers.

Based on these questions, we test the robust of our results in Table 4 by excluding undesirable samples.

5.2.1 Student seating choice

To address the first concern, in the follow-up survey, we asked students two questions about seating arrangements: 1) Were you allowed to choose your own seat last semester? 2) If not, on what basis did the teacher assign the seating arrangement last semester? The second question was presented in multiple-choice format, with answer choices including height, vision, test score, and student personality. Answers from students are summarized in Table A2 and Table A3. We also asked teachers during the follow-up survey to explain the basis on which they assigned seating. When students reported being allowed to choose their own seats, we asked their teachers the reason for this decision.

According to student responses to the first question (Table A2), 82% of students were not allowed to choose where they sat. For the 18% who were allowed to choose their seats, teachers gave two reasons as to why students were allowed to choose where to sit. The first reason was vision problems, as teachers let students sit closer to the blackboard if they could not see clearly. Secondly, some teachers also reported that students would ask to change their seats when they had conflicts with desk mates. This may lead to overestimates of study group effects on student academic performance in Table 4, as we did not include possibilities of potentially worse relationships between students in this data. To test the robustness of the results in Table 4, we excluded students who reported being allowed to choose their own seat and re-estimated the regression using model (3). The results in panel A of Table 5 show that the effects of study groups were qualitatively unchanged, suggesting that our earlier results remained robust.

<Place Table 5 about here>

5.2.2 Indicators of seating arrangements

Student responses to the second question, which asked how their teachers determined seating assignments, are reported Table A3. According to these results, more than half of students (54%) were arranged by their height or vision, which does not bias results directly. As height is mostly influenced by genetic or environmental factors, there is no direct relationship between height and student academic performance (Silventoinen et al., 2007; Dubois et al., 2012). In addition, although there are negative associations between grade and vision because poor vision affects sensory perception, cognition, and school connectedness, teachers arrange for students with poor vision to sit in the front of the classroom, decreasing the negative association (Basch, 2011).

In the rest of the sample classrooms, 11% of students were arranged based on their personality, and 30% of students were arranged based on their math test score. Teachers gave two reasons as to why personality was used as a basis by which to arrange seats, both of which may decrease self-selection bias to some extent. Firstly, teachers often seated outgoing students next to introverted students, as they believed that this arrangement could help students learn from each other and avoid chatting during class. Teachers also often separated students who had a close relationship, as they believed a close relationship could decrease both of their academic performances. Teachers who used test scores as a basis for seating arrangements often did so to encourage students to support one another, usually assigning students with poor performance to sit with better performing desk mates.

To prove the robustness of the main results, we excluded classes that fell into the above two categories from our sample and re-estimated the impact of study groups on student academic performance. According to the results presented in panel B of Table 5, the study group coefficients remained positive and significant, although the results were weaker, at the 5% level. This finding suggests that our previous findings were robust.

5.2.3 Consistency of seating arrangements

To address the potential bias from inconsistent seating arrangements, in the follow-up survey, we asked students another two questions: 1) How often did the class change seats throughout the semester? 2) On what basis did a class change seats throughout the semester? Based on the responses to these two questions, which are summarized in Table A4 and Table A5, we find that there was a low probability for students to change their relative distances to each other. Regarding the first question, 21 classes (19%) did not change seating arrangements during the semester. Among the 93 classes that did change seating arrangements, 31 classes (27%) changed seating arrangements more than once a month throughout the semester (Table A4). The rest of the classes changed seating arrangements less than or equal to once a month (54%). Regarding how seats were changed, there was one class in which seating was changed by groups, and therefore the relative distance between students who sat next to each other did not change (Table A5). In addition, more than half of classes rearranged student seating by row (1%) or column (59%) only. 8% of classes changed seats by row and column together.

For classes that changed seats during the first semester (93 classes), we used student responses to both questions to estimate how long the students maintained the same relative distance to classmates throughout the academic year (Table A6). According to the estimated results, in 66% of the 93 classes, students did not change in relative distance from classmates until the 18th academic week, since for the most part they still sat near the same classmates even when their location in the classroom changed. However, as there are only 16 academic weeks per semester in rural primary schools, and given that most classes use the same seating arrangement strategy for the whole academic year, we could assume that the relative distance between these students did not change throughout the academic year.

In the remaining 34% of classes, relative distance between any two given students changed after 9 or 10 weeks. To prove the robustness of the previous findings, we excluded the classes in which relative distance changed more frequently and re-evaluated the impact of study groups on student academic performance in the remaining samples. The results of this check (Table 5, panel C) were qualitatively similar to previous findings. Taken together, these checks confirmed that our finding of positive and significant effects of study groups on student academic performance was robust.

5.3 Peer effect on student academic performance across terciles

Our results show that being part of a higher-achieving study group boosted the performance of all students. However, existing theoretical studies and empirical research suggest that peer effects may operate nonlinearly depending on the ability of individual students (Burke and Sass, 2013; Han and Li, 2009, 2009; Hoxby and Weingarth, 2005). One example of these nonlinear effects is the "Shining Light model," which posits that a single exemplary student with exceptional academic outcomes can inspire others to improve their academic achievement (Burke and Sass, 2013; Han and Li, 2009; Hoxby and Weingarth, 2005). According to this model, outstanding students may cause other classmates to improve their achievement.

To explore whether this model was supported by our data, we tested the heterogenous effects of study groups by dividing students into terciles based on their initial ranking within their study group. For a given study group, student performance was designated as "low" if the math score of that student fell within the bottom tercile of the study group, "middle" if their math score lay between the 33^{rd} and 67^{th}

percentiles, and “high” if their math score fell within the top tercile. Peer group average score was estimated by weighting step distance.

When dividing students into terciles according to math scores within a study group, we found different results as compared to the overall sample of students. As shown in Table 6, increasing peer group average score had a significant effect on low-performing and middle-performing students, but the effect was not significant for high-performing students.[2,3,4] These results provide a strong argument in favor of distributing top students relatively evenly across classrooms at the elementary school level, rather than isolating them from other students.

<Place Table 6 about here>

6. Mechanism analysis

As shown in the existing literature reviewed in Section 2, peer effects from study groups may affect student academic performance by influencing the math learning attitudes of students (Gest et al., 2008; Ramirez et al., 2013; Ryan and Deci,

[2] The results do not substantively change when using direct distance to predict math score. For the sake of brevity, we have omitted these tables, but they are available upon request.

[3] We perform an additional robustness check by designating students as a “low” type if her score falls within the bottom quintile of the study group, as a “middle” type if her score lies between the 20th and 80th percentiles, and as a “high” type if she falls in the top quintile. The results do not substantively change when using quintile as cut off. For the sake of brevity, we have omitted these tables, but they are available upon request.

[4] We also divided students into different groups based on their academic ranking within classroom. It shows a similar result and can be found in Table A1.

2000). Additionally, peer effects may influence the math learning attitudes of different students in different ways. This section explores the ways in which the impacts of peer effects on math attitudes differ first across student terciles, and then across levels of student progress during the semester.

We first ran the same regressions as in equation (3) but substituted math test score with math learning attitudes, as measured in the baseline and follow-up surveys. We then conducted a heterogenous analysis of the effects of study group average score on math learning attitude, as identification of peer effects on math learning attitudes may depend on the distribution of student math score across different quantiles within a group. As in the previous section, we estimated study group average score using step distance as an instrument.

The results of this analysis are reported in Table 7. We found that students in higher-performing study groups had higher self-concept and cooperation times at baseline compared to classmates in lower-performing study groups. This was especially true for students in the middle and bottom terciles (column 3 and 5, row 1~3). Throughout the semester, intrinsic motivation increased by 0.108 standard deviations on average across all terciles for students in higher performing study groups by the end of academic year (at the time of follow-up survey), with the biggest gain occurring among students in the middle tercile (0.224 standard deviations, column 3, row 5 and 7). In terms of academic self-concept, we found that middle- and bottom-tercile students in higher-performing study groups exhibited higher self-concept at the beginning of semester, and bottom-tercile students in these groups were also less distracted during class (column 2 and 6, row 2 and 3).

<Place Table 7 about here>

Belonging to a higher-performing study group also increased the self-concept of students by 0.234 standard deviations in the middle tercile and reduced their anxiety by 0.137 standard deviations at the end of academic year (column 1 and 2, row 7). These improvements in motivation and self-concept appeared to translate into increased time investment and effort in schoolwork, as the studying time ratio per day of middle-ranked students increased by 2.6% by the end of academic year when they belonged to a higher-performing study group (column 7, row 7). However, for students who ranked in the bottom tercile, the positive effects of belonging to a higher-performing study groups diminished at follow-up.[5]

7. Heterogeneous effects

As shown by the past research summarized in Section 2, several study group characteristics can act as sources of potential heterogeneity in the effects of study groups on student academic performance. These characteristics may include study group structure (including diversity and cohesiveness), gender composition, or student

[5] The results in this section do not substantively change when using direct distance to predict math score. For the sake of brevity, we have omitted these tables, but they are available upon request.

seat distribution in the classroom (Hamilton et al., 2012; Van den Berg et al., 2012; Whitmore, 2005; Johannisson, 2000). Analysis along this line has important policy implications that can help teachers to maximize academic outcomes by taking advantage of peer influence. In this section, we explore whether the effects of study groups on student academic performance differ based on these characteristics, again using step distance to estimate study group average score.

7.1 Study group structure

We first examined the heterogeneity of peer effects on student academic performance as differentiated by study group structure (i.e., diversity and cohesiveness). We divided students into two subgroups based on study group diversity at baseline: students with less varied study groups (standard deviation <=0.655)[6] and students with more varied study groups (standard deviation >0.655). Similarly, we also divide students into two subgroups based on their cohesiveness: those with less cohesive study groups (cohesiveness <=0.5),[7] and those with more cohesive study groups (cohesiveness >0.5).

Panel A of Table 8 demonstrates that the average study group score has significant positive effects on the academic performance of students in study groups with low or high diversity, but the effect is greater for those in low-diversity study groups (0.131 standard deviation vs 0.119 standard deviation) (row 1, columns 1-2).

---

[6] 0.655 is the median of student group diversity for all students.

[7] 0.5 is the median of student group cohesiveness for all students.

However, in terms of cohesion, only students in more cohesive study groups experience a positive and significant effect on academic performance, and the effect on students in less cohesive study groups is insignificant (row 1, columns 3-4). One possible explanation for this result is that, within a group, high diversity may raise communication costs about problem-solving and can also diminish productivity if cohesiveness is low (Hamilton et al., 2012). Also, as noted in Amason and Schweiger（1994), although a certain amount of diversity is necessary for improving the quality of strategic decision making, it can also increase the likelihood of group conflict that may impede cooperation among team members. However, combining this finding with the non-linear effects between different terciles of students, these findings argue against strictly tracking student performance, as lower-ranked students can feel more motivated when they are in a high-achieving study group. Therefore, arranging the best students near middle-tercile students rather than the weakest students may be a rational seat distribution.

<Place Table 8 about here>

7.2 Gender

To explore the heterogeneous effects of gender composition of study groups, we first classified students by gender, and then further divided them into two subgroups based on gender composition: one composed of students in study groups with a low proportion of female members (<=50%),[8] and the other composed of students in study groups with a high proportion of female members (>50%). The results displayed in Panel B of Table 8 show that belonging to a higher performing study group significantly improved academic performance of male students by 0.132standard deviations but had no significant impact on female students (column 1 and 2, row 3). The average study group score also has a positive and more significant impact on students in study groups with a lower proportion of female members, a result that holds for both male and female students (column 3 and 4, row 3and 5).

One possible reason for this is that when studying with partners in a group, males may be more cooperative (Cárdenas et al., 2014). Therefore, gender is an important index for seating arrangements. Rather than assigning gender-homogenous study groups, this finding suggests that teachers should consider assigning higher-performing males to work with students with poor grades, as the lower-performing students may benefit more from microenvironments with a higher proportion of males.

8. Conclusion

Peers play an integral part in education production. However, the presence of endogenous effects, exogenous effects, and correlated effects in the interactions between peers has historically made identifying the causality of peer effects difficult (Manski, 1993). To estimate the causal effects of study groups on student academic

[8] 50% is the median study group female ratio for all students.

performance, we constructed instruments for study group measures by weighting the average score of study groups with direct or step distances between students. Our estimations revealed that study groups impacted student academic performance both positively and significantly, especially for lower-ranked students. This result supports the “Shining Light” theory that low-performing students can benefit from the example of high-performing students, as well as other empirical findings from recent studies (Berthelon et al., 2019; Card and Giuliano, 2016; Chin and Kwon, 2022; Hoxby and Weingarth, 2005). Furthermore, the peer effect is notably more significant within study groups characterized by higher cohesiveness or among male students.

Additionally, our study offers insight into how peer effects operate. Unlike earlier research identifying anxiety as a negative consequence of classroom competition, we did not find any instances of academic anxiety as a result of study groups (Posselt and Lipson, 2016; Sommet et al., 2013; Wilkinson and Pickett, 2009). On the contrary, we found that studying in a higher-achieving study group for one academic year helped relieve anxiety for middle-tercile students. One plausible explanation is that competition for grades among classmates is widespread in classrooms, as teachers and parents often evaluate students according to distribution of grades, which may be a significant source of academic anxiety (Posselt and Lipson, 2016; Tian et al., 2017). As we sorted students into different ranks based on their relative scores within their study groups, among whom rankings are rarely announced, the groups themselves may have been less likely to produce excessive anxiety.

Moreover, our data indicated that academic self-concept (ASC) played an important role in mediating peer effects on academic performance, especially for bottom- and middle-tercile students, whose ASC was greater at the beginning of the semester if they belonged to a higher-performing study group. One likely explanation for this finding is the "big-fish-little-pond effect," which occurs when the target of comparison is included into the mental representation of the self or viewed as similar to the self (Koivuhovi et al., 2022). In other words, bottom- and middle-tercile students may have seen themselves as similar to their higher-performing study mates, leading to an increase in ASC. This beneficial peer effect on ASC persisted until the end of the academic year for middle-tercile students.

We found that intrinsic motivation also played a mediating role between study groups and academic performance. Along with increased ASC, students who began the semester with higher-performing study partners had increased intrinsic motivation at both the beginning and end of the academic year. Unlike ASC, however, this positive impact of study groups on student intrinsic motivation lasted throughout the year, as these students continued to show high intrinsic motivation in the follow-up survey. Our findings suggest that these gains in ASC and intrinsic motivation led low-ranking students to devote more time to learning, ultimately resulting in greater math exam score gains.

We acknowledge several limitations in this study. Firstly, we lacked direct evidence to demonstrate that relative distance was consistent throughout the academic year, as we did not collect new information about seat distribution midway through the academic year. A second limitation of our study was the lack of data regarding previous networks between students. For instance, if some students set up a study group before the start of the academic year, when seats are usually re-assigned, effects from these previous networks would not have been identified. Third, we also lacked information about the specific time spent with each study partner, as students may have been influenced more by certain study partners if they spend more time studying together. This will be considered in our future research in this area. Fourth, our sample schools were all located in economically impoverished areas of China, which may raise the concern of the external validity of this study.

Despite these limitations, by using relative distance between students to investigate the role of peers, our research makes a novel contribution to the current peer effects literature. Specifically, this study reveals that relative distance between students is a viable instrumental variable to reduce the endogeneity of study group formation. Moreover, relative distance between students can determine the intensity of their peer relationships. According to the results from the first stage of regression in the IV analysis, a close relative distance was associated with a higher probability of becoming a study partner. This finding helps fill the information gap in the peer effects literature regarding interaction intensity. By exploring several possible channels by which peers may impact student academic performance, we reveal that intrinsic motivation is the most important mediating variable, which has not been reported in past studies.

Our findings also have implications for educators and researchers alike. First, as our study shows, students are more likely to build study relationships with classmates who sit nearby. Therefore, if educators want to improve the academic outcomes of students with poor class performance, it may be beneficial to seat them with more academically advanced students. Second, our findings reveal fruitful directions for future research, as we show that natural study groups (as opposed to manipulated groups, which are more commonly used in peer effect studies) contain abundant information. For instance, diversity and cohesiveness vary between groups, and the peer effect operates differently in each group, with students benefiting most from membership in groups with low diversity and high cohesiveness. This conclusion provides theoretical support for the future distribution and management of groups in similar studies.

Reference

Aghion, P., Howitt, P., Brant-Collett, M., García-Peñalosa, C., 1998. Endogenous growth theory. MIT press.

Altermatt, E.R., Pomerantz, E.M., Ruble, D.N., Frey, K.S., Greulich, F.K., 2002. Predicting changes in children's self-perceptions of academic competence: a naturalistic examination of evaluative discourse among classmates. Developmental Psychology 38, 903. https://doi.org/10.1037/0012-1649.38.6.903

Amason, A.C., Schweiger, D.M., 1994. Resolving the paradox of conflict, strategic decision making, and organizational performance. International journal of conflict management. https://doi.org/10.1108/eb022745

Anderson, M.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, 1481–1495. https://doi.org/10.1198/016214508000000841

Anderson, M.L., Lu, F., 2017. Learning to Manage and Managing to Learn: The Effects of Student Leadership Service. Management Science 63, 3246–3261. https://doi.org/10.1287/mnsc.2016.2483

Ashcraft, M.H., 2002. Math anxiety: Personal, educational, and cognitive consequences. Current directions in psychological science 11, 181–185. https://doi.org/10.1111/1467-8721.00196

Back, M.D., Schmukle, S.C., Egloff, B., 2008. Becoming friends by chance. Psychological Science 19, 439. https://doi.org/10.1111/j.1467-9280.2008.02106.x

Basch, C.E., 2011. Healthier students are better learners: A missing link in school reforms to close the achievement gap. Journal of school health 81, 593–598.

Benhabib, J., Spiegel, M.M., 1994. The role of human capital in economic development evidence from aggregate cross-country data. Journal of Monetary Economics 34, 143–173. https://doi.org/10.1016/0304-3932(94)90047-7

Berthelon, M., Bettinger, E., Kruger, D.I., Montecinos-Pearce, A., 2019. The Structure of Peers: The Impact of Peer Networks on Academic Achievement. Res High Educ 60, 931–959. https://doi.org/10.1007/s11162-018-09543-7

Boehnke, K., 2008. Peer pressure: A cause of scholastic underachievement? A cross-cultural study of mathematical achievement among German, Canadian, and Israeli middle school students. Social Psychology of Education 11, 149–160. https://doi.org/10.1007/s11218-007-9041-z

Bramoullé, Y., Djebbari, H., Fortin, B., 2020. Peer Effects in Networks: A Survey. Annual Review of Economics 12, 603–629.

Bramoullé, Y., Djebbari, H., Fortin, B., 2009. Identification of peer effects through social networks. Journal of econometrics 150, 41–55. https://doi.org/10.1016/j.jeconom.2008.12.021

Brock, W.A., Durlauf, S.N., 2001. Chapter 54 - Interactions-Based Models, in: Heckman, J.J., Leamer, E. (Eds.), Handbook of Econometrics. Elsevier, pp. 3297–3380. https://doi.org/10.1016/S1573-4412(01)05007-3

Brunner, M., Keller, U., Dierendonck, C., Reichert, M., Ugen, S., Fischbach, A., Martin, R., 2010. The structure of academic self-concepts revisited: The nested Marsh/Shavelson model. Journal of Educational Psychology 102, 964–981. https://doi.org/10.1037/a0019644

Burke, M.A., Sass, T.R., 2013. Classroom Peer Effects and Student Achievement. Journal of Labor Economics 31, 51–82.

Byrne, B.M., 1984. The General/Academic Self-Concept Nomological Network: A Review of Construct Validation Research. Review of educational research 54, 427–456.

Card, D., Giuliano, L., 2016. Can Tracking Raise the Test Scores of High-Ability Minority Students? American Economic Review 106, 2783–2816. https://doi.org/10.1257/aer.20150484

Cárdenas, J.-C., Dreber, A., von Essen, E., Ranehill, E., 2014. Gender and cooperation in children: Experiments in Colombia and Sweden. PloS one 9, e90923. https://doi.org/10.1371/journal.pone.0090923

Carrell, S.E., Sacerdote, B.I., West, J.E., 2013. From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. Econometrica 81, 855–882. https://doi.org/10.3982/ECTA10168

Chin, S., Kwon, E., 2022. Learning with Differing-Ability Peers: Evidence from a Natural Experiment in South Korea. The BE Journal of Economic Analysis & Policy. https://doi.org/10.1515/bejeap-2021-0306

Deci, E.L., 1973. Paying people doesn’t always work the way you expect it to. Hum. Resour. Manage. 12, 28–32. https://doi.org/10.1002/hrm.3930120205

Epple, D., Romano, R.E., 2011. Chapter 20 - Peer Effects in Education: A Survey of the Theory and Evidence, in: Benhabib, J., Bisin, A., Jackson, M.O. (Eds.), Handbook of Social Economics. North-Holland, pp. 1053–1163. https://doi.org/10.1016/B978-0-444-53707-2.00003-7

Foster, G., 2006. It’s not your peers, and it’s not your friends: Some progress toward understanding the educational peer effect mechanism. Journal of Public Economics 90, 1455–1475. https://doi.org/10.1016/j.jpubeco.2005.12.001

Froiland, J.M., Oros, E., 2014. Intrinsic motivation, perceived competence and classroom engagement as longitudinal predictors of adolescent reading achievement. Educational Psychology 34, 119–132. https://doi.org/10.1080/01443410.2013.822964

Gest, S.D., Rulison, K.L., Davidson, A.J., Welsh, J.A., 2008. A reputation for success (or failure): The association of peer academic reputations with academic self-concept, effort, and performance across the upper elementary grades. Developmental Psychology 44, 625–636. https://doi.org/10.1037/0012-1649.44.3.625

Hamilton, B.H., Nickerson, J.A., Owan, H., 2012. Diversity and productivity in production teams, in: Advances in the Economic Analysis of Participatory and Labor-Managed Firms. Emerald Group Publishing Limited.

Han, L., Li, T., 2009. The gender difference of peer influence in higher education. Economics of Education Review 28, 129–134. https://doi.org/10.1016/j.econedurev.2007.12.002

Hanushek, E.A., Kain, J.F., Rivkin, S.G., 2004. Disruption versus Tiebout improvement: the costs and benefits of switching schools. Journal of Public Economics 88, 1721–1746. https://doi.org/10.1016/S0047-2727(03)00063-X

Hare, A.P., Bales, R.F., 1963. Seating position and small group interaction. Sociometry 480–486. https://doi.org/10.2307/2786150

Hinke, S., Leckie, G., Nicoletti, C., 2019. The Use of Instrumental Variables in Peer Effects Models. Oxf Bull Econ Stat 81, 1179–1191. https://doi.org/10.1111/obes.12299

Hoxby, C.M., Weingarth, G., 2005. Taking race out of the equation: School reassignment and the structure of peer effects. Citeseer.

Jackson, C.K., Bruegmann, E., 2009. Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers. American Economic Journal: Applied Economics 1, 85–108. https://doi.org/10.1257/app.1.4.85

Jackson, M.O., 2008. Average distance, diameter, and clustering in social networks with homophily, in: International Workshop on Internet and Network Economics. Springer, pp. 4–11. https://doi.org/10.1007/978-3-540-92185-1_3

Koivuhovi, S., Marsh, H.W., Dicke, T., Sahdra, B., Guo, J., Parker, P.D., Vainikainen, M.-P., 2022. Academic self-concept formation and peer-group contagion: Development of the big-fish-little-pond effect in primary-school classrooms and peer groups. Journal of Educational Psychology 114, 198. https://doi.org/10.1037/edu0000554

Lavy, V., Paserman, M.D., Schlosser, A., 2012. Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers In the Classroom. The Economic Journal 122, 208–237. https://doi.org/10.1111/j.1468-0297.2011.02463.x

Lee, J., 2009. Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. Learning and individual differences 19, 355–365. https://doi.org/10.1016/j.lindif.2008.10.009

Li, L., Liu, Y., 2020. An integrated model of principal transformational leadership and teacher leadership that is related to teacher self-efficacy and student academic performance. Asia Pacific Journal of Education 1–18.

Lu, F., Anderson, M.L., 2015. Peer effects in microenvironments: The benefits of homogeneous classroom groups. Journal of Labor Economics 33, 91–122. https://doi.org/10.1086/677392

Lyle, D.S., 2007. Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point. Review of Economics and Statistics 89, 289–299. https://doi.org/10.1162/rest.89.2.289

Mankiw, N.G., Romer, D., N. Weil, D., 1992. A contribution to the empirics of economic growth. The Quarterly Journal of Economics 107, 407–437.

Manski, C.F., 1993. Identification of Endogenous Social Effects: The Reflection Problem. The Review of Economic Studies 60, 531–542. https://doi.org/10.2307/2298123

Marmaros, D., Sacerdote, B., 2006. How do friendships form? The Quarterly Journal of Economics 121, 79–119.

Marsh, H.W., Abduljabbar, A.S., Parker, P.D., Morin, A.J.S., Abdelfattah, F., Nagengast, B., Möller, J., Abu-Hilal, M.M., 2015. The Internal/External Frame of Reference Model of Self-Concept and Achievement Relations: Age-Cohort and Cross-Cultural Differences. American Educational Research Journal 52, 168–202. https://doi.org/10.3102/0002831214549453

Marsh, H.W., Hau, K.-T., 2003. Big-Fish–Little-Pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools. American psychologist 58, 364. https://doi.org/10.1037/0003-066X.58.5.364

McAndrew, F.T., 1993. Environmental psychology. Thomson Brooks/Cole Publishing Co.

McKeown, S., Stringer, M., Cairns, E., 2016. Classroom segregation: where do students sit and how is this related to group relations? Br Educ Res J 42, 40–55. https://doi.org/10.1002/berj.3200

Min, S., Yuan, Z., Wang, X., Hou, L., 2019. Do peer effects influence the academic performance of rural students at private migrant schools in China? China Economic Review 54, 418–433. https://doi.org/10.1016/j.chieco.2019.02.004

Mora, T., Oreopoulos, P., 2011. Peer effects on high school aspirations: Evidence from a sample of close and not-so-close friends. Economics of Education Review 30, 575–581. https://doi.org/10.1016/j.econedurev.2011.01.004

National Bureau of Statistics, 2019. SHAANXI STATISTICAL YEARBOOK 2019. China Statistical Press, SHANNXI.

OECD, 2014. PISA 2012 technical report.

People's Daily Online, 2014. China's first poverty alleviation day, list of 592 national-level poverty-stricken countiesPeople's Daily Online [WWW Document]. http://politics.people.com.cn/n/2014/1017/c1026-25854065.html. URL http://politics.people.com.cn/n/2014/1017/c1026-25854065.html (accessed 6.8.23).

Pitsia, V., Biggart, A., Karakolidis, A., 2017. The role of students' self-beliefs, motivation and attitudes in predicting mathematics achievement: A multilevel analysis of the Programme for International Student Assessment data.

Learning and Individual Differences 55, 163–173. https://doi.org/10.1016/j.lindif.2017.03.014
Posselt, J.R., Lipson, S.K., 2016. Competition, anxiety, and depression in the college classroom: Variations by student identity and field of study. Journal of College Student Development 57, 973–989. https://doi.org/10.1353/csd.2016.0094
Ramirez, G., Chang, H., Maloney, E.A., Levine, S.C., Beilock, S.L., 2016. On the relationship between math anxiety and math achievement in early elementary school: The role of problem solving strategies. Journal of experimental child psychology 141, 83–100. https://doi.org/10.1016/j.jecp.2015.07.014
Ramirez, G., Gunderson, E.A., Levine, S.C., Beilock, S.L., 2013. Math anxiety, working memory, and math achievement in early elementary school. Journal of Cognition and Development 14, 187–202. https://doi.org/10.1080/15248372.2012.664593
Rhodes, G., Halberstadt, J., Brajkovich, G., 2001. Generalization of Mere Exposure Effects to Averaged Composite Faces. Social Cognition 19, 57–70. https://doi.org/10.1521/soco.19.1.57.18961
Rivera, M.T., Soderstrom, S.B., Uzzi, B., 2010. Dynamics of Dyads in Social Networks: Assortative, Relational, and Proximity Mechanisms. Annu. Rev. Sociol. 36, 91–115. https://doi.org/10.1146/annurev.soc.34.040507.134743
Romer, P.M., 1990. Endogenous technological change. Journal of political Economy 98, S71–S102.
Ryan, R.M., Deci, E.L., 2017. Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness. The Guilford Press, 370 Seventh Avenue, Suite 1200, New York, NY 10001.
Santor, D.A., Messervey, D., Kusumakar, V., 2000. Measuring peer pressure, popularity, and conformity in adolescent boys and girls: Predicting school performance, sexual attitudes, and substance abuse. Journal of youth and adolescence 29, 163–182. https://doi.org/10.1023/A:1005152515264
Silventoinen, K., Bartels, M., Posthuma, D., Estourgie-van Burk, G.F., Willemsen, G., van Beijsterveldt, T.C., Boomsma, D.I., 2007. Genetic regulation of growth in height and weight from 3 to 12 years of age: a longitudinal study of Dutch twin children. Twin Research and Human Genetics 10, 354–363. https://doi.org/10.1375/twin.10.2.354
Simons, J., Dewitte, S., Lens, W., 2000. Wanting to have vs. wanting to be: The effect of perceived instrumentality on goal orientation. British journal of Psychology 91, 335–351. https://doi.org/10.1348/000712600161862
Sommet, N., Pulfrey, C., Butera, F., 2013. Did my MD really go to university to learn? Detrimental effects of numerus clausus on self-efficacy, mastery goals and learning. PloS one 8, e84178. https://doi.org/10.1371/journal.pone.0084178

Thien, L.M., Darmawan, I., Ong, M.Y., 2015. Affective characteristics and mathematics performance in Indonesia, Malaysia, and Thailand: what can PISA 2012 data tell us? Large-scale Assessments in Education 3, 1–16. https://doi.org/10.1186/s40536-015-0013-z

Tian, L., Yu, T., Huebner, E.S., 2017. Achievement goal orientations and adolescents' subjective well-being in school: The mediating roles of academic social comparison directions. Frontiers in psychology 8, 37. https://doi.org/10.3389/fpsyg.2017.00037

Van den Berg, Y.H., Segers, E., Cillessen, A.H., 2012. Changing peer perceptions and victimization through classroom arrangements: A field experiment. Journal of abnormal child psychology 40, 403–412. https://doi.org/10.1007/s10802-011-9567-6

Van Den Berg, Y.H.M., Cillessen, A.H.N., 2015. Peer status and classroom seating arrangements: A social relations analysis. Journal of Experimental Child Psychology 130, 19–34. https://doi.org/10.1016/j.jecp.2014.09.007

Whitmore, D., 2005. Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment. American Economic Review 95, 199–203. https://doi.org/10.1257/000282805774670158

Wilkinson, R.G., Pickett, K.E., 2009. Income inequality and social dysfunction. Annual review of sociology 493–511. https://doi.org/10.1146/annurev-soc-070308-115926

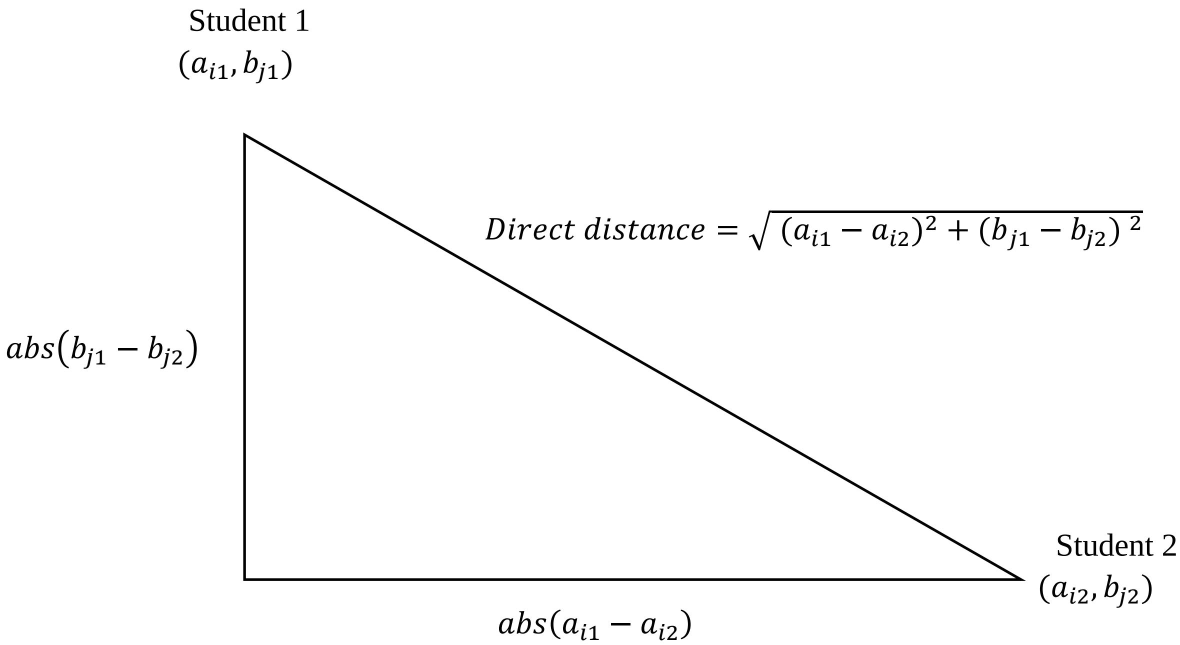

Figure 1. Calculation of direct distance and step distance

Table 1. Student and study group math-related outcomes at baseline and follow-up survey.

		Mean	Std. Dev.	Max.	Min.
Study group measures					
[1]	Average math score of study group at baseline (excluding self)	0.171	0.651	1.907	-3.198
[2]	Average math score of study group at baseline (excluding self, weighted by direct distance)	0.256	0.876	3.855	-5.201
[3]	Average math score of study group at baseline (excluding self, weighted by step distance)	0.254	0.874	3.818	-5.189
Student outcomes at baseline survey					
[4]	Standardized math test score	0.035	0.987	2.023	-4.463
[5]	Anxiety	0.012	0.687	1.920	-1.468
[6]	Self-concept	-0.013	0.675	1.628	-2.011
[7]	Intrinsic motivation	-0.011	0.739	1.394	-2.396
[8]	Instrument motivation	-0.004	0.758	1.034	-3.269
[9]	Frequency of cooperation with classmates	3.057	1.219	5.000	1.000
[10]	Frequency of distraction in class	2.728	1.337	5.000	1.000
[11]	Studying time ratio per day	0.315	0.165	1.000	0.000
Student outcomes at follow-up survey					
[12]	Standardized math test score	0.042	0.983	2.000	-3.611
[13]	Anxiety	0.029	0.739	1.906	-1.433
[14]	Self-concept	-0.029	0.737	1.643	-1.926
[15]	Intrinsic motivation	-0.024	0.785	1.450	-2.311
[16]	Instrument motivation	-0.010	0.800	0.979	-3.470
[17]	Frequency of cooperation with classmates	3.004	1.210	5.000	1.000
[18]	Frequency of distraction in class	2.786	1.282	5.000	1.000
[19]	Studying time ratio per day	0.275	0.154	1.000	0.000
[20]	Observations		2,956		
[21]	Number of classes		114		

Source: Authors' survey

Table 2. Descriptive statistics.

		Mean	Std. Dev.	Max.	Min.
Demographic and family characteristics at baseline					
[1]	Age (years)	10.949	1.044	17.250	5.917
[2]	Gender (1=male; 0=female)	0.527	0.499	1.000	0.000
[3]	Boarding (1=yes; 0=no)	0.380	0.485	1.000	0.000
[4]	Number of family members	4.949	1.460	15.000	2.000
[5]	Mother graduated from junior high school (1=yes; 0=no)	0.436	0.496	1.000	0.000
[6]	Father graduated from junior high school (1=yes; 0=no)	0.506	0.500	1.000	0.000
[7]	Standardized family asset value	-0.007	1.625	4.890	-2.156
[8]	Fourth grade	0.247	0.431	1.000	0.000
[9]	Fifth grade	0.369	0.483	1.000	0.000
[10]	Sixth grade	0.384	0.486	1.000	0.000
Teacher characteristics at baseline					
[11]	Age (years)	36.905	9.712	59.083	22.667
[12]	Gender (1=female; 0=male)	0.492	0.500	1.000	0.000
[13]	Teacher graduated from junior college (1=yes, 0=no)	0.575	0.494	1.000	0.000
[14]	Teaching experience (years)	15.555	11.606	42.000	0.000
Study partner numbers and average distance at baseline					
[15]	Number of students in the class	31.679	11.721	52.000	5.000
[16]	Study partner numbers	4.683	2.621	10.000	1.000
[17]	Average distance between student and study partner (direct distance)	2.714	1.063	8.062	1.000
[18]	Average distance between student and study partner (step distance)	3.352	1.379	11.000	1.000
[19]	Observations		2,956		
[20]	Number of classes		114		

Source: Authors' survey

Table 3. Estimates of the probability of a study relationship between two students (Logit).

VARIABLES		Dependent variable: study relationship=1	
		(1)	(2)
[1]	Direct distance	-0.035***	
		(0.001)	
[2]	Step distance		-0.026***
			(0.001)
[4]	Pseudo R^2	0.182	0.182
[5]	Observations	97,732	
[6]	Number of students	2,956	
[7]	Number of classes	114	

Note: Robust standard errors in parentheses (adjusted for clustering at the class level).The shared personal characteristics between students being controlled for included whether both students were the same gender (yes/no); whether both were boarding students (yes/no); whether both students had a father or mother who completed junior high school (yes/no); absolute difference of sibling numbers between the two students; absolute difference of family asset value between the two students; and absolute difference of baseline math scores. All estimates include county fixed effects. *** p<0.01, ** p<0.05, * p<0.1.

Table 4. Effect of study group average score on student academic performance.

VARIABLES		Dependent variable: Standardized math test score at follow-up survey		
		(1)	(2)	(3)
		Ordinary least squares (OLS)	Instrumental Variable (IV)	
[1]	Study group baseline average score (excluding self)	0.084***		
		(0.025)		
[2]	Study group baseline average score (excluding self, direct distance)		0.114***	
			(0.035)	
[3]	Study group baseline average score (excluding self, step distance)			0.113***
				(0.035)
[4]	Student baseline math score	0.564***	0.558***	0.558***
		(0.020)	(0.021)	(0.021)
[5]	Observations		2,956	
[6]	Adjusted R-squared	0.408	0.408	0.408
[7]	First-stage statistics			
[8]	R-squared		0.586	0.587
[9]	Partial R^2 excluded instrument		0.459	0.460
[10]	*F*-test		349.754	363.034

Note: Robust standard errors in parentheses (adjusted for clustering at the class level).The student characteristics being controlled for included age (in years), gender (1, male; 0, female), boarding (1, yes; 0, no), grade (1, student from grade 5/grade 6; 0, otherwise), number of family members, the education level of parents (1, father or mother of the student has graduated from junior high school; 0, otherwise), and standardized asset value. All estimates include county fixed effects. *** p<0.01, ** p<0.05, * p<0.1.

Table 5. Robustness analysis

VARIABLES	Dependent variable: Standardized math test score at follow-up survey	
	(1)	(2)
	Instrumental Variable (direct distance)	Instrumental Variable (step distance)
Panel A: Excluding students who chose their own seats		
[1] Study group baseline average score (excluding self, step distance)	0.085** (0.041)	0.085** (0.042)
[2] Observations	2362	2362
[3] Adjusted R-squared	0.380	0.380
Panel B: Excluding students seated based on math test score or personality		
[4] Study group baseline average score (excluding self, step distance)	0.082** (0.045)	0.082** (0.045)
[5] Observations	1853	1853
[6] Adjusted R-squared	0.433	0.433
Panel C: Excluding classes in which relative distance between students was not constant throughout the academic year		
[7] Study group baseline average score (excluding self, step distance)	0.091** (0.041)	0.089** (0.041)
[8] Observations	2085	2085
[9] Adjusted R-squared	0.402	0.402

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Each regression controls for all student and teacher characteristics mentioned in the note of Table 4, baseline standardized math exam scores, and county fixed effects. All standard errors account for clustering at the class level.

Table 6. Effect of study group average score on student academic performance, as differentiated by student baseline score ranking within study group (IV estimation).

	VARIABLES	Dependent variable: Standardized math test score at follow-up survey		
		(1)	(2)	(3)
		Student is in bottom tercile (0/1)	Student is in middle tercile (0/1)	Student is in top tercile (0/1)
[1]	Study group baseline average score (excluding self, step distance)	0.201* (0.105)	0.141* (0.081)	-0.041 (0.081)
[2]	Student baseline math score	0.413*** (0.061)	0.603*** (0.050)	0.782*** (0.063)
[3]	Observations	916	1,050	990
[4]	Adjusted R-squared	0.279	0.399	0.369

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Each regression controls for all student and teacher characteristics mentioned in the note of Table 4, baseline standardized math exam scores, and county fixed effects. All standard errors account for clustering at the class level.

Table 7. Effect of study group average score on math learning attitude (IV estimation across tercile subgroups).

VARIABLES	Dependent variable: math learning attitude						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Academic anxiety	Self-concept	Intrinsic motivation	Instrumental motivation	Cooperation frequency	Distraction frequency	Studying time ratio
Panel A: Effect of baseline study group average score on baseline student math learning attitudes							
[1] Baseline study group average score (all samples)	-0.059	0.048	0.120***	0.025	0.297***	-0.070	0.001
	(0.037)	(0.038)	(0.043)	(0.042)	(0.091)	(0.108)	(0.011)
[2] Baseline study group average score (bottom tercile)	-0.077	0.177**	0.247**	0.03	0.584***	-0.332*	0.014
	(0.076)	(0.082)	(0.096)	(0.105)	(0.192)	(0.188)	(0.024)
[3] Baseline study group average score (middle tercile)	-0.093	0.128**	0.253***	0.101	0.401***	-0.09	0.018
	(0.085)	(0.059)	(0.097)	(0.099)	(0.129)	(0.175)	(0.021)
[4] Baseline study group average score (top tercile)	-0.07	-0.004	0.042	-0.058	0.151	0.278	-0.006
	(0.085)	(0.099)	(0.094)	(0.084)	(0.140)	(0.219)	(0.019)
Panel B: Effect of baseline study group average score on follow-up student math learning attitudes							
[5] Baseline study group average score (all samples)	-0.016	0.034	0.108**	0.051	-0.059	0.072	0.008
	(0.037)	(0.035)	(0.049)	(0.043)	(0.099)	(0.105)	(0.008)
[6] Baseline study group average score (bottom tercile)	0.067	-0.068	0.004	0.025	-0.271	0.206	-0.018
	(0.080)	(0.074)	(0.100)	(0.126)	(0.208)	(0.211)	(0.019)
[7] Baseline study group average score (middle tercile)	-0.137*	0.234***	0.224**	0.037	-0.075	0.174	0.026*
	(0.083)	(0.085)	(0.094)	(0.091)	(0.148)	(0.182)	(0.016)
[8] Baseline study group average score (top tercile)	0.077	-0.084	0.023	-0.004	-0.028	0.142	0.005
	(0.087)	(0.080)	(0.084)	(0.078)	(0.174)	(0.153)	(0.017)

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Each regression controls for student and teacher characteristics mentioned in the note of Table 4, baseline standardized math exam scores and county fixed effects. All standard errors account for clustering at the class level. There are 2,956 students, in which 916 students are in the bottom tercile, 1,050 students are in the middle tercile, and 990 students are in the top tercile. Study group baseline average score is estimated by step distance and excluding students herself or himself.

Table 8. Heterogeneity analysis

VARIABLES	Dependent variable: Standardized math test score at follow-up survey			
Panel A: Study group structure	(1)	(2)	(3)	(4)
	Low diversity	High diversity	Low cohesive	High cohesive
[1] Study group baseline average score	0.131***	0.119**	0.033	0.179***
	(0.048)	(0.055)	(0.048)	(0.051)
[2] Observations	1483	1473	1480	1476
Panel B: Student gender and study group female ratio	(1)	(2)	(3)	(4)
	Male	Female	Low female ratio	High female ratio
[3] Study group baseline average score	0.132***	0.079	0.136***	0.078
	(0.044)	(0.060)	(0.042)	(0.064)
[4] Observations	1558	1398	1578	1378
Panel C: Study group female ratio by student gender	(1)	(2)	(3)	(4)
	Male		Female	
	Low female ratio	High female ratio	Low female ratio	High female ratio
[5] Study group baseline average score	0.134***	-0.658	0.363*	0.079
	(0.043)	(0.583)	(0.195)	(0.064)
[6] Observations	1519	39	59	1339

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Each regression controls for all student and teacher characteristics mentioned in the note of Table 4, baseline standardized math exam scores, and county fixed effects. All standard errors account for clustering at the class level. Study group baseline average score is estimated by step distance and excluding the student herself or himself.