Unintended Consequences of Best Intentions: Examining Spillover Effects in Targeted Supplementary Education Interventions

Yujuan Gao, University of Florida, yujuan.gao@ufl.edu

Yue Ma, Stanford University, yma3@stanford.edu

Conner Mullally, University of Florida, connerm@ufl.edu

Scott D Rozelle, Stanford University, rozelle@stanford.edu

[Behavioral & Institutional Economics] prepared for presentation at the 2025 AAEA & WAEA Joint Annual Meeting in Denver, CO; July 27-29, 2025

Unintended Consequences of Best Intentions: Examining Spillover Effects in Targeted Supplementary Education Interventions

Yujuan Gao, Yue Ma, Conner Mullally, and Scott Rozelle*

June 19, 2025

Abstract

This study examines spillover effects of targeted educational interventions through a field experiment in 130 rural Chinese boarding schools, comparing computer-assisted learning (CAL) and traditional workbooks. Results reveal significant negative spillovers of workbook interventions on nontarget students' performance, particularly affecting those closely connected to targeted students. Effects intensify with increased exposure and peer interaction. The key mechanism appears motivational: Observing peers receiving supplementary workbook resources in class reduces students' confidence in the value of their academic efforts for future careers. CAL interventions, conducted outside classrooms, show no such spillovers, highlighting the importance of considering unintended consequences in competitive, resource-limited environments.

JEL Codes: I21, C93, D62

---

*Yujuan Gao is an PhD candidate of Food and Resource Economics at University of Florida (yujuan.gao@ufl.edu). Yue Ma is a PhD student and research fellow at Stanford University. Conner Mullally is an Associate Professor of Food and Resource Economics at University of Florida. Scott Rozelle is a Senior Fellow at Stanford University. We thank the participants and organizers of the 2025 ASSA Meeting, the 2024 APPAM Fall Research Conference, and the 2024 Stanford Center on China's Economy and Institutions (SCCEI) Young Researcher Workshop. We also appreciate valuable feedback from Jared Gars and Xinde "James" Ji at the University of Florida. We are especially grateful to the field researchers and enumerators from the Center for Experimental Economics in Education at Shaanxi Normal University for their dedication. The authors have no financial relationships related to this research to disclose. This randomized controlled trial is registered in the American Economic Association's registry for randomized controlled trials under ID AEARCTR-0003086. The data and code used in this analysis will be made available upon publication.

1 Introdcution

The design of economic policy interventions inherently involves fundamental trade-offs between efficiency and equity under resource constraints. To maximize effect, interventions typically target specific high-need groups. However, social interactions between targeted and nontarget populations generate externalities that reshape welfare impacts beyond direct beneficiaries. Recent experimental evidence documents substantial indirect benefits across multiple domains. Early childhood interventions generate spillover effects on neighboring children's skills and health through interactions between parents or children (Kusumawardhani, 2022; List et al., 2023), while medical treatments create positive externalities for maternal mental health and sibling academic performance (Daysal et al., 2022). Cash transfers to eligible households increase consumption among ineligible village residents (Angelucci & De Giorgi, 2009; Egger et al., 2022), and information disseminated through central community members enhances village-wide vaccination rates (Banerjee et al., 2019). In agricultural settings, social networks of subsidized farmers experience spillovers in technology adoption, yields, and beliefs (Beaman et al., 2021; Carter et al., 2021; Vasilaky & Leonard, 2018). However, resource competition can generate offsetting negative externalities, as evidenced by control households experiencing adverse effects in high treatment concentration areas of agricultural training programs (Duflo et al., 2023).

Educational training programs offer an ideal setting to examine externality effects due to their inherently social nature. Research has shown that even brief social activities, such as interactions during freshman orientation, can create significant peer effects through network formation (Brade, 2024). This is because in schools and workplaces, regular interaction allows individuals to easily observe their peers' resources and behaviors. These social interactions fundamentally shape both skill acquisition and productivity development through multiple channels. Research shows that individual productivity is significantly influenced by coworkers' characteristics and performance—a phenomenon known as productivity spillovers

(Battu et al., 2003; Cornelissen et al., 2017; Guryan et al., 2009). These spillovers can be both positive and negative: Knowledge sharing and skill complementarity create positive externalities (Azoulay et al., 2010; Jaravel et al., 2018), while competitive pressures can generate negative effects that may limit individual growth and overall productivity (Breza et al., 2018; Cornelissen et al., 2017).

We study spillover effects of a supplementary educational intervention targeting boarding students in rural Chinese primary schools. This setting offers unique advantages for studying peer effects. Boarding students reside on campus and rely primarily on school resources after class while interacting regularly with non-boarding peers during school hours. Additionally, rural Chinese primary schools feature fixed classroom assignments where students maintain consistent seating throughout the semester as teachers rotate between rooms. This structure facilitates stable peer networks and enables information diffusion between boarding and non-boarding students. Our cluster-randomized controlled trial spans 130 primary schools, including over 3,000 boarding students across 352 math classes in grades 4-6. The experiment compares three conditions: (i) computer-assisted learning (CAL), (ii) traditional workbook instruction (pencil-and-paper learning treatment), and (iii) a control group. While Ma et al. (2024) did not find positive estimates of the CAL program and workbook program on the test scores of boarding students, our study examines how these targeted interventions indirectly affect their 6,000 non-boarding classmates. This design addresses common challenges in estimating causal peer effects, such as simultaneity, correlated unobservables, and endogenous peer group formation (Manski, 1995). By introducing exogenous variation through targeted interventions while leveraging naturally formed social networks, we can identify genuine spillover effects.

The findings reveal distinct spillover effects of the two interventions. While the CAL intervention shows no significant spillover effects on nontarget (non-boarding) students, the workbook intervention produces notable negative spillover effects on nontarget students who frequently collaborate with targeted (boarding) peers. In addition, the magnitude of these spillovers intensifies as the number of targeted peers increases. However, nontarget students who don't interact with targeted peers remain unaffected by both interventions. Furthermore, for nontarget students with connections to targeted students, the negative spillover effects strengthen as treatment intensity among targeted peers increases. The order in which targeted peers are listed by nontarget students correlates with the strength of the negative spillover effects. Specifically, stronger negative spillovers are observed when the targeted peers are listed as the first peers. Interestingly, these negative spillover effects do not diminish with greater social distance between nontarget students and targeted peers. This suggests that the impact of the workbook intervention extends beyond immediate social connections, potentially affecting the broader classroom dynamic.

We then explore potential mechanisms that drive the deleterious effects of interventions on nontarget students. We first examine how additional training to target peers affects motivation of nontarget student regarding effort invested in learning mathematics. Ryan and Deci (2017) propose that individuals typically experience two types of reward-based motivation. The first is intrinsic motivation, where people engage in activities for personal satisfaction or to better integrate with their social and physical environment. The second, instrumental motivation, is driven by future objectives and pursuits with practical value, such as improving academic performance, enhancing career prospects, increasing financial gains, or achieving job promotions (Eccles & Wigfield, 2002; Gagné & Deci, 2005; Simons et al., 2004). Our analysis reveals that workbook interventions significantly influence how nontarget students view the value of their academic work. When students observe their peers receiving supplementary workbook resources in the classroom, they tend to devalue their own academic efforts. Seeing peers receive supplementary resources appears to diminish students' belief that their effort on regular coursework will build valuable career skills. This trend is particularly prominent for female students. In contrast, computer-assisted learning, which takes place outside the classroom, does not create these negative motivational effects. This difference suggests that the visibility of supplementary resources in shared learning spaces plays a crucial role in generating indirect effects.

Our study reveals two key effects on teacher behavior. While both interventions maintained consistent patterns of classroom time allocation and student assistance, the workbook program had an unintended effect on teacher perceptions. Specifically, teachers showed systematic bias in their evaluations, overestimating performance of students who had close connections to workbook program participants while underestimating those without such connections. This observation aligns with the halo effect (Nisbett & Wilson, 1977; Thorndike, 1920), where students' association with a supplementary program influences how teachers perceive their overall performance. This tendency among students in the treatment group aligns with findings from Ma et al. (2024), which demonstrated that two interventions led teachers to assign significantly higher class rankings to boarding students. Our findings suggest that supplementary educational programs should include specific guidelines for teacher evaluation practices to minimize such unintended assessment biases. Our findings also challenge the initial hypothesis that workbook interventions would intensify classroom competition and create negative spillovers. We found no evidence that nontarget students became less willing to study or collaborate with peers who received workbooks, indicating that cooperative learning patterns remained stable. Additionally, we did not observe increased test score dispersion within groups or classes—a common indicator of heightened academic competition (Abramo et al., 2012; Fallucchi et al., 2021; Hoxby, 2000). In conclusion, the workbook intervention appears to have primarily affected students' academic motivation while leaving both resource allocation and competitive dynamics unchanged in the classroom.

This study contributes to several strands of the literature. First, it speaks to the growing body of research on intervention externalities in human capital development. The existing literature on educational interventions presents a complex picture of spillover effects. While several studies have documented positive externalities on nontarget children (Islam et al., 2021; Kusumawardhani, 2022; List et al., 2023), others have observed either null or negative spillover effects (Becker et al., 2022; Pedersen et al., 2017). Our study adds a crucial dimension to this discourse by examining spillover effects in a resource-constrained, highly competitive learning environment in rural China, highlighting the importance of considering education resource access in program design. Our study is also closely aligned with the work of Angelucci et al. (2019), who used a field experiment on snack choice in school lunchrooms. They found that while peers' actions of choosing grapes had a positive spillover effect on children's take-up of grapes, seeing that peers were incentivized to choose grapes had a negative spillover effect. This is consistent with our observation of negative spillovers from the paper-pencil workbook program, especially among students with close social ties, suggesting that perceived inequalities in resource allocation or attention can have detrimental effects on performance and motivation. We go further by comparing different types of interventions (computer-assisted learning versus traditional workbooks). In addition, we employ a longitudinal approach, building on the work of Abramitzky et al. (2021) who has identified long-term positive spillovers on educational attainment, aspirations, and schooling returns. This approach allows us to capture the temporal evolution of intervention effects and their interaction with changing social dynamics, thereby providing a more comprehensive understanding of spillover effects in educational settings over extended periods.

Our paper also contributes to the literature on peer effects in the general workplace, with important implications in two main areas: organizational management and the design of supplementary training programs. In terms of organizational management, we build upon the work of Cornelissen et al. (2017) and Bentsen et al. (2019), who found small but significant peer effects on wages across various occupations. Our study extends these findings by demonstrating how peer effects operate in a supplementary intervention context, where the stakes and dynamics may fundamentally alter existing social networks, as observed by Banerjee et al. (2024) in formal credit markets. Unlike Duflo et al. (2023), who focused on resource competition in agricultural settings, we explore how interventions act as signals that alter teacher expectations and peer perceptions. Our analysis provides novel insights into supervisors' attitudes when introducing targeted training, offering valuable lessons for maintaining equitable, competitive environments in organizations. Regarding the design of supplementary training programs, our research challenges the common practice in developing countries of training a few workers and relying on organic information diffusion within social networks (Banerjee et al., 2019; Beaman et al., 2021; Chandrasekhar et al., 2022). While this approach could efficiently use limited resources if workers learn from each other, our findings reveal a potential downside: trained workers could impose negative externalities on untrained ones, particularly by reducing motivation in a competitive environment. This aligns with Duflo's observations of negative spillovers in high treatment concentration areas (Duflo et al., 2023), but our study provides a more nuanced understanding of the mechanisms at play. These findings also complement recent work by Anwar et al. (2024), who found that peer composition in training cohorts significantly influences program effectiveness, with treatment effects on reducing arrests being strongest when participants were surrounded by peers with lower arrest histories. By examining these complex interactions between interventions, peer effects, and social networks, our research presents a more comprehensive view of how targeted training initiatives can produce both intended and unintended consequences, with important implications for designing effective and equitable training programs in resource-constrained settings.

2 Background and Conceptual Framework

2.1 Rural-Urban Educational Disparities in China

The disparity in educational resources between rural and urban China has led to significant differences in learning experiences and outcomes. This inequality stems from the unequal distribution of economic and technological resources, creating a persistent urban-rural divide in educational quality and opportunities. Despite government initiatives to address educational disparities through increased rural investment (UNESCO, 2019), significant structural imbalances in education expenditure remain. As of 2019, the per capita educational budget allocation for rural primary schools was 11,127 CNY (1,590 USD), compared to 13,455 CNY (1,922 USD) for urban primary schools (Wei et al., 2022). The consequences of this resource inequality are far-reaching, with rural students experiencing inferior education quality, evidenced by poorer achievement, higher dropout rates, and lower university entrance rates (Wang & Li, 2017; Yang et al., 2014). This disparity also widens the digital skills gap (Chiao & Chiu, 2018) and has long-term implications for human capital accumulation (Londoño de la Cuesta, 1996) and income returns (Herrendorf & Schoellman, 2018; Wang & Li, 2017). Crucially, it hinders the process of structural change necessary for economic development (Laitner, 1999; Ngai & Pissarides, 2024).

Examining educational outcomes in resource-scarce rural environments is critical for tackling disparities, enhancing resource distribution, and guiding policy choices. Additional training can potentially enhance student learning by providing supplementary resources, but it may also intensify existing inequalities. Our study concentrates on identifying the externalities associated with supplementary training. Through analysis of these contexts, we aim to develop interventions that strategically allocate resources, maximizing educational advantages while minimizing the risk of exacerbating inequalities.

2.2 The Schooling Environment in China

The structure of primary education in rural China presents a unique environment for studying the impact of student networks and resource access on educational outcomes. In contrast to many Western educational systems, Chinese rural primary schools typically feature a fixed classroom structure where students remain in assigned seats throughout the semester while teachers rotate between rooms. This arrangement naturally facilitates the formation of stable peer groups and encourages collaborative learning among proximate classmates. The social network of students in these settings is predominantly confined to their immediate classmates, with limited opportunities for interaction outside the classroom.

In addition, a distinctive feature of many rural Chinese primary schools is the prevalence of boarding systems, which accommodate students from remote areas. This system creates two distinct student populations within the same school: boarders, who reside at the school from Sunday afternoon to Friday afternoon and primarily utilize school resources for after-class learning, and non-boarders, who return home daily and potentially access different educational resources. The boarding arrangement facilitates the implementation of additional learning interventions, such as Computer-Assisted Learning (CAL) programs and supplementary workbook exercises, for resident students during times when non-boarders have left campus. This dual system of boarding and non-boarding students within the same classroom environment also offers a rich setting for studying not only the formation of peer networks but also the dynamics of competition and cooperation between these groups.

2.3 Peer-to-Peer Information Diffusion in Academic Settings

The diffusion of information among peers in educational settings is a complex process that involves several interconnected mechanisms. At the foundation of this process is direct verbal communication, where students actively share knowledge, explain concepts, and discuss study materials with one another (Webb, 1989). This explicit exchange of information is complemented by observational learning, a more subtle yet equally powerful mechanism. As posited by Bandura and Adams' social learning theory (Bandura & Adams, 1977), students often adopt study strategies, problem-solving approaches, and even attitudes towards learning by observing and emulating their more successful peers. This observational learning can lead to the passive spread of effective learning materials and methods within groups (Alatas et al., 2016; Noriega-Campero et al., 2018).

The interplay between direct communication and observational learning is often amplified in collaborative learning environments, such as study groups or project teams. These structured settings facilitate a more intense exchange of ideas and knowledge through group discussions and shared problem-solving efforts. Importantly, the effectiveness of these information diffusion mechanisms is not uniform across all peer relationships. The strength of social ties significantly influences both the likelihood and depth of information exchange, with stronger ties generally facilitating more frequent and substantive knowledge sharing (Gee, Jones, & Burke, 2017; Gee, Jones, et al., 2017; Granovetter, 1973). The mixed classroom environment of boarding and non-boarding students provides an ideal setting to study peer effects and information diffusion.

2.4 Competition between Peers

In many educational contexts, the significant emphasis placed on student rankings by parents and teachers has fostered intensely competitive learning environments. This competitive atmosphere often engenders a zero-sum mentality among students, who come to view academic success as a finite resource for which they must compete. Theoretical research has long suggested that such competition could reduce students' inclination to assist others (Drago & Garvey, 1998) and potentially even foster acts of sabotage among rivals (Lazear, 1989). Recent empirical studies have further corroborated these theories. By analyzing university administrative data from China, Chen and Hu (2024) have demonstrated that competitive environments not only impede mutual assistance but also promote unfriendly behaviors among roommates.

Supplementary training for peers can either intensify or reduce such competition. If supplementary training for peers significantly improves their performance, it increases the variability in human capital accumulation among study groups and increases competition directly. In addition, individuals may experience anxiety and pressure when their peers make progress in academic performance. Comparing oneself to those perceived as more successful can lead to feelings of envy and inferiority, potentially causing anxiety and pressure. As a result, ranking differences may become more pronounced and further negatively impact student human capital accumulation (Ashcraft, 2002; Buunk & Ybema, 2003; Ramirez et al., 2016). Furthermore, feelings of unfairness and low expectations regarding individual ranking may arise when students perceive themselves as not receiving supplemental materials while their peers do. According to Self-Determination Theory (SDT) by Ryan and Deci (2017), this sense of inequity can further lead to a loss of study motivation, causing students to be less willing to exert effort or time into their studies.

However, if supplementary training for peers fosters a more collaborative environment, students might allocate more time to group study. In addition, a student's academic success due to interventions can positively influence their peers. This influence may occur through role modeling or extrinsic motivation derived from a student's academic achievements, as noted by Jackson and Bruegmann (2009). Given these competing possibilities, we need to causally identify how spillover effects operate in competitive learning environments.

2.5 Teacher Perceptions

Besides interactions between students, teachers' evaluations and perceptions may change when only certain students receive additional tutoring from external sources, which can link into the "Halo effect". Halo Effect Theory, first introduced by psychologist Edward Thorndike in 1920 (Thorndike, 1920), describes a cognitive bias where an individual's overall impression of a person influences their evaluation of that person's specific traits or characteristics. The Halo Effect can significantly influence teacher evaluations and student outcomes in educational settings. For example, prior knowledge of a student's performance can create bias in grading, potentially leading to inflated grades for students with good reputations (Malouff et al., 2013). In addition, initial teacher perceptions can also create a halo that negatively affects their evaluation of students. These perceptions may be influenced by factors such as gender stereotypes (Carlana, 2019; Lavy et al., 2012) or stereotyping toward immigrants (Alesina et al., 2024).

In our study, we examined potential halo effects in teacher evaluations, particularly concerning non-boarding students who have peers participating in the intervention program. Ma et al. (2024) demonstrated that specific interventions can enhance teachers' objective evaluation of boarding students. Building on this finding, we hypothesized that teachers' observations of positive outcomes in boarding students might influence their assessment of non-boarding peers through a "spillover" effect. This spillover could manifest in two ways: either positively, with improved perceptions of non-boarding students associated with successful boarding peers, or negatively, where teachers might unconsciously evaluate non-boarding students less favorably in comparison to their peers receiving the intervention.

3 Experimental Design and Intervention

3.1 Experimental Design

We designed a field experiment to evaluate the effectiveness of targeted supplementary education interventions and to assess potential spillover effects. The study was conducted within rural primary schools in Shaanxi Province in Northwestern China. A total of 130 schools located in nine impoverished counties were selected to participate in the experiment. Within each school, one class from the fourth, fifth, and sixth grades, each containing a minimum of four boarding students, was randomly chosen, leading to a total sample of 352 school-grades (each represented by one class). As outlined in Figure 1, the experiment consisted of four phases. First, we administered a baseline survey among students, teachers, and principals in October 2017 (near the start of the school year). Second, we designated each of 27 county-grades (nine counties, each with students from three grades) in our sample of 352 school-grades as a strata or block. We then randomly allocated classes within these strata to one of three different treatment conditions after collecting baseline data. Third, in the first half of November 2017, the 4,024 boarding students in the final pool of participants were randomly assigned to one of the three conditions: 1,345 to the CAL treatment group, 1,289 to the workbook treatment group, and 1,390 to the control group. Fourth, the research team returned to the same classes in June 2018 (the last month of the academic year) and administered a follow-up survey. Table A1 summarizes the randomization schedule.

While not receiving supplementary training themselves, the study still surveyed 6,414 non-boarding students who shared classrooms with the participating boarding students. These non-boarding students were distributed across different class types: 2,061 in CAL treatment classes, 2,093 in workbook classes, and 2,260 in control classes.

3.2 Intervention

The randomized controlled trial (RCT) comprised three groups: (1) computer-assisted learning (CAL) with online remedial math tutoring, (2) traditional learning using pencil and paper workbooks, and (3) a control group. Only boarding students participated in the two treatment arms, as they remained at school after classes and could access educational resources exclusively on campus. This design generated exogenous variation in both supplemental CAL and traditional learning, allowing us to assess whether increased investment in educational technology (EdTech) enhances academic performance, particularly during after-school hours when time allocation is more flexible. The CAL intervention was conducted in the computer lab at school. Boarding students in the CAL group were required to use the CAL software after school for one 40-minute weekly session on Sunday afternoons. During each session, those students played animation-based math games designed to help them review and practice the basic math concepts taught in their regular math classrooms. For the workbook group, the boarding student used paper materials in the classroom, which included the same math question in the CAL software after class during the same period when the CAL group was using the software in the computer lab. Non-boarding students typically left school after class and were not subject to group interventions. Neither the teachers nor the schools were informed of the details of the interventions design.

Ma et al. (2024) have reported two key findings regarding the effect of the interventions on boarding students: for the average student, neither traditional supplementary learning (workbook) nor computer-assisted learning (CAL) affected math test scores. However, when examining class grades assigned by math teachers, both supplemental programs improved boarding students' class grades, with no statistically significant differences between the two interventions (see Table A2).

The experimental design also makes it possible to compare the spillover effects on non-boarding students of the two interventions. Spillover effects in educational interventions can vary significantly between technology-based and traditional treatments. The novelty of new technology might create more initial interest and discussion among peers, amplifying short-term spillovers. However, access barriers to technology can limit spillovers in ways that traditional treatments may not experience. For example, spillovers might be limited by access to devices or internet connectivity outside of the treatment group. Also, online knowledge delivered by technology treatments cannot be shared through face-to-face interactions, which are crucial for certain types of knowledge transfer (Alajmi & Al-Qallaf, 2022). By comparing these two intervention types, we can gain valuable insights into the nature and extent of spillover effects in different educational contexts.

4 Data

4.1 Math Test

In this study, math scores are the key measure of a student's academic outcomes. Students took a 35-minute standardized exam in math at the baseline and the follow-up. The exam was grade-appropriate, tailored to the national and provincial-level mathematics curriculum, and constructed by trained psychometricians using a multi-stage process. We standardized math scores into z-scores using each grade's mean and standard deviation for baseline and follow-up surveys. Figure B1 shows the distribution of test scores for boarding and non-boarding students. While there are subtle differences in the score distributions, the overall patterns are quite similar at baseline. Non-boarding students demonstrate a slight advantage on average, as evidenced by their distribution being shifted marginally to the right.

4.2 Mediators

To explore the underlying mechanisms driving student performance, we examined both student self-reported attitudes and teacher-observed behaviors at baseline and in the follow-up survey. For student attitudes, we used four validated scales from the 2012 Programme for International Student Assessment OECD (2014), which are widely applied in educational research (Lee, 2009; Pitsia et al., 2017; Thien et al., 2015). These scales assess math anxiety, academic confidence, intrinsic motivation, and instrumental motivation. Students rated each item on a four-point scale from "strongly disagree" to "strongly agree". Using the GLS weighting procedure described in Anderson (2008), we condensed these responses into four single measures, where positive values indicate above-average levels for each construct. We also assessed students' willingness to collaborate with a single item: "Do you like studying in a group?". Detailed scale items are available in Appendix C. To complement student self-reports, we included teacher evaluations of student classroom performance. These evaluations include average weekly frequencies of student distraction and interruption of classmates, how often students utilize their full mathematical potential in assignments, and the duration of teacher assistance received. This combination of self-reported attitudes and teacher observations provides a comprehensive view of factors potentially influencing student academic performance.

4.3 Study Partner List

To evaluate interactions between non-boarding students and their boarding peers, we utilized study partner lists collected during both baseline and follow-up surveys (see Figure B2). Each student was asked to identify up to ten classmates from the same classroom with whom they most frequently studied or discussed math. From these lists, we constructed a study group for each surveyed student, comprising all the classmates they nominated. These data allow us to track changes in social networks between baseline and follow-up surveys and identify pairwise links between boarding and non-boarding students. Our data shows that non-boarding students reported an average of five study partners at baseline and six in the follow-up survey, with approximately 25% of these being boarding students.

4.4 Distance between Students

There is ample evidence that when students are in close physical proximity to one another in an academic context, it is possible that they can influence each other's academic performance (Marmaros & Sacerdote, 2006; Rivera et al., 2010). To identify the spatial diffusion of the interventions and analyze whether spillovers are localized and diminish with increasing distance, we leveraged data from the final segment of the baseline survey, a graphical representation illustrating the distribution of classroom seating arrangements crafted by enumerators. Utilizing the data derived from this distribution in Figure B3, the relative distances between pairs of students were computed through a series of steps. Initially, a coordinate plane was established within the seating distribution table, with the first row of the classroom serving as the horizontal axis and the first column as the vertical axis. The distance between adjacent desks on each axis was standardized as one desk. Subsequently, coordinates (a_i, b_j) were assigned to students based on their seat order relative to the origin, defined as the first seat in the first row. Finally, the disk distance between two students with coordinates (a_{i1}, b_{j1}) and (a_{i2}, b_{j2}) was determined.

4.5 Socioeconomic Status of Student

We incorporated students' socioeconomic and demographic information as covariates in our model, using variables from the baseline student and teacher surveys. Student variables included gender, parents' education, boarding status, relative household wealth, and family size. Specially, to measure relative household wealth, we created a family asset index from survey data on household items. The survey asked about ownership of various household items, including computers, internet-connected devices, bicycles, microwaves, refrigerators, air conditioners. We summed the items for each household, counting binary items as 1 or 0. We then standardized these sums across the sample. Teacher characteristics comprised age, gender, and education level. Table 1 summarizes student and household demographics and baseline math test scores. Table 2 compares sample characteristics between treatment and control groups for non-boarding students, using an ordinary linear regression model. The dependent variable is workbook or CAL treatment, with student and household characteristics as independent variables. We control for class size, number of boarding peers, and include strata fixed effects, with robust standard errors clustered at the class level. Results show significant differences between workbook and control groups regarding the presence of boarding peers at baseline. For CAL, notable distinctions emerge in teacher gender compared to the control group. These variables are controlled in our treatment effect identification model to isolate exogenous spillover effects.

5 Empirical Approach

One primary goal of this study is to estimate the spillover effect of two interventions on nontarget student academic performance. To do so, we use a linear regression model among non-boarding students:

$$\begin{aligned}\text{Score}_{ic,\text{endline}} = \beta_0 &+ \beta_1 \text{CAL}_{c,\text{baseline}} + \beta_2 \text{Workbook}_{c,\text{baseline}} \\ &+ \beta_3 \text{Boardingnum}_{c,\text{baseline}} + \beta_4 \text{Classsize}_{c,\text{baseline}} + \beta_5 X_{ic,\text{baseline}} \\ &+ \beta_6 \text{Score}_{ic,\text{baseline}} + \pi_s + \epsilon_{ic}\end{aligned} \tag{1}$$

where $Score_{ic,endline}$ is the standardized mathematics score of $student_{ic}$ at the time of the follow-up survey; $CAL_{c,baseline}$ and $Workbook_{c,baseline}$ represent the treatment status of $class_c$;

$Boardingnum_{c,baseline}$ is the number of boarding students in $class_c$ at baseline; $Classsize_{c,baseline}$ is the size of $class_c$; and $X_{ic,baseline}$ is a vector of controls at baseline survey of $student_{ic}$, including gender (1 = male; 0 = female), grade, number of family members, parents' education level (1 = father or mother has graduated from junior high school; 0 = otherwise), and standardized asset value of $student_{ic}$ at the time of the baseline survey, as well as teacher characteristics of age (in years), gender (1 = male; 0 = female), and education level (1 = teacher has graduated from junior high school; 0 = otherwise). $Score_{ic,baseline}$ is the standardized mathematics score of studentic at the time of the baseline survey. π_s is a set of county-grade (strata) fixed effects, with 27 county-grades (9 counties and 3 grades). ϵ_{ic} is a random error term. Rather than including classroom fixed effects, which could introduce exclusion bias in peer effect estimation(Caeyers & Fafchamps, 2024), I opted to adjust standard errors at the classroom level while controlling for observable classroom characteristics. This approach allows me to account for within-classroom error correlation without introducing the mechanical negative relationship between an individual's outcome and their peers' outcomes that can arise with fixed effects

We also estimated Equation (1) separately for two distinct groups: non-boarding students with boarding peers and those without boarding peers. This segregation allowed us to isolate the potential influence of direct connections to treated students.

6 Main Results

6.1 Spillover Effect on Academic Performance

Table 3 presents the results of estimating Equation (1). The results reveal no spillover impact from the Educational Technology (EdTech) intervention on non-boarding students. However, a discernible adverse spillover effect is identified in the case of the workbook intervention, with non-boarding students experiencing a decrease of 0.087 standard deviations in math test scores compared to the control group after one academic year (column 2, row 2) and the difference being significant at 5%. To assess the robustness of our findings, we conducted additional analyses using alternative measures of academic performance. We replaced absolute scores with two types of relative rankings: within-class rankings, which are typically of greatest interest to teachers and parents, and within-study group rankings, which reflect relative academic performance in a student's immediate social network. The results of these analyses, presented in Table A2 (for within-class rankings) and Table A3 (for within-study group rankings), demonstrate the consistency of our results in Table 3.[1]

The observed negative spillover effects on non-boarding students can be attributed to peer interactions rather than direct access to the interventions, as non-boarding students could only be exposed to the information of treatments through their boarding peers. Furthermore, the negative impact on non-boarding students cannot be explained by relative performance decline within classes, since neither intervention significantly improved boarding students' academic performance. This indicates that the negative spillovers are primarily driven by the social interactions between boarding and non-boarding students, with the workbook intervention showing stronger negative effects (-0.087 SD) compared to the CAL program.

To examine the underlying mechanism more directly, we conduct an additional analysis of peer effects between paired boarding and non-boarding students within the same classroom (see Appendix Table A7). This paired-student analysis reveals that under normal conditions, non-boarding students benefit from high-performing boarding peers, with a one standard deviation increase in a boarding student's score associated with a 0.076 standard deviation increase in their paired non-boarding peer's performance. However, the treatment interventions fundamentally alter this relationship. While the CAL program maintains the cooperative dynamic, the workbook treatment creates a more competitive environment. The significant interaction term (0.052, $p<0.05$) combined with the negative main treatment effect (-0.095, $p<0.05$) indicates that workbook-treated boarding students' gains are partially

[1] As an additional robustness check, we removed 535 students (5.2% of the sample) who switched from non-boarding to boarding status during the second semester. Results remained consistent with our main findings, as shown in Table A4.

offset by negative effects on their non-boarding peers, consistent with zero-sum competitive dynamics.

One possible reason for the difference in spillovers between the two interventions is the environment in which they were conducted. The CAL program took place in a computer lab, where information diffusion is less effective compared to the classroom setting used for the workbook treatment. In classrooms, students are in closer proximity, which naturally facilitates information sharing. Additionally, the implementation process may have played a significant role. Both interventions were carried out on Sunday afternoons with boarding students, while non-boarding students were at home. In the classroom setting, instructors were required to collect the workbooks after each weekly session; however, some students may have retained their workbooks if they hadn't finished, potentially leading to more diffusion than in the CAL program. The paired-student analysis supports this interpretation, showing that the workbook treatment specifically disrupts the typically positive peer learning dynamics, transforming collaborative relationships into competitive ones.

6.2 Intensity of Treatment

The intervention was implemented across two academic semesters in Chinese primary schools, starting two months into the first semester. While the planned treatment time was 1,120 minutes (28 weeks $\times$ 40 minutes per session), the actual average engagement time was lower: 590 minutes for the Computer-Assisted Learning (CAL) program and 636 minutes for the workbook program. This discrepancy was attributed to several external factors: technical challenges in the CAL program (including unreliable internet, login issues, and power outages), student absences due to illness or personal reasons, and institutional interruptions such as local traditional holidays and weather-related closures. These combined factors resulted in varying treatment intensities across the participant population.

Higher treatment doses may facilitate greater spillover effects to connected students (Babcock & Hartman, 2010), as intensely treated peers might share more resources, strategies, and skills. To examine this relationship, we modified our analysis by replacing the binary treatment indicators $CAL_{c,baseline}$ and $Workbook_{c,baseline}$ with continuous variables of treatment intensity. We measured treatment intensity as the average time boarding students in treated classes spent on CAL or workbook activities, divided into 10 quantiles. This approach allows us to assess whether the magnitude of spillover effects correlates with the degree of treatment exposure among boarding peers.

Our analysis reveals significant variations in spillover effects based on intervention type and intensity. Regression results in Table 4 demonstrate that increased workbook usage by boarding peers correlates with stronger negative spillover effects on non-boarders. Specifically, a 10% increase in workbook time by boarding classmates corresponds to a 0.014 standard deviation reduction in academic performance for all non-boarders, and a 0.018 standard deviation reduction for those with boarding peers (column 1 and 2, row 2). Figure 2 further illustrates these effects across quantiles. For the CAL intervention (Panel A), impacts vary non-uniformly across quantiles, with students having boarding peers showing slightly more stable negative effects. However, wide confidence intervals suggest considerable uncertainty in these estimates. The workbook intervention (Panel B) displays more consistent patterns, with students having boarding peers generally showing null or more negative effects compared to those without. Notably, the workbook effects exhibit a larger range (approximately -0.2 to 0.2) compared to CAL effects (roughly -0.1 to 0.1). These findings provide strong support for the dose-dependent nature of spillover effects in educational interventions. The more pronounced negative spillover at higher intensity levels of treatment suggests that social connections play a crucial role in transmitting these effects. They also highlight the importance of considering not only the presence of an intervention but also its intensity when evaluating spillover effect.

However, since variations in treatment intensity were determined by endogenous factors rather than random assignment (unlike the treatment arms), we will interpret those intensity-based findings with appropriate caution to avoid drawing causal conclusions.

6.3 Number and Order of Boarding Peers

To explore how network characteristics influence spillover effects, we modified our analysis to account for both network density and connection strength. Network density, represented by the number of treated peers, can amplify spillovers as more nodes become "activated" (Centola, 2010). Simultaneously, stronger connections between individuals, typically characterized by more frequent interactions, may magnify spillover effects by providing increased opportunities for information and behavior transmission (Granovetter, 1973). To capture these dynamics, we introduced interactions between the two treatments (CAL and workbook) and network-related variables into Equation (1). These variables include the number of boarding peers and whether the first nominated peer is a boarding student. Although students were not explicitly instructed to nominate study partners based on interaction frequency, we assume that peers nominated earlier in the list interact more frequently with the respondent. This analysis focuses on non-boarding students who nominated at least one study partner and is conducted only for those with boarding peers. By examining these interactions, we aim to understand how network structure and connection strength modulate the spillover effects of our educational interventions. The Equation is as following:

$$\begin{aligned}
\text{Score}_{ic,\text{endline}} = \beta_0 &+ \beta_1 \text{CAL}_{c,\text{baseline}} + \beta_2 \text{Workbook}_{c,\text{baseline}} \\
&+ \beta_3 \text{Network}_{ijc,\text{baseline}} + \beta_4 (\text{CAL}_{c,\text{baseline}} \times \text{Network}_{ijc,\text{baseline}}) \\
&+ \beta_5 (\text{Workbook}_{c,\text{baseline}} \times \text{Network}_{ijc,\text{baseline}}) \\
&+ \beta_6 \text{Boardingnum}_{ic,\text{baseline}} + \beta_7 \text{Classsize}_{c,\text{baseline}} \\
&+ \beta_8 X_{ic,\text{baseline}} + \beta_9 \text{Score}_{ic,\text{baseline}} + \pi_s + \epsilon_{ic}
\end{aligned} \tag{2}$$

where $Network_{ijc,baseline}$ denotes network-related variables for $student_{ic}$ and all other items are consistent with Equation (1). To ensure that treatment coefficients represent average spillover effects, the number of boarding peers and distance with boarding peers have been de-meaned by non-boarding students before creating interaction terms. Our analysis reveals that network characteristics significantly moderate the spillover effects of educational interventions, as shown in Table 5. For the workbook intervention, we find that for each additional treated peer, the negative spillover effect increases by 0.031 standard deviations (SD) ($p < 0.10$), and when the first-nominated peer is a boarding student (treated), the negative spillover effect is 0.120 SD larger ($p < 0.05$). In contrast, the CAL program shows no significant variation in spillover effects across different network configurations. These findings can be interpreted through social comparison theory: Stronger ties facilitate greater information exchange (Levin & Cross, 2004), closer peers serve as more relevant comparators (Zell & Alicke, 2010), and students with more treated friends may perceive a disadvantage in their own educational opportunities (Fiske, 2010). The differential effects between workbook and CAL interventions suggest that the visibility of the intervention may play a role in triggering social comparison processes. These results underscore the importance of considering network structures and social dynamics when implementing and evaluating educational interventions, as they can significantly influence spillover effects.

6.4 Spatial Fade-out

The relative distance between treated peers, represented by path length in network terms, may influence the persistence of peer effects as they spread through the network. Since information spread and the strength of peer influence declines with social distance (Aral & Nicolaides, 2017; List et al., 2023), we expect to observe a spatial pattern: the spillover effects of the workbook should diminish as we broaden the relative distance between non-boarding students and their boarding peers. To further explore this pattern and shed light on the relationship between spillover effects and distance, we calculated the mean value of desk distance to boarding peers, ranging from 0 to 12, with values of 9-12 recorded as 8 to ensure adequate statistical power. This analysis only includes samples of individuals who have interacted with boarding peers at baseline.

We examine heterogeneity by distance using Equation (2), where cross terms represent interactions between treatments and relative distance. According to the results in Table 5 (column 3), we did not observe a significant difference in treatment effect with increasing average distance. Therefore, we cannot conclude that the spillovers of the two interventions are localized and decrease as the distance to a treated peer increases. This unexpected finding challenges conventional theories that peer influence weakens with increased social distance, as we find no significant attenuation of workbook intervention spillover effects as physical distance between students increases. Several factors could explain this result: The contained classroom environment facilitates interaction regardless of seating distance, the high visibility of workbooks makes them salient to all students, and effects may spread through social networks rather than physical proximity. Additionally, the motivational impacts we identified may operate at a classroom-wide level through general awareness rather than direct peer transmission. This suggests that in close-knit classroom environments, the visibility and psychological impact of targeted resources may matter more than physical distance in shaping outcomes.

7 Identifying the Social Network Mechanism

Understanding the operation of social networks in educational settings is crucial to explaining how supplementary education interventions, which directly influence boarding students' academic performance, may indirectly affect their non-boarding peers. In the context of limited educational resources in rural primary schools in China, supplementary education interventions would make the learning environment more competitive (Chen & Hu, 2024; Ramirez et al., 2016). Furthermore, the interventions may crowd out existing educational resources for non-boarding students, perhaps through preference changes among teachers.

To investigate the spillover effects on competiton between students and teachers' perceptions of students following the introduction of CAL and workbook interventions, we refined our analytical approach in Equation (1). Instead of using math scores as the primary outcome variable, we substituted various mediators collected during the follow-up survey. These mediators include measures of peer relationships, student-teacher interactions, academic self-perception, and other relevant social and academic indicators that described in the Data section. We also included related baseline variables as covariates where applicable. We will estimated Equation (2) and present results separately depending on whether non-boarding students have established connections with boarding peers at baseline. This segregation allowed us to isolate the potential influence of direct connections to treated students.

7.1 Competition between Peers

Assessing the Program's Impact on Classroom Competitiveness

We first tested the competition hypothesis by examining whether supplementary education interventions created a more competitive learning environment. We analyzed the impact of interventions on non-boarding students' willingness to work in groups, test score dispersion within classes or study groups, and the likelihood of having boarding peers at follow-up. Following Abramo et al. (2012), Fallucchi et al. (2021), and Hoxby (2000), we used test score dispersion in the class [2] as a proxy for competitiveness in the learning environment. Our results in Table 6 revealed no negative spillover effects from the workbook intervention on these measures for non-boarding students with boarding peers. Contrarily, the Computer-Assisted Learning (CAL) intervention increased these students' likelihood of studying in groups. For non-boarding students without links to treated students, the workbook intervention also increased their willingness to study in groups but had no significant effect on test score dispersion or number of boarding peers. These findings consistently indicate

[2] Test score dispersion is measured by the standard deviation of test scores within each study group or classroom.

that the supplementary education interventions did not negatively impact the collaborative atmosphere or significantly increase competitiveness within the classroom, suggesting that increased competition is unlikely to be the mechanism behind the observed negative spillover effects on non-boarding students.

Academic Anxiety and Motivation of Non-boarding Students

Prior research discussed in Conceptual Framework suggests that improved academic performance of boarding students may influence the anxiety levels and motivation of their non-boarding peers (Jackson & Bruegmann, 2009; Ramirez et al., 2016; Ryan & Deci, 2017). To further investigate these potential channels, we employed Equation (1), replacing the original outcomes with mediators related to the math learning process. These mediators include academic anxiety, confidence, and both intrinsic and instrumental motivation.

Table 7 presents the results of our analysis, revealing that the spillover effects of the workbook intervention vary depending on the level of interaction between non-boarding and boarding students. For non-boarding students who study alongside boarding peers, the intervention reduces their instrumental motivation by 0.051 points (p¡0.10, column 4, row 2). Conversely, no negative spillover effects on academic motivation were observed for non-boarding students who lack direct interaction with boarding peers. Interestingly, the intervention decreases anxiety scores by 0.1 points (p¡0.05, column 1, row 8) for non-boarding students with less interaction with boarding counterparts. While modest, this effect suggests unexpected benefits in stress alleviation for this group.

A closer examination of instrumental motivation, as shown in Table 8, reveals significant negative spillover effects on two out of four items for non-boarding students who have boarding peers. These items are: (1) "Studying mathematics is very important to me because it will enhance my future job skills" (-0.042 points, p¡0.05); and (2) "Mathematics is a relatively important subject for me because I will need to use it in my future studies" (-0.041 points, p¡0.10). These effects indicate that the intervention may be inadvertently altering non-boarding students' perceptions of the long-term values of mathematical learning. This shift could be attributed to three main factors. First, non-boarding students may perceive that their educational needs are being deprioritized. Second, they may question the value of regular classroom instruction when compared to supplementary programs. Third, they may reassess their academic standing and future opportunities relative to boarding peers who receive additional support.

7.2 Teacher Perceptions

The Halo Effect Theory, initially proposed by Thorndike (1920), suggests that teachers' overall impressions of students can influence their evaluations of academic performance. Ma et al. (2024) observed that both the CAL program and workbook intervention improved teachers' rankings of boarding students. Building on this observation, we hypothesize that receiving supplementary training may serve as a positive signal to teachers. Consequently, we aim to investigate whether the cognitive bias associated with the Halo Effect could lead to positive spillover effects on teachers' perceptions and expectations of non-boarding students who have connections with intervention recipients. If present, these altered perceptions could potentially impact teacher-student interactions and, ultimately, classroom management. To test those hypothesis, we measured teacher evaluation bias. This bias is calculated by comparing teachers' evaluations of student rankings to the students' actual score rankings within classroom during follow-up survey. We then employed this post-intervention evaluation bias as the outcome variable in Equation (1). To account for pre-existing biases and ensure a valid comparison, we included the corresponding bias from the baseline survey as a control variable in our analysis.

Table 9 presents our analysis of teacher evaluation bias, revealing significant effects of the workbook intervention on teachers' perceptions. These effects are associated with students' connections to boarding peers. For non-boarding students with boarding peers, teachers tend to overestimate academic performance, ranking these students approximately 1 position higher ($p < 0.05$) than their actual test score ranking (column 2, row 2). Conversely, for non-boarding students without boarding peers, teachers tend to underestimate performance, ranking these students about 1.6 positions lower ($p < 0.05$) than their actual test score ranking (column 3, row 2). These findings suggest that the workbook intervention induces the Halo Effect, leading to a positive bias in teacher evaluations of non-boarding students who interact more frequently with intervention recipients, while creating a negative bias for non-boarding students without such connections.

Did this bias affect how teachers allocated their time and managed their students? To answer this, we further investigate the spillover effects on teachers' perceptions by examining their post-intervention assessments of non-boarding students in-class behavior and the level of assistance provided to these students by using Equation (1). The outcome variables include teachers' evaluations of students' distractibility, tendency to interrupt, demonstration of math ability, and estimated weekly time spent assisting each student. To ensure accurate comparisons, we replaced baseline test scores with corresponding pre-intervention survey variables where applicable. Our analysis, presented in Table 10, reveals no negative spillover effects from the intervention. Notably, for non-boarding students without boarding peers, the workbook treatment leads teachers to perceive an increased likelihood (6.8%, $p < 0.1$) that these students are exerting their full effort in studying mathematics (column 3, row 8). This finding suggests that while the intervention may have led teachers to overestimate the academic performance of students connected with treated peers, it did not significantly alter teachers' evaluations of other in-class behaviors or reduce the time invested in their students. These results indicate that the intervention primarily influenced academic assessment perceptions without affecting the teachers' overall approach to student interaction or classroom management.

To test whether these spillover effects operate through direct peer-to-peer interactions rather than general classroom dynamics or teacher behavioral changes, we examine how individual peers' academic performance directly affects non-boarding students when only boarding peers receive treatment (see Appendix A5). The analysis reveals that non-boarding students normally benefit from high-performing peers regardless of boarding status, with positive peer effects of similar magnitude for both boarding and non-boarding peers. However, the workbook treatment fundamentally disrupts peer relationships throughout the classroom social network. Notably, the workbook intervention creates negative effects not only when boarding peers receive treatment, but also affects relationships among non-boarding peers who did not receive any treatment themselves. This suggests that the workbook intervention generates classroom-wide competitive dynamics that extend beyond direct treatment recipients to alter the entire peer learning ecosystem. When some students receive intensive workbook training, it appears to shift the collaborative norm within the classroom, making all peer relationships more competitive rather than cooperative. In contrast, the CAL treatment maintains cooperative peer learning dynamics across all student pairs.

8 Heterogeneity on Gender

In rural Chinese primary schools, educational disparities between female and male students naturally arise from gender discrimination in the social environment and gender inequality in family values (Hannum, 2005; Hannum et al., 2021; Wu, 2024). Therefore, it is worthwhile to investigate gender-specific responses to peers receiving supplementary education in the context of such an uneven distribution of educational resources. We analyze non-boarding students separately by gender using Equation (1). Panel A of Table 11 reveals that workbook interventions have a more pronounced adverse spillover effect on female students who have boarding peers, decreasing their scores by 0.108 SD (p¡0.01), while showing no significant impact on male students (column 1 and 2, row 2). We also estimated spillover effects by different genders for non-boarding students without boarding peers, presented in Panel B. These findings indicate that spillover effects from both interventions are insignificant when non-boarding students are not exposed to boarding peers, suggesting that negative spillovers can be mitigated in such circumstances.

We further analyzed whether instrumental motivation remains the primary channel through which negative spillover operates, using Equation (1) but replacing the outcome variable with instrumental motivation and its first two items, which have been significantly impacted by the workbook intervention for the entire population. Analyzing separately the subgroups of non-boarding students with and without boarding peers, we found that for non-boarding students with boarding peers, the workbook intervention reduced female students' instrumental motivation by 0.119 points (p¡0.05, column 3, row 2). Specifically, it affected two items: "Studying mathematics is very important to me because it will enhance my future job skills" (-0.098 points, p¡0.1) and "Mathematics is a relatively important subject for me because I will need to use it in my future studies." (-0.07 points, p¡0.1)(column 5 and 7, row 2) In contrast, the workbook intervention had no significant effect on instrumental motivation for non-boarding students without boarding peers. According to Table A6, when comparing the above treatment effects by gender, we find significant differences between female and male students in the workbook treatment, but only among students with boarding peers (Panel A). Specifically, the difference is significant for instrumental motivation (0.121, p¡0.05) and math effort worthwhile (0.104, p¡0.05). In contrast, among students without boarding peers (Panel B), we observe no significant gender differences in treatment effects.

These results further corroborate the presence of spillover effects propagating through interactions between boarding and non-boarding peers. Additionally, they highlight that future-oriented motivation emerges as a significant channel for these effects. Most importantly, they indicate that even if the workbook interventions are distributed equally among boarding students regardless of gender, non-boarding female students might still feel at a greater comparative disadvantage than their male counterparts, where the latter are more accustomed to receiving preferential treatment in education. Because of these existing disparities, female students may be more acutely aware of any changes in educational resources within their peer group. When they see their boarding peers (both male and female) receiving additional resources through workbook interventions, it may highlight and exacerbate their sense of disadvantage. In addition, observing peers receive additional resources might lead female students to question their own abilities and potential, especially if they perceive these resources as necessary for success, which could result in a decline in academic motivation.

9 Conclusion

Our findings reveal distinct patterns in how supplementary educational interventions affect nontarget students. Traditional workbook interventions negatively impact peers who don't receive the intervention, while EdTech interventions, conducted outside the classroom, show no spillover effects. These negative spillovers appear linked to direct classroom interactions with intervention recipients. The magnitude of negative effects increases with both "treatment dose" (time spent on workbooks) and exposure to treated peers. Specifically, the spillover intensity correlates with both the number of treated peers and the frequency of interactions with them. These effects manifest primarily through decreased future-oriented motivation among nontarget students. Importantly, neither intervention altered teachers' attention to non-boarding students or classroom competition levels. In stead, we observed shifts in teacher perceptions, with teachers showing preference for nontarget students who frequently study with their treated peers. In resource-constrained environments, boys - who typically receive more familial attention - demonstrate greater resilience to these negative spillovers.

Our work has several important policy implications. First, in the realm of educational intervention design, our results underscore the critical need for policymakers and program designers to consider not only the direct effects on target individuals but also the potential spillover effects on nontarget peers. The stark contrast between the null effects of EdTech interventions and the negative spillovers from traditional workbook interventions highlights the importance of intervention method and implementation context. Second, for organizational management, our study emphasizes the necessity of carefully evaluating how targeted training or resource allocation might affect team dynamics and individual motivation. The observed changes in teacher perceptions and the decline in future-oriented motivation among nontarget students call for strategies that maintain perceived equity in resource distribution. Third, in terms of human capital development, our research challenges the assumption that selective interventions in resource-constrained environments will naturally yield positive outcomes for all. Instead, it suggests that skill development and knowledge transfer strategies should be designed with a holistic view of the entire group or organization, not just the direct recipients. Policymakers and leaders should consider implementing more inclusive intervention strategies or, when selective interventions are necessary, develop complementary measures to mitigate potential negative spillovers on nontarget individuals. These considerations are particularly crucial in resource-constrained settings where the equitable distribution of educational resources can have far-reaching implications for long-term human capital development and social mobility.

We acknowledge several limitations in our study. Firstly, the boarding status of 535 students (5.2% of the sample) changed from non-boarding to residing on campus during the second semester of the academic year. Despite not receiving the intervention, their learning environment underwent a significant transformation. Fortunately, we identified and excluded these students, conducted a robustness check, and confirmed that their exclusion did not impact our main findings (Table A4). Secondly, information dissemination extends beyond the confines of the student study group and even the classroom. Students may exchange information with classmates they do not frequently study with, or even with peers from the same school but different classrooms or neighborhoods. Due to the absence of a comprehensive map detailing their social networks, we are unable to measure spillovers through these potential channels, leaving room for future research to explore this aspect.

Tables

Table 1 – Descriptive Statistics

Variable	Mean	SD	Min	Max	N
Fourth grade	0.304	0.460	0	1	6414
Fifth grade	0.350	0.477	0	1	6414
Sixth grade	0.345	0.476	0	1	6414
Gender (1=male; 0=female)	0.530	0.499	0	1	6414
Mother graduated from primary school	0.434	0.496	0	1	6414
Father graduated from primary school	0.516	0.500	0	1	6414
Family asset index	0.275	1.638	−1.874	3.561	6414
Number of family members	5.036	1.534	2	17	6414
Math test score at baseline	0.028	0.979	−4.294	1.973	6414
Has boarding peers at baseline	0.608	0.488	0	1	6414
Teacher's gender (1=male; 0=female)	0.525	0.499	0	1	340
Teacher's age (years)	37.115	9.074	22.168	60.167	340
Teacher graduated from two-year college	0.632	0.482	0	1	340
Class size	40.255	13.312	5	76	340
Number of non-boarding students	27.311	12.159	1	61	340

Note: The family asset index is derived from survey data on household items. It includes ownership of various household items, including computers, internet-connected devices, bicycles, microwaves, refrigerators, and air conditioners. We summed the items per household, with binary items counted as 1 or 0. These sums were then standardized across the sample to create the index.
Source: Authors' survey.

Table 2 – Comparisons of Sample Characteristics Between Treatment and Control Groups

	(1) Workbook-Control	(2) CAL Program-Control
Student's gender (1=male; 0=female)	0.013	0.004
	(0.013)	(0.013)
Mother graduated from junior high school	0.012	0.021
	(0.017)	(0.019)
Father graduated from junior high school	-0.023	0.010
	(0.017)	(0.017)
Family asset index	-0.010	-0.005
	(0.006)	(0.005)
Number of family members	0.006	0.009
	(0.007)	(0.007)
Teacher's gender (1=male; 0=female)	0.026	0.190**
	(0.090)	(0.095)
Teacher's age (years)	0.004	-0.006
	(0.005)	(0.005)
Teacher graduated from two-year college	0.002	0.060
	(0.102)	(0.098)
Class size	-0.005	0.000
	(0.003)	(0.003)
Number of boarding peers	0.000	-0.007
	(0.007)	(0.007)
Standardized score at baseline survey	0.010	0.014
	(0.012)	(0.012)
Exposure to boarding peers	-0.055**	-0.005
	(0.025)	(0.024)
Constant	0.492**	0.581**
	(0.218)	(0.226)
Observations	4,353	4,320
R-squared	0.085	0.082
P-value (F-test)	0.382	0.353

Note: Cluster and class fixed effects are included; Standard errors are clustered at the grade level and robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3 – Treatment Effect on Non-boarding Student Academic Performance

	(1) Whole Sample	(2) Has Boarding Peers	(3) No Boarding Peers
CAL Program	0.018	0.006	0.047
	(0.038)	(0.041)	(0.053)
Workbook Treatment	-0.048	-0.087**	0.014
	(0.038)	(0.041)	(0.053)
Control Mean	0.021	0.071	-0.063
Observations	6,414	3,898	2,516
R-squared	0.422	0.422	0.435

Note: The model controls for outcomes at baseline. Student and teacher characteristics are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at the class level).
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4 – Effect of Treatment Intensity on Non-boarding Student Academic Performance

	(1) Whole Sample	(2) Has Boarding Peers	(3) No Boarding Peers
CAL Program	-0.002	-0.002	0.000
	(0.006)	(0.006)	(0.008)
Workbook Treatment	-0.014**	-0.018***	-0.008
	(0.005)	(0.006)	(0.007)
Control Mean	0.021	0.071	-0.063
Observations	6,414	3,898	2,516
R-squared	0.423	0.422	0.435

Note: The model controls for outcomes at baseline. Student and teacher characteristics are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at the class level).

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5 – Heterogeneity on Social Network Characteristics

	Dependent Variable: Math Test Score		
	(1) Number of Boarding Peers	(2) Order of Nominated Boarding Peers	(3) Desk Distance with Boarding Peers
CAL Program	0.017	0.038	0.005
	(0.038)	(0.044)	(0.041)
Workbook Treatment	-0.047	-0.018	-0.082**
	(0.037)	(0.040)	(0.041)
Number of Boarding Peers	0.019		
	(0.013)		
CAL*Number of Boarding Peers	-0.012		
	(0.018)		
Workbook*Number of Boarding Peers	-0.031*		
	(0.018)		
First Nominated Peer is Boarding		0.047	
		(0.041)	
CAL*First Nominated Peer is Boarding		-0.088	
		(0.057)	
Workbook*First Nominated Peer is Boarding		-0.120**	
		(0.053)	
Desk Distance with Boarding Peers			0.011
			(0.011)
CAL*Desk Distance with Boarding Peers			-0.016
			(0.015)
Workbook*Desk Distance with Boarding Peers			0.005
			(0.017)
Control Mean	0.021	0.029	0.071
Observations	6,414	6,343	3,894
R-squared	0.426	0.420	0.422

Note: Social network characteristics were collected from baseline, including number, order of boarding peers and distance with boarding peers. The number of boarding peers and distance with boarding peers have been de-meaned by non-boarding students. The model controls for outcomes at baseline. Student and teacher characteristics are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at the class level). Missing values in column 2 occur because 71 students in our sample have no friends. We only include samples of whoever interacted with boarding peers in the regression in column 3.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 6 – Treatment Effect on Non-boarding Student Cooperation and Competition

	(1) Like studying in a group?	(2) Test score dispersed in the group	(3) Test score dispersed in the class	(4) Number of boarding peers
Panel A: Has Boarding Peers				
CAL Program	0.053*	-0.025	-0.023	0.014
	(0.028)	(0.020)	(0.018)	(0.174)
Workbook Treatment	0.036	-0.006	-0.001	-0.107
	(0.025)	(0.019)	(0.018)	(0.179)
Control Mean	0.596	0.780	0.898	6.008
Observations	3,892	3,897	3,898	3,898
R-squared	0.044	0.168	0.326	0.147
Panel B: No Boarding Peers				
CAL Program	0.043	-0.031	-0.002	-0.024
	(0.028)	(0.028)	(0.023)	(0.243)
Workbook Treatment	0.058**	-0.013	0.009	-0.329
	(0.027)	(0.028)	(0.021)	(0.224)
Control Mean	0.548	0.758	0.893	5.640
Observations	2,512	2,424	2,516	2,516
R-squared	0.036	0.181	0.409	0.160

Note: The model controls for outcomes at baseline. Student and teacher characteristics are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at the class level). Missing values in columns 1 and 2 are due to some students not answering the questions.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 7 – Treatment Effect on Non-boarding Student Anxiety and Motivation

	(1) Anxiety	(2) Confidence	(3) Intrinsic Motivation	(4) Instrument Motivation
Panel A: Has Boarding Peers				
CAL Program	-0.002	0.019	0.017	-0.007
	(0.034)	(0.029)	(0.036)	(0.030)
Workbook Treatment	-0.020	0.011	0.004	-0.051*
	(0.033)	(0.029)	(0.034)	(0.031)
Control Mean	0.024	-0.019	-0.011	0.092
Observations	3,850	3,832	3,841	3,870
R-squared	0.231	0.338	0.226	0.138
Panel B: No Boarding Peers				
CAL Program	0.028	0.053	-0.007	0.013
	(0.048)	(0.046)	(0.047)	(0.048)
Workbook Treatment	-0.102**	0.049	-0.002	-0.027
	(0.050)	(0.046)	(0.042)	(0.043)
Control Mean	0.034	-0.018	-0.006	0.057
Observations	2,484	2,475	2,485	2,497
R-squared	0.232	0.298	0.241	0.132

Note: The model controls for outcomes at baseline. Student and teacher characteristics are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at the class level). Missing values are due to some students not answering the questions.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 8 – Treatment Effect on Non-boarding Student Motivation Items (Students with Boarding Peers Only)

	(1) Math effort worthwhile	(2) Math for future job	(3) Math for future studies	(4) Math knowledge for future job
CAL Program	-0.004	-0.001	0.001	0.001
	(0.022)	(0.022)	(0.022)	(0.023)
Workbook Treatment	-0.042**	-0.041*	-0.017	-0.024
	(0.021)	(0.023)	(0.020)	(0.023)
Control Mean	0.495	0.417	0.427	0.469
Observations	3,873	3,878	3,884	3,890
R-squared	0.093	0.070	0.068	0.074

Note: The model controls for outcomes at baseline. Student and teacher characteristics are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at the class level). The observations include only non-boarding students who have boarding peers. Missing values are due to some students not answering the questions. Column headers represent abbreviated versions of the following statements: (1) Making some efforts in mathematics is worthwhile, as it will be helpful to me in the future; (2) Studying mathematics is very important to me because it will enhance my future job skills; (3) Mathematics is a relatively important subject for me because I will need to use it in my future studies; (4) I will learn a lot of knowledge from mathematics, and this knowledge will be helpful for me to find a job in the future.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 9 – Treatment Effect on Evaluation of Ranking from Teachers Comparing to Actual Ranking

	(1) Whole Sample	(2) Has Boarding Peers	(3) No Boarding Peers
CAL Program	-0.247	0.036	-1.028
	(0.306)	(0.435)	(0.760)
Workbook Treatment	-0.019	0.993**	-1.602**
	(0.298)	(0.445)	(0.669)
Control Mean	0.508	-0.975	0.574
Observations	5,967	3,618	2,349
R-squared	0.597	0.615	0.579

Note: The model controls for outcomes at baseline. Student and teacher characteristics are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at the class level). Missing values are due to a lack of teacher evaluations from some classes.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 10 – Treatment Effect on Teacher's Evaluation of Non-Boarding Students' Class Performance and their Assistance

	(1) Distraction Frequency	(2) Interrupting Classmates	(3) Math Ability Utilization	(4) Help from Teachers (Min)
Panel A: Has Boarding Peers				
CAL Program	-0.004	0.007	-0.019	-0.699
	(0.038)	(0.028)	(0.034)	(1.702)
Workbook Treatment	0.048	0.049	-0.030	0.117
	(0.034)	(0.031)	(0.032)	(1.618)
Control Mean	0.534	0.295	0.589	17.644
Observations	3,814	3,813	3,812	3,763
R-squared	0.169	0.143	0.152	0.191
Panel B: No Boarding Peers				
CAL Program	0.054	-0.005	0.040	-0.473
	(0.041)	(0.037)	(0.042)	(1.722)
Workbook Treatment	-0.005	0.017	0.068*	0.749
	(0.039)	(0.042)	(0.039)	(1.775)
Control Mean	0.577	0.353	0.478	17.692
Observations	2,459	2,459	2,460	2,423
R-squared	0.185	0.142	0.152	0.215

Note: The model controls for outcomes at baseline. Student and teacher characteristics are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at the class level). Missing values are due to a lack of teacher evaluations from some classes. Column headers represent: (1) Distraction Frequency at Class, (2) Frequency of Interrupting Classmates, (3) Frequency to Give Full Play to Mathematics Ability, (4) Getting Help from Teachers (Minutes).

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 11 – Heterogeneity in Gender

	(1)		(2)		(3)		(4)	
	Math test score		Instrumental motivation		Math effort worthwhile		Math for future job	
	Female	Male	Female	Male	Female	Male	Female	Male
Panel A: Has boarding peers								
CAL Program	-0.019	0.021	-0.042	0.035	-0.018	0.003	-0.022	0.058
	(0.050)	(0.050)	(0.051)	(0.042)	(0.038)	(0.035)	(0.041)	(0.040)
Workbook	-0.108**	-0.073	-0.119**	-0.008	-0.098***	0.000	-0.070*	-0.045
	(0.050)	(0.052)	(0.049)	(0.043)	(0.036)	(0.033)	(0.041)	(0.043)
Control Mean	0.098	0.046	0.059	0.018	3.475	3.409	3.289	3.288
Observations	1,872	2,026	1,857	2,013	1,864	2,020	1,869	2,022
R-squared	0.408	0.442	0.147	0.137	0.117	0.091	0.071	0.079
Panel B: No boarding peers								
CAL Program	0.037	0.056	-0.013	0.032	-0.021	0.002	0.021	0.041
	(0.070)	(0.059)	(0.072)	(0.060)	(0.050)	(0.050)	(0.056)	(0.050)
Workbook	-0.074	0.081	-0.008	-0.063	-0.021	-0.058	0.032	-0.005
	(0.061)	(0.064)	(0.059)	(0.061)	(0.047)	(0.047)	(0.051)	(0.051)
Control Mean	-0.049	-0.074	0.042	-0.060	3.451	3.378	3.299	3.210
Observations	1,142	1,374	1,136	1,362	1,141	1,371	1,140	1,367
R-squared	0.434	0.457	0.135	0.144	0.110	0.097	0.079	0.073

Note: The model controls for outcomes at baseline. Student and teacher characteristics are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at class level). Missing values are due to some students not answering the questions. Dependent variables: (1) Standardized math test score, (2) Instrumental motivation, (3) "Making some efforts in mathematics is worthwhile, as it will be helpful to me in the future", (4) "I will learn a lot of knowledge from mathematics, and this knowledge will be helpful for me to find a job in the future".
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Figures

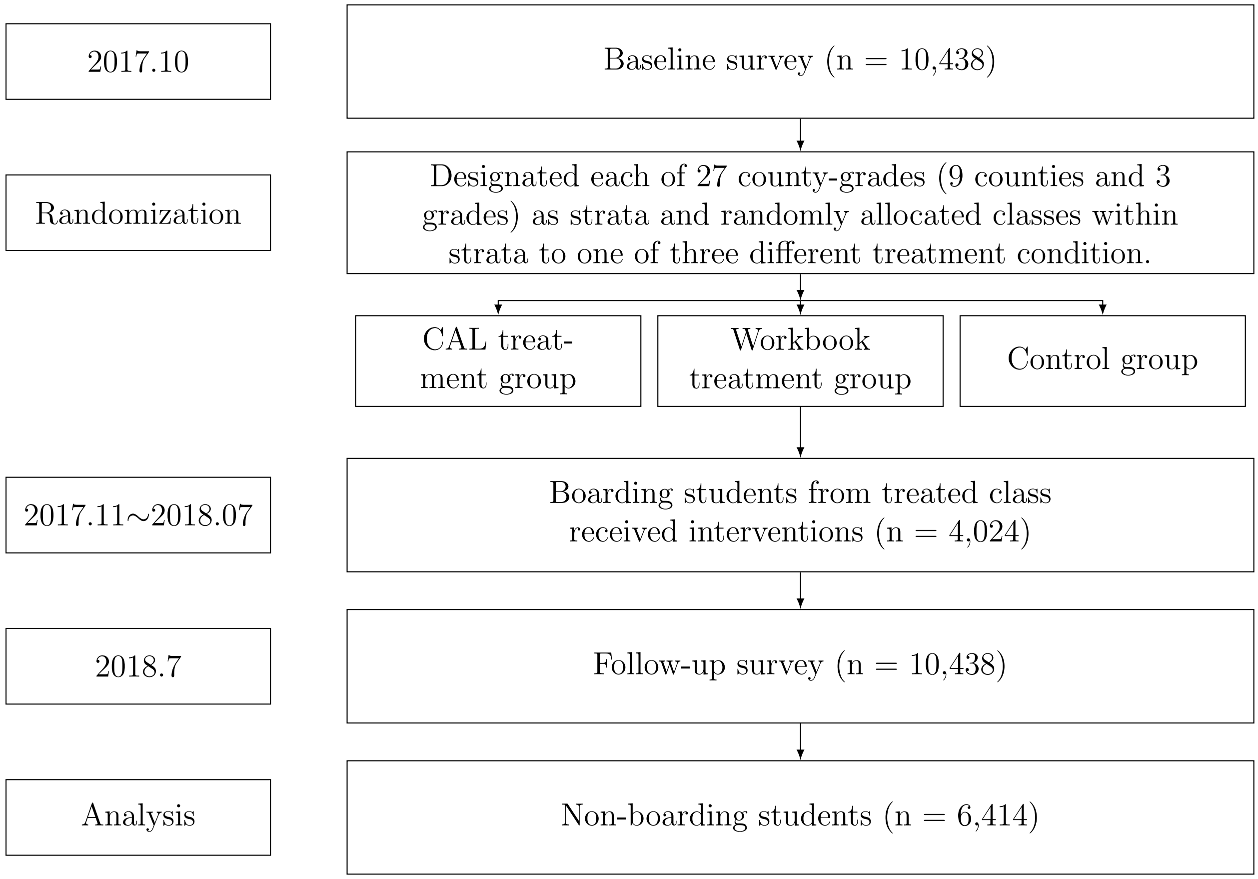
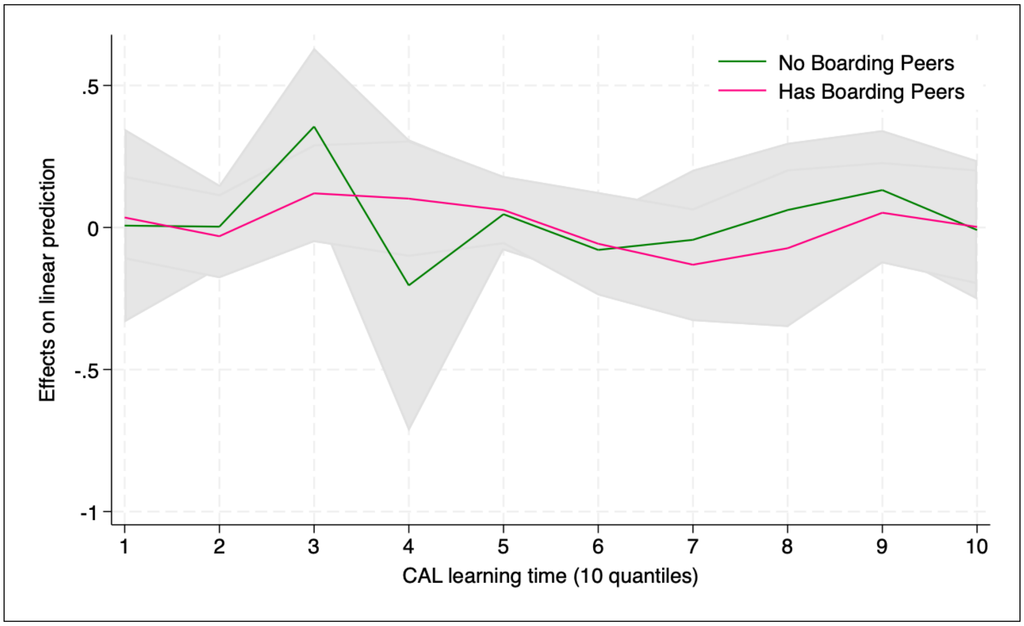


Figure 1 – Experimental Timeline and Design

Panel A: Average marginal effects of CAL program



Panel B: Average marginal effects of workbook program

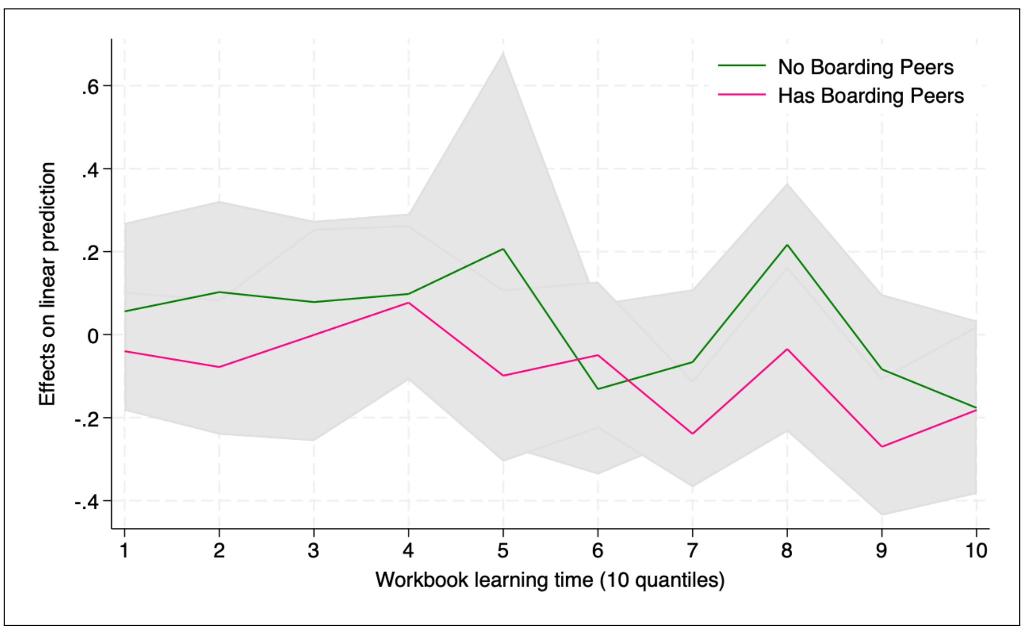


Figure 2 – Average Marginal Effects by Treatment Intensity

Note: These figures show the average marginal effects of CAL (top) and Workbook (bottom) by treatment intensity. The x-axis represents learning time in 10 quantiles, and the y-axis shows the effects on linear prediction. Green lines represent effects for students with no boarding peers, while pink lines represent effects for students with boarding peers. The gray shaded areas indicate confidence intervals. *Source:* Authors' Survey.

Appendix A. Appendix tables

Table A1 – Randomization Schedule of the Program

	CAL	Workbook	Control
Boarding	40-minute/week after school in the computer lab	40-minute/week after school in the classroom	No interventions
4,024	1,345	1,289	1,390
Non-Boarding	Leaving school after class, no interventions		
6,414	2,061	2,093	2,260

Note: The table shows the randomization schedule of the program with the number of participants in each category.

Table A2 – CAL Program Effects on Math Test Scores and Class Grades

	Test Scores		Class Grades	
	(1)	(2)	(3)	(4)
CAL Program	0.033	0.032	1.743*	1.758*
	(0.039)	(0.039)	(0.919)	(0.922)
Workbook Treatment	-0.026	-0.029	1.531*	1.603*
	(0.046)	(0.046)	(0.877)	(0.876)
Additional Controls	No	Yes	No	Yes
R^2	0.432	0.436	0.300	0.308

Note: Table is created based on results from Ma et al. (2024). All columns control for baseline measures (test score or class grade). Even-numbered columns include additional covariates: liking math (scale 1-100), student age (years), gender, father graduated junior high, mother graduated junior high, teacher characteristics (teacher gender, teacher experience in years, teacher attended college) and class characteristics (number of boarding students in the class, class size). Cluster-robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A3 – Treatment Effect on Non-boarding Student Ranking within Class

	(1) Whole Sample	(2) Has Boarding Peers	(3) No Boarding Peers
CAL Program	-0.096	-0.226	0.270
	(0.183)	(0.284)	(0.424)
Workbook Treatment	-0.092	-0.882***	1.017**
	(0.196)	(0.323)	(0.403)
Control Mean	19.693	19.586	19.872
Observations	6,414	3,898	2,516
R-squared	0.455	0.469	0.444

Note: The model controls for outcomes at baseline. Student and teacher characteristics are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at the class level). Missing values are due to some students not answering the questions.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A4 – Treatment Effect on Non-boarding Student Ranking within Group

	(1) Whole Sample	(2) Has Boarding Peers	(3) No Boarding Peers
CAL Program	-0.048	-0.021	-0.068
	(0.051)	(0.069)	(0.066)
Workbook Treatment	-0.114**	-0.134*	-0.050
	(0.051)	(0.071)	(0.059)
Control Mean	3.181	3.548	2.573
Observations	6,414	3,898	2,516
R-squared	0.394	0.359	0.425

Note: The model controls for outcomes at baseline. Student and teacher characteristics are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at the class level).
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A5 – Peer Effects Between Paired Students

	(1) Peers are Boarding Students	(2) Peers are Non-boarding Students
Score of paired student	0.083***	0.071***
	(0.020)	(0.017)
CAL group (1=yes; 0=no)	-0.000	0.010
	(0.041)	(0.042)
CAL $\times$ Score of paired student	-0.015	0.022
	(0.029)	(0.022)
Workbook group (1=yes; 0=no)	-0.096**	-0.100**
	(0.045)	(0.043)
Workbook $\times$ Score of paired student	0.067**	0.049*
	(0.031)	(0.026)
Standardized score at baseline	0.551***	0.538***
	(0.023)	(0.024)
Score of paired student at baseline	-0.009	-0.010
	(0.015)	(0.011)
Control Mean	0.100	0.0978
Observations	7,854	14,154
R-squared	0.434	0.411

Note: The dependent variable is the standardized math test score of non-boarding students at follow-up. Column (1) shows results when non-boarding students' peers are boarding students, and Column (2) shows results when non-boarding students' peers are other non-boarding students. The analysis examines how non-boarding students' performance is affected by their peers' performance within the same classroom. Only boarding students received treatment (CAL or workbook interventions). Student and teacher characteristics for both the non-boarding student and their peers are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at the class level).
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A6 – Treatment Effect on Non-boarding Students, Excluding Those Who Changed Boarding Status

	(1) Whole Sample	(2) Has Boarding Peers	(3) No Boarding Peers
CAL Program	0.015	0.003	0.044
	(0.040)	(0.043)	(0.055)
Workbook Treatment	-0.044	-0.091**	0.024
	(0.038)	(0.043)	(0.053)
Control Mean	0.029	0.081	-0.053
Observations	5,882	3,513	2,369
R-squared	0.417	0.415	0.434

Note: The model controls for outcomes at baseline. Student and teacher characteristics are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at the class level). 535 students (57.25% of the sample) changed from non-boarding to residing on campus during the second semester of the academic year and were excluded during analysis.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A7 – Compare Treatment Effect Between Male and Female

	(1) Math test score	(2) Instrumental motivation	(3) Math effort worthwhile	(4) Math for future job
Panel A: Has boarding peers				
CAL Program	-0.026	-0.046	-0.025	-0.019
	(0.050)	(0.053)	(0.039)	(0.043)
Workbook	-0.111**	-0.119**	-0.099***	-0.068*
	(0.050)	(0.048)	(0.036)	(0.041)
Gender (1=male; 0=female)	0.008	-0.046	-0.056*	-0.016
	(0.038)	(0.039)	(0.032)	(0.034)
CAL*Gender	0.062	0.082	0.032	0.076
	(0.058)	(0.064)	(0.046)	(0.057)
Workbook*Gender	0.047	0.121**	0.104**	0.038
	(0.059)	(0.057)	(0.046)	(0.050)
Control Mean	0.071	0.038	3.441	3.288
Observations	3,898	3,870	3,884	3,891
R-squared	0.422	0.132	0.093	0.062
Panel B: No boarding peers				
CAL Program	0.047	-0.006	-0.020	0.019
	(0.075)	(0.075)	(0.052)	(0.057)
Workbook	-0.055	-0.005	-0.023	0.034
	(0.067)	(0.060)	(0.049)	(0.050)
Gender (1=male; 0=female)	0.032	-0.043	-0.038	-0.057
	(0.060)	(0.051)	(0.045)	(0.049)
CAL*Gender	0.001	0.027	0.019	0.021
	(0.084)	(0.074)	(0.060)	(0.068)
Workbook*Gender	0.124	-0.069	-0.038	-0.039
	(0.079)	(0.073)	(0.060)	(0.066)
Control Mean	-0.063	-0.013	3.412	3.249
Observations	2,516	2,497	2,511	2,506
R-squared	0.436	0.131	0.095	0.070

Note: The model controls for outcomes at baseline. Student and teacher characteristics are controlled. All estimates include strata fixed effects. Robust standard errors are in parentheses (adjusted for clustering at class level). Missing values are due to some students not answering the questions.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Appendix B. Appendix figures

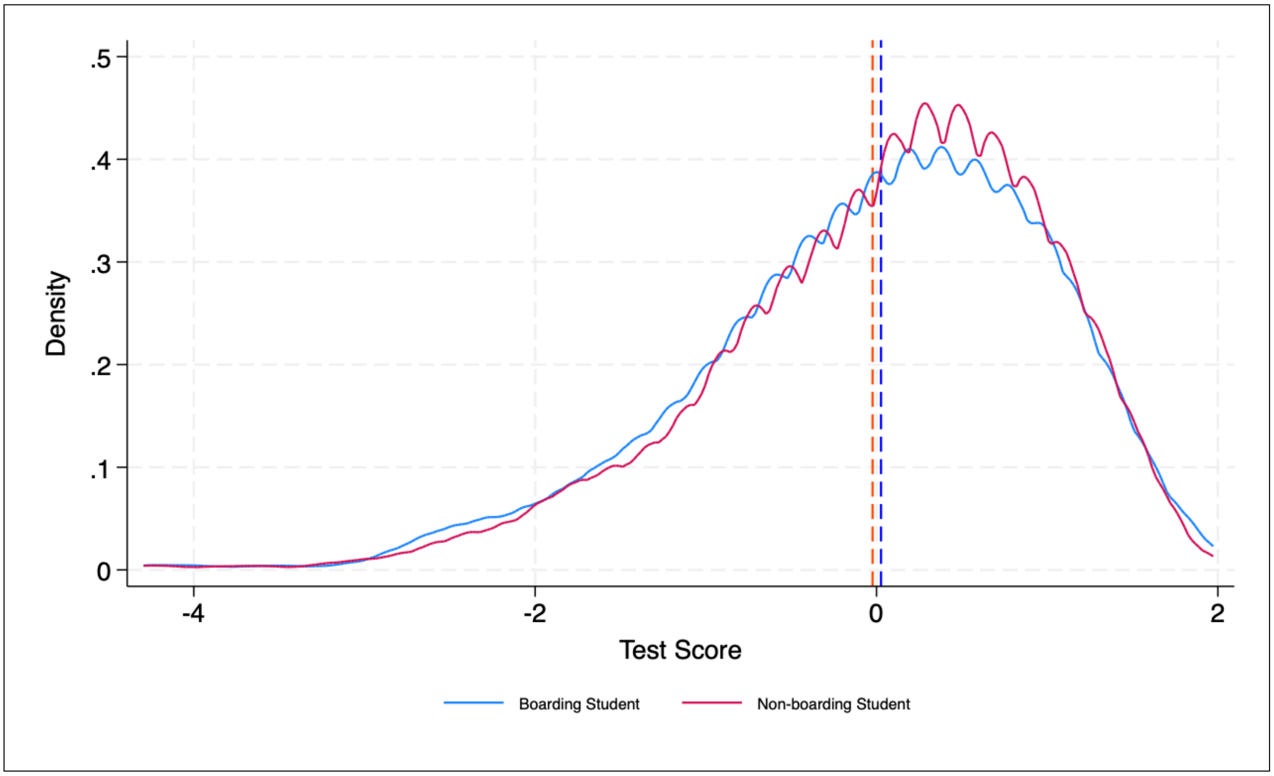


Figure B1 – Distribution of Standardized Math Test Score at Baseline between Boarding and Non-Boarding Students

Note: This figure shows the distribution of standardized math test scores at baseline for boarding and non-boarding students. The x-axis represents the standardized score, and the y-axis represents the density.

同学1 (classmates 1)	**同学2 (classmates 2)**	**同学3 (classmates 3)**	**同学4 (classmates 4)**	**同学5 (classmates 5)**
1.	**2.**	**3.**	**4.**	**5.**
同学6 (classmates 6)	**同学7 (classmates 7)**	**同学8 (classmates 8)**	**同学9 (classmates 9)**	**同学10 (classmates 10)**
6.	**7.**	**8.**	**9.**	**10.**

Figure B2 – Study Partner List

Note: This figure displays the form used to collect study partners' names.

Class Id: **Teachers Name:**

Podium

Door

1		2		3		4		5		6		7		8
Name	ID	Name	ID	Name	ID	Name	ID	Name	ID	Name	ID	Name	ID	Name

Figure B3 – Seat Distribution Table

Note: This figure displays the form used to record seat distribution within the classroom. The first column starts from door, and the first row starts from podium.

Appendix C. Items of Math Learning Attitude scales

Anxiety included 5 items: a) I often worry that it will be difficult for me in mathematics classes; b) I get very tense when I have to do mathematics homework; c) I get very nervous doing mathematics problems; d) I feel helpless when doing a mathematics problem; e) I worry that I will get poor grades in mathematics.

Self-concept included 5 items: a) I am just not good at mathematics; b) I get good grades in mathematics; c) I learn mathematics quickly; d) I have always believed that mathematics is one of my best subjects; e) In my mathematics class, I understand even the most difficult work.

Intrinsic motivation included 4 items: a) I enjoy reading about mathematics; b) I look forward to my mathematics lessons; c) I do mathematics because I enjoy it; d) I am interested in the things I learn in mathematics.

Instrumental motivation included 4 items: a) Making an effort in mathematics is worth it because it will help me in the work that I want to do later on; b) Learning mathematics is worthwhile for me because it will improve my career prospects and chance; c) Mathematics is an important subject for me because I need it for what I want to study later on; d) I will learn many things in mathematics that will help me get a job.

References

Abramitzky, R., Lavy, V., & Pérez, S. (2021). The long-term spillover effects of changes in the return to schooling. *Journal of Public Economics*, *196*, 104369. https://doi.org/10.1016/j.jpubeco.2021.104369

Abramo, G., Cicero, T., & D'Angelo, C. A. (2012). The dispersion of research performance within and between universities as a potential indicator of the competitive intensity in higher education systems. *Journal of Informetrics*, *6*(2), 155–168. https://doi.org/10.1016/j.joi.2011.11.007

Alajmi, B. M., & Al-Qallaf, C. L. (2022). Fostering knowledge-sharing behavior through social capital: The implications of face-to-face and online interactions. *Global Knowledge, Memory and Communication*, *71*(4), 274–292. https://doi.org/10.1108/GKMC-01-2021-0007

Alatas, V., Banerjee, A., Chandrasekhar, A. G., Hanna, R., & Olken, B. A. (2016). Network structure and the aggregation of information: Theory and evidence from indonesia [Publisher: American Economic Association]. *The American Economic Review*, *106*(7), 1663–1704. Retrieved September 30, 2024, from https://www.jstor.org/stable/43861109

Alesina, A., Carlana, M., La Ferrara, E., & Pinotti, P. (2024). Revealing stereotypes: Evidence from immigrants in schools. *American Economic Review*, *114*(7), 1916–1948. https://doi.org/10.1257/aer.20191184

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*(484), 1481–1495. https://doi.org/10.1198/016214508000000841

Angelucci, M., & De Giorgi, G. (2009). Indirect effects of an aid program: How do cash transfers affect ineligibles' consumption? [Publisher: American Economic Association]. *The American Economic Review*, *99*(1), 486–508. Retrieved November 18, 2024, from https://www.jstor.org/stable/29730193

Angelucci, M., Prina, S., Royer, H., & Samek, A. (2019). Incentives and unintended consequences: Spillover effects in food choice [Publisher: American Economic Association]. *American Economic Journal: Economic Policy*, *11*(4), 66–95. https://doi.org/10.1257/pol.20170588

Anwar, S., Baird, M., Engberg, J., & Smart, R. (2024). Do sectoral training programs reduce arrests?: Evidence from a low-income targeted training program RCT. *Journal of Human Resources*, 1122–12629R4. https://doi.org/10.3368/jhr.1122-12629R4

Aral, S., & Nicolaides, C. (2017). Exercise contagion in a global social network. *Nature Communications*, *8*(1), 14753. https://doi.org/10.1038/ncomms14753

Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences [Publisher: SAGE Publications Inc]. *Current Directions in Psychological Science*, *11*(5), 181–185. https://doi.org/10.1111/1467-8721.00196

Azoulay, P., Zivin, J. S. G., & Wang, J. (2010). SUPERSTAR EXTINCTION. *QUARTERLY JOURNAL OF ECONOMICS*, *125*(2), 549–589.

Babcock, P., & Hartman, J. (2010, December). *Networks and workouts: Treatment size and status specific peer effects in a randomized field experiment* (w16581). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w16581

Bandura, A., & Adams, N. E. (1977). Analysis of self-efficacy theory of behavioral change. *Cognitive Therapy and Research*, *1*(4), 287–310. https://doi.org/10.1007/BF01663995

Banerjee, A., Breza, E., Chandrasekhar, A. G., Duflo, E., Jackson, M. O., & Kinnan, C. (2024). Changes in social network structure in response to exposure to formal credit markets. *Review of Economic Studies*, *91*(3), 1331–1372. https://doi.org/10.1093/restud/rdad065

Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2019). Using gossips to spread information: Theory and evidence from two randomized controlled trials. *The Review of Economic Studies*, *86*(6), 2453–2490. https://doi.org/10.1093/restud/rdz008

Battu, H., Belfield, C. R., & Sloane, P. J. (2003). Human capital spillovers within the workplace: Evidence for great britain*. *Oxford Bulletin of Economics and Statistics*, *65*(5), 575–594. https://doi.org/10.1111/j.1468-0084.2003.00062.x

Beaman, L., BenYishay, A., Magruder, J., & Mobarak, A. M. (2021). Can network theory-based targeting increase technology adoption? *American Economic Review*, *111*(6), 1918–1943. https://doi.org/10.1257/aer.20200295

Becker, M., Kocaj, A., Jansen, M., Dumont, H., & Lüdtke, O. (2022). Class-average achievement and individual achievement development: Testing achievement composition and peer spillover effects using five german longitudinal studies. *Journal of Educational Psychology*, *114*(1), 177–197. https://doi.org/10.1037/edu0000519

Bentsen, K. H., Munch, J. R., & Schaur, G. (2019). Education spillovers within the workplace. *Economics Letters*, *175*, 57–59. https://doi.org/10.1016/j.econlet.2018.11.025

Brade, R. (2024). Short-term events, long-term friends? freshman orientation peers and academic performance. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4793902

Breza, E., Kaur, S., & Shamdasani, Y. (2018). The morale effects of pay inequality*. *The Quarterly Journal of Economics*, *133*(2), 611–663. https://doi.org/10.1093/qje/qjx041

Buunk, B. P., & Ybema, J. F. (2003). Feeling bad, but satisfied: The effects of upward and downward comparison upon mood and marital satisfaction. *British Journal of Social Psychology*, *42*(4), 613–628. https://doi.org/10.1348/014466603322595301

Caeyers, B., & Fafchamps, M. (2024). Exclusion bias and the estimation of peer effects. *Journal of Human Resources*, 1120–11337R2. https://doi.org/10.3368/jhr.1120-11337R2

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias*. *The Quarterly Journal of Economics*, *134*(3), 1163–1224. https://doi.org/10.1093/qje/qjz008

Carter, M., Laajaj, R., & Yang, D. (2021). Subsidies and the african green revolution: Direct effects and social network spillovers of randomized input subsidies in mozambique. *American Economic Journal: Applied Economics*, *13*(2), 206–229. https://doi.org/10.1257/app.20190396

Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, *329*(5996), 1194–1197. https://doi.org/10.1126/science.1185231

Chandrasekhar, A. G., Duflo, E., Kremer, M., Pugliese, J. F., Robinson, J., & Schilbach, F. (2022). Blue spoons: Sparking communication about appropriate technology use. *NBER Working Paper 30423*. https://doi.org/http://www.nber.org/papers/w30423

Chen, S., & Hu, Z. (2024). How competition shapes peer effects: Evidence from a university in china. *Review of Economics and Statistics*. https://doi.org/https://doi.org/10.1162/rest_a_01471

Chiao, C., & Chiu, C.-H. (2018). The mediating effect of ICT usage on the relationship between students' socioeconomic status and achievement. *The Asia-Pacific Education Researcher*, *27*(2), 109–121. https://doi.org/10.1007/s40299-018-0370-9

Cornelissen, T., Dustmann, C., & Schönberg, U. (2017). Peer effects in the workplace. *American Economic Review*, *107*(2), 425–456. https://doi.org/10.1257/aer.20141300

Daysal, N. M., Simonsen, M., Trandafir, M., & Breining, S. (2022). Spillover effects of early-life medical interventions. *The Review of Economics and Statistics*, *104*(1), 1–16. https://doi.org/10.1162/rest_a_00982

Drago, R., & Garvey, G. T. (1998). Incentives for helping on the job: Theory and evidence. *Journal of Labor Economics*, *16*(1), 1–25. https://doi.org/10.1086/209880

Duflo, E., Keniston, D., Suri, T., & Zipfel, C. (2023). Chat over coffee? diffusion of agronomic practices and market spillovers in rwanda. *National Bureau of Economic Research, No. w31368*. https://doi.org/10.3386/w31368

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals [Publisher: Annual Reviews]. *Annual Review of Psychology*, *53*, 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Egger, D., Haushofer, J., Miguel, E., Niehaus, P., & Walker, M. (2022). General equilibrium effects of cash transfers: Experimental evidence from kenya. *Econometrica*, *90*(6), 2603–2643. https://doi.org/10.3982/ECTA17945

Fallucchi, F., Fatas, E., Kölle, F., & Weisel, O. (2021). Not all group members are created equal: Heterogeneous abilities in inter-group contests. *Experimental Economics*, *24*(2), 669–697. https://doi.org/10.1007/s10683-020-09677-5

Fiske, S. T. (2010). Envy up, scorn down: How comparison divides us. *American Psychologist*.

Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, *26*(4), 331–362. https://doi.org/10.1002/job.322

Gee, L. K., Jones, J., & Burke, M. (2017). Social networks and labor markets: How strong ties relate to job finding on facebook's social network. *Journal of Labor Economics*, *35*(2), 485–518. https://doi.org/10.1086/686225

Gee, L. K., Jones, J. J., Fariss, C. J., Burke, M., & Fowler, J. H. (2017). The paradox of weak ties in 55 countries. *Journal of Economic Behavior & Organization*, *133*, 362–372. https://doi.org/10.1016/j.jebo.2016.12.004

Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, *78*(6), 1360–1380.

Guryan, J., Kroft, K., & Notowidigdo, M. J. (2009). Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics*, *1*(4), 34–68. https://doi.org/10.1257/app.1.4.34

Hannum, E. (2005). Market transition, educational disparities, and family strategies in rural china: New evidence on gender stratification and development. *Demography*, *42*(2), 275–299. https://doi.org/10.1353/dem.2005.0014

Hannum, E., Liu, X., & Wang, F. (2021). Estimating the effects of educational system consolidation: The case of china's rural school closure initiative. *Economic Development and Cultural Change*, *70*(1), 485–528. https://doi.org/10.1086/711654

Herrendorf, B., & Schoellman, T. (2018). Wages, human capital, and barriers to structural transformation. *American Economic Journal: Macroeconomics*, *10*(2), 1–23. https://doi.org/10.1257/mac.20160236

Hoxby, C. (2000). Peer effects in the classroom: Learning from gender and race variation. *NBER Working Paper 7867*. https://doi.org/10.3386/w7867

Islam, A., Vlassopoulos, M., Zenou, Y., & Zhang, X. (2021). Centrality-based spillover effects. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3880728

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*(4), 85–108. https://doi.org/10.1257/app.1.4.85

Jaravel, X., Petkova, N., & Bell, A. (2018). Team-specific capital and innovation [Publisher: American Economic Association]. *The American Economic Review*, *108*(4), 1034–1073. Retrieved September 23, 2024, from https://www.jstor.org/stable/26527997

Kusumawardhani, P. N. (2022). Spillover effects of investment in early childhood education and development (ECED) centers: Evidence from indonesia. *Education Economics*, *30*(6), 590–611. https://doi.org/10.1080/09645292.2021.2019196

Laitner, J. (1999). Structural change and economic g rowth. *R EVIEW OF ECON OM IC STU D IES*.

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*(559), 208–237. https://doi.org/10.1111/j.1468-0297.2011.02463.x

Lazear, E. P. (1989). Pay equality and industrial politics [Publisher: University of Chicago Press]. *Journal of Political Economy*, *97*(3), 561–580. Retrieved September 30, 2024, from https://www.jstor.org/stable/1830455

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*(3), 355–365. https://doi.org/10.1016/j.lindif.2008.10.009

Levin, D. Z., & Cross, R. (2004). The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer. *Management Science*, *50*(11), 1477–1490. https://doi.org/10.1287/mnsc.1030.0136

List, J., Momeni, F., Vlassopoulos, M., & Zenou, Y. (2023). Neighborhood spillover effects of early childhood interventions. *Centre for Economic Policy Research*.

Londoño de la Cuesta, J. L. (1996). *Poverty, inequality, and human capital development in latin america, 1950-2025*. World Bank.

Ma, Y., Fairlie, R., Loyalka, P., & Rozelle, S. (2024). Isolating the "tech" from EdTech: Experimental evidence on computer-assisted learning in china. *Economic Development and Cultural Change*, *72*(4), 1923–1962. https://doi.org/10.1086/726064

Malouff, J. M., Emmerton, A. J., & Schutte, N. S. (2013). The risk of a halo bias as a reason to keep students anonymous during grading [Publisher: SAGE Publications Inc]. *Teaching of Psychology*, *40*(3), 233–237. https://doi.org/10.1177/0098628313487425

Manski, C. F. (1995). *Estimation problems in the social sciences*. Harvard University Press, Cambridge, Massachusetts.

Marmaros, D., & Sacerdote, B. (2006). How do friendships form? *The Quarterly Journal of Economics*, *121*(1), 79–119.

Ngai, L. R., & Pissarides, C. A. (2024). Structural change in a multisector model of growth.

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *35*(4).

Noriega-Campero, A., Almaatouq, A., Krafft, P., Alotaibi, A., Moussaïd, M., & Pentland, A. (2018). The wisdom of the network: How adaptive networks promote collective intelligence. https://doi.org/10.48550/arXiv.1805.04766

OECD. (2014). *PISA 2012 technical report*. www.oecd.org/pisa/pisaproducts/PISA%2020%2012%20Technical%20Report_Chapter%203.pdf

Pedersen, D. E., Swenberger, J., & Moes, K. E. (2017). School spillover and college student health. *Sociological Inquiry*, *87*(3), 524–546. https://doi.org/10.1111/soin.12161

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

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

Rivera, M. T., Soderstrom, S. B., & Uzzi, B. (2010). Dynamics of dyads in social networks: Assortative, relational, and proximity mechanisms. *Annual Review of Sociology*, *36*(1), 91–115. https://doi.org/10.1146/annurev.soc.34.040507.134743

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory : Basic psychological needs in motivation, development, and wellness*. Guilford Publications. https://doi.org/10.1521/978.14625/28806.

Simons, J., Vansteenkiste, M., Lens, W., & Lacante, M. (2004). Placing motivation and future time perspective theory in a temporal perspective. *Educational Psychology Review*, *16*(2), 121–139. https://doi.org/10.1023/B:EDPR.0000026609.94841.2f

Thien, L. M., Darmawan, I. G. N., & 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), 3. https://doi.org/10.1186/s40536-015-0013-z

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of applied psychology*, *4*(1), 25–29.

UNESCO. (2019). *United nations educational, scientific and cultural organization: China.* http://uis.unesco.org/en/country/cn

Vasilaky, K. N., & Leonard, K. L. (2018). As good as the networks they keep? improving outcomes through weak ties in rural uganda. *The University of Chicago*, *66*(4). https://doi.org/https://doi.org/10.1086/697430

Wang, Y., & Li, H. (2017). An empirical study on the impact of educational gap on income gap. *Proceedings of the 2017 International Conference on Education Science and Economic Management (ICESEM 2017)*. https://doi.org/10.2991/icesem-17.2017.127

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research*, *13*(1), 21–39. https://doi.org/10.1016/0883-0355(89)90014-1

Wei, Y., Zhu, L., & Ji, C. (2022, July). *Analysis report on disparities in per capita expenditure on compulsory education between urban and rural areas in china*. https://ciefr.pku.edu.cn/cbw/kyjb/a4d8114b72d24d9c9fbdaa4d11ab8cd8.htm

Wu, Y. (2024). Disparities in primary and secondary education for girls in rural china. *Frontiers in Educational Research*, *7*(1). https://doi.org/10.25236/FER.2024.070125

Yang, J., Huang, X., & Liu, X. (2014). An analysis of education inequality in china. *International Journal of Educational Development*, *37*, 2–10. https://doi.org/10.1016/j.ijedudev.2014.03.002

Zell, E., & Alicke, M. D. (2010). The local dominance effect in self-evaluation: Evidence and explanations. *Personality and Social Psychology Review*, *14*(4), 368–384. https://doi.org/10.1177/1088868310366144