Title
Education Resources and Mental Health in the Developing World

ISBN
9798290614410

Author
Jiang, Jiting

Publication Date
2025-06-15

Peer reviewed|Thesis/dissertation

Education Resources and Mental Health in the Developing World

By

JITING JIANG
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Agricultural & Resource Economics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

____

Travis Lybbert, Chair

____

Steve Boucher

____

Jenna Stearns

____

Scott Rozelle

Committee in Charge

2025

Contents

List of Figures

List of Tables

Dissertation Abstract

Education plays a pivotal role in shaping both individual and societal growth, yet evaluations of educational investments have largely overlooked mental health, a critical dimension of human capital. This dissertation includes three empirical chapters that investigate how educational resources relate to mental health across key educational stages—primary school, junior high, and college—in developing country contexts, focusing specifically on China.

In the first empirical chapter, I examine the effects of providing high-quality, in-class libraries through a school-level randomized control trial (RCT) in rural Chinese primary schools. The intervention significantly reduces attention and behavioral risks, specifically symptoms of Attention Deficit Hyperactivity Disorder (ADHD), by about 1.8 percentage points, a 30 percent reduction. Low-performing students at baseline benefit most from the intervention. Interestingly, the protective treatment effects are less attributable to direct improvements in reading habits or learning activities, and may be more likely from indirect gains such as improved classroom environments and shifts in general perceptions and behaviors of students and teachers.

The second empirical chapter evaluates the impact on mental health outcomes in adulthood of China's rapid, large-scale expansion of college access (HEE) in the late 1990s. Using a cohort difference-in-differences approach, I find that greater access to higher education significantly increases college attainment, but it results in lower levels of mental well-being in adulthood—measured by self reported depression symptoms. These impacts are more pronounced among those from lower socioeconomic backgrounds and those who gained access because of HEE. A closer examination of potential mechanisms suggests that these effects are not driven by college education itself nor by pre-existing differences unrelated to college. Rather, the psychological costs for this group likely result from a combination of complex factors, including challenges faced during college that have lasting effects and post-graduation difficulties that fall short of the elevated aspirations tied to earning a college degree.

Finally, in my last chapter, I explore the link between visual acuity and multidimensional mental health among junior high school students in rural China. Using survey data and standardized assessments, I find that uncorrected visual impairment is significantly correlated with poorer mental health, as measured by the Strengths and Difficulties Questionnaire. Conversely, access to and

regular use of eyeglasses among visually impaired students is associated with better mental health, highlighting the potential of simple health interventions, such as vision correction, to support student mental health in low-resource settings.

Collectively, this dissertation advances our understanding of the relationships between educational investments and mental health in developing contexts. By demonstrating how different educational inputs across various developmental stages can significantly shape mental health outcomes—often distinctly from traditional academic measures—it underscores the importance of incorporating mental health considerations into educational policy and research. Ultimately, recognizing these subtle and sometimes unintended consequences is crucial for designing educational interventions that foster holistic human capital development, especially among disadvantaged populations.

Acknowledgment

I am deeply grateful to all the people who have supported me throughout my Ph.D. journey. First and foremost, I would like to thank my dissertation committee. I am especially thankful to Travis Lybbert for his outstanding mentorship and support. From brainstorming ideas to providing feedback on countless drafts, he has always offered thoughtful guidance, clear communication, and continued encouragement when I needed it most. The space he gave me to explore different paths was truly invaluable. I am also grateful to Scott Rozelle for generously providing access to key resources—data, experiments, collaboration opportunities with the Stanford team at SCCEI, and guidance on writing and publishing. His support played a crucial role in shaping this dissertation. Additionally, I thank Jenna Stearns who has been an inspiration in what it means to be a strong empirical researcher. Her high standards and to-the-point feedback early in the process helped sharpen my thinking and improved the rigor of my work. Thank you as well to Steve Boucher for his kind encouragement and helpful comments on my prospectus and dissertation drafts.

To my family—thank you for being my foundation. Teng Bian, my dear husband, thank you for your unconditional love and unwavering support through all the highs and lows. Thank you for always believing in me and bring out the best in me. To my daughter Emily—thank you for the joy and light you bring into my life every single day. And to my parents and my aunt, thank you for your love, trust, and steady support from afar. I couldn't have done this without you.

I'm also fortunate to have had a community of peers and friends who made this journey more collaborative, insightful, and joyful. Huan and Yiwei—thank you for the research collaboration and for modeling good research practices. Carol—thank you for the friendship and for being someone whose insights and clarity I've always appreciated. Eunju—thank you for our regular check-ins and mutual encouragement that helped me stay on track. And Shanchao—thank you for guiding me through the job search and sharing your experiences. I am also grateful to the department and the staff at Davis for always being patient, helpful, and supportive.

Lastly, I would like to acknowledge those who supported me before I joined UC Davis. I am especially thankful to my professors at Tufts University—Kyle Emerick, Jeff Zabel, and Ujjayant Chakravorty—for inspiring me to pursue a Ph.D. I am also grateful to Qingsong He from my undergraduate years for his early support and belief in my potential.

Chapter 1

Introduction

Education is fundamental to both individual development and societal advancement. For individuals, education not only promotes employment and earnings, but also contributes to health and other non-labor market outcomes (Card, 1999; Lochner and Moretti, 2004; Grossman, 2006; Oreopoulos and Salvanes, 2011). For societies, education fosters a more skilled workforce, reduces poverty, and drives long-term economic growth and social mobility (Barro, 2001; Hanushek and Woessmann, 2008; Chetty et al., 2014). Recognizing these substantial returns beliefs, countries worldwide-both developed and developing-have invested heavily in expanding access and improving learning (Kremer, Brannen and Glennerster, 2013; UNESCO, 2020). Despite this progress, identifying and implementing effective, evidence-based educational investments remains crucial, particularly in developing contexts where resources are constrained and educational disparities persist.

Traditional evaluations of educational investment/interventions have focused on academic performance, labor market success, and physical health, while neglecting a critical dimension of human development: mental health. Mental health encompasses emotional, psychological, and social wellbeing, influencing every parts of an individual's life. Mental health problems have become increasingly pervasive and costly globally (WHO, 2010, 2022*a*,*b*). Additionally, mental health often unfolds through complex and nuanced pathways (Currie and Stabile, 2006; Bleakley, 2010; Kieling et al., 2011; Fazel et al., 2014*a*; Kinderman et al., 2015; Karunamuni, Imayama and Goonetilleke, 2021). These facts underscore the need to incorporate mental health into evaluations of educational initiatives. Effective interventions should not only prioritize academic and physical outcomes but also address diverse mental health needs across the entire educational spectrum from early childhood education through higher education. Each stage presents its own developmental requirements and institutional complexities that must be considered for a more holistic understanding of educational impacts.

This dissertation thus investigates an overarching yet underexplored question: what is the relationship between educational resources and mental health in developing contexts? More specifically, it explores several critical sub-questions. Are certain types of educational inputs particularly effective in shaping mental health outcomes? Are these inputs the same as those that affect academic performance such as test scores? Are particular subgroups of individuals more likely to benefit from specific educational interventions? Could widely accepted inputs—such as increased access to education—carry underappreciated short- or long-term costs to mental health? How might these relationships vary across different educational stages, from primary school to college? And finally, do effects differ depending on the dimension of mental health considered, such as attention and behavior as well as depression?

The remainder of this dissertation consists of three empirical chapters, each chapter focusing on a distinct educational stage and investigating how educational resources relate to mental health outcomes. The analyses employ a range of methodological approaches, including causal inference with observation data, field-based randomized control trials (RCT), and correlational analysis. I also use multiple data sources, including publicly available social surveys, original survey data from field experiments, standardized assessments, population census data as well as national educational yearbooks.

In chapter two, "The Impact of In-Class Libraries on Students' Attention and Behavior: Evidence from Chinese Rural Primary Schools", I study the causal relationship between reading resources provision and student mental health—specifically, the risks of Attention Deficit Hyperactivity Disorder (ADHD). The intervention is a recent school-level RCT that provided curated, in-class libraries with independent reading materials to fourth and fifth graders in rural China. The analysis uses two rounds of survey data collected by the field team, including rich demographic and performance information on students, their primary caregivers, and teachers. I find that, on average, in-class libraries significantly reduced the rate of ADHD among primary school students in the sample by about 1.8 percentage points, a 30 percent reduction at a 1% statistical significance level. Low-performing students at baseline benefit most from the intervention. Notably, these findings are in contrast to the commonly reported null effects of similar interventions on academic outcomes like test scores, emphasizing the importance of considering mental health outcomes as a key dimension in the evaluation of educational interventions.

The protective treatment effects on ADHD risks, interestingly, seem less attributable to direct improvements in reading habits or learning activities among students, teachers, parents, and peers. Instead, they may be linked to indirect improvements, such as enhancements to classroom environments and changes in the general perceptions and behaviors of both students and teachers, supported by the social psychology literature (McTaggart, 2009; Niemiec and Ryan, 2009; Morsink et al., 2022). Such classroom-based interventions may offer a promising and culturally acceptable alternative to individual-level diagnosis and treatment for mental health in resource-constrained settings like rural China, where mental health issues are often overlooked, misunderstood, and stigmatized.

Chapter three, titled "The Effect of Higher Education Expansion on Mental Health: Evidence from China" looks at college education. It examines the relationship between increased access to college, as a special form of educational resource, and long-term mental health outcomes. Leveraging the large-scale and rapid expansion of higher education (HEE) in China that began in 1999, I identify causal effects of HEE on mental health by employing cohort difference-in-differences (DID), a quasi-experimental design. In this study, mental health is measured using self-reported survey data, focusing on recent symptoms of depression and subjective well-being. I find that increased college access significantly raises educational attainment. But it also has unintended consequences for mental health that persist well into an individual's 30s—roughly a decade after their expected college graduation. These effects are particularly pronounced among those from lower socioeconomic backgrounds and are primary driven by compliers—those who attended college as a result of the expansion.

Why might expanded access to higher education lead to less favorable mental health for individuals in their 30s? I explore several different explanations. The findings suggest that these observed effects are unlikely to be driven by difficulties in completing college, by pre-existing differences in mental health profiles unrelated to college, or by college education being inherently detrimental to well-being. Instead, a more plausible explanation, though not directly testable with the available data, is that individuals newly gaining access to college through expansion may experience greater psychological challenges during and after college, driven by a complex interplay of factors. For example, many may have entered college with weaker academic preparation and limited family support, while also facing greater exposure to stressors such as peer comparison—challenges that can make it harder to navigate college and social experiences, with mental health effects persisting into later life. They might also obtain relatively modest labor market returns and encounter additional barriers even after having earned a college degree, falling short of the elevated aspirations often associated with higher education. For example, many compliers were funneled into second- or third-tier colleges, where lower educational quality and weaker labor market signaling can result in less favorable employment outcomes. In such, unmet expectations may lead to feelings of inadequacy, dissatisfaction, or continual pressure to succeed, all of which can contribute to longer-term stress and reduced well-being.

The final chapter of my dissertation, "The Association Between Visual Impairment, Educational Outcomes, and Mental Health: Insights from Eyeglasses Usage among Junior High School Students in Rural China", focuses on the junior high school stage and investigates the links between visual impairment, eyeglasses usage and mental health. To assess mental health in a multidimensional way, students complete the official Chinese simplified version of the Strengths and Difficulties Questionnaire (SDQ). Approximately 9.3% of students in the sample are classified as high risk for mental health problems. The analysis presents a significant positive association between impaired vision and poorer mental health, and eyeglasses usage is associated with better mental health among visually impaired students. These findings suggest providing vision correction resources may have meaningful effects beyond academic performance, extending to students' psychological well-being.

Overall, this dissertation contributes to a deeper understanding of the relationship between educational resources and mental health in developing contexts. It explores different types of educational resources and multiple dimensions of mental health across three distinct educational stages, investigating both short-term and long-term effects using a range of empirical methods and diverse data sources. Across all chapters, I show educational inputs play an important role in shaping individuals' mental well-being over time, especially for vulnerable populations (such as low-performing students, those from poor socioeconomic backgrounds, and students with visual impairments). The findings also have important policy implications. They highlight the need to recognize mental health as a key outcome of educational investment, alongside more commonly studied metrics such as test scores, since educational resources may affect mental health in ways that differ from their impact on cognitive outcomes. Moreover, well-targeted educational interventions—especially those designed to meet the specific needs of disadvantaged students—may be particularly effective in supporting children's holistic development in low-resource contexts.

However, the nature and magnitude of these relationships are nuanced and context-dependent. The results vary across different population groups, depend on what types of resources are provided, how and when they are delivered, and differs by the specific dimensions and timeframes of mental health being considered. Explaining these patterns is complex, as they likely involves subtle pathways such as changes in classroom environments, peer dynamics, academic and social preparedness, and labor market experiences. Future research should build on this work by incorporating diagnostic mental health measures, expanding assessment to include additional dimensions such as emotional resilience, and employing rich datasets alongside rigorous research designs tailored to mental health outcomes and mechanisms. Replication across diverse settings is also essential to strengthen external validity and ensure policy relevance. More broadly, this research calls for greater recognition of the relationship between education and mental health in both research and policy conversations. Recognizing the subtle, multifaceted, and sometimes unintended consequences of educational investments is critical to designing policies that support holistic human capital development.

Chapter 2

The Impact of In-Class Libraries on Students' Attention and Behavior: Evidence from Chinese Rural Primary Schools

2.1 Introduction

Children spend a substantial amount of their time in school, making it an essential setting for educational initiatives to improve learning outcomes. Several types of educational interventions are shown to be effective in improving test scores in developing countries. Examples include remediation education programs, computer-assisted learning initiatives, teacher incentive projects, and school management reforms (Banerjee et al., 2007; Glewwe, Kremer and Moulin, 2009; Barrera-Osorio and Linden, 2009; Muralidharan and Sundararaman, 2011; Duflo, Hanna and Ryan, 2012; Pradhan et al., 2014).

However, providing additional resources to the school systems that are in need—a common type of educational intervention—has generally shown limited effectiveness in improving student's test scores and cognitive learning outcomes (Kremer, Brannen and Glennerster, 2013; Ganimian and Murnane, 2016). Neither teacher nor non-teacher inputs alone, despite low levels of resources, appear to improve test scores across different geographic areas (Glewwe et al., 2004; Glewwe, Kremer and Moulin, 2009; Duflo, Dupas and Kremer, 2015). Research on reading resource provision projects also often reports null results (Borkum, He and Linden, 2012; Abeberese, Kumler and Linden, 2014). In the context of China, Yi et al. (2019) studied the one-year treatment effects of an in-class library project and showed little effects on reading and academic achievement, consistent with prior studies[1].

While substantial evidence highlights the limited academic impacts of resource-provision educational programs (as mentioned above), much less is known about their effectiveness in promoting children's well-being, an equally crucial aspect of their overall development beyond academic outcomes. Would these programs share similar null results for mental health? Not necessarily. While academic performance and mental health are connected, mental health represents a distinct dimension of human capital development. It encompasses emotional, psychological, and social well-being, and often unfolds through complex and nuanced pathways (Currie and Stabile, 2006; Bleakley, 2010; Kieling et al., 2011; Fazel et al., 2014*a*; Kinderman et al., 2015; Karunamuni, Imayama and Goonetilleke, 2021). The distinctive nature of mental health underscores the need to evaluate educational interventions through a broader lens that accounts for their potential impacts on students' well-being, not just their academic achievements. For example, programs such as the in-class library project Yi et al. (2019) studied may have unintended protective effects on children's mental health, even if they do not yield significant academic benefits.

In this paper, I examined the effects of an in-class library intervention similar to the one studied in Yi et al. (2019), which provided in-need independent reading resources for fourth and fifth graders in rural Jiangxi, China. However, the intervention I evaluated was implemented in different schools and during different years than her study (see Sections 2 and 3 for details). Additionally, I focused on its impact on attention and behavior, specifically the risks of Attention Deficit Hyperactivity Disorder (ADHD) symptoms[2].

ADHD, one of the most common mental health challenges worldwide, especially among children aged between 3-17, poses particular difficulties for children in the school environment (Daley and

[1]The intervention that Yi et al. (2019) studied was implemented in 2016 and provided curated independent reading resources to primary school students in approximately 40 rural primary schools in Jiangxi Province in China.

[2]Note that, the impacts of in-class libraries on students' attention and behavior were not specifically pre-registered and should therefore be interpreted as exploratory.

Birchwood, 2010; Rogers and Tannock, 2018; Morsink et al., 2017; Smith et al., 2020)[3]. It is often characterized by difficulty paying attention, hyperactivity, and impulsive behavior, which could seriously interfere with students' everyday behaviors and learning performance. Previous literature has shown that children with ADHD are not only likely to continue exhibiting symptoms of ADHD into adolescence, but they also tend to experience other disorders simultaneously, like anxiety and reading difficulties (Mannuzza and Klein, 2000; Sexton et al., 2011; Bitsko, 2022). Many long-term consequences of ADHD are also discussed, including low schooling attainment, continuing mental health problems, and poor labor market outcomes later in life (Currie and Stabile, 2006; Frazier et al., 2007; Fletcher and Wolfe, 2008; Erskine et al., 2016). The prevalence and consequences of ADHD are even more severe in developing countries like China (Patel et al., 2008; Fazel et al., 2014*a*; Danielson et al., 2018; Lola et al., 2019; Pang et al., 2021).

Although in-class libraries might not initially appear to address attention and behavioral challenges, they hold significant theoretical and practical relevance. Theoretically, in social psychology, the classroom is often viewed as an effective intervention medium for students at risk of or diagnosed with ADHD (Cordell and Cannon, 1985; MacAulay, 1990; Daley and Birchwood, 2010; Rogers and Tannock, 2018; Morsink et al., 2022). For example, a simple in-class library may create a more stimulating and supportive learning environment that meets students' fundamental needs for autonomy and relatedness, fostering improved attention, behavior, and overall well-being (McTaggart, 2009; Niemiec and Ryan, 2009; Morsink et al., 2022). Practically, classroom-level interventions are especially important in resource-constrained contexts like rural China, where mental health challenges are prevalent and often exacerbated by pervasive cultural stigma that frames mental health issues as moral failings and discourages direct, individual-level diagnosis and treatment (Patel et al., 2008; Hong et al., 2011; Fazel et al., 2014*b*). In contrast, indirect school-level interventions, like the in-class library project, could offer distinct advantages by providing a non-invasive, cost-effective, and scalable way to support children's development (Lee, 2009; Jordans et al., 2013; Fazel et al., 2014*a*; Morsink et al., 2022; Kieling et al., 2011; Fazel et al., 2014*b*; Pang et al., 2021; Jiang et al., 2022).

To empirically test whether in-class libraries affect student attention and behavior, particularly

[3] For example, according to the Centers for Disease Control and Prevention (CDC), there were 6 million (9.8%) children aged 3-17 years in the US ever diagnosed with ADHD using data from 2016 to 2019. Boys are twice more likely to be diagnosed with ADHD than girls. Though with varied rates, ADHD is prevalent across the world.

regarding ADHD symptoms, I used the home version of the ADHD rating scale-IV (ADHD RS-IV) to assess ADHD symptoms. The scale, reliable and valid for evaluating ADHD in China (Su et al., 2015), involved primary caregivers rating the frequency of 18 ADHD symptoms over the past six months. Specifically, each item was scored on a scale of 0 (never or rarely), 1 (sometimes), 2 (often) or 3 (very often). The total score was calculated by summing the scores over the 18 items. Higher scores indicate higher risks for poor attention and behavior outcomes. I then used the cutoff point score of 26 on the total score to identify students at risk of ADHD, following previous literature validated in the Chinese context (Su et al., 2015)[4]. In other words, students with total scores above 26 were considered at risk of ADHD. Based on this criterion, the rate of ADHD in my sample was around 6%.

The in-class library treatment specifically involved setting up free in-class libraries for fourth and fifth graders in the treatment schools, with each classroom equipped with over 70 carefully selected independent reading books[5]. The treatment assignment was at the school level across three counties in southern Jiangxi province, China, during the 2017-2018 academic year. Specifically, around 3000 students from 84 classes in 39 schools were in the control group with no in-class library provided, while about 4000 students from 98 classes in 38 schools were assigned to receive the treatment with an in-class library. Access to high-quality independent reading resources is particularly meaningful for rural students in China, who often lack adequate materials and face challenges with reading skills, behaviors, and attitudes. Two rounds of survey data were collected: the baseline survey was implemented in May 2017 before the treatment started, and the follow-up survey was conducted about eight months after the treatment implementation in May 2018. The surveys included detailed demographic and performance information on students, primary caregivers, and subjective teachers (for additional details, refer to the Data section).

On average, in-class libraries significantly reduced the rate of ADHD among primary school students in the sample by about 1.8 percentage points, a 30 percent reduction at a 1% statistical significance level[6]. Importantly, low-performing students in the baseline reading comprehension test

[4] I also used different ADHD at-risk cutoffs for robustness checks.

[5] The independent books were carefully selected by Chinese education experts to make sure they were age-appropriate, covering a broad range of topics outside the school curriculum. Besides the book provision, no further instructions or training were in place to motivate teachers or hold them accountable for using these resources, but the project did require teachers and/or students to manage the books after the receipt.

[6] Considering potential treatment noncompliance, I also estimate the local average treatment effect (LATE) using an instrumental variable (IV) framework. The LATE findings align with the main intention-to-treatment (ITT)

seemed to benefit most from the intervention, with 6.7 percentage points lower ADHD rate than relatively high-performing students[7]. A data-driven approach using random causal forests was also employed for the heterogeneity analysis, yielding consistent insights. Multiple testing corrections were conducted when necessary (see the Results section for more details).

Based on the positive impacts of the intervention on ADHD risks among primary school students in rural China, I further investigated the mechanisms through which in-class libraries helped prevent developing impairments in attention and behavior. Specifically, I considered the possible reading-related roles of the student (own), teacher, parent, and peer-related factors through the intervention. The protective treatment effects on ADHD did not seem to work through students' more efficient usage of and learning from the reading resources on their own or with teachers, parents, and peers. I then discussed potential remaining mechanisms, focusing on environmental and indirect influences. For example, in-class libraries likely improved the classroom environment by meeting students' basic psychological needs such as autonomy and relatedness in a way that positively affected students' attention and behavior.

The paper, despite its exploratory nature, contributes to the literature on educational program evaluation by examining a widely implemented type of intervention that provides additional resources to the school system, focusing specifically on mental health outcomes. While existing evidence often centers on learning outcomes such as test scores, and frequently reports null effects, a notable research gap persists regarding the impact of these programs on children's well-being, an equally critical aspect of their overall development beyond academics. My findings demonstrate the potential for school-level reading-resource-provision interventions to significantly impact one prevalent mental health challenge, namely ADHD. This highlights the need for further research to evaluate and identify effective educational interventions, particularly those at the school and classroom levels, to support children facing mental health challenges, especially in developing regions. Moreover, this study suggests that the mechanisms driving these well-being effects may be indirect and environmental, rather than direct and active changes in reading-related behaviors. This insight contributes to the broader literature on social psychology by emphasizing the role of classroom environments in shaping student psychology and behavior (Fraser, 2012; MacAulay, 1990; Dorman,

results (see in the Results section).

[7] Low-performing students in reading are those with baseline reading test scores ranked in the bottom 20 percent. Other subgroup analyses, including gender, were also conducted but did not reveal statistically significant differences.

2001; Dörnyei and Muir, 2019; Cordell and Cannon, 1985; Daley and Birchwood, 2010; Rogers and Tannock, 2018; Morsink et al., 2022).

2.2 Institutional Context and Treatment Design

2.2.1 Institutional Context

Attention Deficit Hyperactivity Disorder (ADHD) is one of the most common and challenging child mental health disorders, which can have potentially long-lasting negative effects on children's lives, including academic performance, economic well-being, and overall life satisfaction (Currie and Stabile, 2006; Fletcher and Wolfe, 2008; DuPaul and Stoner, 2014; Thomas et al., 2015; Erskine et al., 2016). Rural places are reported to experience greater challenges in these aspects. Pang et al. (2021) documented that, in rural China, the prevalence of ADHD is about 7.5% among rural school children, with even higher rates among students with lower grade levels and cognitive ability, compared to the worldwide ADHD prevalence of 7.2%. Several unique factors in rural China and similar developing regions may exacerbate attention and behavioral problems. These include the high-pressure educational system with its high-stakes exams, extensive rural-to-urban migration leading to the "left-behind" children phenomenon, and gender discrimination manifesting as a preference for sons (Jiang et al., 2022).

Despite the well-recognized role of adequate reading ability for academic achievement and long-term well-being worldwide, the reading performance, behaviors, and attitudes of school children in rural China are very poor. According to Gao et al. (2021), compared to students from other countries that participated in the Progress in International Reading Literacy Study (PIRLS), the reading achievement of rural Chinese students (based on a representative sample) is among the lowest. Students also show low levels of independent reading quantity and reading confidence in the absence of any treatment (Gao et al., 2018)[8]. Partly due to the stringent school and examination system in China, students, parents, and even teachers commonly hold negative attitudes toward independent reading, emphasizing that all attention and efforts should be placed on major subjects like Chinese language and mathematics to optimize chances of achieving high scores (Yi et al.,

[8]Based on my sample, rural students at the baseline spent little time reading (with less than 40% of students reading for 30 minutes at home daily), did not enjoy reading (about 7%), and lacked confidence in their ability to read (about 14%).

2019)[9].

Several key factors may be contributing to the current status quo. First, there is a lack of adequate reading resources that are suitable, interesting, and readily available for school children at school or at home[10]. Though most rural schools do appear to have school libraries built, especially after the governmental library support program since 2003 (the Ministry of Education of China, 2003), students tend to report that many school libraries are often locked, do not allow borrowing books for home reading, and lack interesting books appropriate for their ages (Wang, 2012). Second, when reading resources are provided, pushback often comes from schools, teachers and parents, likely due to biased beliefs about independent reading and a lack of incentive to use them (Wang et al., 2020). For example, even with a nationwide education initiative from the central government to promote reading literacy and independent reading among rural school children in place, independent reading resources are not always welcomed either in school or at home. Teachers and parents tend to believe that reading outside the curriculum is useless and distracts students from getting high academic scores. Third, in China, reading programs and library facilities are not viewed as central components of the education system. Teachers, in most cases, are not evaluated on promoting students' reading in any way, so they typically lack incentives to make full use of the provided reading resources or systematically teach how to read in school. Students may simply not have the reading skills to make sense of available reading resources.

2.2.2 Treatment Design

To understand whether providing an in-class library could protect youth from developing impairments in attention and behaviors, I leveraged an RCT study in three rural counties in Jiangxi province in China. The intervention, implemented by a Chinese nongovernmental organization (NGO) together with the educational bureau, provided an in-class library free of charge to all selected classes of grades 4 and 5 in about half of the primary schools in September 2017, at the start of the school year of 2017. The other half of the schools did not receive treatment but were

[9] For example, in my baseline sample, about 65% (75%) of students (parents) perceived no positive effect of independent reading, and about 7% (13%) of students (parents) perceived strictly negative effects of independent reading.

[10] According to my baseline sample, about 85% of students reported having a school library, while only about 45% (10%) students said that they had a classroom library (over 25 books at home). Note that the classroom libraries reported in the baseline usually did not have the high quality as the intervention provided.

otherwise similar (see Figure 2.1)[11].

The freely provided in-class library came with one sturdy bookshelf made of solid wood and was stocked with the same set of 70+ independent reading books for school children. These age-appropriate high-quality books were carefully selected by a panel of Chinese education experts[12], and were intended to cover a broad range of reputable titles beyond what was normally taught in the school curriculum with varied difficulty levels, such as the winning entries for children's literature and natural science. Once installed in the treated schools, the in-class libraries belonged to the school, and the teachers and/or students were responsible for managing them. A local coordinator would help inspect and replace books that were found damaged or lost every semester. Beyond the provision of hardware and books, no additional instructions or training was provided to schools. In addition, there were no explicit requirements for school principals or teachers to promote the usage of the library, unless they felt like doing so.

2.3 Randomization, Data, and Balance

2.3.1 Randomization

Randomization was carried out at the school level within each of the three counties in the southern part of Jiangxi province in China[13]. To sample the schools, an official list of all rural schools (in total 458 schools) from the counties was first obtained from county education bureaus, and a sample of 120 schools was then randomly selected based on the number of schools in each county, ensuring the sample reflected the overall school distribution across the three counties[14]. 43 out of the 120 schools were later dropped since they had already been provided with free in-class libraries in the previous years, according to the intervention roll-out plan. Of the remaining 77

[11] Although the treatment group has 38 schools and the control group has 39 schools, there are quite some differences in the number of classes (98 in the treatment group vs 84 in the control group) and the number of students (3784 in the treatment group vs 2861 in the control group). The main regression results are robust regardless of whether the class size is included as a control variable (see Appendix Table 2.28).

[12] The final book list was selected by a specialist group that consisted of more than 20 educators, writers, publishers, and librarians.

[13] The three counties in Jiangxi have been nationally designated as poverty counties since 2012, and more than 80% of the population are rural residents (National Bureau of Statistics of China, 2016). Educational resources in the sample counties are also very poor. For example, there are far from an adequate number of primary schools in place, and many rural children have to travel long distances to school on a daily basis. The teacher-to-student ratio is at the magnitude of 1 to 20 (National Bureau of Statistics of China, 2016).

[14] Rural schools that had fewer than 100 students in total were not included in the official list.

schools that were never treated before, 38 schools (49.35%) are in County A, 24 schools (31.17%) are in County B, and 15 schools (19.48%) are in County C. Within each county, about half of the schools were then assigned to the treatment group with free in-class library provision and the other half to the control group with no free in-class library, right before the start of the 2017 school year (August 2017). In each of these selected schools, at most two classes from both the fourth and fifth grades were randomly selected to receive the in-class libraries (due to financial constraints)[15]. All students from the sampled classes were then included. The final sample consists of 6645 students of grades 4 and 5 from 182 classes in 77 schools.

2.3.2 Data

The primary data sources are two rounds of surveys conducted by the field team. The baseline survey was collected before the treatment was assigned (May 2017), and the follow-up survey was around eight months after the treatment implementation (May 2018). Specifically, the surveys included all sample students, primary caregivers, and subjective teachers to collect their basic background information and key performance indicators.

To the sampled students at baseline, the research team administered a standardized 30-minute reading comprehension test, a widely used IQ test, as well as a detailed student survey questionnaire. The reading comprehension test used questions from the 2011 version of Progress in International Reading Literacy Study (PIRLS), which is an international assessment of reading comprehension in the fourth grade and has been conducted every five years since 2001 with about 50 countries worldwide (Mullis et al., 2012). Closely timed and proctored by the enumerators, the test had three short passages to read, each followed by multiple choice questions and writing questions to fill out. The wording was carefully translated into Chinese according to the PIRLS translation guidelines by the team and reviewed by a panel of experts and teachers. I normalized the reading test scores using the grade-specific score distributions. The IQ test was also administered using the Raven's Standard Progressive Matrices for children (Raven IQ test), as a measure of cognitive ability (Raven, 1941; Borghans et al., 2016)[16].

[15] If there are one or two classes in each grade, then all classes would be included; while if more than two classes exist in each grade, two classes would be randomly selected into the sample.

[16] The Raven IQ test is a widely used nonverbal IQ test evaluating general human intelligence and abstract reasoning, which comprises 5 sets of 12 items each. A final IQ score was assigned from the scores on these 60 questions according to an established norm; a score lower than 85 indicates low cognitive ability (Pang et al., 2021).

Lastly, a detailed student survey questionnaire was used to collect the following four aspects of information about students: (1) individual demographic characteristics and family background, (2) student self-reported reading resources and usage at school, classroom and home, as well as (3) student reading interests, confidence, perceptions and behaviors. Table 2.1 presents the main measures related to these aspects. In particular, the student survey questionnaire included two internationally established metrics from 2011 PIRLS called the Student Like Reading Scale (SLR) and the Student Confident in Reading Scale (SCR) with proper translation to measure students' reading interests and confidence in rural China context. The Student Like Reading Scale (SLR) consists of 8 statements and 24 possible points on a four-point Likert scale where 0 = disagree a lot, 1 = disagree a little, 2 = agree a little, and 3 = agree a lot[17]. I later calculated the total scores by summing each question's scores and categorized students whose total scores were below 11 as "Do not like reading", following the official guide. Similarly, the Student Confident in Reading Scale (SCR) comprises 7 statements (in total 21 points) on the same four-point Likert scale, and the suggested score cutoff for "Do not feel confident in reading" is set at 10. The Appendix Table 2.10 and Table 2.12 provide more details of the two standard reading scales.

The baseline questionnaire was also directly administered to the Chinese language subject teachers and asked about their demographics and perceptions towards independent student reading. For example, teachers were asked "what effect of independent reading do you perceive for students' language or math test performances?" and chose between "no effect", "negative effect", or "positive effect". A baseline primary caregiver survey questionnaire was also used to collect information on their basic demographics like educational attainment and occupations, family asset, their perception of independent reading on students' academic performance, as well as their general or reading-specific interactions with the children at home. Here, the primary caregiver is defined as the person at home who is most responsible for taking care of the student, usually the mother or paternal grandparent. They were asked to fill out the survey, which was brought home and returned the following day by students.

Around eight months after the intervention, the team revisited the sampled classes and conducted a follow-up survey, in a similar form as the baseline survey but with several changes. Importantly, one key module on attention and behavior was newly added to the follow-up survey, which I

[17]To keep the scores consistent, I also reversely coded some questions as instructed.

later used as key outcomes (see Appendix Table 2.8 for more details of the scale). Specifically, the home version of the Attention Deficit Hyperactivity Disorder rating scale-IV (ADHD RS-IV) was included in the primary caregiver questionnaire and asked the student's primary caregiver to rate the frequency of 18 ADHD symptoms over the past six months on a four-point Likert scale where 0 = rarely or never, 1 = sometimes, 2 = often, and 3 = very often. Higher total scores indicate higher risks for ADHD symptoms. Following Su et al. (2015), I define the student as at risk of ADHD if the sum of his or her ADHD scores is above the cutoff point of 26 out of 54 points. I show in the results section that the main findings are not sensible to the specific ADHD score cutoff I use. It is also important to note that this ADHD scale is used mainly to screen for ADHD, and is not used to make a clinical diagnosis of ADHD.

Also, note that there were minor revisions on the SLR and SCR scales used in the follow-up survey compared to the baseline survey. To highlight, the 2018 follow-up survey employed the 2016 PIRLS version, while the 2017 baseline survey used the 2011 version. The 2016 version of the SLR scale has 10 statements and a cutoff of 14 points, compared to 8 statements and a cutoff of 9 points in the 2011 version, while the 2016 version of SCR has 6 statements and a cutoff of 9 points, compared to 7 statements and a cutoff of 10 points in the 2011 version. The Appendix Table 2.11 and Table 2.13 provide further details of the two new versions of the SLR and SCR scales.

2.3.3 Balance, Attrition and Non-compliance

In this subsection, I present the balance test results in Table 2.1, which shows no statistically meaningful differences between the assigned treatment and control groups in terms of various baseline characteristics such as student characteristics, family background, as well as reading-related measures in resource availability, attitudes, behaviors and skills[18]. A very small group of students who were included in the baseline survey, for some reason, did not take the follow-up survey a year later and there is no evidence of a correlation between students' attrition and treatment status[19].

[18]It is interesting to learn that at baseline, approximately 45% of both treated and control schools had a library or reading room in the classroom. However, these pre-existing classroom libraries in rural Chinese schools are typically of poor quality, heavily focused on academic reading, and provide limited access. I later ran an interaction term model and found that the positive impacts on ADHD were not driven by the presence of these libraries before the intervention (see Appendix Table 2.27).

[19]Specifically, among the total sampled 6645 students from baseline, 443 students (about 6.7%) did not complete the second round of the survey. By treatment status, the attritted student share over one year was also very similar at 6.5% - 7% level (see Figure 2.1).

There also appears to be non-compliance to the treatment assignment in that some treated schools that should have installed in-class libraries somehow fail to do so, according to the students' self-reported availability of in-class libraries. Note that, the accounting of non-compliance relies on students' reports which can suffer from measurement errors and should be interpreted with caution[20]. Nevertheless, I assigned the actual intervention status the value of one if more than 50% of students within the same classroom report having an in-class library, which I call "built in-class library" and zero if otherwise. About 60% of the classrooms assigned with an in-class library were reported by students to have built an in-class library. Given the non-compliance, I later scaled the treatment effects to estimate the local average treatment effect (LATE) in addition to the ITT estimate.

2.4 Econometric Model

I estimated the following equation using the ordinary least squares regression (OLS):

$$Y_{isc} = \beta_0 + \beta_1 T_{sc} + \beta_2 X_{isc0} + \mu_c + \epsilon_{isc} \tag{2.1}$$

where Y_{isc} measures the outcome of interest for student i in school s from county c, T_{sc} is the binary treatment indicator showing whether the student is assigned to treatment, and X_{isc0} are baseline control variables including student and family characteristics[21]. County fixed effects, denoted as μ_c, were also controlled since the randomization was performed within county strata. β_1 is the main coefficient of interest, which measures the treatment's intention-to-treat (ITT) causal effects on the outcomes of interests, that is, the effect of assigning an in-class library on outcomes. To gauge the local average treatment effect (LATE), i.e., the treatment effect on compliers, I also scaled the ITT effects by treatment compliance (later shown in the results section) using an instrumental variables (IV) approach. As I will show, the treatment assignment variable correlates strongly with the actual treatment status. Standard errors were all clustered at the school level, the level of

[20] A better option of gauging the treatment compliance would be to have some objective observation data on classroom library status, such as whether there is an in-class library, whether the books are suitable for students, and whether students can easily access the books.

[21] I included the following set of baseline controls: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year and whether the student is from low-income families with no refrigerator at home.

treatment assignment.

For the main results, I ran the model on students' ADHD symptoms. I also employed other relevant outcomes in the analysis to consider the potential channels. The outcomes to test can be roughly grouped into four categories, including (1) the student's own behavioral and learning changes, (2) the teacher's responses, (3) the parent's activities, and (4) other students' interactions at school. I will explain more in the results section about these specific outcomes and why they are relevant. For the different outcomes I tested, I also performed multiple comparison adjustments using the Benjamini–Hochberg procedure, setting a false discovery rate of 5% (Benjamini and Hochberg, 1995).

Finally, I tested whether the effects of in-class libraries differed among subgroups of students, such as male students, left-behind students, and students with poor baseline reading skills, by employing interaction terms of the intervention and the subgroup indicators in the regression (with other terms remaining the same). Specifically, in the main tables, I defined students as left behind if both their parents migrated out for more than nine months this year. Students with poor baseline reading skills (or low-performing students) were those whose baseline reading comprehension test scores were ranked in the bottom 20 percent.

2.5 Results

This section presents the estimated treatment effects. I first showed the ITT and LATE results on ADHD. I then analyzed the heterogeneity of treatment effects, especially by students' baseline reading skills. The last subsection discussed the mechanisms by examining the roles of key relevant parties involved in the human capital production process[22].

2.5.1 Treatment Effects on ADHD

I first estimated the effect of in-class libraries on the rates of ADHD among students[23]. According to Table 2.2, I found that the intervention significantly reduced the rate of ADHD among primary school students in the sample by about 1.8 percentage points, a reduction of 30 percent, at a 1% statistical significance level (see panel A, column 2). After adjusting for multiple comparisons,

[22] These relevant parties include students, teachers, parents, peers, and classroom environment.

[23] Following Su et al. (2015), I used the ADHD score cutoff of 26 to define the student as at risk of ADHD.

the effects remained significant at 5%. The main results remained similar with or without student and household characteristics (see panel A, column 1)[24]. The LATE estimate was also consistent with the ITT estimate, either with or without baseline controls and county fixed effects (see panel B, columns 1 and 2). The first-stage results and F-stats shown in the last two rows in the table suggested that the random treatment assignment is strongly correlated with the actual intervention status (as expected)[25]. To test the robustness of the ADHD risk cutoffs I used, I ran a series of regressions, similar to those previously described, applying a range of score cutoffs from 1 to 53 (the total scale being 54 points). The estimates were shown to be robust in Appendix Figure 2.3[26].

Despite the share of students at risk of ADHD (what I am most interested in), I also presented both the ITT and LATE results for ADHD total scores and the two ADHD subscales scores, i.e., ADHD inattention scores and ADHD hyperactivity and impulsivity scores, in Table 2.2 from columns (3) to (8). The coefficients on main treatment indicators all appeared to be negative, which indicated that in-class libraries could help improve students' mental status, though not all estimates were statistically significant[27]. To provide some suggestive evidence of what aspects of ADHD were affected, I also analyzed each of the 18 ADHD symptom questions (shown in the Appendix Tables 2.16 and 2.17). For example, in Appendix Table 2.17, students were significantly less likely to leave seats in the classroom (column 2), act as if "driven by a motor" (column 5), or talk excessively (column 6). Note that, these results were exploratory and should be interpreted with caution, since they were not intended to be analyzed at the item level and the results no longer remained significant after adjusting for multiple comparisons after the Benjamini–Hochberg correction (Benjamini and Hochberg, 1995).

In addition to ADHD, I presented the parallel set of main results using a self-reported complete behavioral screening questionnaire called The Strengths and Difficulties Questionnaire (SDQ) in

[24]Without baseline controls and county FE, I found that the project significantly reduced the prevalence of ADHD among primary school students in the sample by about 1.9 percentage points, a reduction of 30 percent, at 1% statistical significance level (see Table 2.2, panel A, column 1).

[25]The first-stage results suggested that 65% of students were treated, and the F-stats was as large as 60, well beyond 10 (a rule of thumb).

[26]I also visualized the cumulative distribution function (CDF) of ADHD scores between treatment and control groups. As shown in Appendix Figure 2.4, I observed a clear leftward shift in the treatment groups, which implies a reduction in ADHD risks (with lower scores meaning better outcomes).

[27]Specifically, for estimates with no control added, in-class libraries seemed to reduce ADHD total scores by 5.6 points as well as ADHD hyperactivity and impulsivity scores by 2.9 points at 10% level, not for ADHD inattention scores. All estimates with the whole set of controls included were however no longer significant, though the ratio of coefficients and standard errors were relatively large.

the Appendix Table 2.18 through 2.20. The abnormal share of peer problems (one of the four SDQ subscales) seemed to be improved by the treatment but not for other dimensions, which are emotional symptoms, conduct problems, and hyperactivity/inattention, as well as the total scores (see Appendix Table 2.18, panels A and B, column 5 and columns 1 to 4)[28]. Together with the results for the rate of ADHD reported by caregivers, it could suggest that the intervention had a nuanced effect on aspects of behavior and mental health.

2.5.2 Treatment Effect Heterogeneity

I then studied treatment heterogeneity on the ADHD measures depending on student gender, left-behind status, and baseline reading skills. In panel A of Table 2.3, I found that in-class libraries could benefit low-performing students in reading at baseline regarding ADHD (row 4). Also, when compared with students with not very poor baseline reading comprehension test scores, these low-performing students (with scores ranked in the bottom 20% of the distribution) had larger improvements in attention and behavior as a result of the in-class library project (row 1). These patterns were shared across all four ADHD measures (columns 1 to 4), and all remained significant at the 5% level after multiple testing corrections[29]. Specifically, for low-performing students in reading at baseline, the total treatment effect in reducing the rate of ADHD was around 7.3 percentage points, 6.7 percentage points more gain than relatively high-performing students (column 1). Similarly, the intervention significantly reduced the ADHD total scores by about 2 points among students with poor baseline reading skills (column 2) and reduced each of ADHD two subscale scores by roughly 1 point (columns 3 and 4). The overall protective treatment effect seemed to be driven by low-performing students. It is worth noting that at baseline, this student subgroup was more likely to have poor subject test scores, held misconceptions of reading, did not like or feel confident in reading, and spent less time on reading-related activities (see Appendix Figure 2.2). In panels B and C, I also reported the treatment effects on female students and left-behind students accordingly. Male students and students not left behind seemed to benefit from the project, though the interaction terms were mostly not statistically significant (columns 1

[28] The ITT and LATE estimates on SDQ abnormal peer problem remain significant at 10% level after adjusting for multiple comparisons (see Appendix Table 2.18, panels A and B, column 5).

[29] The protective treatment effects also seemed most evident for students with poor baseline reading scores when using the more subjective measure, i.e., SDQ (see Appendix Table 2.19 and 2.21, column 5).

to 4).

Moreover, I complemented the standard heterogeneity analysis with a novel data-driven approach, called random causal forests (RCF). These forests were formed to maximize variation in treatment effects across splits in the sample, as opposed to minimizing prediction error as in classic random forest algorithms. This approach yielded two key findings[30]. First, I plotted the distribution of estimated treatment effects from RCF (see Appendix Figure 2.5). It is reassuring that its mean is closely aligned with the average treatment effects derived from regression estimates. Second, I reported the analysis of feature importance and identified the baseline reading score as the most significant factor in determining treatment effect heterogeneity (see Appendix Figure 2.6). This insight emphasizes the baseline reading score's critical role in understanding how different students respond to the intervention.

2.5.3 Mechanisms

This section examines the mechanisms through which in-class libraries improved primary school students' attention and behavior as measured by ADHD, especially for those with initial poor reading skills. I started by ruling out that students actively increased borrowing, reading, and learning from independent reading in the context of the intervention. I also checked other possible reading-related roles of teachers, parents, and peers in students' disruptive behaviors and found no significant effects as well. I then discussed possible remaining mechanisms, focusing on environmental and indirect influences. For example, the presence of in-class libraries may have contributed to improvements in the overall classroom environment. This pathway aligns with Self-Determination Theory (SDT), a prominent framework in social psychology that highlights the role of supportive environments in fostering motivation and well-being (Deci and Ryan, 2013, 2004).

Students' responses and behaviors

According to the estimation results in Table 2.4 (panels A and B), in-class libraries, on average, did not incentivize more students to borrow reading books from the in-class library (column 1), nor to bring books home for reading frequently (column 2). A similar pattern was observed in students'

[30] To implement, I utilized the "CausalForestDML" from the "econML" library in Python, which was developed by Microsoft Research for estimating heterogeneous treatment effects using machine learning methods.

reading behaviors. That is, students from treatment groups did not read independent books after school more often (column 3), nor did they significantly increase the amount of time spent on reading in school or at home combined (column 4). The number of books students completed this semester was also not different between the treatment and control groups (column 5). For low-performing students in reading at baseline, I did not find their borrowing or reading behaviors increased in a meaningful way (see Appendix Table 2.22).

In addition to borrowing and reading behaviors, I also evaluated the possible learning outcomes from the in-class library project on primary school students, whose results were shown in Table 2.5. Specifically, I broke it down into three categories, which were (1) student reading interests and confidence, (2) student reading beliefs, and (3) student reading skills. As shown in all panels, I did not find much supporting evidence that the provision of suitable children's books in the classroom improved students' learning outcomes regarding their reading interests (column 1), confidence (column 2), and skills (columns 5 and 6), which was not so surprising given that students were not utilizing the books more[31,32,33]. While these findings were based on student self-reported surveys, it is possible that the library was used in less measurable ways, such as casual browsing during school hours rather than intensive reading at home. Testing this hypothesis would require detailed observational data on library usage, which unfortunately I did not have for 2018.

Teachers, primary caregivers and peers' behaviors

As suggested by the existing literature (Mullis et al., 2012), teachers play a very important role in students' behaviors and overall learning. Thus it makes sense to examine whether teachers increased their interactions with students because of the treatment, which reduced students' disruptive behaviors. Here, I mainly focused on reading-related interactions since they were plausibly

[31] In addition to the dummy variables for student reading interests and confidence as I presented in Table 2.5 columns 1 and 2 (i.e., "Do not like reading" and "Do not feel confident in reading"), I also showed in the Appendix Table 2.23 and Table 2.24 about student responses to the total scores as well as different components of the student like reading scale and student reading confidence scale. Similarly, I did not see much gain in these outcomes.

[32] I did find that the in-class library reduced negative perceptions about independent reading on language tests among students with poor baseline reading scores. This student group at baseline is more likely to share wrong beliefs about how independent reading might negatively affect their academics (see Appendix Table 2.25, column 3). The intervention could have served as a strong signal to challenge their misconceptions.

[33] Yi et al. (2019) found significant positive effects of a similar in-class library project on student affinity toward reading and student reading habits, in contrast to what I showed here. The differences could be due to the different institutions involved in project implementation, since both the educational bureau and the NGO collaborated to organize the second round of project rollout relevant to my sample while only the NGO was responsible for implementation as studied in their paper.

most likely to be affected by the project, though these extra activities or efforts were not explicitly required in the intervention. From Table 2.6, I found that in-class libraries did not motivate teachers to encourage students to borrow books (column 1) or recommend more independent reading (column 2), nor did they assign home assignments relevant to independent reading to students (column 3) or were more likely to lead regular reading classes (column 4)[34,35]. Moreover, to understand if there is a meaningful impact of the brief orientation meeting in half of the treated schools driving the results, I regressed the rate of ADHD risks across the three groups and tested for differences between the treated groups with and without the orientation meeting. As expected, the orientation meeting did not have a significant additional impact (difference: 0.006; p-value: 0.5112).

The behaviors of primary caregivers (usually parents) and peers are also critical. However, as shown in Table 2.7, I did not find that primary caregivers invest more money or time after the intervention. The effect estimates for spending buying books for their children or spending quality time studying with children were relatively small and statistically insignificant (columns 1 and 2). Similarly, for the role of peers, I showed that in-class libraries did not significantly promote interaction among peers in reading classes (column 3), nor did students spend more time discussing with peers or reading together with friends (columns 4 and 5)[36,37].

Possible environmental and indirect influences

While the previous subsections empirically reveal null effects on reading-related behavior changes among students, teachers, peers, and caregivers, the mechanisms underlying the observed protective effects on ADHD outcomes may extend beyond these active and direct changes. This section explores potential remaining mechanisms that could explain the in-class library's impact on students' attention and behavior, particularly among those with initially poor reading skills. The discussion focuses on environmental and indirect factors within the classroom, with the caveat that these

[34] If anything, teachers' activities and efforts related to reading rather seemed to be crowded out by the intervention, since the estimates from column 1 to 4 were mostly negative though insignificant.

[35] The intervention also did not improve the perception of Chinese teachers toward the role of independent reading in language tests (see Appendix Table 2.26, column 1).

[36] If anything, those students with poor baseline reading skills even discussed less with peers about the contents that they didn't understand from independent reading (Appendix Table 2.28, column 4).

[37] Considering the possibility that students may interact better with each other in activities outside of reading. In Appendix Table 2.29, I used two other aggregate indicators, which are the Student Bullying Scale (columns 1 and 2) and the Students' Sense of School Belonging Scale (columns 3 and 4) from PIRLS, as outcomes. It appeared that treatment did not improve or harm the overall school experiences since all estimates were small and insignificant.

remaining pathways could not be empirically tested in this study.

Classroom environment A potential mechanism at work is that the presence of an in-class library might alleviate ADHD symptoms by improving the classroom environment in a passive manner. Environments that better satisfy basic psychological needs can enhance motivation and well-being, as outlined by Self-Determination Theory (SDT) (Morsink et al., 2022). SDT posits that motivation, essential for well-being, depends on fulfilling the psychological needs of autonomy, relatedness, and competence (Deci and Ryan, 2013, 2004). In educational settings, autonomy allows students to experience as self-driven and freely chosen. Relatedness fosters a sense of belonging and support, and competence gives students a feeling of effectiveness in school activities. Importantly, Rogers and Tannock (2018) found that children with ADHD often perceive classrooms as controlling, feel incompetent, and report strained teacher relationships. Addressing these needs could positively impact ADHD symptoms and related behaviors (Morsink et al., 2022).

In-class libraries may have supported these psychological needs, even though direct measures of internal motivation are unavailable. By offering diverse reading options, the library could promote autonomy, enabling students to make independent choices about engagement. It also offered materials not directly linked to academic reading, reinforcing this sense of autonomy. The library might also enhance relatedness by signaling investment in students' learning experiences, fostering a sense of support and community. These subtle, passive influences could be particularly impactful for students with low baseline reading scores. However, its impact on competence remains less clear, as competence typically involves active skill development, such as significant improvements in reading confidence or performance, which were not evident in this study (see Table 2.5).

Indirect influences Another potential mechanism is that the in-class library intervention may have indirectly influenced students' and teachers' perceptions and non-reading-related behaviors through its setup and management process. The establishment and management of the library might have signaled an increased emphasis on organization, structure, order, and quiet, reflective practices within the classroom. These changes could subtly shift how students and teachers perceive the classroom atmosphere, gradually reshaping norms and behaviors unrelated to reading activities. For example, students might develop a sense of responsibility or self-regulation inspired by the library's structured design, accessibility, and maintenance. Similarly, teachers might adopt more positive or proactive classroom management strategies in response to these changes. Over time, such adjustments could contribute to improved attention and behavior among students.

2.6 Discussion and Conclusion

This paper conducted an exploratory analysis to understand the impact on school children's attention and behavior of in-class libraries, a common type of intervention that provides educational inputs to school systems that usually lack. Specifically, I documented the positive effect of the treatment on rural students' ADHD rate. On average, in-class libraries significantly reduced the rate of ADHD among rural primary school students in the sample by about 1.8 percentage points, a reduction of 30 percent. The improvement from the treatment was more evident for students with poor reading skills initially. But in-class libraries didn't significantly increase students' reading time alone or with teachers, parents, and peers, nor did it improve reading comprehension.

These findings imply, on one hand, that in-class libraries did not significantly impact students' attention and behavior through active participation, such as reading more frequently or for longer periods. This limited effect may be attributed to deeply rooted biases as well as a lack of attention or interest in independent reading, which are challenging to overcome. On the other hand, the results may suggest that the pathways of influence are more indirect and passive. One plausible explanation for the project's beneficial impact on ADHD symptoms is an improved classroom environment that aligns with Self-Determination Theory (SDT). That is, the improvement in classroom environments better meets students' basic psychological needs such as autonomy and relatedness leading to better motivation and behaviors. However, due to data limitations, I was unable to conclusively test this hypothesis, so it should be considered as suggestive. Nonetheless, if this line of interpretation holds, the effectiveness of in-class libraries in improving the classroom environment could be very context-specific and depend on how salient in-class libraries are to students. For example, US classrooms may already be packed with books and other educational materials. The addition of a library would have limited effects. While in rural China, the scarcity of quality independent reading materials combined with a highly controlling classroom climate allows these in-class libraries to stimulate the classroom environment, promote autonomy and relatedness, and thereby improve student behavior.

Building on this understanding of the intervention's impact on ADHD and classroom dynamics, it is pertinent to relate these findings to broader educational outcomes like test scores, which are extensively studied in the existing literature. As I showed in Appendix Table 2.30, though ADHD improved, there did not seem to be main effects on standardized reading, Chinese, and math test scores (see panels A and B). The overall null effects on test scores are consistent with the literature assessing reading resource provision interventions (Borkum, He and Linden, 2012; Gao et al., 2018), including Yi et al. (2019) that evaluated a similar in-class library project in the Chinese context. I argue that the null results on standardized learning outcomes do not contradict my current finding of meaningful positive treatment effects on ADHD for two principal reasons. First, it may take a longer time for in-class libraries to show changes in students' learning outcomes. The improved classroom environment fostering motivation may be more evident in ADHD over a relatively short period, while test scores are just much harder to change, even if ADHD is mainly viewed as a strong correlate of academic performance. In other words, the protective treatment effects on ADHD may help students pay attention to class, follow instructions and complete tasks, which can affect their academic performance in theory but take time to be observed in practice[38]. Second, providing books alone might not be enough to improve test scores. This is especially true if one considers the various factors that may be involved in academic performance besides adequate access to resources, such as learning ability, study strategies, and teacher quality, just to name a few. The positive effects on ADHD of the lean resource provision intervention may not necessarily translate onto test scores because ADHD is a mental health challenge that has unique characteristics, particularly on the non-cognitive side.

Next, I discuss several limitations of the study. First, the impacts of in-class libraries on students' attention and behavior were not specifically pre-registered and should be viewed as exploratory. Future research is suggested to adopt a confirmatory approach using a pre-analysis plan specifically for mental health conditions and tests in different contexts to validate the findings of this study. Second, though I used a well-established ADHD scale reported by primary caregivers and followed the standard coding and interpretation practices, there could be potential concerns for bias. On the one hand, the ADHD measure only serves as a screening scale rather than a diagnostic test. It provides useful preliminary insights into a child's symptoms and can also screen for coexisting conditions like oppositional-defiant disorder, anxiety, and depression. However, a full

[38]The finding of a statistically significant effect of the project on Chinese standardized test scores measured at the endline survey for low-performing students in reading at baseline highlights the possibility of overall impacts on learning outcomes in the longer term (see Appendix Table 2.30, panel C, column 2)

diagnostic evaluation by a pediatrician is essential to differentiate ADHD from other conditions and determine the need for specialist referrals. On the other hand, ADHD RS-IV can be an underestimate of the ADHD symptoms, especially for students whose parents are less committed or poorly educated[39]. Moreover, measurement errors could exist due to the way the household questionnaire was distributed. In this study, no home visits were performed to directly survey students' primary caregivers, which could be more precise but also more expensive and time-consuming[40]. Future studies could consider incorporating more comprehensive diagnostic tests alongside screening scales, as well as integrating multiple sources of information like teacher reports and behavioral observations.

Another limitation concerns the interpretative nature of the results regarding mechanisms and policy implications. The hypothesis connecting the intervention's positive impact on ADHD symptoms to an improved classroom environment and other indirect factors remains indicative, without comprehensive data available for a formal test. Future studies could aim to gather comprehensive data to test these hypotheses more robustly. Moreover, considering students' passive engagement in reading materials, there may be cheaper alternatives such as digital resources that could offer similar benefits at a lower cost. A different approach that actively engages students in reading might also achieve comparable improvements in attention and behavior. Future research should explore these alternatives, comparing the efficacy of digital resources and more interactive reading strategies against traditional methods like in-class libraries.

In conclusion, the study provides promising evidence of the potential benefits on students' attention and behavior of providing reading resources to students at the classroom level in rural China, especially for those with poor reading skills. It contributes to a growing body of research that supports a more holistic approach to educational interventions, one that considers the mental health characteristics of students and their psychological needs. The paper calls for educational policies and interventions that are context-specific, recognizing the unique challenges and opportunities present in different educational settings, especially in classrooms. By continuing to explore and understand these complexities, we can better design and implement educational strategies that

[39] Some previous research suggests that parents tend to give valid and reliable reports of children's ADHD symptoms (Faraone, Biederman and Zimmerman, 2005; Biederman et al., 2006).

[40] Instead, students were asked to take the survey home for their primary caregiver to complete and return to school the next day.

effectively support students' human capital development.

Main Figures and Tables

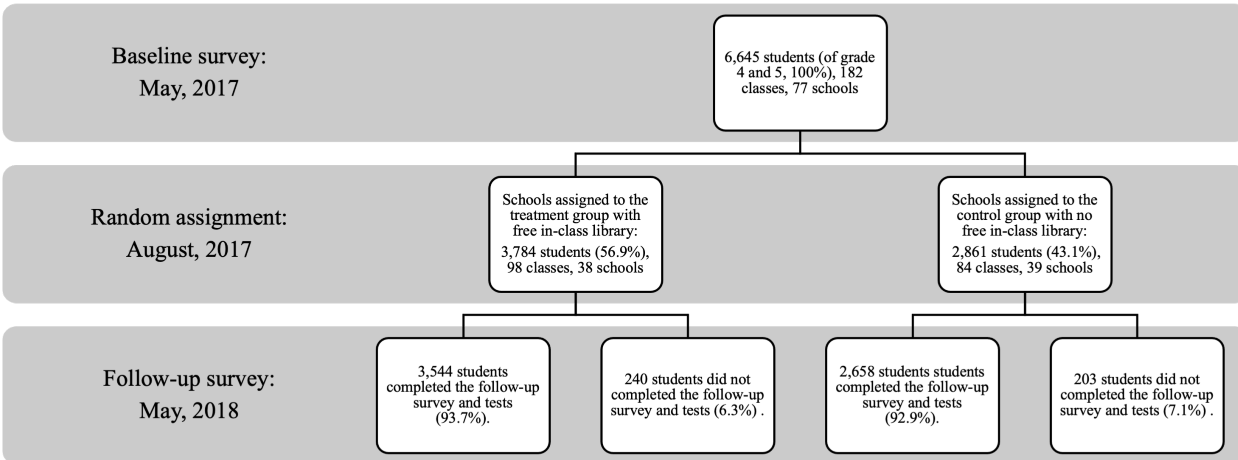

Figure 2.1: Research Design of the In-class Library Project in Chinese Rural Primary Schools

Table 2.1: Balance Table of Baseline Characteristics of Students, by Treatment Group

	(1) Control group		(2) Treatment group		(3) T-test
	N	Mean	N	Mean	P-value
Student Characteristics					
Age (in years)	2922	11.044	3769	11.020	0.599
	[39]	[0.036]	[38]	[0.027]	
Female (1=Yes; 0=No)	2941	0.490	3778	0.492	0.855
	[39]	[0.009]	[38]	[0.007]	
Boarding at school (1=Yes; 0=No)	2936	0.114	3769	0.099	0.768
	[39]	[0.037]	[38]	[0.031]	
Family Characteristics					
At least one parent had 9 years of schooling or higher (1=Yes; 0=No)	2875	0.192	3670	0.186	0.729
	[39]	[0.012]	[38]	[0.010]	
At least one parent had a professional occupation (1=Yes; 0=No)	2912	0.211	3754	0.210	0.958
	[39]	[0.012]	[38]	[0.013]	
Left-behind (1=Both parents not at home; 0=Otherwise)	2851	0.477	3670	0.479	0.940
	[39]	[0.018]	[38]	[0.020]	
Student from low-income families (1=No refrigerator at home)	2937	0.064	3771	0.078	0.134
	[39]	[0.006]	[38]	[0.007]	
Reading Resources					
Have library or reading room at school (1=Yes; 0=No)	2937	0.857	3776	0.864	0.915
	[39]	[0.051]	[38]	[0.046]	
Have library or reading room at classroom (1=Yes; 0=No)	2919	0.480	3744	0.443	0.757
	[39]	[0.079]	[38]	[0.090]	
Student had over 25 books at home (1=Yes; 0=No)	2929	0.111	3765	0.091	0.348
	[39]	[0.018]	[38]	[0.011]	

Note: The number of clusters for the treatment and control group and the standard errors of the group means are shown inside []. Standard errors are clustered at the school level. T-tests' p-values are shown in column (3). ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels.

Table 2.1: Balance Table of Baseline Characteristics of Students, by Treatment Group (continued)

	(1) Control group		(2) Treatment group		(3) T-test
	N	Mean	N	Mean	P-value
Student Reading Attitudes					
Student Like Reading scale (0 to 24, 2011 version)	2859	16.800	3667	16.749	0.834
	[39]	[0.137]	[38]	[0.197]	
Do not like reading (1=Yes; 0=No)	2859	0.068	3667	0.075	0.475
	[39]	[0.005]	[38]	[0.009]	
Student Confident in Reading scale (0 to 21, 2011 version)	2866	13.003	3664	13.110	0.514
	[39]	[0.119]	[38]	[0.113]	
Do not feel confident in reading (1=Yes; 0=No)	2866	0.143	3664	0.125	0.149
	[39]	[0.008]	[38]	[0.009]	
Student perceived negative effects of reading on Chinese (1=Yes; 0=No)	2938	0.066	3763	0.069	0.829
	[39]	[0.007]	[38]	[0.007]	
Student perceived negative effects of reading on Math (1=Yes; 0=No)	2936	0.124	3753	0.116	0.592
	[39]	[0.012]	[38]	[0.008]	
Student Reading Habits					
Student never borrowed books from school (1=Yes; 0=No)	2937	0.767	3776	0.707	0.368
	[39]	[0.045]	[38]	[0.049]	
Student never borrowed books from classroom (1=Yes; 0=No)	2916	0.616	3722	0.669	0.583
	[39]	[0.065]	[38]	[0.071]	
Read for 30 minutes or more at home daily (1=Yes; 0=No)	2939	0.353	3766	0.366	0.696
	[39]	[0.022]	[38]	[0.026]	
Talk about readings with friends (1=Yes; 0=No)	2934	0.405	3767	0.397	0.863
	[39]	[0.030]	[38]	[0.036]	
Read together with friends (1=Yes; 0=No)	2937	0.367	3767	0.366	0.974
	[39]	[0.027]	[38]	[0.030]	
Borrow books that friends read (1=Yes; 0=No)	2935	0.537	3760	0.532	0.899
	[39]	[0.028]	[38]	[0.035]	
Student Achievement in Reading and IQ					
Standardized reading score (0 to 1)	2939	0.008	3777	-0.028	0.552
	[39]	[0.029]	[38]	[0.053]	
Student cognitive ability (in Raven's IQ)	2879	91.856	3692	92.286	0.663
	[39]	[0.693]	[38]	[0.703]	

Note: The number of clusters for the treatment and control group and the standard errors of the group means are shown inside []. Standard errors are clustered at the school level. T-tests' p-values are shown in column (3). ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels.

Table 2.2: Treatment Effects on ADHD Outcomes at Eight Months After the Intervention

	ADHD risks		ADHD scores					
	Rate of ADHD (1=Yes; 0=No)		Total (0 to 54)		Inattention (0 to 27)		Hyperactivity (0 to 27)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Intention-to-treat effect								
Assigned in-class library	-0.019***^{,a}	-0.018***^{,a}	-0.557*	-0.503	-0.202	-0.192	-0.285*	-0.235
	(0.006)	(0.006)	(0.319)	(0.312)	(0.172)	(0.165)	(0.163)	(0.159)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
County FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4,402	4,119	4,402	4,119	4,548	4,255	4,588	4,291
Mean of Dep. Var.	0.063	0.060	12.447	12.325	6.959	6.902	5.617	5.554
Panel B: Local average treatment effect								
Built in-class library	-0.030***^{,a}	-0.027***^{,a}	-0.864*	-0.779*	-0.312	-0.295	-0.446*	-0.366
	(0.010)	(0.010)	(0.482)	(0.464)	(0.264)	(0.251)	(0.244)	(0.234)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
County FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4,399	4,117	4,399	4,117	4,545	4,253	4,584	4,288
Mean of Dep. Var.	0.063	0.060	12.447	12.326	6.960	6.903	5.616	5.554
First Stage	0.645	0.643	0.645	0.643	0.647	0.645	0.643	0.641
F Stats	64.015	62.671	64.015	62.671	64.302	63.115	62.972	62.188

Note: Columns (1) and (2) show the ADHD risks, while the rest present ADHD scores. The cutoff for the prevalence of ADHD follows the paper (Su et al., 2015). For ADHD scores, I include the ADHD total scores and its two subscale scores, which are inattention (IA) scores and hyperactivity-impulsivity (HI) scores. Higher scores indicate a higher risk for poorer mental health outcomes. For several regressions, I include the following set of baseline controls: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, ^{,a} ($^{,a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.3: Heterogeneity Treatment Effects on ADHD Outcomes (ITT)

	ADHD risks	ADHD scores		
	(1) Rate of ADHD (1=Yes; 0=No)	(2) Total (0 to 54)	(3) Inattention (0 to 27)	(4) Hyperactivity (0 to 27)
Panel A: by low-performing students in reading				
Assigned in-class library	-0.067***,a	-1.721**,a	-0.794**,a	-0.816**,a
× low-performing students in reading	(0.025)	(0.704)	(0.376)	(0.338)
Assigned in-class library	-0.006	-0.206	-0.054	-0.095
	(0.006)	(0.312)	(0.162)	(0.165)
Low-performing students in reading	0.124***,a	4.688***,a	2.268***,a	2.310***,a
	(0.020)	(0.507)	(0.276)	(0.241)
Total effects for low-performing students	-0.073***,a	-1.927***,a	-0.848**,a	-0.911***,a
	(0.023)	(0.669)	(0.366)	(0.311)
Baseline controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	4,111	4,111	4,247	4,283
Mean of Dep. Var.	0.061	12.322	6.901	5.553
Panel B: by female students				
Assigned in-class library	0.016	0.219	0.270	-0.058
× female students	(0.016)	(0.510)	(0.275)	(0.245)
Assigned in-class library	-0.026**	-0.614	-0.328	-0.205
	(0.012)	(0.420)	(0.203)	(0.222)
Female students	-0.044***,a	-2.547***,a	-1.116***,a	-1.406***,a
	(0.012)	(0.347)	(0.195)	(0.179)
Total effects for female students	-0.009	-0.395	-0.058	-0.264
	(0.007)	(0.385)	(0.226)	(0.176)
Baseline controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	4,119	4,119	4,255	4,291
Mean of Dep. Var.	0.060	12.325	6.902	5.554
Panel C: by left-behind students				
Assigned in-class library	0.001	0.725	0.250	0.424*
× left-behind students	(0.014)	(0.449)	(0.249)	(0.228)
Assigned in-class library	-0.018*	-0.832**,a	-0.305	-0.428**
	(0.009)	(0.358)	(0.195)	(0.175)
Left-behind students	-0.005	-0.715**	-0.388**	-0.270
	(0.011)	(0.354)	(0.185)	(0.194)
Total effects for left-behind students	-0.017*	-0.107	-0.055	-0.004
	(0.010)	(0.411)	(0.219)	(0.215)
Baseline controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	4,119	4,119	4,255	4,291
Mean of Dep. Var.	0.060	12.325	6.902	5.554

Note: Low-performing students in reading are those with baseline reading test scores ranked in the bottom 20 percent. Left-behind students are those whose fathers and mothers migrated out for more than nine months this year. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, ^{,a} ($^{,a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.4: Treatment Effects on Students Borrowing and Reading Behaviors

	Endline student borrowing behaviors		Endline student reading behaviors		
	(1) Never borrowed books from classroom library (1=Yes; 0=No)	(2) At least monthly bring home books from classroom library (1=Yes; 0=No)	(3) At least monthly read independent books after school (1=Yes; 0=No)	(4) Read independent books daily for 30 minutes or more (1=Yes; 0=No)	(5) The number of independent books read in this semester
Panel A: Intention-to-treat effect					
Assigned in-class library	0.108	0.018	-0.009	-0.019	-0.037
	(0.084)	(0.070)	(0.031)	(0.033)	(0.047)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	5,863	5,867	5,880	5,879	5,860
Mean of Dep. Var.	0.449	0.362	0.744	0.688	0.611
Panel B: Local average treatment effect					
Built in-class library	0.168	0.028	-0.014	-0.029	-0.058
	(0.145)	(0.107)	(0.049)	(0.051)	(0.076)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	5,863	5,867	5,878	5,877	5,858
Mean of Dep. Var.	0.449	0.362	0.744	0.688	0.611
First Stage	0.642	0.642	0.642	0.642	0.641
F Stats	77.821	77.902	77.795	77.824	77.606

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, $,^{a}$ ($,^{a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.5: Treatment Effects on Students Learning from Independent Reading

	Endline reading interests and confidence		Endline reading beliefs		Endline reading skills	
	(1) Do not like reading (1=Yes; 0=No)	(2) Do not feel confident in reading (1=Yes; 0=No)	(3) Perceived negative effect of independent reading on language test (1=Yes; 0=No)	(4) Perceived negative effect of independent reading on math test (1=Yes; 0=No)	(5) Apply what learned from independent reading (1=Yes; 0=No)	(6) Standardized reading score (0 to 1)
Panel A: Intention-to-treat effect						
Assigned in-class library	0.012	0.002	-0.001	0.019	-0.005	-0.024
	(0.012)	(0.017)	(0.005)	(0.013)	(0.020)	(0.049)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,866	5,875	5,875	5,868	5,880	5,886
Mean of Dep. Var.	0.091	0.156	0.024	0.117	0.192	0.033
Panel B: Local average treatment effect						
Built in-class library	0.019	0.003	-0.001	0.029	-0.008	-0.035
	(0.019)	(0.026)	(0.008)	(0.021)	(0.032)	(0.078)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,864	5,873	5,873	5,866	5,878	5,878
Mean of Dep. Var.	0.091	0.156	0.024	0.117	0.192	0.034
First Stage	0.642	0.641	0.642	0.641	0.642	0.642
F Stats	77.760	77.604	77.918	77.573	77.795	77.795

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, ^{,a} ($^{,a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.6: Treatment Effects on Teachers Supporting Students' Independent Reading

	(1) Encourage students to borrow independent books (1=Yes; 0=No)	(2) Recommend several independent readings (1=Yes; 0=No)	(3) Assign independent reading homework (1=Yes; 0=No)	(4) Lead regular reading classes (1=Yes; 0=No)
Panel A: Intention-to-treat effect				
Assigned in-class library	0.001	-0.084	-0.123	-0.174*
	(0.078)	(0.058)	(0.079)	(0.098)
Baseline controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	5,880	5,879	5,880	5,880
Mean of Dep. Var.	0.665	0.863	0.722	0.424
Panel B: Local average treatment effect				
Built in-class library	0.001	-0.131	-0.192	-0.271
	(0.121)	(0.097)	(0.129)	(0.164)
Baseline controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	5,878	5,877	5,878	5,878
Mean of Dep. Var.	0.665	0.863	0.722	0.424
First Stage	0.642	0.642	0.642	0.642
F Stats	77.795	77.872	77.795	77.795

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, $'^{a}$ ($'^{a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.7: Treatment Effects on Primary Caregivers and Peers Behaviors

	Primary caregivers' behaviors		Peers' interactions		
	(1) Spent money buying books (1=Yes; 0=No)	(2) Spend quality time studying with student (1=Yes; 0=No)	(3) Interact with peers in reading classes (1=Yes; 0=No)	(4) Discuss with peers what they don't understand in independent books (1=Yes; 0=No)	(5) Read independent books together with friends regularly (1=Yes; 0=No)
Panel A: Intention-to-treat effect					
Assigned in-class library	-0.016	0.005	0.032	-0.034	-0.031
	(0.029)	(0.018)	(0.074)	(0.031)	(0.020)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	4,417	4,394	5,880	5,879	5,880
Mean of Dep. Var.	0.763	0.730	0.217	0.521	0.139
Panel B: Local average treatment effect					
Built in-class library	-0.024	0.008	0.049	-0.053	-0.048
	(0.045)	(0.029)	(0.114)	(0.048)	(0.032)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	4,414	4,391	5,878	5,877	5,878
Mean of Dep. Var.	0.763	0.730	0.217	0.521	0.139
First Stage	0.643	0.645	0.642	0.642	0.642
F Stats	62.796	62.958	77.795	77.764	77.795

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Low-performing students in reading are those with baseline reading test scores ranked in the bottom 20 percent. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, $,^{a}$ ($,^{a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

2.7 Appendix A: Main Scales Definitions

Table 2.8: Attention Deficit Hyperactivity Disorder (ADHD) Rating Scale-IV

	Never or rarely	Some-times	Often	Very often
ADHD inattention (IA)				
1 Fails to give close attention to details or makes careless mistakes in schoolwork.	0	1	2	3
3 Has difficulty sustaining attention in tasks or play activities.	0	1	2	3
5 Does not seem to listen when spoken to directly.	0	1	2	3
7 Does not follow through on instructions and fails to finish work.	0	1	2	3
9 Has difficulty organizing tasks and activities.	0	1	2	3
11 Avoids tasks (e.g., schoolwork, homework) that require sustained mental effort.	0	1	2	3
13 Loses things necessary for tasks or activities.	0	1	2	3
15 Is easily distracted.	0	1	2	3
17 Is forgetful in daily activities.	0	1	2	3
ADHD hyperactivity and impulsivity (HI)				
2 Fidgets with hands or feet or squirms in seat.	0	1	2	3
4 Leaves seat in classroom or in other situations in which remaining seated is expected.	0	1	2	3
6 Runs about or climbs excessively in situation in which it. is inappropriate.	0	1	2	3
8 Has difficulty playing or engaging in leisure activities quietly.	0	1	2	3
10 Is "on the go" or acts as "driven by a motor".	0	1	2	3
12 Talks excessively.	0	1	2	3
14 Blurts out answers before questions have been completed.	0	1	2	3
16 Has difficulty awaiting turn.	0	1	2	3
18 Interrupts or intrudes on others.	0	1	2	3

Note: There are two subscales in ADHD: one is inattention (IA), and the other is hyperactivity-impulsivity (HI).

Table 2.9: Strengths and Difficulties Questionnaire (SDQ) Mental Health Scale

	Not True	Somewhat True	Certainly True
1 I am restless. [hyperactivity]	0	1	2
2 I get a lot of headaches. [emotion]	0	1	2
3 I get very angry. [conduct]	0	1	2
4 I am usually on my own. [peer]	0	1	2
5 I usually do as I am told. [conduct]	2	1	0
6 I worry a lot. [emotion]	0	1	2
7 I am constantly fidgeting. [hyperactivity]	0	1	2
8 I have one good friend or more. [peer]	2	1	0
9 I fight a lot. [conduct]	0	1	2
10 I am often unhappy. [emotion]	0	1	2
11 Other people of my age generally like me. [peer]	2	1	0
12 I am easily distracted. [hyperactivity]	0	1	2
13 I am nervous in new situations. [emotion]	0	1	2
14 I am often accused of lying or cheating. [conduct]	0	1	2
15 Other children or young people pick on me. [peer]	0	1	2
16 I think before I do things. [hyperactivity]	2	1	0
17 I take things that are not mine. [conduct]	0	1	2
18 I get on better with adults than with people my age. [peer]	0	1	2
19 I have many fears. [emotion]	0	1	2
20 I finish the work I am doing. [hyperactivity]	2	1	0

Notes: [] shows the subscale shorthand to which each question belongs. Specifically, [emotion] stands for the scale of the emotional problems, [conduct] for the conduct problems scale, [hyperactivity] for the hyperactivity scale, and [peer] for the peer problems scale. Higher scores mean worse mental health.

Table 2.10: The PIRLS Students Like Reading Scale (2011)

	Agree a lot	**Agree a little**	**Disagree a little**	**Disagree a lot**
What do you think about reading?				
Tell how much you agree with each of these statements.				
1 I read only if I have to#				
2 I like talking about what I read with other people				
3 I would be happy if someone gave me a book as a present				
4 I think reading is boring#				
5 I would like to have more time for reading				
6 I enjoy reading				
	Every day or almost every day	**Once or twice a week**	**Once or twice a month**	**Never or almost never**
How often do you do these things outside of school?				
7 I read for fun				
8 I read things that I choose myself				

Note: Questions ended with symbol # are reverse coded.

Table 2.11: The PIRLS Students Like Reading Scale (2016)

	Agree a lot	**Agree a little**	**Disagree a little**	**Disagree a lot**
What do you think about reading?				
Tell how much you agree with each of these statements.				
1 I like talking about what I read with other people				
2 I would be happy if someone gave me a book as a present				
3 I think reading is boring#				
4 I would like to have more time for reading				
5 I enjoy reading				
6 I learn a lot from reading				
7 I like to read things that make me think				
8 I like it when a book helps me imagine other worlds				
	Every day or almost every day	**Once or twice a week**	**Once or twice a month**	**Never or almost never**
How often do you do these things outside of school?				
9 I read for fun				
10 I read to find out about things I want to learn				

Note: Questions ended with symbol # are reverse coded.

Table 2.12: The PRILS Students Confident in Reading Scale (2011)

	Agree a lot	**Agree a little**	**Disagree a little**	**Disagree a lot**
How well do you read?				
Tell how much you agree with each of these statements.				
1 I usually do well in reading				
2 Reading is easy for me				
3 Reading is harder for me than for many of my classmates#				
4 If a book is interesting, I don't care how hard it is to read				
5 I have trouble reading stories with difficult words#				
6 My teacher tells me I am a good reader				
7 Reading is harder for me than any other subject#				

Notes: Questions ended with symbol # are reverse coded.

Table 2.13: The PRILS Students Confident in Reading Scale (2016)

	Agree a lot	**Agree a little**	**Disagree a little**	**Disagree a lot**
How well do you read?				
Tell how much you agree with each of these statements.				
1 I usually do well in reading				
2 Reading is easy for me				
3 I have trouble reading stories with difficult words#				
4 Reading is harder for me than for many of my classmates				
5 Reading is harder for me than any other subject#				
6 I am just not good at reading#				

Notes: Questions ended with symbol # are reverse coded.

2.8 Appendix B: Additional Results

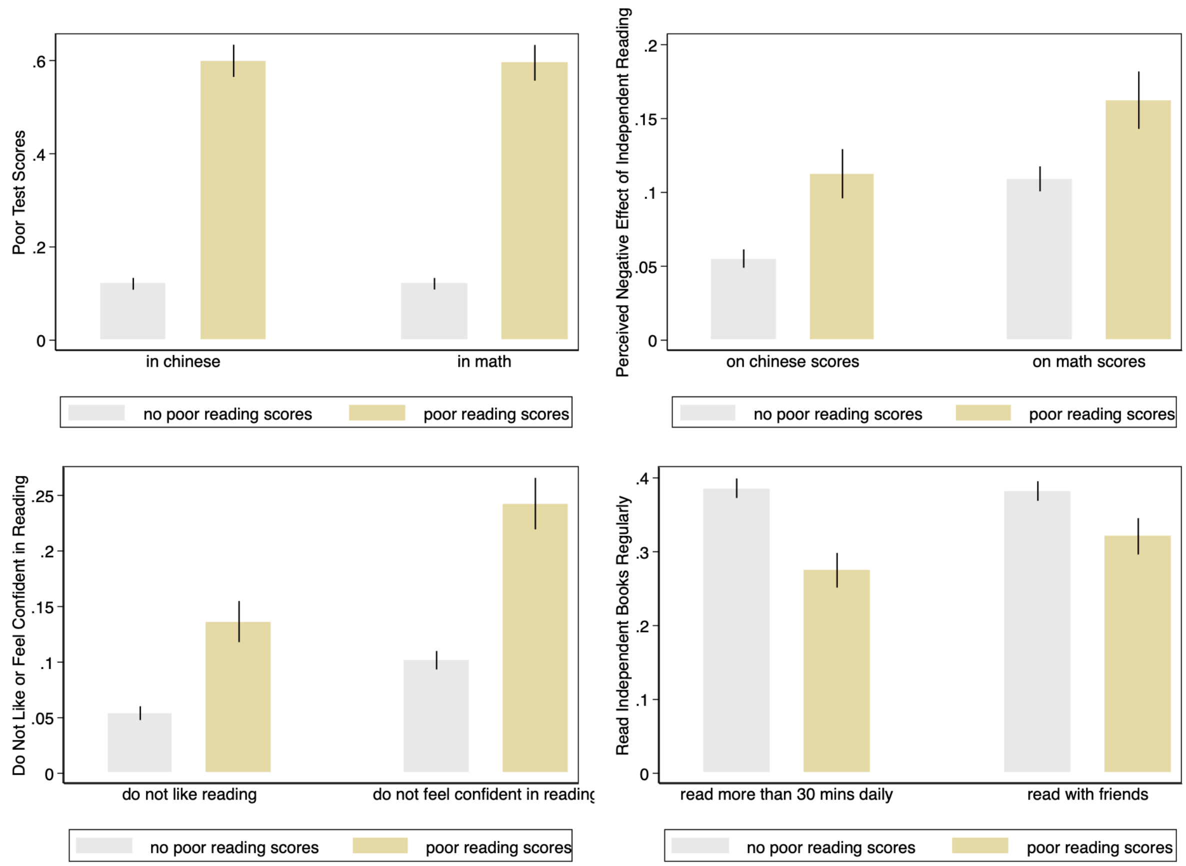

Figure 2.2: Comparisons Between Students with Poor versus No Poor Baseline Reading Scores

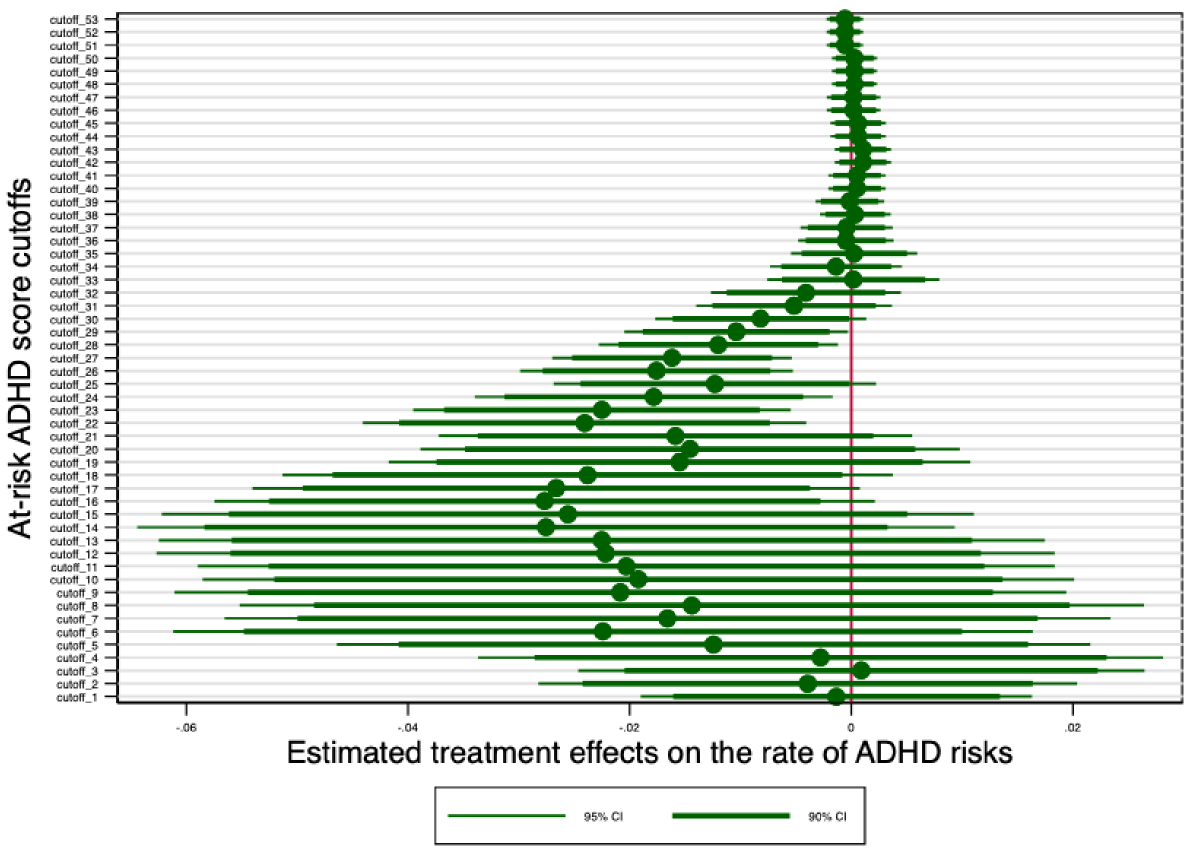

Figure 2.3: Treatment Effects on the Rate of ADHD Risks (varying score cutoffs)

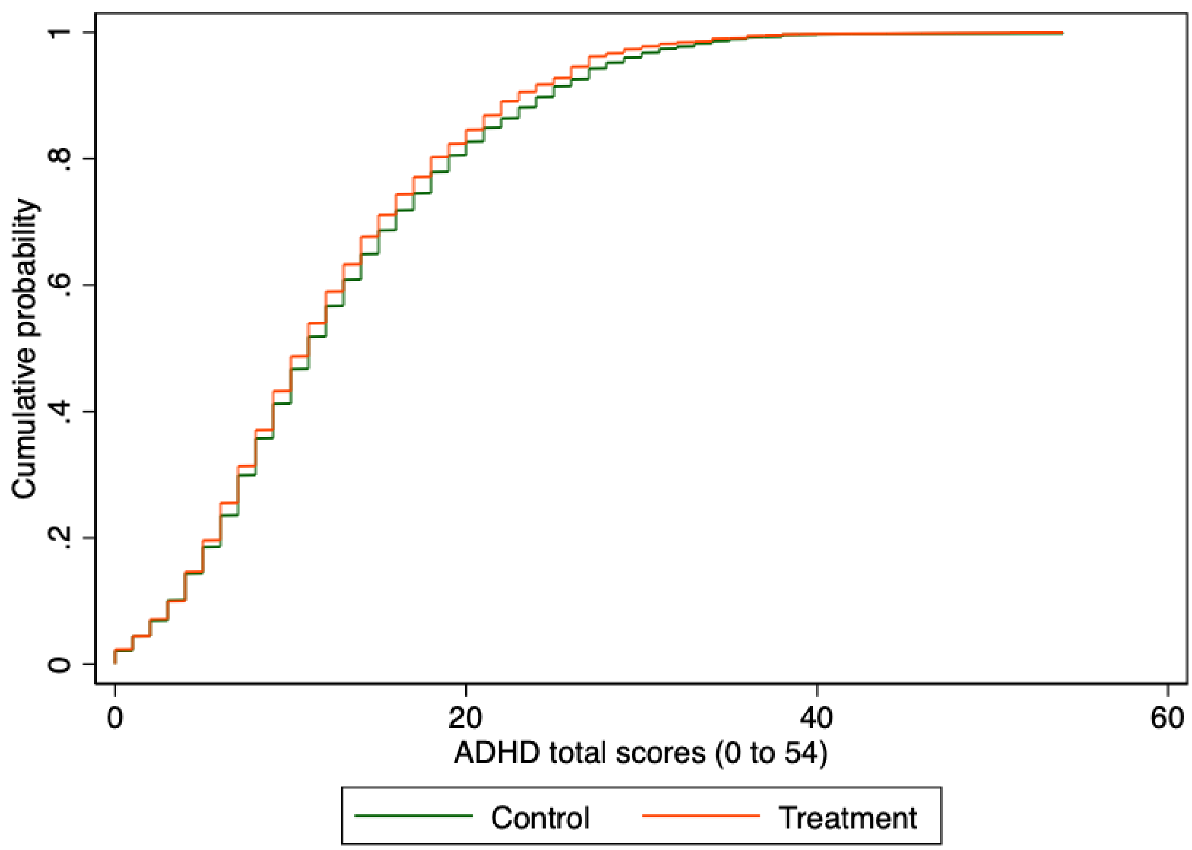

Figure 2.4: Cumulative Distribution of ADHD Scores: Treatment vs Control Groups

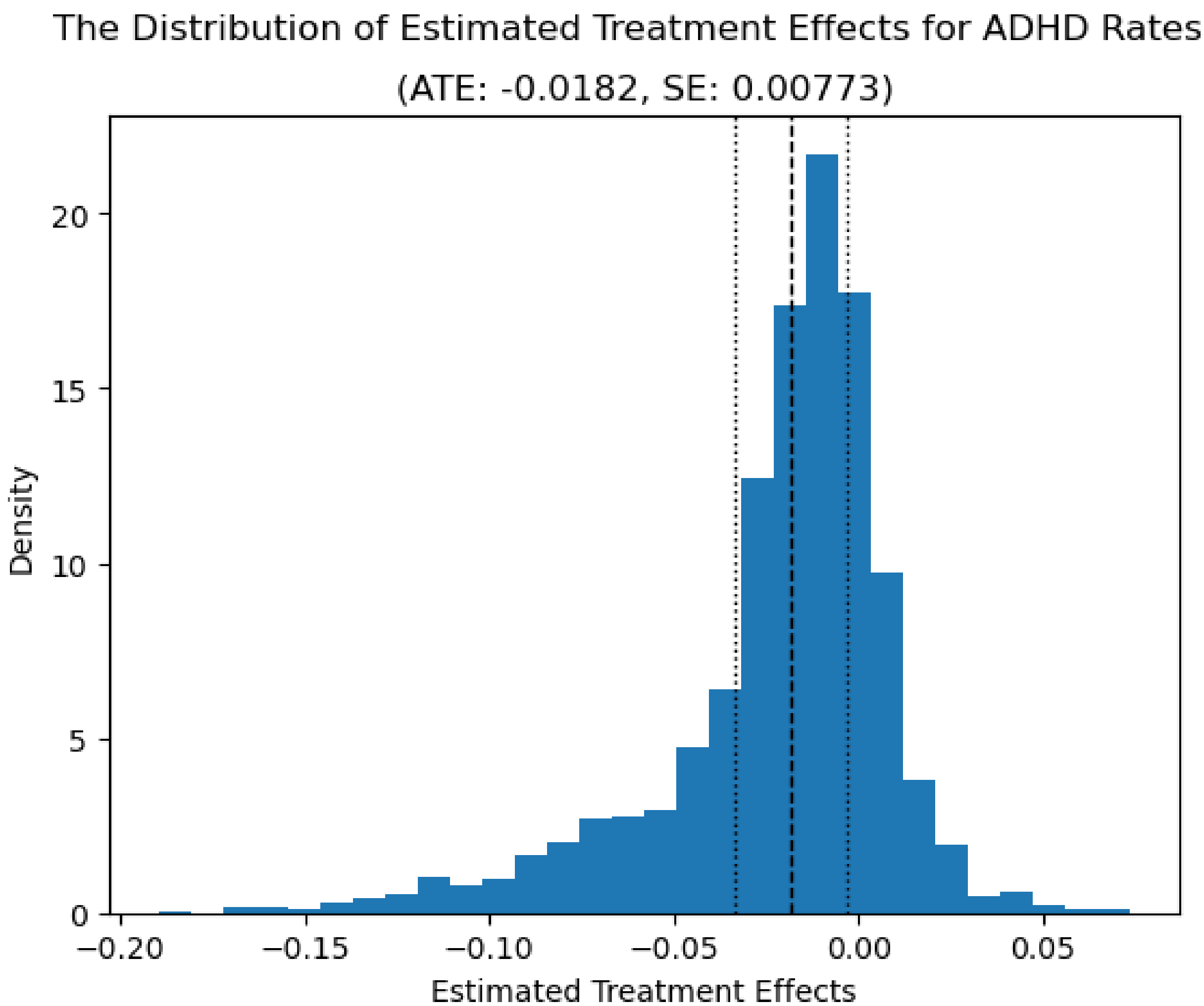

Figure 2.5: The Distribution of Estimated Treatment Effects for ADHD Rates

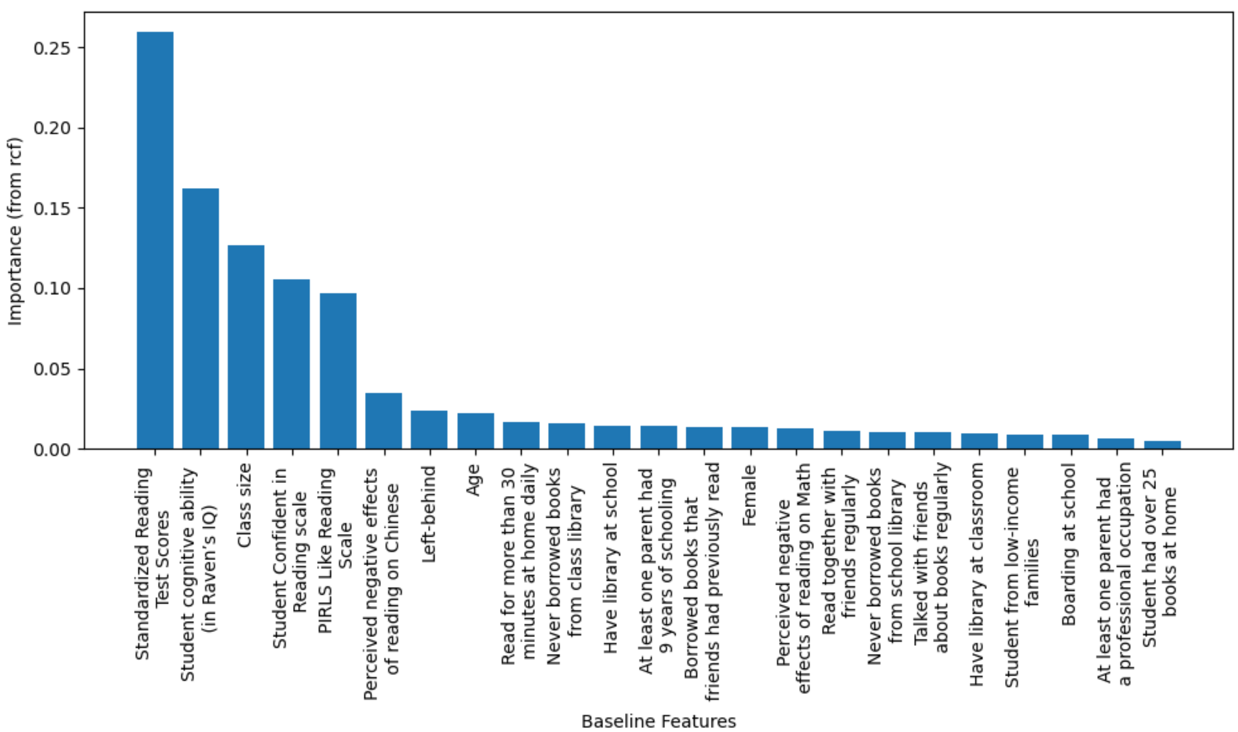

Figure 2.6: Feature Importance from Random Causal Forest Analysis on ADHD

Table 2.14: Treatment Effects on ADHD Outcomes (controlled for class size)

	ADHD risks	ADHD scores		
	(1) Rate of ADHD (1=Yes; 0=No)	(2) Total (0 to 54)	(3) Inattention (0 to 27)	(4) Hyperactivity (0 to 27)
Panel A: Intention-to-treat effect				
Assigned in-class library	-0.017***,a	-0.466	-0.175	-0.206
	(0.006)	(0.312)	(0.165)	(0.160)
Baseline controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	4,119	4,119	4,255	4,291
Mean of Dep. Var.	0.060	12.325	6.902	5.554
Panel B: by low-performing students in reading				
Assigned in-class library	-0.067***,a	-1.719**,a	-0.793**,a	-0.815**,a
× low-performing students in reading	(0.025)	(0.704)	(0.376)	(0.338)
Assigned in-class library	-0.005	-0.183	-0.043	-0.073
	(0.006)	(0.315)	(0.163)	(0.168)
Low-performing students in reading	0.124***,a	4.682***,a	2.266***,a	2.305***,a
	(0.020)	(0.507)	(0.276)	(0.243)
Total effects for low-performing students	-0.072***,a	-1.903***,a	-0.837**,a	-0.888***,a
	(0.023)	(0.661)	(0.362)	(0.308)
Baseline controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	4,111	4,111	4,247	4,283
Mean of Dep. Var.	0.061	12.322	6.901	5.553

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. I also included class size as an additional control. Low-performing students in reading are those with baseline reading test scores ranked in the bottom 20 percent. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, ^{,a} ($^{,a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.15: Heterogeneous Treatment Effects on ADHD Outcomes (by pre-existing classroom library status)

	ADHD risks	ADHD scores		
	(1) Rate of ADHD (1=Yes; 0=No)	(2) Total (0 to 54)	(3) Inattention (0 to 27)	(4) Hyperactivity (0 to 27)
Assigned in-class library	-0.002	-0.452	-0.165	-0.244
× Pre-existing classroom library	(0.012)	(0.555)	(0.309)	(0.284)
Assigned in-class library	-0.016*	-0.291	-0.115	-0.120
	(0.009)	(0.398)	(0.207)	(0.199)
Pre-existing classroom library	0.000	0.041	0.002	0.004
	(0.009)	(0.409)	(0.247)	(0.208)
Total effects for pre-existing classroom library	-0.019**	-0.743*	-0.279	-0.364
	(0.009)	(0.433)	(0.243)	(0.223)
Baseline controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	4,119	4,119	4,255	4,291
Mean of Dep. Var.	0.060	12.325	6.902	5.554

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Pre-existing classroom library indicator was assigned a value of 1 if over 50% of students in the same classroom self-reported the presence of a library at baseline, and 0 if not. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, ^{,a} ($^{,a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.16: Treatment Effects on ADHD Outcomes, each IA symptom

	(1) Makes careless mistakes	(2) Difficulty sustaining attention	(3) Not Listen	(4) Not follow through	(5) Difficulty organizing	(6) Avoids tasks	(7) Loses things	(8) Easily distracted	(9) Forgetful
Panel A: Intention-to-treat effect									
Assigned in-class library	0.008	-0.013	-0.042	-0.021	-0.028	-0.048*	-0.046*	0.013	-0.038
	(0.022)	(0.029)	(0.026)	(0.029)	(0.025)	(0.028)	(0.027)	(0.025)	(0.029)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,472	4,424	4,439	4,460	4,428	4,452	4,453	4,436	4,438
Mean of Dep. Var.	1.037	0.785	0.722	0.692	0.756	0.599	0.748	0.865	0.781
Panel B: Local average treatment effect									
Built in-class library	0.012	-0.021	-0.066	-0.032	-0.042	-0.074*	-0.073*	0.020	-0.058
	(0.035)	(0.045)	(0.040)	(0.045)	(0.038)	(0.040)	(0.041)	(0.039)	(0.046)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,469	4,421	4,437	4,457	4,425	4,449	4,450	4,433	4,435
Mean of Dep. Var.	1.037	0.785	0.722	0.692	0.757	0.599	0.748	0.866	0.782
First Stage	0.643	0.644	0.645	0.644	0.644	0.643	0.644	0.644	0.644
F Stats	62.502	62.969	63.626	62.905	62.633	62.581	62.610	62.502	62.796

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, $'^{a}$ ($'^{a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.17: Treatment Effects on ADHD Outcomes, each HI symptom

	(1) Fidgets/ squirms	(2) Leaves seat	(3) Runs about	(4) Difficulty playing	(5) On the go	(6) Talks excessively	(7) Blurts out answers	(8) Difficulty awaiting turn	(9) Interrupts
Panel A: Intention-to-treat effect									
Assigned in-class library	-0.047	-0.045*	-0.021	-0.001	-0.047**	-0.049*	0.001	-0.012	0.001
	(0.030)	(0.025)	(0.023)	(0.025)	(0.024)	(0.026)	(0.024)	(0.026)	(0.024)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,460	4,429	4,463	4,451	4,428	4,452	4,457	4,435	4,444
Mean of Dep. Var.	0.749	0.488	0.451	0.651	0.608	0.862	0.676	0.581	0.575
Panel B: Local average treatment effect									
Built in-class library	-0.072	-0.070*	-0.033	-0.004	-0.073**	-0.076**	0.003	-0.018	0.002
	(0.045)	(0.037)	(0.035)	(0.039)	(0.035)	(0.038)	(0.037)	(0.040)	(0.038)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,457	4,426	4,460	4,448	4,425	4,449	4,454	4,432	4,441
Mean of Dep. Var.	0.749	0.488	0.451	0.650	0.608	0.862	0.677	0.581	0.575
First Stage	0.643	0.643	0.643	0.643	0.642	0.643	0.644	0.644	0.643
F Stats	62.564	62.311	62.695	62.677	62.051	62.633	62.713	62.921	62.563

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, ^{,a} ($^{,a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.18: Treatment Effects on SDQ Abnormal Measures

		SDQ four subscales			
	(1) Abnormal SDQ total scores (1=Yes; 0=No)	(2) Abnormal emotional problems (1=Yes; 0=No)	(3) Abnormal conduct problems (1=Yes; 0=No)	(4) Abnormal hyperactivity problems (1=Yes; 0=No)	(5) Abnormal peer problems (1=Yes; 0=No)
Panel A: Intention-to-treat effect					
Assigned in-class library	-0.003	-0.005	-0.001	-0.003	-0.028**$^{,a'}$
	(0.009)	(0.008)	(0.014)	(0.007)	(0.011)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	5,876	5,876	5,876	5,876	5,876
Mean of Dep. Var.	0.098	0.081	0.147	0.062	0.101
Panel B: Local average treatment effect					
Built in-class library	-0.004	-0.009	0.001	-0.004	-0.044**$^{,a'}$
	(0.015)	(0.013)	(0.021)	(0.011)	(0.018)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	5,874	5,874	5,874	5,874	5,874
Mean of Dep. Var.	0.098	0.081	0.147	0.062	0.101
First Stage	0.642	0.642	0.642	0.642	0.642
F Stats	77.827	77.827	77.827	77.827	77.827

Note: The Strengths and Difficulties Questionnaire (SDQ) total difficulties scores equals the sum of the four subscales: emotional symptoms, peer problems, conduct problems, and hyperactivity. The cutoff points used for abnormality for SDQ can be obtained from the guidelines at (https://www.sdqinfo.org/a0.html). I include the following set of baseline controls: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, ^{,a} ($^{,a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.19: Heterogeneous Treatment Effects on SDQ Abnormal Measures

		SDQ four subscales			
	(1) Abnormal SDQ total scores (1=Yes; 0=No)	(2) Abnormal emotional problems (1=Yes; 0=No)	(3) Abnormal conduct problems (1=Yes; 0=No)	(4) Abnormal hyperactivity problems (1=Yes; 0=No)	(5) Abnormal peer problems (1=Yes; 0=No)
By low-performing students in reading					
Assigned in-class library	-0.025	0.014	0.009	0.012	-0.054**
× low-performing students in reading	(0.023)	(0.021)	(0.025)	(0.016)	(0.024)
Assigned in-class library	0.001	-0.009	-0.005	-0.006	-0.019*
	(0.010)	(0.009)	(0.013)	(0.007)	(0.010)
Low-performing students in reading	0.075***,a	0.019	0.103***,a	0.004	0.106***,a
	(0.020)	(0.017)	(0.017)	(0.011)	(0.020)
Total effects for low-performing students	-0.025	0.005	0.005	0.006	-0.073***
	(0.022)	(0.019)	(0.025)	(0.017)	(0.026)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	5,865	5,865	5,865	5,865	5,865
Mean of Dep. Var.	0.098	0.081	0.147	0.062	0.101

Note: The Strengths and Difficulties Questionnaire (SDQ) total difficulties scores equals the sum of the four subscales: emotional symptoms, peer problems, conduct problems, and hyperactivity. The cutoff points used for abnormality for SDQ can be obtained from the guidelines at (https://www.sdqinfo.org/a0.html). I include the following set of baseline controls: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, ^{,a} ($^{,a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.20: Treatment Effects on SDQ Scores

		SDQ four subscales			
	(1) SDQ total scores (0 to 40)	(2) Emotional problems scores (0 to 10)	(3) Conduct problems scores (0 to 10)	(4) Hyperactivity problems scores (0 to 10)	(5) Peer problems scores (0 to 10)
Panel A: Intention-to-treat effect					
Assigned in-class library	-0.067	-0.022	-0.037	0.059	-0.066
	(0.223)	(0.076)	(0.075)	(0.076)	(0.062)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	5,876	5,876	5,876	5,876	5,876
Mean of Dep. Var.	12.877	3.291	2.667	3.595	3.325
Panel B: Local average treatment effect					
Built in-class library	-0.109	-0.039	-0.057	0.093	-0.106
	(0.345)	(0.119)	(0.117)	(0.121)	(0.094)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	5,874	5,874	5,874	5,874	5,874
Mean of Dep. Var.	12.875	3.289	2.667	3.595	3.323
First Stage	0.642	0.642	0.642	0.642	0.642
F Stats	77.827	77.827	77.827	77.827	77.827

Note: The Strengths and Difficulties Questionnaire (SDQ) total difficulties consists of the following four subscales: emotional symptoms, peer problems, conduct problems, and hyperactivity. The cutoff points used for abnormality for SDQ can be obtained from the guidelines at (https://www.sdqinfo.org/a0.html). I include the following set of baseline controls: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, $,^{a}$ ($,^{a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.21: Heterogeneous Treatment Effects on SDQ Scores

		SDQ four subscales			
	(1) SDQ total scores (0 to 40)	(2) Emotional problems scores (0 to 10)	(3) Conduct problems scores (0 to 10)	(4) Hyperactivity problems scores (0 to 10)	(5) Peer problems scores (0 to 10)
By low-performing students in reading					
Assigned in-class library	-0.189	-0.018	0.057	-0.008	-0.220**
× low-performing students in reading	(0.309)	(0.123)	(0.118)	(0.120)	(0.107)
Assigned in-class library	-0.076	-0.031	-0.064	0.052	-0.033
	(0.234)	(0.081)	(0.079)	(0.082)	(0.061)
Low-performing students in reading	2.194***^{,a}	0.536***^{,a}	0.672***^{,a}	0.380***^{,a}	0.606***^{,a}
	(0.266)	(0.095)	(0.097)	(0.095)	(0.092)
Total effects for low-performing students	-0.266	-0.049	-0.007	0.044	-0.253**
	(0.323)	(0.126)	(0.117)	(0.115)	(0.115)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	5,865	5,865	5,865	5,865	5,865
Mean of Dep. Var.	12.870	3.290	2.665	3.593	3.322

Note: The Strengths and Difficulties Questionnaire (SDQ) total difficulties consists of the following four subscales: emotional symptoms, peer problems, conduct problems, and hyperactivity. The cutoff points used for abnormality for SDQ can be obtained from the guidelines at (https://www.sdqinfo.org/a0.html). I include the following set of baseline controls: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, ^{,a} ($^{,a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.22: Heterogeneous Treatment Effects on Students Borrowing and Reading Behaviors

	Endline student borrowing behaviors		Endline student reading behaviors		
	(1) Never borrowed books from classroom library (1=Yes; 0=No)	(2) At least monthly bring home books from classroom library	(3) At least monthly read independent books after school (1=Yes; 0=No)	(4) Read independent books daily for 30 minutes or more (1=Yes; 0=No)	(5) The number of independent books read in this semester
By low-performing students in reading					
Assigned in-class library	0.078	0.058	-0.004	0.003	-0.064
× low-performing students in reading	(0.053)	(0.044)	(0.037)	(0.037)	(0.041)
Assigned in-class library	0.091	0.007	-0.005	-0.017	-0.022
	(0.084)	(0.073)	(0.028)	(0.031)	(0.047)
Low-performing students in reading	0.027	-0.109***,^{a}	-0.153***,^{a}	-0.135***,^{a}	-0.116***,^{a}
	(0.034)	(0.027)	(0.025)	(0.023)	(0.028)
Total effects for low-performing students	0.169*	0.065	-0.009	-0.014	-0.086
	(0.096)	(0.070)	(0.050)	(0.048)	(0.059)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	5,851	5,855	5,868	5,868	5,849
Mean of Dep. Var.	0.449	0.362	0.744	0.688	0.611

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Low-performing students in reading are those with baseline reading test scores ranked in the bottom 20 percent. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, ,^{a} (,$^{a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.23: Treatment Effects on SLR Outcomes, each question

		SLR individual questions									
	(1) Student Like Reading Scale (0 to 30)	(2) I read for fun	(3) I read things that I choose myself	(4) I like talking about what I read with other people	(5) I would be happy if someone gave me a book as a present	(6) I think reading is boring#	(7) I would like to have more time for reading	(8) I enjoy reading	(9) I learn a lot from reading	(10) I like to read things that make me think	(11) I like it when a book helps me imagine other worlds
Panel A: Intention-to-treat effect											
Assigned in-class library	-0.245	-0.091	-0.038	-0.030	-0.032	0.020	-0.034	-0.002	-0.021	-0.013	-0.005
	(0.343)	(0.073)	(0.055)	(0.038)	(0.033)	(0.032)	(0.043)	(0.045)	(0.034)	(0.035)	(0.032)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,866	5,879	5,878	5,873	5,878	5,876	5,878	5,878	5,878	5,876	5,876
Mean of Dep. Var.	21.353	1.699	1.667	1.739	2.407	2.469	2.058	2.204	2.484	2.279	2.344
Panel B: Local average treatment effect											
Built in-class library	-0.380	-0.141	-0.059	-0.046	-0.049	0.031	-0.052	-0.003	-0.034	-0.020	-0.008
	(0.550)	(0.120)	(0.087)	(0.060)	(0.053)	(0.049)	(0.069)	(0.070)	(0.055)	(0.055)	(0.050)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,864	5,877	5,876	5,871	5,876	5,874	5,876	5,876	5,876	5,874	5,874
Mean of Dep. Var.	21.354	1.700	1.667	1.739	2.407	2.469	2.058	2.204	2.483	2.279	2.344
First Stage	0.642	0.642	0.642	0.642	0.642	0.642	0.642	0.642	0.642	0.642	0.642
F Stats	77.760	77.786	77.756	77.862	77.816	77.833	77.816	77.816	77.748	77.729	77.746

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, $,^{a}$ ($,^{a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.24: Treatment Effects on SCR Outcomes, each question

		SCR individual questions					
	(1) Student Confident in Reading Scale (0 to 18)	(2) I usually do well in reading	(3) Reading is easy for me	(4) I have trouble reading stories with difficult words#	(5) Reading is harder for me than for many of my classmates	(6) Reading is harder for me than any other subject#	(7) I am just not good at reading#
Panel A: Intention-to-treat effect							
Assigned in-class library	-0.073	-0.042	-0.022	0.020	0.003	-0.018	-0.013
	(0.190)	(0.032)	(0.047)	(0.031)	(0.039)	(0.047)	(0.041)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,875	5,879	5,878	5,879	5,877	5,879	5,878
Mean of Dep. Var.	11.815	1.999	1.917	1.615	1.868	2.248	2.169
Panel B: Local average treatment effect							
Built in-class library	-0.114	-0.065	-0.034	0.031	0.005	-0.028	-0.020
	(0.300)	(0.052)	(0.075)	(0.048)	(0.060)	(0.074)	(0.065)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,873	5,877	5,876	5,877	5,875	5,877	5,876
Mean of Dep. Var.	11.815	1.999	1.917	1.615	1.868	2.248	2.169
First Stage	0.641	0.642	0.642	0.642	0.642	0.642	0.642
F Stats	77.604	77.786	77.748	77.786	77.692	77.786	77.738

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, ^{,a} ($^{,a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.25: Heterogeneous Treatment Effects on Students Learning from Independent Reading

	Endline reading interests and confidence		Endline reading beliefs		Endline reading skills	
	(1) Do not like reading (1=Yes; 0=No)	(2) Do not feel confident in reading (1=Yes; 0=No)	(3) Perceived negative effect of independent reading on language test (1=Yes; 0=No)	(4) Perceived negative effect of independent reading on math test (1=Yes; 0=No)	(5) Apply what learned from independent reading (1=Yes; 0=No)	(6) Standardized reading score (0 to 1)
By low-performing students in reading						
Assigned in-class library	0.002	0.024	-0.036**	0.011	-0.030	0.033
× low-performing students in reading	(0.022)	(0.022)	(0.015)	(0.024)	(0.026)	(0.065)
Assigned in-class library	0.010	-0.006	0.007	0.016	0.001	-0.008
	(0.010)	(0.016)	(0.004)	(0.012)	(0.023)	(0.043)
Low-performing students in reading	0.070***,[a]	0.136***,[a]	0.057***,[a]	0.073***,[a]	-0.010	-1.167***,[a]
	(0.017)	(0.014)	(0.013)	(0.018)	(0.019)	(0.044)
Total effects for low-performing students	0.012	0.019	-0.029*	0.027	-0.029	0.026
	(0.024)	(0.025)	(0.016)	(0.025)	(0.022)	(0.063)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,856	5,863	5,864	5,857	5,868	5,874
Mean of Dep. Var.	0.091	0.156	0.024	0.116	0.193	0.036

Table 2.26: Treatment Effects on the Perceived Effect of Independent Reading on Students' Academic Performance by Chinese Teachers and Primary Caregivers

	Endline chinese teachers' perceived no positive effect of independent reading		Endline primary caregivers' perceived no positive effect of independent reading	
	(1) on language test (1=Yes; 0=No)	(2) on math test (1=Yes; 0=No)	(3) on language test (1=Yes; 0=No)	(4) on math test (1=Yes; 0=No)
Panel A: Intention-to-treat effect				
Assigned in-class library	-0.012	-0.019	-0.031*	0.015
	(0.035)	(0.057)	(0.017)	(0.016)
Baseline student controls	No	No	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	202	197	4,475	4,401
Mean of Dep. Var.	0.059	0.152	0.219	0.785
Panel B: Local average treatment effect				
Built in-class library	-0.015	-0.023	-0.049*	0.023
	(0.046)	(0.075)	(0.026)	(0.026)
Baseline student controls	No	No	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	201	196	4,472	4,398
Mean of Dep. Var.	0.060	0.153	0.219	0.785
First Stage	0.757	0.756	0.642	0.645
F Stats	144.307	143.876	62.384	63.683

Note: For Chinese teachers' outcomes, only the teacher sample is used (without merging with the student survey) and no baseline student controls are added in the analysis. For primary caregivers' outcomes, the student sample is used. The following set of baseline controls are included for primary caregivers' outcomes: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, ^{,a} ($^{,a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.27: Heterogeneous Treatment Effects on Teachers Supporting Students' Independent Reading

	(1) Encourage students to borrow independent books (1=Yes; 0=No)	(2) Recommend several independent readings (1=Yes; 0=No)	(3) Assign independent reading homework (1=Yes; 0=No)	(4) Lead regular reading classes (1=Yes; 0=No)
By low-performing students in reading				
Assigned in-class library	-0.025	-0.022	0.022	0.006
× low-performing students in reading	(0.056)	(0.037)	(0.047)	(0.047)
Assigned in-class library	0.005	-0.078	-0.127	-0.173*
	(0.074)	(0.055)	(0.078)	(0.099)
Low-performing students in reading	-0.031	-0.042*	-0.069**	-0.052
	(0.035)	(0.022)	(0.029)	(0.034)
Total effects for low-performing students	-0.020	-0.100	-0.104	-0.167
	(0.102)	(0.075)	(0.094)	(0.103)
Baseline controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	5,868	5,867	5,868	5,868
Mean of Dep. Var.	0.665	0.863	0.722	0.424

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Low-performing students in reading are those with baseline reading test scores ranked in the bottom 20 percent. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, $,^{a}$ ($,^{a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.28: Heterogeneous Treatment Effects on Primary Caregivers and Peers Behaviors

	Primary caregivers' behaviors		Peers' interactions		
	(1) Spent money buying books (1=Yes; 0=No)	(2) Spend quality time studying with student (1=Yes; 0=No)	(3) Interact with peers in reading classes (1=Yes; 0=No)	(4) Discuss with peers what they don't understand in independent books (1=Yes; 0=No)	(5) Read independent books together with friends regularly (1=Yes; 0=No)
By low-performing students in reading					
Assigned in-class library	0.009	-0.020	-0.023	-0.059*	-0.002
× low-performing students in reading	(0.042)	(0.036)	(0.042)	(0.032)	(0.030)
Assigned in-class library	-0.017	0.008	0.036	-0.021	-0.031
	(0.027)	(0.022)	(0.076)	(0.032)	(0.024)
Low-performing students in reading	-0.120***	0.070**	0.007	-0.033	-0.008
	(0.029)	(0.029)	(0.032)	(0.021)	(0.025)
Total effects for low-performing students	-0.008	-0.012	0.013	-0.080**	-0.033
	(0.047)	(0.029)	(0.078)	(0.035)	(0.023)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	4,409	4,386	5,868	5,867	5,868
Mean of Dep. Var.	0.763	0.730	0.218	0.521	0.139

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Low-performing students in reading are those with baseline reading test scores ranked in the bottom 20 percent. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, $'^{a}$ ($'^{a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.29: Treatment Effects on Student Bullying and Sense of School Belonging Scales

	Endline student bullying		Endline student sense of school belonging	
	(1) Student Bullying scale (0 to 24)	(2) Student bullying weekly (1=Yes; 0=No)	(3) Student Sense of School Belonging Scale (0 to 15)	(4) Student with little sense of school belonging (1=Yes; 0=No)
Panel A: Intention-to-treat effect				
Assigned in-class library	0.182	0.001	0.073	-0.001
	(0.273)	(0.014)	(0.189)	(0.013)
Baseline controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	5,869	5,869	5,875	5,875
Mean of Dep. Var.	16.093	0.210	10.624	0.112
Panel B: Local average treatment effect				
Built in-class library	0.282	0.001	0.114	-0.002
	(0.416)	(0.022)	(0.290)	(0.021)
Baseline controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	5,867	5,867	5,873	5,873
Mean of Dep. Var.	16.093	0.210	10.624	0.112
First Stage	0.642	0.642	0.642	0.642
F Stats	77.954	77.954	77.804	77.804

Note: Both scales are internationally established metrics from PIRLS called the Student Bullying Scale and the Student Sense of School Belonging Scale. More details on the scale questions and cutoffs can be found at PIRLS official website. The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, ^{,a} ($^{,a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Table 2.30: Treatment Effects on Students Standardized Test Scores

	(1) Standardized reading score (0 to 1)	(2) Standardized chinese score (0 to 1)	(3) Standardized math score (0 to 1)
Panel A: Intention-to-treat effect			
Assigned in-class library	-0.024	0.046	0.016
	(0.049)	(0.047)	(0.068)
Baseline controls	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Observations	5,886	2,960	2,926
Mean of Dep. Var.	0.033	0.067	0.027
Panel B: Local average treatment effect			
Built in-class library	-0.035	0.071	0.028
	(0.078)	(0.072)	(0.106)
Baseline controls	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Observations	5,878	2,956	2,922
Mean of Dep. Var.	0.034	0.066	0.028
First Stage	0.642	0.639	0.645
F Stats	77.795	75.992	79.052
Panel C: by low-performing students in reading			
Assigned in-class library	0.033	0.257**	0.133*
× low-performing students in reading	(0.065)	(0.101)	(0.075)
Assigned in-class library	-0.008	0.011	0.011
	(0.043)	(0.041)	(0.061)
Low-performing students in reading	-1.167***,[a]	-0.968***,[a]	-1.064***,[a]
	(0.044)	(0.064)	(0.055)
Total effects for low-performing students	0.026	0.268**	0.144
	(0.063)	(0.104)	(0.089)
Baseline controls	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Observations	5,874	2,953	2,921
Mean of Dep. Var.	0.036	0.069	0.028

Note: The following set of baseline controls are included: student gender, age, grade, whether the student boards at school, whether at least one parent had nine years of schooling or higher, whether at least one parent had a professional occupation, whether both parents not at home for at least three months this year, and whether the student is from low-income families with no refrigerator at home. Low-performing students in reading are those with baseline reading test scores ranked in the bottom 20 percent. Standard errors are clustered at the school level (shown in parentheses). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. For relevant estimates, $,^{a}$ ($,^{a'}$) indicates that the null hypothesis is rejected at a 0.05 (0.1) level after correction for a false discovery rate (Benjamini and Hochberg, 1995).

Chapter 3

The Effect of Higher Education Expansion on Mental Health: Evidence from China

3.1 Introduction

Since the mid-20th century, higher education has expanded rapidly worldwide, beginning in the United States and Western Europe before extending to many developing regions like Africa and Asia (Schofer and Meyer, 2005; Carnoy et al., 2013). These expansions are widely viewed as key drivers of both individual and societal advancement. Existing studies have linked higher education and expansion efforts to improved labor market outcomes, broader economic growth (Hill, Hoffman and Rex, 2005; Kimenyi, 2011; Hout, 2012; Barrow and Malamud, 2015; Jia and Li, 2021), as well as reductions in mortality and improvements in physical health outcomes and behaviors (Grossman, 2006; Oreopoulos and Salvanes, 2011; Barrow and Malamud, 2015). However, as more countries pursue higher education expansion as a strategy for economic growth and social development, questions about the broader consequences of such reforms have gained attention. For example, some studies point to persistent challenges around educational equity and institutional capacity (Schendel and McCowan, 2016; Ferreyra et al., 2017; Luo et al., 2020). One emerging yet often underappreciated area of concern involves mental health. Individuals attending college often navigate this transition during their early adulthood—a period particularly vulnerable to the onset of mental health conditions, with consequences that can persist into later life (Leppink et al., 2016; Ribeiro et al., 2018; Haidar et al., 2018)[1].

In this paper, I provide causal evidence on the long-run mental health effect of increased college opportunities in China, focusing on the large-scale higher education expansion (HEE) in the late 20th century. This unexpected reform increased annual college admissions from 1 million to 1.5 million in 1999 alone, with enrollment numbers increasing (almost linearly) to 4 million by 2003 (see Figure 3.1)[2]. Representing one of the largest global expansions of higher education (Jia and Li, 2021), my findings show that greater exposure to HEE significantly increased college attainment. However, this expansion resulted in a higher frequency of reported depression symptoms and lower levels of subjective well-being in individuals' 30s. These patterns were more profound among individuals from lower socioeconomic backgrounds[3].

This study relies on two data sources. The first includes multiple statistical yearbooks and the 2000 Population Census, which provide information on college enrollment and the college-age population and are used to construct the HEE exposure indicators. The second combines five recent waves of the Chinese General Social Survey (CGSS) from 2010 to 2017, offering detailed individual characteristics, including self-reported mental health outcomes. To measure exposure to HEE, I construct a province-by-cohort intensity measure based on pre-reform enrollment shares and national expansion trends, capturing variation in college opportunities across provinces and birth cohorts.

Leveraging these variations, I use a cohort difference-in-differences (DID) strategy to estimate the effects of HEE (i.e., the intention-to-treat, ITT), following the approach developed by Duflo (2001)[4]. I provide evidence supporting the identification assumption that the trends in these

[1] For example, college students often face heightened psychological stress due to academic pressures, competitive environments, and the need to balance social, financial, and personal responsibilities. These challenges may be exacerbated by the rapid pace of educational reforms.

[2] HEE increased the college admission rates (calculated as the ratio of the number of newly admitted students each year by the total number of college applicants in that year) from 35% to 45% in 1999 (a 30% increase). The first wave of the high expansion rate roughly continued to 2003, doubling the college admission rates compared to the pre-expansion level (see Figure 3.1).

[3] Indicators used for lower socioeconomic background include low family social class when respondent aged 14, father (mother) as an agricultural worker when aged 14, father (mother) with no more than junior high schools.

[4] I provided a detailed explanation in Section 4 on why the primary focus of this paper is ITT rather than the local average treatment effect (LATE); the latter specifically focuses on compliers who respond to HEE. Nonetheless, I also presented empirical results for LATE in Section 5, particularly in comparison to OLS.

examined outcomes in lower-expanded provinces provide a good counterfactual for the trends in higher-expanded provinces in the absence of HEE. The main findings are robust to various additional checks, including using alternative HEE exposure measures and different specification/sample choices.

I also investigate the mechanisms driving the mental health effects of HEE and find out that these impacts were concentrated among individuals who were induced to attend college because of the expansion—the so-called "compliers". In contrast, the effects were not primarily driven by individuals who remained unable to attend college despite increased college opportunities (i.e., the never-goers) or those who would have attended college regardless (i.e., the always-goers). Additional evidence shows that compliers tended to come from relatively disadvantaged backgrounds. These patterns suggest that the psychological impacts of educational expansion may vary significantly across subpopulations. Importantly, the observed mental health outcomes do not appear to be inherent to college education itself, as always-takers were not similarly affected. Nor do they seem to be driven solely by experiences unrelated to college, such as pre-differences in mental health profiles or post-graduation labor market conditions, since individuals with no college attainment (e.g., those with at most a junior high or high school education) did not show worse mental health outcomes following HEE.

I further discuss several mechanisms that may help explain why this group was particularly vulnerable to these effects, though data limitations prevent direct testing of these specific pathways. In particular, I consider stressors they may have encountered in college, with consequences that persist into later life. These include academic unpreparedness, limited family support, peer pressure—factors that may influence self-esteem and perceived well-being over time. Post-graduation challenges may also play a role. Many compliers face relatively modest labor market returns—such as lower-paying jobs and slower career prospects—despite having earned a college degree. When these outcomes fall short of the elevated aspirations often associated with higher education, individuals may experience increased pressure or dissatisfaction, which could affect long-term mental well-being.

It is important to clarify that these findings do not imply college education expansion is inherently detrimental to mental health. Rather, they underscore the heterogeneous and context-dependent nature of mental health-related educational costs, particularly when expansions occur at a rapid pace. Such large-scale policy shifts implemented over a short period can involve complex trade-offs and uneven distributional consequences that merit careful consideration, especially in developing contexts. These results also underscore the importance of including mental health as a relevant outcome when evaluating the broader impacts of higher education policies. Targeted support may be needed for certain groups to ensure fully realizing the intended benefits of higher education expansions.

My study contributes to three strands of literature. First, I add to the literature on mental health determinants by providing evidence of how higher education expansion affects mental health. In the economics of happiness literature, commonly discussed factors for subjective well-being include (1) income and unemployment, (2) race and sex (and other demographic characteristics), (3) physical health, (4) family life, and (5) education (Clark, 2018). Income is the most widely studied factor among these individual characteristics, and the common belief is that individuals with higher levels of income are more likely to have higher levels of subjective well-being. But when it comes to education (and programs that promote education), the existing work has mixed findings (Clark, 2018)[5]. My work contributes to this literature by empirically showing that exposure to increased higher education opportunities does not uniformly improve mental health for all individuals in the long run.

Second, the paper contributes to the literature on the effects of HEE by evaluating its long-term impacts on mental health, a critical yet underexplored dimension of well-being. Most prior studies on HEE have focused on labor market outcomes, such as employment (Wu and Zhao, 2010; Li, Whalley and Xing, 2014), income (Li et al., 2017; Knight, Deng and Li, 2017), firm productivity (Che and Zhang, 2018; Ma, 2024; Feng, Tan and Wang, 2022; Feng and Xia, 2022). Few have looked at other dimensions of well-being like health (Fu et al., 2022), and none have examined mental health, to my best knowledge. Moreover, while most existing research focuses on short-term effects—typically within four years of graduation—I examine the longer-term consequences of HEE, approximately a decade later, when individuals are in their 30s. Therefore, my paper offers a broader perspective by examining whether HEE has long-term impacts, with a particular focus on mental health.

[5]According to Clark (2018), one point why education may not necessarily lead to better mental health could be due to the imbalance between raised expectation in general from education and real outcomes like incomes.

Third, my work offers additional insights to the existing literature on the relationship between education and health. For one thing, most studies focus on the low end of education using compulsory schooling reforms in developed countries, while I look at the college margin leveraging China's planned and top-down college admission system. For another, I also pay attention to mental well-being, while previous work tends to look at mortality (Lleras-Muney, 2005; Clark and Royer, 2013; Buckles et al., 2016), self-rated health status (Oreopoulos, 2006; Kemptner, Jürges and Reinhold, 2011; Brunello, Fabbri and Fort, 2013), and other physical health outcomes (?De Walque, 2007; Albouy and Lequien, 2009). My findings add to the debate on whether education affects mental health at the college margin, where some studies found protective effects of education on mental well-being (Crespo, López-Noval and Mira, 2014; Dursun and Cesur, 2016; Jiang, Lu and Xie, 2020; Kondirolli and Sunder, 2022), while others showed the opposite (Sironi, 2012; Dahmann and Schnitzlein, 2019; Avendano, de Coulon and Nafilyan, 2020), primarily in the context of lower levels of education.

The remainder of this paper is organized as follows. Section 2 reviews the relevant institutional background of China's college admission system and describes important details specific to the 1999 Higher education expansion. Section 3 introduces the dataset, outlines the construction of key measures, and presents summary statistics, including trends in key outcomes based on raw data. Section 4 describes the econometric framework, including the identification strategies and estimation methods employed. Section 5 then presents the results, examining the effects of HEE on college attainment and mental health, and explores potential mechanisms underlying these findings. Robustness checks for the main results are discussed in Section 6. Finally, Section 7 concludes the paper.

3.2 Institutional Background

3.2.1 China's Education System and College Admissions

Following the schooling system common to Western countries, students in China normally enter primary school at age 6 or 7 and spend six years there, followed by three years at junior high school. Nine years of compulsory schooling are required by law, while senior high school and post-secondary education are voluntary. Those who complete academic high school (typically in three years) are eligible to take the national college entrance exam for higher education, such as college and graduate schools. The National College Entrance Exam (NCEE) in China is mandatory for entering into almost all colleges and it takes place on July 6-9 each year (before 2003)[6]. Since it is virtually the only way that Chinese students can be admitted to higher education institutions and highly valued by cultural norms, there is typically more demand for college seats than supply, which was even more true in the 1990s and 2000s (Davey, De Lian and Higgins, 2007; Chen and Kesten, 2016; Fan et al., 2018; Cai et al., 2019; Jia and Li, 2021)[7]. There are also, in general, four tiers of higher education institutions (HEIs) in China to choose from (with selectivity descending in the tier numbers) (Loyalka, 2009). The Tier 1 group consists of about 100 colleges that are the most competitive and of the highest quality[8]. They are mainly four-year public universities under the central government's control, such as the Ministry of Education. Tier 2 universities are also four-year and public, but most are under the jurisdiction of provincial governments. Tier 3 is generally a private four-year institution, while Tier 4 is a three-year vocational college that can be public or private. In general, the last three tiers can only admit students from their own province, while Tier 1 can enroll high school graduates all over the country.

In addition to the high stakes and the selective nature of college admissions, there are several unique administrative features regarding the exam and college enrollment arrangement that are worth mentioning. First, college enrollment is ultimately under the control of the Ministry of Education (MOE) while subject to the overall development goals of the State Council. Provincial governments and universities are responsible for arranging the exam and the whole admission process, making sure to abide by MOE's higher-level plans. One key component of the higher-level plans is that every year before the exam starts in July, each university gets assigned an admission quota (as the cap) by the MOE, restricting the total number of students that can be admitted into its majors. Once the enrollment plan is assigned, universities are required not to make changes but strictly implement the plan[9]. After the exam, every province would announce an admission cutoff score for each tier of schools and track (either science or arts) based on the within-province exam

[6] After 2003, the exam dates changed to June 6-9 instead of July.

[7] For example, in 2014, there were 9.39 million test takers competing for about 7 million college spots. http://www.businessweek.com/articles/2014-06-06/china-girds-for-high-stress-gaokao-weekend.

[8] It is estimated that approximately 5% of candidates enroll into the top-tier universities every year (Jia and Li, 2021).

[9] If universities do have to make the changes in the enrollment plan, they must file a formal request to the central government and ask for permission.

scores distribution and total admission quotas. They then share approximately the quota number of students' application folders to the choices of universities within the specific tier[10]. Universities will make the admission decisions in priority order by admitting the students with higher rankings until the university-major quota is reached. Thus, the MOE plays a central role in determining the total number of college enrollments nationwide and its distribution across provinces.

Second, from the student side, they can only take the exam in their Hukou province (usually the same province where they were born) on the designated days. They also have to fill out an application form indicating a few universities in each university tier and several majors for each of these institutions that they are interested in attending and putting them in order, either before or after the exam. Usually, 20 days after taking the exam, they will receive at most one college-major admission offer based on their NCEE scores and their rankings within Hukou provinces. They then get to decide whether to accept that offer or not. If he/she decides not to accept the offer, he/she will not be able to go to any college that year. The university-major-specific offer (if accepted) is very likely to be the one that students are graduating with, since Chinese students do not often change majors in their college years and it is extremely rare for them to transfer between universities. Also, the college dropout rate is very low in China, unlike in other high-income countries, since college institutions do not "fail" students except in extreme circumstances.

3.2.2 The 1999 Higher Education Expansion in China

The nationwide higher education expansion (HEE) initiated in 1999 is the most large-scale and rapid higher education expansion in Chinese history and one of the world's biggest[11]. College admission rates rose sharply during this period. In 1998, the admission rate (calculated as the ratio of the number of newly admitted students each year to the total number of college applicants in that year) was 35%. By 1999, it had increased to 45%—a 30% relative increase—and further reached 60% by 2003, nearly doubling the pre-expansion level (see the solid line with point markers in Figure 3.1). The yearly number of newly admitted students followed a similar trend. It was at a

[10] About 10–20 percent more student files are allowed to be shared to ensure that there will be enough students to fill the quota given common mismatches between applicants and universities.

[11] In the early 1990s, the higher education total enrollment (per 1000 population) of China was about one-third of the US in the 1920s (Carnoy et al., 2013). Only about 2(3.5) million students attended 3-year or 4-year higher education in 1990(1998), but the number surged to about 11 million in 2003 (that of the US in the 1940s), only around four years after the first year of this expansion (see the dashed line with X-shaped markers in Figure 3.1).

level of 1 million between 1994-1998. But in 1999 alone, half a million more high school graduates were entered, and the number of entrants kept increasing (almost linearly) to 4 million throughout 2003 (see the dotted line with triangle markers in Figure 3.1).

China's HEE expansion was quite unanticipated, initiated by a letter from economist Dr. Tang Min and his wife in November 1998 advocating for increased higher education enrollment to boost economic growth. The proposal was soon adopted by the central government and quickly implemented across all provinces in 1999. The possible driving forces behind the higher education expansion may include an increasingly large demand for a higher quality labor force, the high private and social return to higher education, as well as the capability of colleges to expand at the time (Zhang, 2011). To implement this, the government adopted the top-down and strictly controlled approach by increasing the provincial enrollment quota directly and abruptly. According to Feng and Xia (2022), it is likely that the MOE only raised the growth rate of annual enrollment at the national level, without revising the benchmark enrollment quota distribution over provinces that were determined in the previous years, although the exact allocation rule was not publicly available. I showed evidence in the next section supporting this critical observation.

3.3 Data and Key Variables

3.3.1 Data Sources

The China Education Statistical Yearbooks from 1994 to 2004 were used to examine provincial-year variation in college enrollment. The 2000 Population Census provided data on the population aged 15-19 in each province in 2000 (as a proxy for other years). Both datasets were published by China's National Bureau of Statistics (NBS). I later combined the two to construct the expansion intensity measures, which are further explained in the next subsection.

I also pooled five waves of the Chinese General Social Survey (CGSS) from 2010 to 2017 (i.e., 2010, 2012, 2013, 2015, 2017) for a wealth of individual and family characteristics. The CGSS has been conducted by the National Survey Research Center at Renmin University of China annually since 2010 through face-to-face interviews aiming to systematically monitor the changing relationship between social structure and quality of life in both urban and rural China. The surveys follow a three-stage stratified design using Probability-Proportional-to-Size Sampling (PPS) based on GDP and other key metrics, covering almost all the provinces in mainland China[12].

3.3.2 Key Variables Construction

Individual characteristics and outcomes.— The primary data for individual characteristics and outcomes were drawn from survey questions in the CGSS dataset. Definitions for all main variables are presented in Appendix Table 3.11. Below, I describe several key dependent variables central to this paper. First, for college attainment, I created a dummy variable indicating whether people completed at least an academic college education, based on the highest level of education an individual received and whether they completed that level of schooling[13]. Second, for mental health outcomes, I started with the questions shared in all the five CGSS waves, recording the frequencies of depression in the past four weeks as well as self-reported general health status[14]. Specifically, respondents rated the frequency of mental health issues over the past four weeks on a scale from 1 (always) to 5 (never). General health status was asked as "How do you feel about your current general health? Very unhealthy 1 to very healthy 5". Using these, I created binary variables for better health outcomes, coded as 1 if respondents reported being "very healthy" or "healthy" (or answered "never" or "rarely" for the mental health question).

In addition, I made use of alternative mental health questions regarding subjective well-being (SWB), hope, optimism, happiness, and life satisfaction from a separate module available in the 2017 wave (not shared in all waves). Specifically, the SWB module was based on the subjective well-being scale for Chinese citizens (SWBS-cc20), consisting of 20 questions with ten experience dimensions and could be grouped into two broad categories (Huang and Xing, 2019)[15]. The optimism module used the "revised life orientation test LOT-R" (Scheier, Carver and Bridges, 1994),

[12]In general, 100 administrative counties were selected as primary sampling units (PSUs), 480 communities as the secondary sampling units (SSUs), and 12,000 households/individuals as the third-level sampling units (TSUs). In each sampled household, only one individual aged 18 and above was selected.

[13]Considering the different types of college education, I here only included academic ones, excluding vocational college and adult higher education. I argue this selection makes sense because HEE has the largest effect on the probability of attending a four-year college (Li, Whalley and Xing, 2014).

[14]Other health-related outcomes include the frequencies of health problems affecting daily life in the past four weeks, as well as heights (in meters) and weights (in kilograms). The latter two pieces of information were used to calculate an overweight indicator, defined as a Body Mass Index (BMI)—the ratio of weight to squared height—greater than or equal to 25.

[15]The two broad categories are "Experience of mental and physical health" and "Experience of satisfaction and growth." The ten experience dimensions are experiences of "Adaptation to interpersonal relation," "Mental health," "Goal and personal value," "Psychological balance," "Physical health," "Growth and progress," "Satisfaction and abundance," "Self-acceptance," "Confidence towards society," "Family atmosphere." The scale structure is summarized in Figure 3.8.

and the hope module used the "state hope scale" (Snyder et al., 1996). I later constructed three mental well-being indexes (i.e., SWB overall index, optimism index, and hope index) using the method proposed by Anderson (2008), which weighted down variables in the index proportionally to covariance with other variables. As for happiness evaluation, respondents were asked to rate a score (maximum 10, minimum 0) for their current happiness status and gave a cutoff score above which was considered happy for them. I combined the two questions and treated the person as happy if the happiness score was above (or equal to) his own cutoff score. Overall life satisfaction and marriage satisfaction were also considered in my analysis; both took values from 1 "very satisfied" to 5 "not satisfied at all" in the survey[16]. I created a dummy variable that equals 1 if the answer is "very satisfied" and "satisfied".

HEE exposure measure.— I defined the HEE exposure measure as the individual exposure to changes in college enrollment rate due to HEE in their province and expected college-going year (if after 1999). According to my definition, the HEE exposure would be zero for any cohorts who are expected to go to college before the policy year. In other words, given each individual's province and the expected college-going year after 1999, there was a non-zero rate of HEE-induced increases in college opportunities that the individual was exposed to in that province in that year.

Rather than using actual increases in college enrollment rates, which may be endogenous to both demand and supply factors, I constructed the exposure measure described below to isolate the plausibly exogenous supply-side shock. This approach leverages a key observation: provincial enrollment shares remained relatively stable over the years, even after the expansion. The Ministry of Education likely increased each college's enrollment quota in expansion using a national growth rate set by the State Council out of political and economic development motives. As a result, there was a persistence of enrollment distribution at the province level. The descriptive plot I provided below (see Figure 3.2) shows that the provincial enrollment share in the pre-expansion year of 1998 and post-expansion year of 2000 lay pretty close to the 45-degree line, supporting the critical observation mentioned above.

To construct the exposure measure, I first calculated each province's share of annual enrollments in the pre-expansion years, averaging from 1995 to 1998. Next, I computed the national increases

[16]The two questions go as follows: (1) All things considered, how satisfied are you with your life? (2) How satisfied are you with your marriage life?

in enrollment for each year from 1999 to 2004, adjusting for the linear pre-expansion trend. I then multiplied the two to obtain the predicted provincial increases in enrollment each year. Finally, I divided the predicted changes in enrollment by the total population aged 15 to 19 in each province in the year 2000. I obtained an HEE exposure measure varying only at the province level by averaging over the post-expansion period from 1999 to 2004[17].

Individuals' birth years and current province were used to define their exposure to higher education expansion. Specifically, I calculated the expected college-going years for cohorts based on their birth year as well as the typical timing of education in China. I assumed students started primary school at age six and went to college at eighteen. For example, the theoretical college year for the cohort born in 1980 was 1998 while the 1982 cohorts were expected to attend college in 2000. For provincial information, only their current provinces of residence were available and I could not identify the exact provinces where respondents attended college. While this is not ideal, I argue that migration is less of a concern in China compared to countries like the United States. This is largely due to the restrictive Hukou registration system, which limits interprovincial movement and often ties individuals to their localities[18]. In the subsection of the robustness checks, I also restricted my sample to respondents whose current province (at the survey time) was the same as the one they had at age 18; fortunately, the results are relatively robust.

3.3.3 Analytic Sample and Summary Statistics

The cohorts analyzed in this study were born between 1976 and 1986 and were expected to attend college between 1994 and 2004. Specifically, the 1981–1986 birth cohorts were defined as the treatment group, while the 1976-1980 cohorts as the control group. The 1981 cohort was the first to be affected because they were supposed to graduate from high school and took the college exam in 1999 when the massive HEE began. The youngest 1986 cohort was supposed to attend college in 2004, which was the last year in the first wave of massive HEE. The 1976 cohort was chosen so that I had the same cohort windows as the treatment group. Out of 31 provinces in mainland China, I excluded "Chongqing" because it was part of Sichuan before 1997 and thus had

[17] Without averaging over the post-expansion period, I got the predicted changes in college enrollment rate at the province and year level.

[18] Hukou system is also called the household registration system. It classifies people as rural and urban at birth, usually according to the mother's Hukou status.

fewer pre-expansion periods than the rest[19]. The final analysis sample consists of 9339 individuals from 30 provinces.

Table 3.1 presents the summary statistics of the sample. For respondents, the mean years of schooling were 10 to 11 years and only 14% of them attended/completed college. Most of them were from the Han ethnicity (90%), the largest ethnic group in China, and only a small share stayed religious (10%). Around half were female and had rural Hukou. In their 30s, the percentage of obtaining good general, mental, and physical health was 78%, 73%, and 88%, correspondingly; about 20% had obesity problems. As for family characteristics, most parents did not complete junior high school (thus categorized as having low levels of schooling), and half of the parents of the studied sample worked in agriculture when the respondents were 14.

Figure 3.3 shows the raw trends in the prevalence of college attainment as well as mental health scores between high expanded regions (provinces with the top third of HEE intensities) and low expanded regions (provinces with the bottom two-thirds of HEE intensities). This figure conveys the intuition of the empirical strategy discussed in the next section[20]. HEE intensities for all provinces (averaged over post-reform years) are displayed in Appendix Table 3.12. Appendix Figure 3.9 then visualizes the overall HEE intensity distribution for all provinces and all post-reform years, where the mean(median) value is 0.203(0.146).

3.4 Empirical Analysis

To estimate the long-term effects of HEE on mental health, I followed the cohort difference-in-differences (DID) approach. The design exploits both temporal and geographical variations in the individual exposure to changes in college enrollment rate driven by the rapid and unexpected nature of HEE. This section first discusses the identification strategy and then presents the estimation procedures.

[19] Excluding Sichuan together with Chongqing was considered as part of the robustness checks.

[20] One might notice that the 1996 cohort deviates slightly from the overall trend in college attainment and mental health scores for the high expanded group. However, I found no evidence of a pilot program or policy change during that period that specifically affected this group differently from the rest. For now, I interpreted this as a temporary fluctuation.

3.4.1 Identification Strategy

My identification strategy for estimating HEE effects built on two sources of variations. First, provinces had different HEE expansion intensities during the reform. My paper shows evidence that differences in expansion intensities stem from the initial provincial enrollment allocation, which was established long before the expansion and remained stable over time. This initial allocation was largely driven by geographical and political factors rather than economic and demographic conditions, where the latter are more directly related to college education and individual well-being (Feng and Xia, 2022; Ma, 2024)[21]. As a result, provinces that initially had larger pre-reform enrollment shares tended to increase more in enrollment during the reform. Second, within the same province, different birth cohorts were exposed differently depending on how their college-going age overlapped with the expansion year, which was likely to be exogenous given the unexpected nature of HEE as described in Section 2.2. Those who were expected to attend colleges after (before) 1999 were (not) exposed to the 1999 higher education expansion and comprised the treatment (control) group.

Following Duflo (2001), I started the analysis by constructing the interaction terms of the province-specific expansion intensity (i.e., the predicted changes in enrollment rate specific to the province) multiplied by the treated cohorts dummies as the key explanatory variables to illustrate the intuition for the research design and test for pre-trends regarding key outcomes. The detailed regression specification was presented in the next subsection. Intuitively, I estimated a DID model that compares the mental health outcomes (and college attainment) of those exposed to HEE to non-exposed ones between provinces with different expansion intensities. The identifying assumption is that the trends in these examined outcomes observed in relatively lower-expanded provinces provide a good counterfactual for the trends in higher-expanded provinces in the absence of HEE. I provided graphical evidence in support of this assumption in the results section. Given the possibility of province-year confounding factors, I also used the region-by-cohort trend (region-by-cohort FE in robustness checks) to account for regional-specific differential time differences in the outcomes.

It is important to note that the primary focus of this paper is on ITT effects, which captures

[21] Ma (2024) shows that the initial college endowment is not correlated with the trend in economic development and population by plotting the changes in log GDP (population) between 1982 and 2005 over log college enrollment of 1982 for each city in China.

the effects of the college expansion, including those who, despite increased opportunities, still do not enroll (never-takers)—a segment comprising more than half of the total cohort population. Another relevant estimate is the local average treatment of effects (LATE) which focuses on the effects of college education, only relevant for individuals who respond directly to expanded access by attending college (compliers). While both ITT and LATE estimates provide valuable insights into different aspects of HEE, I argue that in this context, ITT offers a broader perspective on the societal impact of educational expansion policies, which is the central focus of this paper. Nonetheless, in Section 5, I also presented the LATE estimates of the causal relationship between college attainment and mental health outcomes within the context of HEE.

With advancements in the DID methodology—particularly in addressing staggered treatment timing and heterogeneous treatment effects—it is important to note that my setup minimizes concerns about staggered rollout issues. This is because all provinces received treatment simultaneously. For example, college expansion occurred across all provinces for the 1999 cohort, though with different treatment intensities. This contrasts with a scenario where province A started expanding college access in 1999 while province B started the expansion in 2000.

3.4.2 Estimation Strategy

Building on the cohort DID specification described earlier, I leveraged additional variation in expansion intensities from the post-expansion period by directly using the HEE exposure measure, which varied at the province year level. The idea here is that within a province, later exposed cohorts might well face different exposure intensities than early exposed cohorts. For example, cohorts who were expected to attend college in the year 2002 in the province of Zhejiang may be more intensively exposed to HEE than cohorts who were expected to attend college in the year 2000 in the same province.

I then estimated the HEE effects on mental health outcomes using this exposure measure, applying the following equation as my main specification:

$$Y_{ipcw} = \beta_0 + \beta_1 Expansion_{ipc} + \beta_2 X_{ipcw} + \theta_p + \delta_c + \gamma_w + \lambda_r \times cohort_c + \epsilon_{ipcw} \tag{3.1}$$

where Y_{ipcw} denotes the college attainment and mental health outcomes of individual i in province p expected to go to college in year c from wave w. HEE exposures are directly measured by HEE intensity in province p for cohort c. X_{ipcw} is a vector of individual-level controls, including gender, father's education, the logarithm of real GDP per capita (and government expenditure per capita) of local provinces when the individual is age 18, and the number of hospital beds (per 10,000 persons). θ_p is the province fixed effect that controls for all the cohort-invariant but province-specific (un)observables that may affect an individual's education or mental health. δ_c is the cohort fixed effect where common trends in education or mental health across provinces are captured. γ_w is the wave fixed effects controlling for any variation between survey waves. Unobservable heterogeneous cohort trends that may be correlated with the expansion intensities remain a crucial concern to the cohort DID strategy. To alleviate some of the concerns, I also added the region by cohort trend to allow for regional-specific differential trends in the outcomes ($\lambda_r \times cohort_c$). The parameters of β_1 capture the ITT effects of HEE on education and mental health outcomes in the treated periods. Standard errors are clustered at the province level.

I also presented results using the standard cohort DID specification described in the previous subsection as robustness checks. An individual's HEE exposure status is alternatively measured by an interaction between (mean) HEE intensity in his province and a dummy variable for being treated cohorts. The treated group here are cohorts expected to go to college between 1999 and 2004, with the remaining cohorts being the control group; all else are defined as in equation (1).

$$\begin{aligned} Y_{ipcw} = \beta_0 + \beta_1 AvgExpansion_{ip} \times I(1999 \leq c \leq 2004) + \beta_2 X_{ipcw} \\ + \theta_p + \delta_c + \gamma_w + \lambda_r \times cohort_c + \epsilon_{ipcw} \end{aligned} \tag{3.2}$$

A by-cohort specification was employed to visualize the pre-trend and heterogeneous treatment effect by cohort. Instead of having a single indicator for all cohorts expected to go to college after 1999, each cohort now has its dummy variable indicator. The cohorts who expected to go to college in 1998 serve as the baseline, and the parameters of $\beta_{1,\gamma}$ capture the effects of HEE on education and mental health outcomes for each cohort.

$$Y_{ipcw} = \beta_0 + \sum_{\gamma=1994}^{2004} \beta_{1,\gamma} AvgExpansion_{ip} \times I(c = \gamma) + \beta_2 X_{ipcw} \tag{3.3}$$
$$+\theta_p + \delta_c + \gamma_w + \lambda_r \times cohort_c + \epsilon_{ipcw}$$

To estimate the LATE , I used a two-stage least squares (2SLS) estimator. In the first stage, I regressed the college attainment indicator on expansion intensities with additional controls as specified in equation (1). In the second stage, I regressed the outcome variables on the predicted values from the first stage, including the same set of controls. Essentially, LATE is derived by scaling the impact of college expansion on the outcome (ITT) by the effect of college expansion on college attainment.

3.5 Results

This section reports and discusses the results. I started by presenting the results graphically, followed by the main table results which examined the effect of HEE on mental health outcomes. I then provided additional analyses on socioeconomic heterogeneity, labor market outcomes and family life decisions, offering insights into mechanisms. Last, I systematically discussed the possible mechanisms that might help explain the observed effects on mental health.

3.5.1 Main Results on College Education and Mental Health

Figure 3.4 shows the by-cohort ITT estimates of $\beta_{1,\gamma}$ from equation (3) that compares the prevalence of college attainment across provinces with different HEE expansion intensities for each cohort expected to attend college before and after the introduction of the program. It is clear that none of the coefficients in the pre-period are statistically significant at the 5% level. But starting in 1999 and onwards, the college attainment trends in high expanded provinces are significantly higher than in lower expanded provinces at 5% level. The differences roughly stayed the same during the post-expansion years that I looked at. I also estimated the equation (1) specification on college attainment and found that individuals who got more exposed to HEE at the time were more likely to attain college (the coefficient, 0.081, is positively significant at 1% level).

Table 3.2, Panel A presents the main results of the paper, examining the ITT effects of the HEE on mental health outcomes. Specifically, column (1) analyzes whether an individual self-reports having good mental health, based on a question about the frequency of depression in the past four weeks. Column (2) extends the analysis by including whether the individual self-reports general health status as good, which is closely related to mental health. Interestingly, instead of seeing positive HEE effects on mental (and general) health, I found the exact opposite. Provinces with larger increases in college opportunities for eligible cohorts show worse mental (and general) health, compared to provinces with smaller expansions. The adverse effects of the HEE on mental health are particularly evident in Figure 3.5 which shows the by-cohort estimates on good mental health based on equation (3).

What's more, I used additional mental health measures from the 2017 survey wave by applying the baseline specification to other mental health measures, including subjective well-being (SWB), hope, optimism, happiness and life satisfaction. These alternative mental health measures offer an opportunity not only to assess the reliability of the primary outcome measure but also to gain deeper insights into which specific aspects of mental health are affected. The results are presented in Table 3.3[22]. Similar to the negative HEE effects on depression frequency observed in Table 3.2 using the pooled sample, Table 3.3 shows that exposure to college opportunities leads to significantly more frequent depression (column 1) and worse overall subjective well-being (column 2) in the 2017 wave. For indexes of optimism, hope and happiness (columns 3 to 5), the estimates are also negative, though not statistically significant at the 5% level. In contrast, for life evaluation outcomes, I saw insignificant but positive estimates (shown in columns 6 and 7). This differing pattern for life satisfaction is not entirely surprising, particularly given the substantial economic development in recent years in China. Besides, research by Hovi and Laamanen (2021) suggested that life evaluation or satisfaction measures might capture different dimensions of well-being compared to the rest; for example, Kahneman and Deaton (2010) showed that high income could buy life satisfaction but

[22]According to the recent economics of well-being literature, subjective well-being can be loosely speaking divided into three types - (i) cognitive, (ii) hedonic and (iii) eudaimonic (Hovi and Laamanen, 2021; OECD, 2013). For example, OECD (2013) lists that the broad definition of subjective well-being encompasses three elements: "(i) Life evaluation – a reflective assessment of a person's life or some specific aspect of it, (ii) Affect – a person's feelings or emotional states, typically measured with reference to a particular point in time, and (iii) Eudaimonia – a sense of meaning and purpose in life, or good psychological functioning." Arguably, I also grouped the available well-being measures into two broad groups; one on life satisfaction and the other on overall emotional well-being, including both types (ii) and (iii).

not happiness.

In addition to these ITT effects, I reported and compared the relationship between college attainment and mental health outcomes using OLS and 2SLS, as shown in Panels B and C of Table 3.2, respectively. The OLS results in Panel B, where mental health outcomes were regressed on a college attainment dummy variable, show that people who graduated college tend to have better mental and general health (though not significant at the 5% level). These results largely align with the positive correlations commonly seen in the education and health literature (Cutler and Lleras-Muney, 2006, 2012). However, it is important to note that these OLS estimates can rarely be interpreted as causal, since people often self-select into different levels of education. Panel C of Table 3.2 reports the 2SLS estimates, using HEE as an instrument variable (IV) for college completion. The results reveal a strikingly different relationship between college attainment and mental health compared to the OLS findings. Specifically, the coefficients point in the opposite direction, challenging the positive associations observed in the correlational analysis discussed earlier[23].

There may be some concerns regarding weak instruments, as indicated by first-stage F-statistics (about 9.768) slightly below the conventional threshold of 10[24]. However, I argue that this weak IV issue is less problematic in this context. Specifically, weak instruments typically result in larger standard errors and estimates biased toward the OLS results. Table 3.2 indeed shows much larger 2SLS standard errors than those from OLS. Yet, despite this, the 2SLS results remain statistically significant at the 5% level. Moreover, the 2SLS estimates reverse the direction of the OLS estimates, suggesting that any bias towards OLS results would likely lead to an underestimation of the negative effects in the 2SLS estimates. This further strengthens confidence in the robustness of the findings.

3.5.2 Additional Results on Heterogeneity, Labor Market and Family Outcomes

This section presents additional findings to help identify potential mediators of the observed HEE effects on mental health, by examining socioeconomic heterogeneity, labor market outcomes and family life decisions. A deeper exploration of possible mechanisms is discussed in the next

[23]Table 3.2 also includes estimation results for physical health (column 3, based on a question asking the frequency of health problems affecting daily life in the past four weeks) and obesity (column 4, based on BMI calculated from current height and weight). These results similarly show no positive effects of the HEE and reveal consistent opposing patterns between the OLS and 2SLS results.

[24]Recent research also indicates that the common standard for acceptable instrument strength—a first-stage F-statistic of 10—may be too low and that a higher threshold is necessary (Keane and Neal, 2024).

subsection.

Table 3.4 presents the heterogeneous effects of HEE on mental health by family background. Notably, the adverse HEE effects on mental health are more severe for the groups with poor socioeconomic status. Specifically, I estimated the heterogeneous effect of HEE on mental health by pre-determined family background, including low family social class when the respondent is aged 14, father as an agricultural worker when aged 14, father (mother) with no more than junior high schooling. Consistent through all subgroup indicators, the effect on groups with poorer family backgrounds is much more negative in values, and most are also significantly different than their counterparts.

Table 3.5 shows the HEE effects on labor market outcomes and family life factors using the baseline specification. The results indicate that HEE positively affects log annual personal income at the 10% level (column 3) but does not significantly impact labor force participation and high-skilled occupation take-up, though with positive estimates (columns 1 and 2). For family-related outcomes, larger exposure to HEE leads to reduced/delayed marriage (columns 4 and 5) and childbearing (column 6) among individuals in their 30s. These findings highlight the broader life changes from HEE, from labor market participation to family life decisions[25].

3.5.3 Mechanisms

What explains the causal effects of HEE on mental health in individuals in their 30s? This section systematically explores potential mechanisms. First, these effects may be driven by those unable to access college despite the expanded opportunities (never-takers). Second, individuals induced to attend college by HEE (compliers) could have encountered unique challenges during their additional college-related experiences. Third, college education itself might affect mental health through mediating factors such as earnings, work conditions, health behaviors and lifestyle changes, affecting all attendees, including always-takers—those who would have attended college regardless. I showed the evidence below that the observed HEE effects on mental health were not driven by non-college-goers/never-takers or by a general causal effect of college education on mental health. Instead, the results point to the unique vulnerabilities of compliers going through college,

[25] An important caveat to note is that the labor market outcomes used here are all based on self-reports which could suffer from nontrivial measurement errors.

particularly those from disadvantaged backgrounds. The final part of this section then discusses several specific pathways that help explain why poor compliers were more affected.

If Never-Takers Lacking College Access Drive the Effects?

One possibility is that the observed HEE effects on mental health are driven by those who still cannot attend college, given the increased college opportunities from HEE. If this were the case, I would expect worse mental health outcomes only for those who attained high school. In contrast, individuals with at most junior high school education, who likely had limited prospects for college, and those who attained college or beyond would not be expected to experience similar effects. Table 3.6 provides suggestive evidence based on reported schooling information, that helps rule out this explanation.

Specifically, I run the baseline specification on mental health outcomes for three education subgroups: those with at most junior high school (column 1), completed high school (column 2), and attended college or beyond (column 3). An additional column (column 4) aggregates all non-college-goers. The results in column 5 specifically show that college-goers, rather than non-college-goers, experience worse mental health outcomes, suggesting that the negative effects are not driven by those excluded from college opportunities[26].

If Compliers Drive the Effects?

Another explanation to consider is that compliers may be more vulnerable to psychological challenges associated with college-related experiences due to certain underlying factors—for example, less advantaged family backgrounds and lower levels of academic preparedness. This hypothesis implies that compliers would primarily drive the observed effects, which is partially supported by the 2SLS results among compliers from Panel C of Table 3.2. To further investigate, I characterized compliers based on available data and examined whether they disproportionately contribute to the negative HEE effects on mental health.

First, I predicted compliers for the post-HEE cohorts and compared their characteristics with those of non-complier groups (as shown in Table 3.7). Specifically, predicted compliers (column 4)

[26] A caveat is that the subgroup indicator is likely affected by the treatment which could lead to biases, and thus should be interpreted cautiously.

tend to have better socioeconomic backgrounds than predicted never-goers (column 8) but worse backgrounds than predicted always-goers (column 6)[27]. T-test results for differences between compliers and always-goers (column 9) and between compliers and never-goers (column 10) are also reported. I also included the mean characteristics of the overall post-HEE cohorts to get a sense of how each group compares to the average (column 2). Details of the prediction methodology for the different groups are provided in the Appendix B.

Second, I checked whether compliers drive the observed HEE effects on mental health. In other words, do always-goers also experience negative mental health effects from HEE, similar to compliers? If both groups show adverse effects, it could suggest that college education itself causally influences mental health. However, if only compliers are negatively affected while always-goers are not, this would support the hypothesis that the effects are driven by characteristics unique to compliers. To test this, I compared outcomes for predicted compliers in the post-HEE period with matched individuals from the pre-HEE period[28].

Table 3.8 shows the results based on predicted always-goers and compliers. Compliers experienced worse mental health following HEE (columns 1 to 3), whereas always-goers show no significant negative effects (columns 4 to 6). As a robustness check, I compared the groups based on college-going predictions using only pre-expansion data. Assuming defiers are negligible, those predicted to go to college based on pre-expansion data are likely to be always-goers. For this group, I also found no evidence of negative mental health effects from HEE (Appendix Table 3.13). Moreover, Appendix Table 3.14 provides further evidence that the negative effects among compliers are concentrated among those from lower socioeconomic status (SES) backgrounds. Taking together, I showed suggestive evidence that compliers (especially those from disadvantaged backgrounds) likely drove the observed HEE effects on mental health[29].

[27]Overall, compliers are better off than the treated sample average (column 2), which makes sense since never-goers take up a large portion of the sample.

[28]For matching, I used a 1-to-1 nearest-neighbor propensity score matching (PSM) method without replacement. The variables included in the propensity score estimation are gender, ethnicity, religion, political affiliation, and parents' educational levels and job types when the respondent was 14 years old.

[29]Keep in mind that these findings should be interpreted with caution, as the results rely on good group predictions and good matches to the pre-HEE cohorts. Limitations include the restricted set of covariates available for prediction and matching, the convenient probability cutoff, and the risks of a lack of common support.

If College Education Causally Affects Mental Health?

A final explanation considers the mechanisms by which college education may causally impact mental health outcomes, for which I provided evidence against. First, as shown earlier, the observed HEE effects on mental health are primarily driven by compliers rather than always-goers. This argues against the explanation that college education worsens mental health for everyone who goes to college. The heterogeneous analysis by socioeconomic status further challenges this story. Second, based on the tests examining whether HEE affects factors such as labor market outcomes (Table 3.5) and work stress (Appendix Table 3.15), I did not see meaningful effects that would be associated with worse mental health. For marriage market outcomes, larger exposure to HEE overall leads to reduced/delayed marriage (columns 4 and 5) and childbearing (column 6) in one's 30s shown in Table 3.5. While the sign could match up with the belief that being married (or having a child) is systematically correlated with better well-being, I argue that this mechanism alone is not sufficient. It does not explain why individuals from lower socioeconomic status (SES) backgrounds are disproportionately affected, especially since they are not necessarily less likely to marry or have children compared to their higher SES counterparts—either on average or as a result of HEE (see Table 3.9 and Appendix Table 3.16)[30].

Why Are Poorer Compliers Disproportionately Affected?

The results above suggest that the mental health effects of HEE are not driven by never-takers or by college education per se. Instead, they highlight the unique vulnerabilities of compliers—those who enrolled due to the expansion—particularly individuals from disadvantaged backgrounds. What, then, explains the disproportionate impact on this group? One possibility is that these individuals may struggle to graduate. However, this explanation seems less plausible given the structure of the Chinese college system. Unlike in the U.S., admission to Chinese colleges is extremely competitive, but graduation is relatively straightforward. Nearly all admitted students successfully graduate, with dropout rates as low as 1% to 3% (Jia and Li, 2021). It is also unlikely that the observed effects are solely due to compliers having inherently worse mental health profiles

[30] For brevity, I showed here the effects of HEE on marriage and child status by father as an agricultural worker when respondents aged 14 (see Table 3.9). Results remain similar using other family background measures (Appendix Table 3.16).

unrelated to college education. If that were the case, individuals with even lower educational attainment, such as those with at most a junior high school education, would be expected to exhibit worse mental health outcomes, which is not what I see in Table 3.6.

However, it is likely that these individuals faced greater psychological costs compared to their always-taker peers, during college and after graduation, leading to less favorable mental health outcomes in their 30s. Several underlying pathways may help explain this pattern, although data limitations prevent direct testing of these specific mechanisms. One possibility is that they often enter college with weaker academic preparation and fewer family resources, making them more vulnerable to academic stress and unfavorable social comparisons. Surrounded by higher-achieving peers, these individuals may experience a decline in self-esteem or a diminished sense of self-worth—factors strongly linked to poorer mental health outcomes later in life (Hautala and Lehti, 2024).

In addition, compliers—particularly those from disadvantaged backgrounds—may confront less favorable labor market outcomes and additional post-graduation difficulties, falling short of their high aspirations from earning a college degree. In China, a college degree is widely viewed as a pathway to success, higher income, and elevated social status. When these expectations go unmet, individuals may internalize a sense of failure or inadequacy or ongoing pressure to "catch up", ultimately leading to lower life satisfaction and poorer mental health.

Different factors contribute to these less favorable outcomes experienced by some individuals after graduation. Under HEE, many individuals were funneled into second- or third-tier universities, which typically offer lower educational quality and confer weaker labor market advantages compared to top-tier institutions in China. As a result, these graduates are more likely to obtain jobs with modest paying and limited prestige, which can constrain their career advancement and upward mobility. Moreover, many compliers—especially those from poor and often rural backgrounds—face additional post-graduation challenges, such as hukou-related barriers and difficulties migrating into urban social environments. These disappointing outcomes and persistent structural barriers—when contrasted with the high societal expectations associated with a college degree—may contribute to chronic stress and help explain why the mental health impacts of HEE appear to be concentrated within this group.

3.6 Robustness

This section presents the robustness checks for the main results. The set of checks can be grouped into (i) alternative HEE exposure measures, (ii) different specification choices, and (iii) different choices of samples. Additional tests are shown in the appendix.

3.6.1 Alternative HEE Exposure Measures

Using interaction terms.— I first showed that the results remained robust when using alternative HEE exposure measures with interaction terms, as discussed in Section 4.1. The interaction term is defined as the province-specific expansion intensity multiplied by the treated cohorts' dummies. The estimation equation follows equation (2) in Section 4.2, with results shown in Panel A of Table 3.10. Consistent with previous findings in Table 3.2, Column (3) in Table 3.10 shows that HEE significantly worsens an individual's mental health regarding depression frequency.

Alternative denominator for exposure measure.— To assess the sensitivity of the HEE exposure measure constructed in equation (1), I replaced the denominator—the population aged 15 to 19 in 2000—with the population aged 20 to 24 in 2000. The results remain robust, as shown in Panel B of Table 3.10.

3.6.2 Different Specification Choices

Using region by cohort FE.— In the main specification, I used region by cohort trend to absorb regional-specific differential trends in the outcome that could be potentially correlated with HEE. In consideration of nonlinear time-varying changes in outcomes that are regionally specific, I also tried using region by cohort FE (see Table 3.10, Panel C), which did not affect the main results much.

Controlling for Hukou status.— The Hukou household registration system is unique in China and plays a very important role in social mobility. The system rigidly divides people into either rural or urban subgroups and strongly favors urban counterparts in the sense that they get to receive higher quality social welfare benefits regarding education resources and health care. The Hukou status one can get may well be associated with their education choices and lifelong health status. Here I showed the results were robust if I added control of Hukou status (see Table 3.10, Panel D).

3.6.3 Different Sample Choices

Restricted to those not migrated after 18.— I also tried to impose stricter assumptions on migration regarding residence and/or Hukou status changes to address some of the concerns related to the data limitation that I did not observe birth province and birth Hukou status. Specifically, I combined the information on their current residence (and Hukou) status information with the change history to create sample indicators of those who never changed residence or Hukou status since age 18. I later restricted my sample to those who did not change their residence county (Table 3.10, panel E) as well as to those who did not change their Hukou status and stayed in their current county (Table 3.10, panel F)[31]. The results are relatively robust for the two restricted samples. If anything, the higher education expansion seems to have a strong effect on increasing college attainment and worsening mental health.

Excluding provinces.— Also, to be sure that my results are not driven by a small number of provinces, I re-estimated the main specifications multiple times, each time excluding one province. Results on college completion and good mental health are shown in Figure 3.6. It is clear that among all 31 trials, most estimates are aligned closely and have the expected sign (with positive estimates on college and negative estimates on mental health). They are all statistically significant at 95% level since the confidence intervals do not seem to include zero[32].

With more pre-expansion periods.— In the main analysis, I focused on cohorts that were expected to attend college between the years 1994 and 2004 and showed no trend in the pre-period, i.e., from years 1994 to 1998 (in total four years). I now extended the pre-period back to the year 1990 and re-checked the pre-trends according to cohorts with theoretical college-going years between 1990 and 1998 (in total eight years). As expected, I did not see clear converging or diverging pre-trends in the key outcomes, shown in Figure 3.7.

[31] The two are more restricted than necessary since I only need individuals not to migrate to other provinces (not another county).

[32] Excluding Beijing seems to lead to a bit different estimates. But the sign is less of a concern, since it even strengthens the findings.

3.7 Conclusion and Discussion

Despite the global expansion of higher education, few studies have examined its impact on mental health in developing contexts. This paper analyzes the effects of a large-scale and rapid college expansion program starting in 1999 in China, providing evidence on how education expansion policies affect mental well-being in the long term. Using a cohort difference-in-differences strategy following (Duflo, 2001), I find that greater exposure to HEE significantly increased college attainment. However, the expanded access through HEE also led to a higher frequency of reported depression symptoms and lower levels of subjective well-being later in life, especially for the group with poorer socioeconomic background (despite no clear worsening in their labor market outcomes).

To better understand these patterns, I explore the channels through which HEE may have influenced long-term mental health. The evidence suggests that compliers who get to attend college after expansion but not before likely drive the negative HEE effects on mental health. In contrast, always-takers, who would have attended college regardless, and never-takers, who remain unable to attend despite the increased opportunities, are less likely to drive these outcomes. I also show evidence that compliers were often more socially disadvantaged than their college peers.

Given these results, it is unlikely that the observed effects are causally driven by college education itself. The absence of similar effects among always-takers, as well as the lack of evidence through common mediating factors such as income, support this interpretation. I also argue that the impacts are unlikely to be solely attributable to pre-existing differences in mental health profiles or to post-college labor market conditions. This view is further supported by the finding that individuals with no college attainment, such as those with at most a (junior) high school education, did not exhibit worse mental health outcomes from HEE.

Instead, HEE likely exposed compliers to greater psychological costs linked to their college-related experiences, though the available data do not allow for direct testing of specific pathways. These challenges may begin during college, with mental health consequences that persist into their 30s. Contributing factors could include weaker academic preparation, limited family resources, and greater exposure to stressors such as peer comparison. These experiences can lead to declines in self-esteem or even self-worth over time—factors strongly linked to poorer mental health outcomes later in life. After graduation, additional difficulties may arise. For example, many compliers were funneled into second- or third-tier colleges, where lower educational quality and weaker labor market signaling can lead to less favorable employment outcomes, falling short of rising aspirations tied to earning a college degree. Also, those from disadvantaged backgrounds may face more barriers after graduation, such as difficulties with urban migration and integration, contributing to long-term psychological stress and dissatisfaction.

There are several limitations to this paper that are worth mentioning. First, given the data constraints, I can only examine a few mental health outcomes from self-reports and not at all psychiatry measured, which at best reflect the risk of certain conditions. Mental health reporting varying by education may also raise concerns about biases in the estimates (Bertrand and Mullainathan, 2001; Spitzer and Weber, 2019; Spitzer, 2020). This may not be a big issue in this study, though, since highly educated people tend to report having better mental health. Thus, if anything, my negative HEE estimates on mental health are viewed as lower bounds, and true effects can be even more adverse. In addition, the mental health outcomes I used are far from comprehensive and may not be able to cover various aspects of mental well-being. For example, some papers (Kahneman and Deaton, 2010; Hovi and Laamanen, 2021; Clark and Lee, 2021) suggest that eudaimonia and emotional well-being might be other important dimensions of well-being that are not often well captured by life evaluation measures[33]. This can be an issue since mental health is multidimensional, and HEE may be responsible for some dimensions but not others. Future work could try to replicate the results by employing better and broader mental health measures.

Second, while the evidence in this study supports the hypothesis that individuals who attended college only because of HEE faced higher psychological costs in their college-related experiences, the analysis is not yet able to definitively confirm this. A more direct and refined categorization of college-goers and compilers, ideally with data about academic performance or ability, can better test whether these individuals are indeed driving the effects. If they are, detailed information on their college and post-graduation experiences would be essential to directly assess the psychological costs induced by HEE and better identify the mechanisms at play, such as the tier of college attended, quality of peer networks, field of study, and early career trajectories. Besides, the lack of more precise and objective measures of labor market outcomes, as well as detailed data on health

[33] A sample question for eudaimonia is as follows: "Overall, to what extent do you feel the things you do in your life are worthwhile." Emotional well-being instead focuses on the frequencies of experiencing different emotional states, such as nervousness, down, and calm.

behaviors and lifestyles, limits the ability to fully investigate potential mediating factors linking education and mental health. Future research should focus on identifying improved data sources and exploring ways to confirm the underlying mechanisms.

Overall, my empirical findings from a developing context suggest that while large-scale and rapid college expansion increases access to higher education, it also possibly introduces underappreciated long-term costs, particularly to well-being, for relatively disadvantaged students who newly gain college access under such expansions. This vulnerability to costs is unlikely inherent to college itself or from experiences unrelated to college. Rather, it is more likely shaped by the challenges of navigating newly available college-related experiences during and after college, challenges that reflect a complex interplay of factors, including demographic characteristics, family background, academic preparation, peer network, college quality, and labor market conditions among others. As such, while college expansion can be a valuable investment on average and for many, its effects, when considering mental health as an important outcome, are heterogeneous and context-dependent, particularly in large-scale, rapid reform settings.

Note that these findings do not imply that higher education is inherently harmful or that expansion should be avoided. Rather, they highlight the importance of incorporating mental health outcomes into program evaluation and recognizing the heterogeneity in individuals' experiences and well-being, particularly when educational expansions happen at a rapid pace. To ensure that all are well-positioned to benefit from higher education, there is likely a need for additional support for certain groups—such as enhanced academic and financial assistance, access to mental health resources, and guidance on educational and career pathways. Future research should continue to explore how to design inclusive and adaptive programs that maximize the long-term benefits of expansion while mitigating the challenges faced by vulnerable groups.

Main Figures and Tables

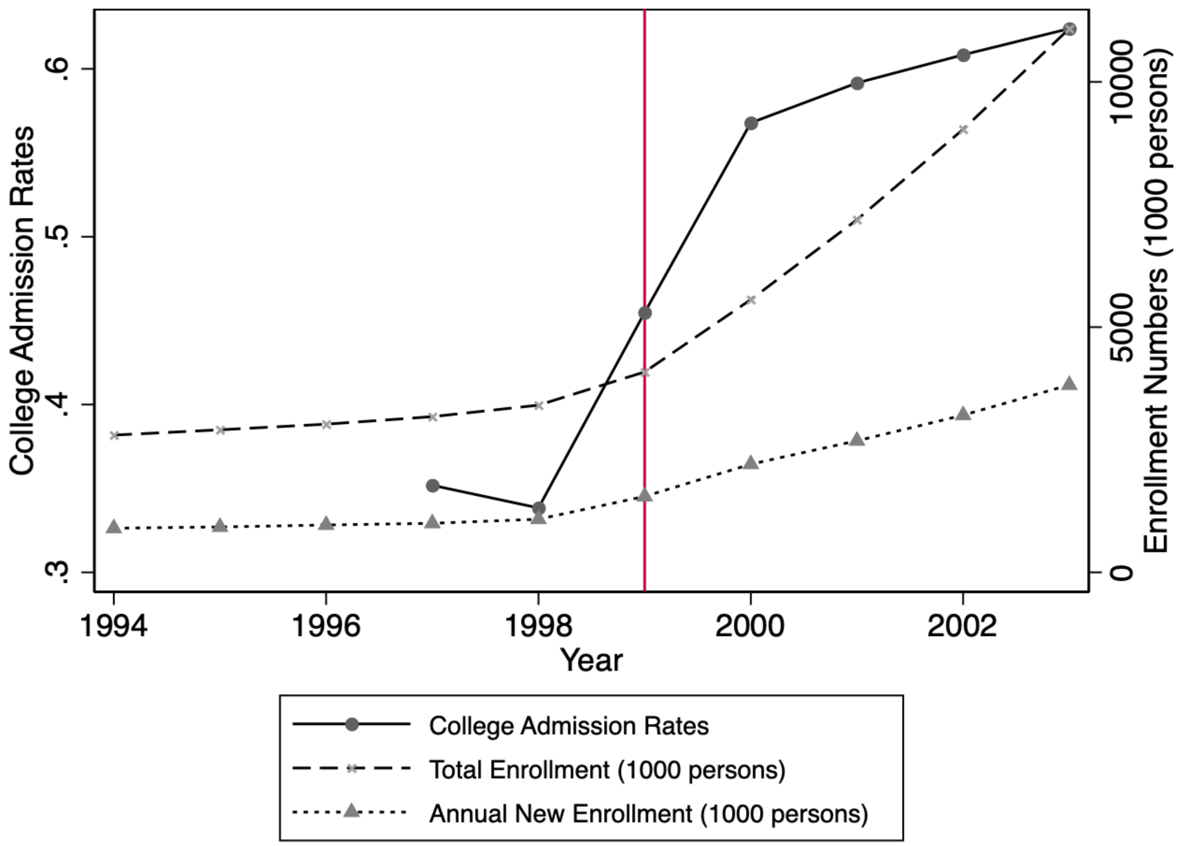

Figure 3.1: Higher Education Expansion in China about Enrollment

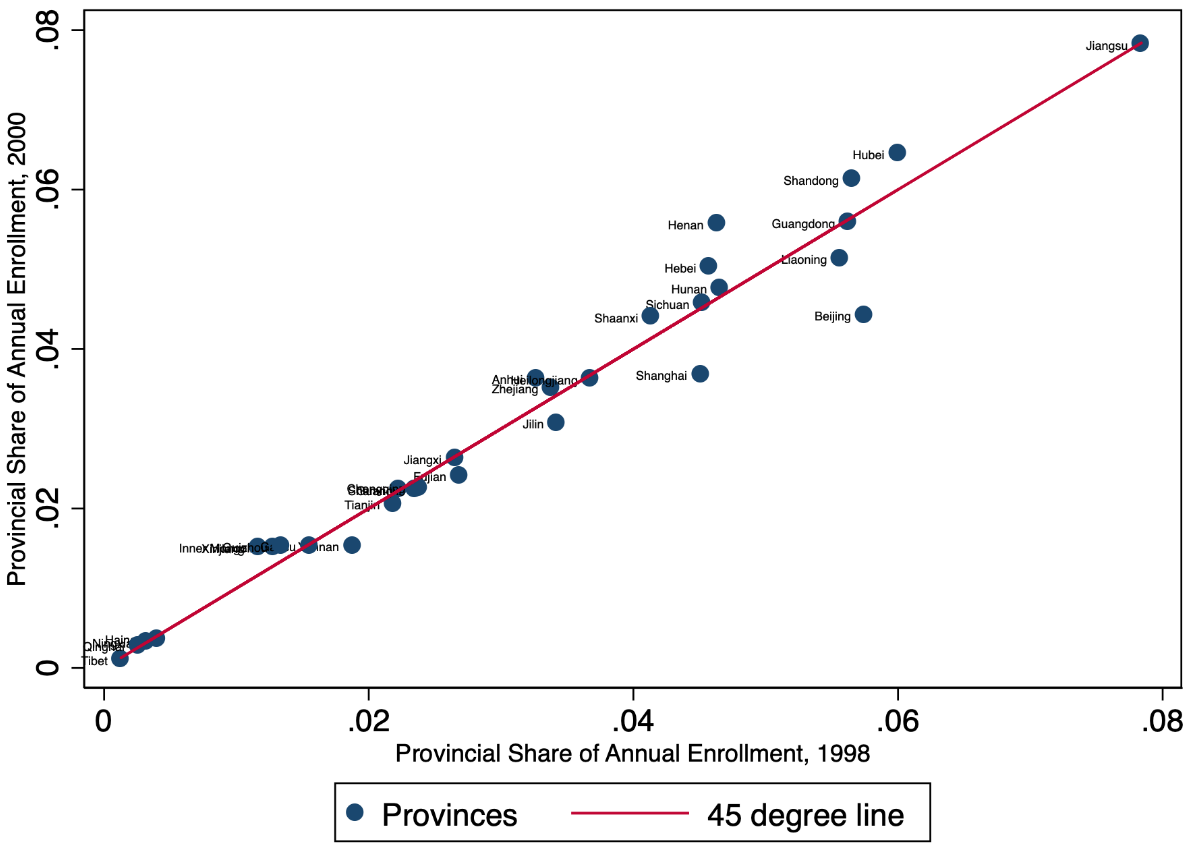

Figure 3.2: Enrollment Expansion Pattern across Provinces

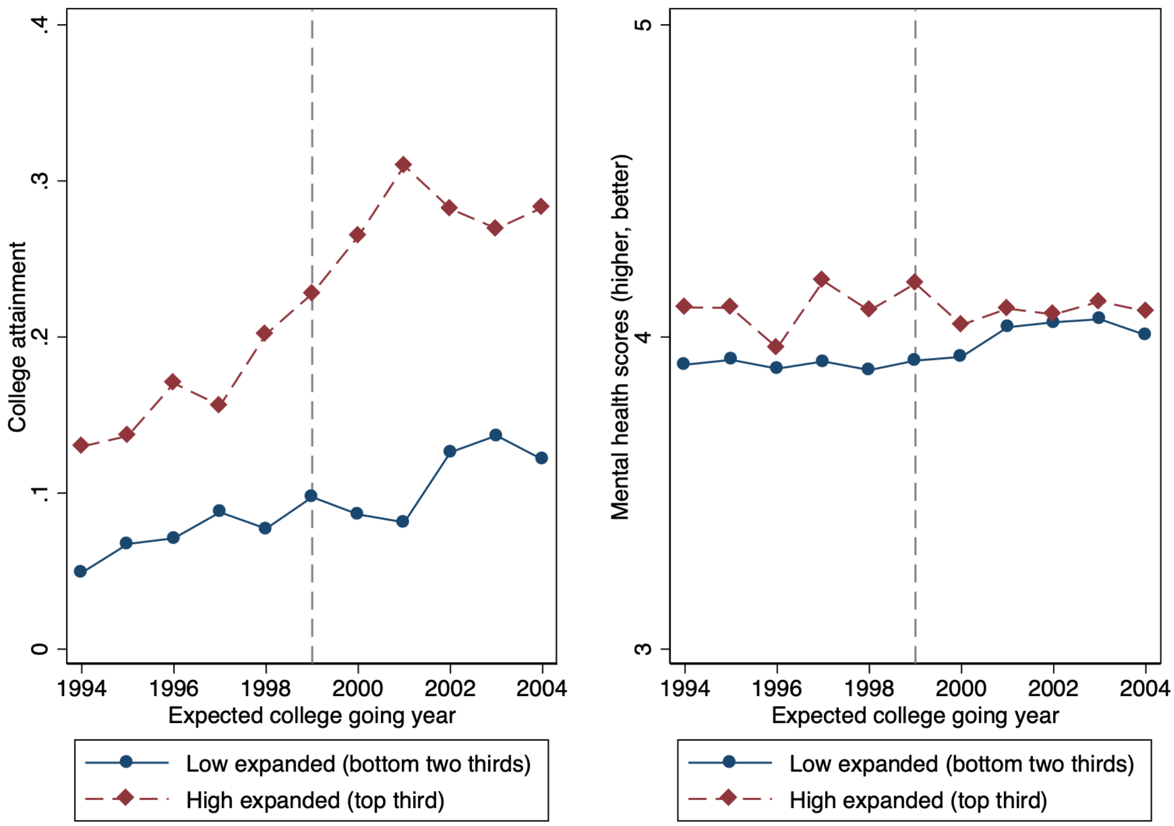

Figure 3.3: Raw Trends in the Prevalence of College Attainment and Mental Health Scores by High Expanded and Low Expanded Provinces

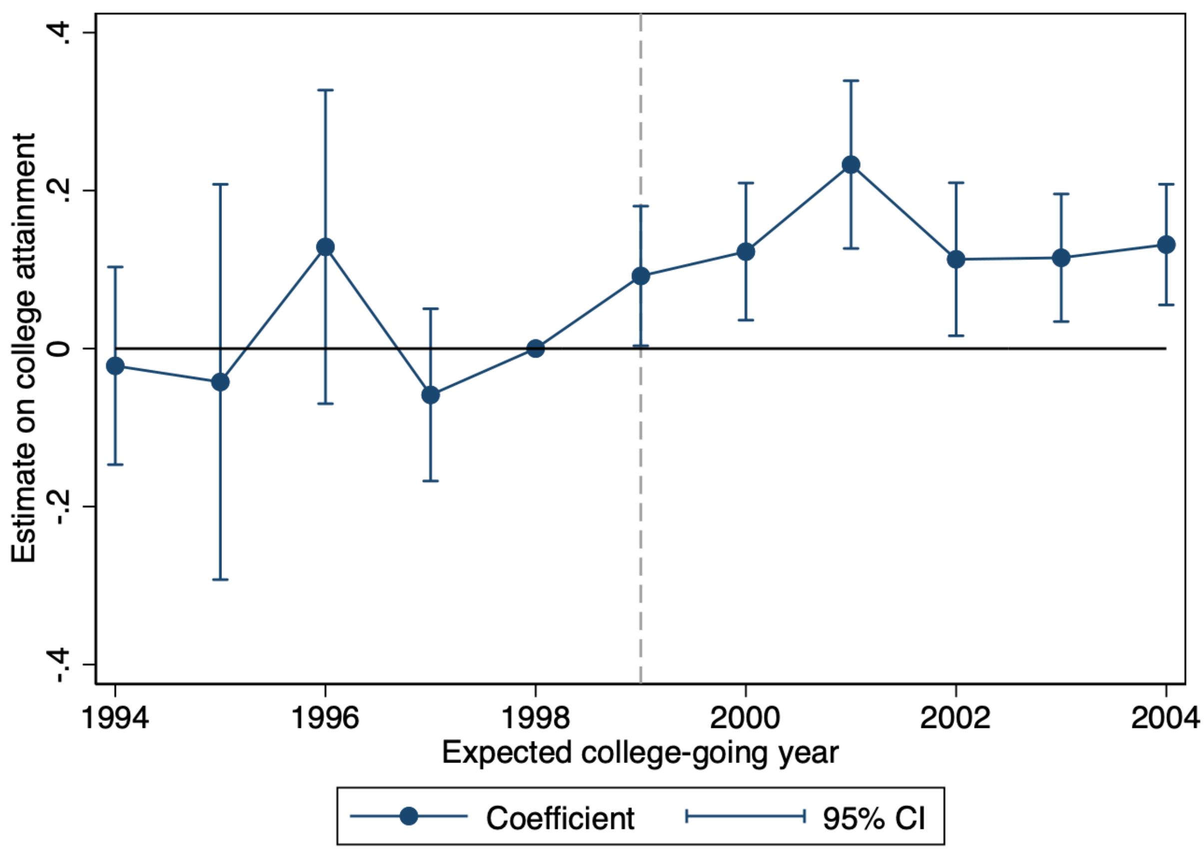

Figure 3.4: HEE Effects on College Attainment of Different Cohorts

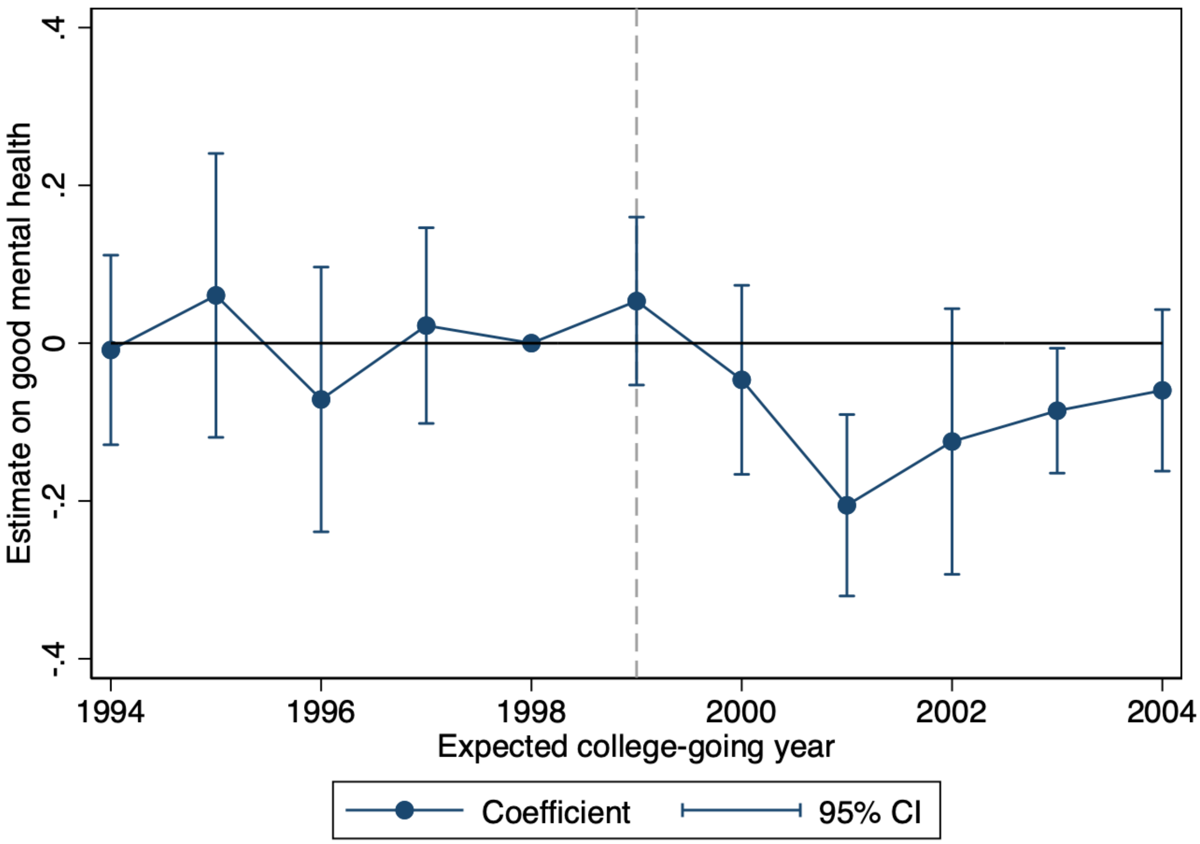

Figure 3.5: HEE Effects on (Good) Mental Health of Different Cohorts

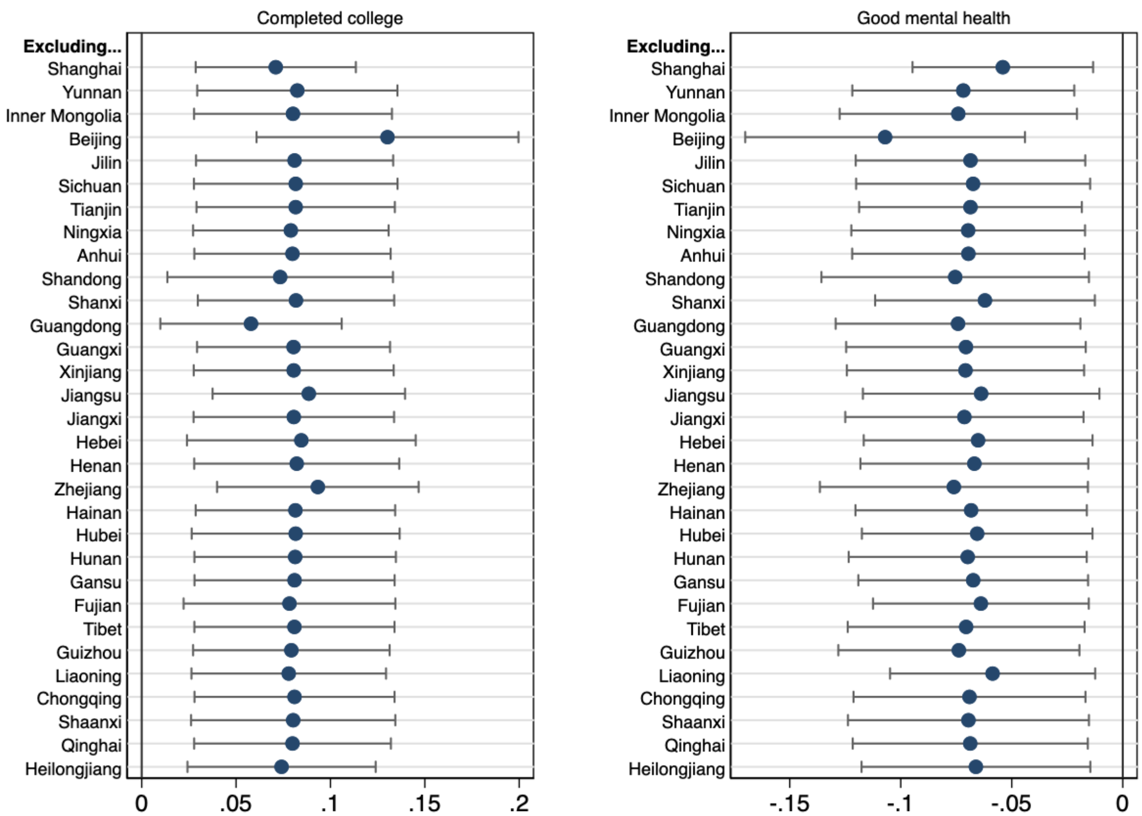

Figure 3.6: Robustness to Excluding Provinces

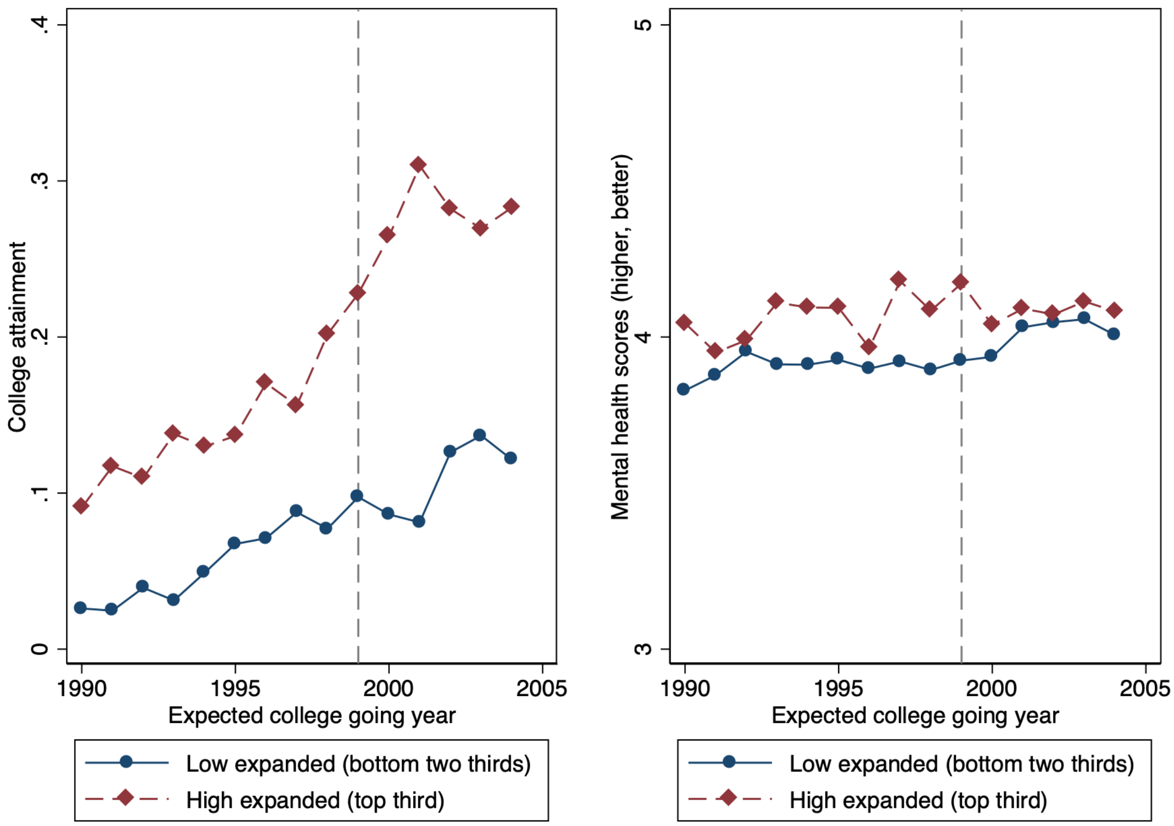

Figure 3.7: Robustness to Longer Pre-periods

Table 3.1: Sample Summary Statistics

	Mean	Std. Dev.	N	Min	Max
College attained	.138	0.345	9339	0	1
Total years of schooling	10.453	4.259	9339	0	16
Health measures					
Good general health	.779	0.415	9339	0	1
Good mental health	.726	0.446	9339	0	1
Good physical health	.875	0.331	9339	0	1
Obese	.201	0.401	9312	0	1
Individual characteristics					
Birth year	1981	3.159	9339	1976	1986
Female	.530	0.499	9339	0	1
Rural Hukou	.510	0.500	9339	0	1
Han ethnicity	.903	0.296	9339	0	1
Religious	.114	0.318	9339	0	1
Family characteristics					
Father low schooling	.777	0.416	9339	0	1
Father as Ag worker	.532	0.499	9339	0	1
Mother low schooling	.860	0.347	9339	0	1
Mother as Ag worker	.584	0.493	9339	0	1

Table 3.2: The Effects of HEE on Mental Health (and Physical Health) Outcomes

	(1) Good mental health	(2) Good general health	(3) Good physical health	(4) Obese
Panel A: ITT effects				
HEE exposure	-0.069**	-0.091**	-0.029	0.017
	(0.026)	(0.035)	(0.024)	(0.020)
Panel B: OLS estimates				
College attained	0.006	0.030*	0.023*	-0.029***
	(0.017)	(0.017)	(0.012)	(0.008)
Panel C: 2SLS estimates				
College attained	-0.853**	-1.123*	-0.364	0.210
	(0.328)	(0.571)	(0.298)	(0.276)
Fstat	9.768	9.768	9.768	9.785
Controls	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Survey wave FE	Yes	Yes	Yes	Yes
Region by cohort trend	Yes	Yes	Yes	Yes
Observations	9339	9339	9339	9312
Mean of dependent var	0.726	0.779	0.875	0.201

Note: Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.3: The Effects of HEE on Alternative Mental Health Measures (2017 survey wave)

	General emotional well-being					Life evaluation	
	(1) Good mental health	(2) SWB overall index	(3) Optimism index	(4) Hope index	(5) Happiness Happiness	(6) Life satisfaction	(7) Marriage satisfaction
HEE exposure	-0.147***	-0.491***	-0.145	-0.160	-0.194	0.128	0.099
	(0.037)	(0.175)	(0.191)	(0.173)	(0.138)	(0.095)	(0.087)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region by cohort trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1991	662	662	662	662	662	591
Mean of dependent var	0.665	0.090	0.068	0.212	0.677	0.760	0.898

Note: Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.4: Heterogeneous Effects of HEE on Mental Health by Family Background

	Good mental health											
	by low SES at 14			by father ag worker			by father low education			by mother low edcuation		
	(1) Low	(2) Not low	(3) Diff	(4) Ag	(5) Not ag	(6) Diff	(7) Low	(8) Not low	(9) Diff	(10) Low	(11) Not low	(12) Diff
HEE exposure	-0.160***	-0.019	-0.140**	-0.198**	0.020	-0.218***	-0.089**	-0.046	-0.043	-0.091***	-0.016	-0.075
	(0.055)	(0.029)	(0.062)	(0.077)	(0.030)	(0.083)	(0.043)	(0.044)	(0.054)	(0.03)	(0.067)	(0.066)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region by cohort trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4999	4264	9263	4967	4372	9339	7256	2083	9339	8031	1308	9339
Mean of dependent var	0.710	0.746	0.726	0.723	0.729	0.726	0.721	0.741	0.726	0.723	0.743	0.726

Note: Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.5: The Effects of HEE on Labor Market and Family Outcomes

	Labor market outcomes			Family related outcomes		
	(1) Currently working	(2) High-skilled occupation	(3) Log annual own income	(4) Never married	(5) Married before 25	(6) Child any (= 1)
HEE exposure	0.023	0.007	0.151*	0.137**	-0.072	-0.109**
	(0.024)	(0.042)	(0.085)	(0.056)	(0.045)	(0.049)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Region by cohort trend	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9337	7473	7553	9339	9339	9339
Mean of dependent var	0.820	0.293	10.122	0.116	0.592	0.804

Note: Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.6: HEE Effects on Mental Health for Different Educational Attainment Groups

	Good mental health				
	(1)	(2)	(3)	(4)	(5)
	at most junior high	only academic high	college and beyond	no college	diff(college vs. no college)
HEE exposure	-0.047	-0.002	-0.209***	-0.036	-0.174**
	(0.062)	(0.099)	(0.059)	(0.045)	(0.074)
Controls	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes
Survey wave FE	Yes	Yes	Yes	Yes	Yes
Region by cohort trend	Yes	Yes	Yes	Yes	Yes
Observations	4479	850	1290	8048	9339
Mean of dependent var	0.697	0.761	0.755	0.721	0.726

Note: Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.7: Characteristics of Compliers

	Treated sample		Compliers		Always-goers		Never-goers		T-test differences (compliers vs.)	
	(1) N	(2) Mean/SE	(3) N	(4) Mean/SE	(5) N	(6) Mean/SE	(7) N	(8) Mean/SE	(9) always-goers	(10) never-goers
Low social class at 14	4871	0.509	396	0.402	408	0.304	3128	0.562	0.098***	-0.161***
		(0.500)		(0.025)		(0.023)		(0.009)		
Father ag worker at 14	4,904	0.501	397	0.300	410	0.149	3153	0.649	0.151***	-0.349***
		(0.500)		(0.023)		(0.018)		(0.009)		
Mother ag worker at 14	4,904	0.546	397	0.360	410	0.183	3153	0.741	0.177***	-0.380***
		(0.498)		(0.024)		(0.019)		(0.008)		
Father low education	4,904	0.731	397	0.602	410	0.249	3153	0.818	0.353***	-0.216***
		(0.443)		(0.025)		(0.021)		(0.007)		
Mother low education	4,904	0.819	397	0.673	410	0.354	3153	0.903	0.319***	-0.230***
		(0.385)		(0.024)		(0.024)		(0.005)		

Note: Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.8: Heterogeneous Effects of HEE on Mental Health by Predicted Compliers/Always-goers

	Good mental health					
	Compliers vs. Not			Always-goers vs. Not		
	(1) Compliers	(2) Not compliers	(3) Diff	(4) Always-goers	(5) Not always-goers	(6) Diff
HEE exposure	-0.156**	-0.024	-0.133*	0.035	-0.056	0.090
	(0.069)	(0.038)	(0.077)	(0.077)	(0.059)	(0.105)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Region by cohort trend	Yes	Yes	Yes	Yes	Yes	Yes
Observations	765	7628	8395	807	7588	8395
Mean of dependent var	0.771	0.721	0.725	0.779	0.719	0.725

Note: Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.9: Heterogeneous Effects of HEE on Marriage Outcomes by Family Background

	Never Married			Child any (= 1)		
	(1) Father ag worker	(2) Father not ag worker	(3) Diff	(4) Father ag worker	(5) Father not ag worker	(6) Diff
HEE exposure	0.070	0.119**	-0.049	-0.070	-0.100***	0.030
	(0.059)	(0.052)	(0.041)	(0.082)	(0.035)	(0.077)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Region by cohort trend	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4967	4372	9339	4967	4372	9339
Mean of dependent var	0.062	0.177	0.116	0.881	0.717	0.804

Note: Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.10: The Effects of HEE on Education and Mental Health (and Physical Health) Outcomes (alternative specifications)

	(1) Completed college	(2) Good general health	(3) Good mental health	(4) Good physical health	(5) Obese
Panel A: using alternative HEE exposure measure					
HEE exposure	0.134**	-0.045	-0.071**	-0.023	0.042
	(0.062)	(0.049)	(0.032)	(0.028)	(0.033)
Panel B: using different denominator to calculate HEE exposure					
HEE exposure	0.084***	-0.100**	-0.070**	-0.031	0.015
	(0.027)	(0.037)	(0.026)	(0.025)	(0.021)
Panel C: using region by cohort FE					
HEE exposure	0.073***	-0.104**	-0.061**	-0.025	0.030
	(0.026)	(0.039)	(0.026)	(0.026)	(0.022)
Panel D: controlling for hukou status					
HEE exposure	0.071**	-0.094**	-0.070***	-0.031	0.017
	(0.027)	(0.035)	(0.025)	(0.025)	(0.020)
Panel E: restricted sample stayed in current county					
HEE exposure	0.108***	-0.067	-0.104***	-0.054	0.032
	(0.029)	(0.042)	(0.025)	(0.039)	(0.021)
Panel F: restricted sample with same hukou status and stayed in current county					
HEE exposure	0.130***	-0.091*	-0.081**	-0.053	0.018
	(0.018)	(0.045)	(0.038)	(0.035)	(0.025)

Note: Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

3.8 Appendix A: Additional Results

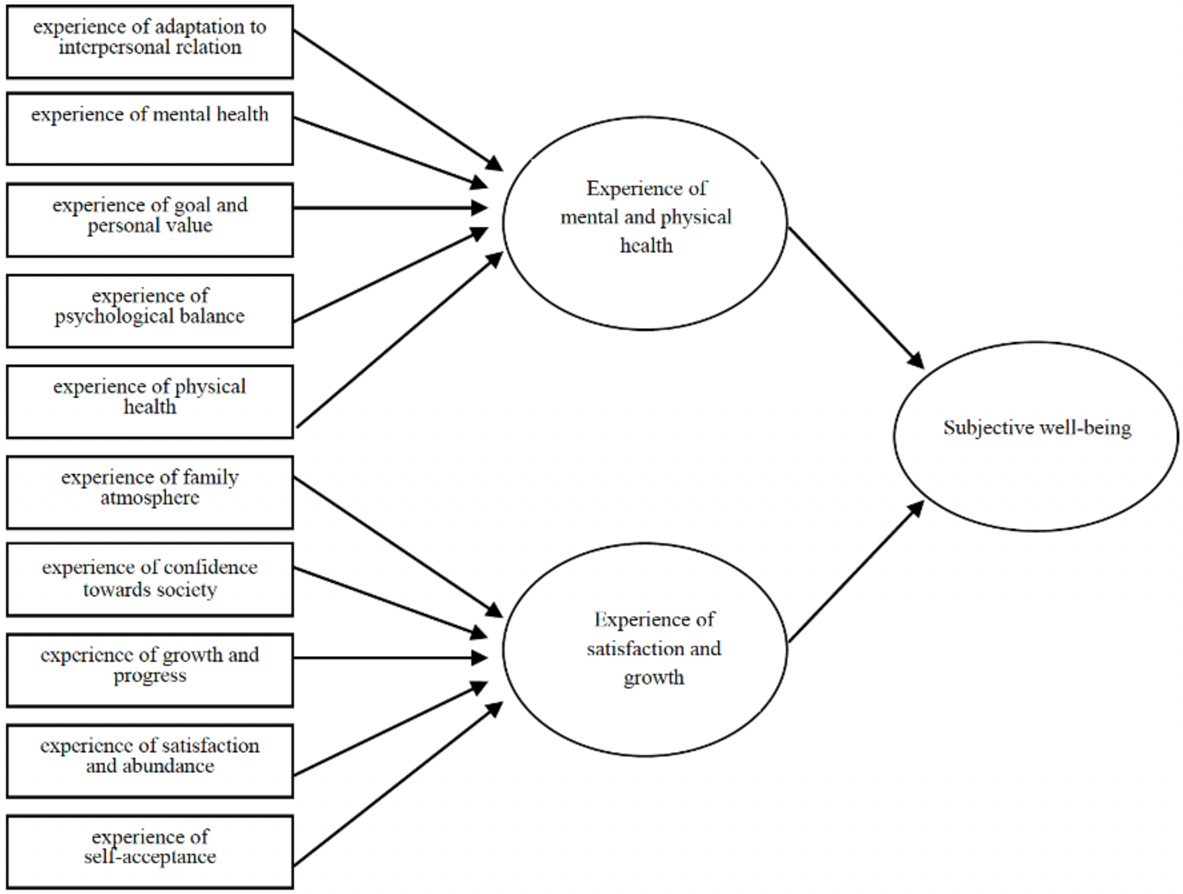

Figure 2. The structural model of Chinese people's subjective well-being

Figure 3.8: The SWB Scale for Chinese Citizens SWBS-cc20

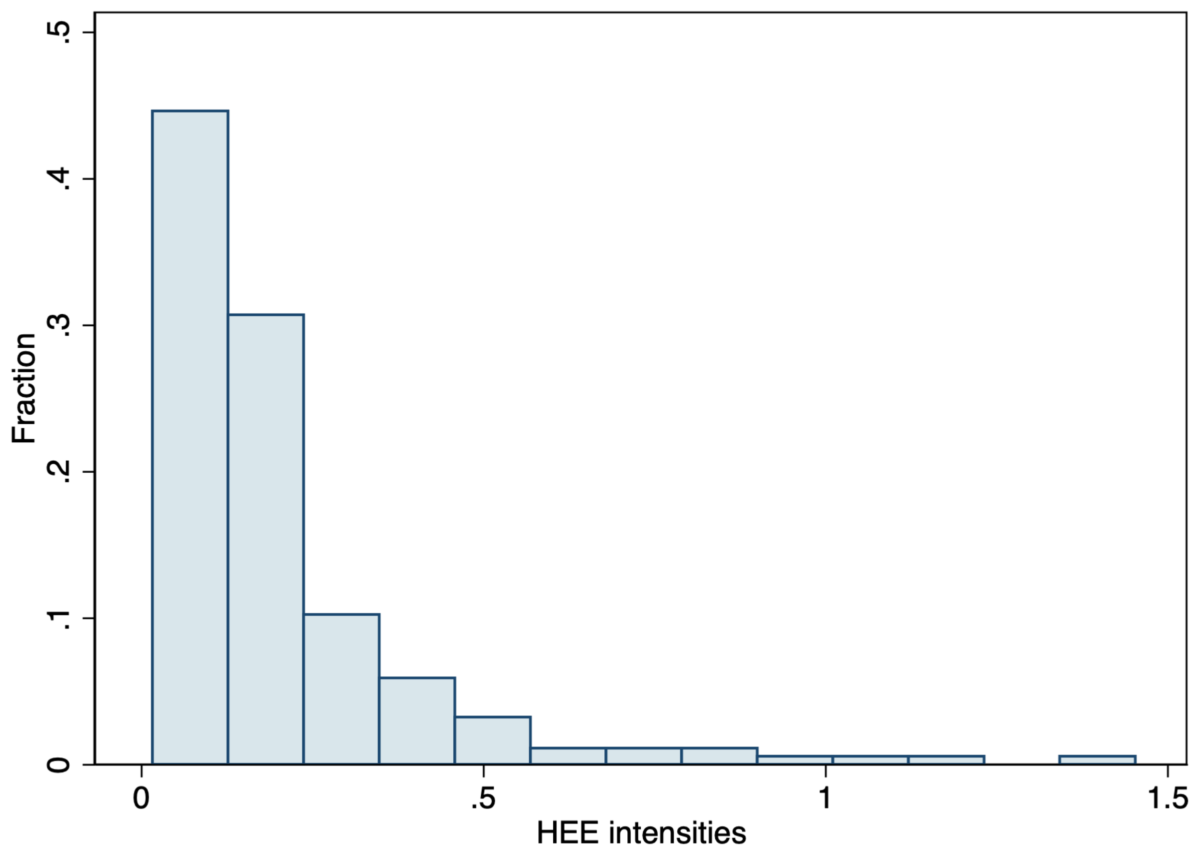

Figure 3.9: Distribution of HEE Intensities

Table 3.11: Variable Definitions

Variables	Definitions	Main questions
Completed college	= 1 if college (or above) is the highest educational level one have completed so far	What is your current highest level of education?
Good mental health	= 1 if the frequency of feeling depressed in the past four weeks are either "rare" or "never"	In the past four weeks, how often did you feel depressed? (always 1; often 2; sometimes 3; rarely 4; never 5)
Good physical health	= 1 if the frequency of having health problems in the past four weeks are either "rare" or "never"	In the past four weeks, how often have health problems affected your work or other daily activities? (always 1; often 2; sometimes 3; rarely 4; never 5)
Good general health	= 1 if respondent reported own health as either "relatively healthy" or "very healthy"	How do you feel about your current general health? (very unhealthy 1; less healthy 2; so-so 3; relatively healthy 4; very healthy 5)
Obese	= 1 if BMI >= 25	What is your current height? & What is your current weight?
SWB overall index	based on the subjective well-being scale for Chinese citizens (SWBS-cc20), using the index construction method proposed by Anderson (2008), directly on scores	The Subjective Well-being module consists of 20 questions with ten experience dimensions and can be grouped into two broad categories (Huang and Xing, 2019). The scale structure is summarized in Figure **??**.
Optimism index	based on the "revised life orientation test LOT-R" Scheier, Carver and Bridges (1994), using the index construction method proposed by Anderson (2008), directly on scores	See the reference paper for detailed scale questions

Continued on next page

Table 3.11 – continued from previous page

Variables	Definitions	Main questions
Hope index	based on "state hope scale" (Snyder et al., 1996), using the index construction method proposed by Anderson (2008), directly on scores	See the reference paper for detailed scale questions
Happiness	= 1 if the self-reported happiness scores are higher than the cutoff score that they reported as happy for them	Please rate your current happiness (maximum 10, minimum 0). & What is the lowest happiness score that you consider to be happy (maximum 10, minimum 0)?
Life satisfaction	= 1 if respondent reported either "very satisfied" or "satisfied"	All things considered, how satisfied are you with your life? (very satisfied 1;satisfied 2; neither satisfied nor unsatisfied 3; unsatisfied 4; not satisfied at all 5)
Marriage satisfaction	= 1 if respondent reported either "very satisfied" or "satisfied"	How satisfied are you with your marriage life? (very satisfied 1;satisfied 2; neither satisfied nor unsatisfied 3; unsatisfied 4; not satisfied at all 5)
Currently working	= 1 if reported as currently working, either in agriculture or non-agricultural sector	What is your working experience and work status? (currently working in non-agriculture sector 1 currently working in agriculture, once had non-agricultural work 2 currently working in agriculture, never worked in non-agriculture sector 3 currently not working, only worked in agriculture 4 currently not working, once had non-agricultural work 5 never worked before 6)

Continued on next page

Table 3.11 – continued from previous page

Variables	Definitions	Main questions
High-skilled occupation	= 1 if the occupation falls into either of the following three majors, which are (1) Managers (2) Professionals (3) Technicians and Associate Professionals	Own occupation code based on ISCO-88, later grouped into the following nine majors: (1) Managers (2) Professionals (3) Technicians and Associate Professionals (4) Clerical Support Workers (5) Services and Sales Workers (6) Skilled Agricultural, Forestry and Fishery Workers (7) Craft and Related Trades Workers (8) Plant and Machine Operators and Assemblers (9) Elementary Occupations
Log annual own income	log own income, adjusted for inflation, in 2010 dollars	What is your own income of last year (relative to the survey year)
Never married	= 1 if current marital status is single (among all)	What is your current marital status: (single 1; cohabitation 2; first married with spouse 3; remarried with spouse 4; separated but not divorced 5; divorced 6; widowed 7)
Married before 25	= 1 if the difference between first marriage year and birth year among all <= 25	In which year did you get married for the first time?
Child any (>= 1)	= 1 if the number of children reported greater or equal to one (among all)	How many children do you have (including stepchildren, adopted children, or deceased children)?
Low social class at 14	= 1 if the reported social class level is lower than (or equal to 3)	What social class level you think your family was when you were 14 years old? (the highest level 10, the lowest level 0)
Father ag worker at 14	= 1 if worked in agriculture either full time or part time	When you were 14 years old, what was your father's employment status?

Continued on next page

Table 3.11 – continued from previous page

Variables	Definitions	Main questions
Mother ag worker at 14	= 1 if worked in agriculture either full time or part time	When you were 14 years old, what was your father's employment status?
Father low education	= 1 if received no more than junior high school	What is your father's highest level of education?
Mother low education	= 1 if received no more than junior high school	What is your mother's highest level of education?

Table 3.12: Predicted Changes in College Enrollment Rate at the province level

	province	intensity	rank1
Beijing	4	0.819	1
Shanghai	1	0.626	2
Tianjin	7	0.455	3
Liaoning	27	0.342	4
Jilin	5	0.290	5
Jiangsu	15	0.261	6
Shaanxi	29	0.251	7
Chongqing	28	0.228	8
Heilongjiang	31	0.226	9
Sichuan	6	0.202	10
Hubei	21	0.197	11
Zhejiang	19	0.165	12
Jiangxi	16	0.150	13
Hunan	22	0.148	14
Shanxi	11	0.145	15
Gansu	23	0.143	16
Fujian	24	0.139	17
Shandong	10	0.138	18
Xinjiang	14	0.127	19
Anhui	9	0.127	20
Hebei	17	0.124	21
Qinghai	30	0.113	22
Ningxia	8	0.112	23
Henan	18	0.112	24
Guangdong	12	0.112	25
Inner Mongolia	3	0.107	26
Hainan	20	0.0909	27
Yunnan	2	0.0904	28
Guangxi	13	0.0899	29
Guizhou	26	0.0809	30
Tibet	25	0.0667	31

Table 3.13: Heterogeneous Effects of HEE on Mental Health by Predicted College-goers

	Good mental health		
	(1)	(2)	(3)
	Predicted college-goers	Predicted not-college-goers	Diff
HEE exposure	-0.069	-0.058	-0.011
	(0.047)	(0.062)	(0.069)
Controls	Yes	Yes	Yes
Province FE	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Survey wave FE	Yes	Yes	Yes
Region by cohort trend	Yes	Yes	Yes
Observations	863	7532	8395
Mean of dependent var	0.757	0.721	0.725

Note: Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.14: Heterogeneous Effects of HEE on Mental Health (within predicted complier group)

	Good mental health											
	by low SES at 14			by father ag worker			by father low education			by mother low education		
	(1) Low	(2) Not low	(3) Diff	(4) Ag	(5) Not ag	(6) Diff	(7) Low	(8) Not low	(9) Diff	(10) Low	(11) Not low	(12) Diff
HEE exposure	-0.272***	-0.081	-0.190*	-0.432	-0.092	-0.341	-0.202	-0.143	-0.059	-0.224**	0.016	-0.240
	(0.085)	(0.097)	(0.107)	(0.330)	(0.107)	(0.371)	(0.124)	(0.151)	(0.224)	(0.100)	(0.165)	(0.206)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region by cohort trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	328	434	765	227	535	767	481	281	767	524	237	767
Mean of dependent var	0.735	0.797	0.770	0.736	0.787	0.771	0.753	0.808	0.771	0.758	0.806	0.771

Note: Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.15: The Effect of HEE on Work Stress

	(1) Often do heavy physical labor	(2) Often feel stressed at work	(3) Often work on weekends
HEE exposure	-0.061	-0.224*	-0.125
	(0.121)	(0.110)	(0.188)
Controls	Yes	Yes	Yes
Province FE	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Survey wave FE	Yes	Yes	Yes
Region by cohort trend	Yes	Yes	Yes
Observations	229	229	228
Mean of dependent var	0.21	0.306	0.474

Note: Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.16: Heterogeneous Effects of HEE on Marriage Outcomes by Family Background (more)

	by Father Low Education			by Mother Low Education		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent var:	Low	Not low	Diff	Low	Not low	Diff
Never married	0.104*	0.095*	0.009	0.067	0.123*	-0.056
	(0.059)	(0.054)	(0.033)	(0.05)	(0.062)	(0.036)
Child any (= 1)	-0.107*	-0.045	-0.063	-0.057	-0.116	0.059
	(0.054)	(0.048)	(0.056)	(0.055)	(0.069)	(0.066)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Region by cohort trend	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

3.9 Appendix B: Compliers Prediction

To determine whether the adverse effects of HEE on mental health are driven by compliers with distinct traits, such as family background and academic ability, I employed the following steps to predict compliers based on the available data. Specifically, I used all the pre-HEE cohort data to model college-going behavior without expansion using LASSO linear regression. I used 10-fold cross-validation to select the optimal shrinkage parameter and select covariates. I then predicted the probability of going to college for the overall sample (based on the OLS post-estimation after LASSO). To ease the interpretation, I transformed the continuous college-going probabilities from prediction into a binary indicator for the post-HEE cohorts by assigning the cutoff that the percentage of college graduates for the post-HEE cohorts matches the actual percentage for the pre-HEE cohorts[34]. Next, I compared the predicted college-going behavior for those cohorts who expected to have a college-going year later than 1999 to their actual college-going status. I defined compliers are those who predicted not to go to college but actually go. Similarly, always-goers (never-goers) are those who (do not) go to college in both predicted and actual cases. Last, I run a balance test among the three different predicted groups. See Table 3.7 for the results.

[34]The cutoff I used is 0.35. That is, an individual is predicted to go to college after HEE if the predicted college-going probability is at least greater than 0.35. Here I assumed there are no meaningful differences between the pre-and post-HEE samples regarding baseline characteristics I used for prediction.

Chapter 4

The Association Between Visual Impairment, Educational Outcomes, and Mental Health: Insights from Eyeglasses Usage among Junior High School Students in Rural China[1]

Xiaodong Pang[1] , **Huan Wang**[2] , **Yiwei Qian**[3] , **Sabrina Zhu**[2] , **Yuwei Adeline Hu**[2] , **Scott Rozelle**[2] , **Nathan Congdon**[5,6] , **Jiting Jiang**[4]

[1]This chapter is based on joint work with the aforementioned authors and was published in *Scientific Reports (2024).* The author affiliations are as follows: [1]College of Education, Capital Normal University, Beijing, China [2]Stanford Center on China's Economy Institutions (SCCEI), Freeman Spogli Institute for International Studies, Stanford University, Stanford, California, USA [3] Research Institute of Economics and Management, Southwestern University of Finance and Economics, Sichuan, China [4] Agricultural and Resource Economics, University of California, Davis, California, USA [5] School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, Northern Ireland, United Kingdom [6] Orbis International, New York, NY, USA

4.1 Introduction

Visual impairment (vision acuity $< 6/12$) is among the most common health problems worldwide, comprising half of all disabilities among young people (Bourne et al., 2017). Uncorrected refractive error is currently the leading cause of visual disability among school-aged children (WHO, 2023). Although refractive error can be managed safely and inexpensively with properly fitted eyeglasses (Ma et al., 2015), a significant number of children in need of refractive correction remain untreated (Ma et al., 2014, 2022; Li et al., 2010; Yi et al., 2015; Congdon et al., 2008; Zhang et al., 2009). For example, between 64% and 85% of refractive error cases in rural China remained uncorrected by eyeglasses in a population-based study (Maul et al., 2000). Even in high-income countries, such as Australia, one out of every four children in need of eyeglasses do not have them (Robaei et al., 2006).

Visual impairment has been associated with mental illness such as depression and anxiety among children, with the strongest associations observed when refractive error remains uncorrected (Yi et al., 2015; Li et al., 2022; Guan et al., 2018). One possibility for this association is that students who struggle to see clearly in class may experience depression and anxiety symptoms due to learning difficulties (Pinquart and Pfeiffer, 2012). Students with poor vision may also feel excluded by classmates without visual impairment who are able to participate more readily in classroom activities, leading to an increase in peer-related anxiety in visually impaired students (George and Duquette, 2006; Ralejoe, 2021). It is important to note that children's mental health issues are multidimensional and may affect them in different ways. These considerations call for a comprehensive assessment of children's mental health and its relationship with visual impairment.

Despite increasingly strong evidence of an association, the precise mechanisms linking the use of eyeglasses to children's mental health are yet to be fully understood. Nevertheless, it is evident that supplying eyeglasses for vision correction can enhance academic performance (Ma et al., 2014, 2018). Moreover, improved educational outcomes from wearing eyeglasses could lead to better mental health (Dudovitz et al., 2016).

Visual impairment caused by uncorrected refractive error may be especially likely to exacerbate student mental health problems in countries with competitive educational systems and limited resources. These competitive educational systems appear to place even adolescents without vision impairment at higher risk of mental health problems (Li et al., 2010; Sun et al., 2011). Because of the potential synergy between academic stress, visual impairment, and mental health conditions such as anxiety and depression, studies in countries with documented high levels of student academic stress and uncorrected visual impairment are of particular importance.

Rural China is such a setting. Approximately half of the 12.8 million children worldwide with impaired vision due to uncorrected refractive error live in China (Resnikoff et al., 2008). In 2019, 72% of junior high school students in China were found to be myopic (Wang et al., 2021). Despite this high prevalence, less than a quarter of students needing eyeglasses in rural and urban migrant Chinese settings own them (Ma et al., 2014; Ma, 2024; Li et al., 2010; Wang et al., 2021). In addition, even when free eyeglasses are provided to visually impaired children, many do not wear them regularly, due mainly to concerns that use of glasses might weaken their eyes (Ma, 2024; Du et al., 2023).

The aim of this paper is to address existing gaps in the literature and to determine the relationship between impaired vision, eyeglasses usage, mental health, and educational outcomes among junior high school students in rural China. We chose to study junior high school students because the prevalence of uncorrected refractive error and unmet need for eyeglasses peaks at this age, as does academic pressure as they prepare for the highly-competitive national high school entrance examination (Ma, 2024; Li et al., 2020; Essau et al., 2008; Wang et al., 2015). To achieve the study aim, we assessed visual acuity, eyeglasses wear, academic achievement, and mental health status in a population sample of 19,425 junior high school students in rural western China with data gathered in 2019. We had four objectives. First, we examined the prevalence of student visual impairment by trained medical and research professionals. Second, we examined the rates of student mental health problems according to their Strengths and Difficulties Questionnaire (SDQ) scores. Third, we investigated the associations between students' visual impairment, visual impairment corrected by eyeglasses, and mental health. Fourth, we examined heterogeneous effects by analyzing how the use of eyeglasses among students with high or low academic performance influenced both their mental health and educational aspirations.

4.2 Methods

Ethics approval for this study was granted by the Institutional Review Boards at Stanford University (Stanford, CA, USA), Queen's University Belfast (Belfast, United Kingdom), and the Zhongshan Ophthalmic Center, Sun Yat-Sen University (Guangzhou, China). Permission was given by the Ningxia Board of Education and principals from the surveyed junior high schools, and informed consent was provided by at least one parent of all participating children. The principles of the Declaration of Helsinki were followed throughout the study.

4.2.1 Study Location and Sampling

We drew our sample from an ongoing trial involving junior high school students in Ningxia, a landlocked, predominantly rural autonomous region in northwest China. Ningxia has an area of 66,400 square kilometers and a population of 6.62 million, 62% of whom are ethnic Han and 38% of which belonging to China's Muslim Hui minority. Ningxia's per capita gross domestic product was $8,174 in 2018, some 20% below the national average (China Statistics Press, 2018). The initial survey was carried out in October 2019.

A two-step selection protocol was used to obtain a representative sample of all rural junior high school students in Ningxia. In the first step, the sample schools were selected from a list of all 273 junior high schools in Ningxia, acquired from local education bureaus in the six prefectures. To more closely focus on the rural population, students from the 122 schools in Yinchuan, the capital city of the autonomous region, were excluded from the sample. Schools with fewer than 40 children in 7th grade were also excluded due to logistical constraints. All other schools were included in the study, for a total of 124 participating institutions. In the second step, in each of the sampled junior high schools, at most two classes per grade were randomly selected from among all 7th and 8th grade classes (with children largely falling between the ages of 12 to 15 years). The final sample consisted of 20,375 students from 474 classes at 124 schools.

4.2.2 Data Collection

Data were collected by teams of trained enumerators using a survey instrument consisting of two components: a questionnaire and a visual acuity (VA) assessment.

Questionnaire Survey

Enumerators distributed questionnaires to students to collect information about their demographic characteristics, academic performance, mental health, and vision. The demographic survey covered details such as student gender, ethnicity, rural residential status of the family, only child status, student grade, boarding at school, attendance of extracurricular classes, parental divorce, parental out-migration for work, grandparent(s) as the primary caregiver, parents' age and years of schooling, household assets index, educational and career aspirations, and ownership of glasses. We also conducted a 30-minute standardized mathematics test to assess student academic performance. The test was designed with input from local education experts to ensure coherence with China's national curriculum. All scores were normalized by grade level.

An official Chinese simplified version of the Strengths and Difficulties Questionnaire (SDQ) screening instrument for children was used to determine students' mental health status. The SDQ has a strong internal consistency (Cronbach's alpha coefficient = 0.81) and high levels of reliability (Pearson's correlation coefficient = 0.71) (Yao et al., 2009; Lai et al., 2010). The questionnaire includes 20 items on behavioral strengths and difficulties, scored on a 3-point scale (0 = not true, 1 = somewhat true, and 2 = certainly true). Measured items are grouped into four subscales (0-10 points, with higher score indicating greater difficulties): emotional problems, conduct, hyperactivity/inattention, and peer relationships.

We use both SDQ score and SDQ abnormality to measure students' mental health. The SDQ score, including SDQ total score, Internalizing SDQ score, and Externalizing SDQ score, was used in our data analysis. For SDQ abnormality, we scored the SDQ questionnaire according to the guidance provided on the SDQ official website; the abnormal score ranges are: 20–40 for total difficulties, 7–10 for emotional problems, 5–10 for conduct, 7–10 for hyperactivity, and 6–10 for peer problems (Goodman, 1997). We mainly used amalgamated scales that are sometimes used instead of the four separate difficulties subscales described above.30 These alternative scales are called the 'internalizing' (comprising the emotional and peer problem subscales) and 'externalizing' subscales (comprising the conduct and hyperactivity subscales). Referencing Yu et al. (2022), we defined the student with abnormal internalizing problems if the students obtained abnormal scores on either emotional problems or peer problems subscales. We defined the student with abnormal externalizing problems if the students obtained abnormal scores on either conduct problems subscales or on hyperactivity subscales.

Visual Acuity (VA) Assessment

VA assessment was conducted at each school by an optometrist and a trained research assistant. For each student, each eye was tested individually at 4m without refraction using Early Treatment Diabetic Retinopathy Tumbling-E study charts (Precision Vision, La Salle, IL) in an illuminated indoor space. 4 During the testing process, students moved from the 6/60 line to each successively lower line if they were able to identify the orientation of at least four out of five optotypes, to a maximum vision of 6/3. VA in an eye was identified as the lowest line for which the student was able to correctly determine at least four out of five optotypes. If the student could not read the 6/60 line, the chart was positioned at a distance of 1m, and the resulting visual acuity was divided by 4. The assessment included cycloplegic automated refraction with subjective refinement to ascertain prescriptions for eyeglasses eligibility (the myopia cutoff is 0.5 diopters [D]). The eye examination team was trained by Zhongshan Ophthalmic Center (ZOC) at Sun Yat-Sen University.

4.2.3 Statistical Analysis

First, we presented summary statistics encompassing student characteristics, mental health measured through SDQ, educational aspirations, and academic performance assessed by standardized math tests across four groups: the total sample, students with normal vision, students with vision impairment wearing glasses, and students with visual impairment without eyeglasses. We then performed t-tests to analyze disparities in demographics, mental health, educational aspirations, and academic performance between students with visual impairment wearing glasses and those with normal vision, as well as between students with visual impairment without glasses and those with normal vision. Second, we conducted several multivariate Ordinary Least Squares (OLS) regressions to examine associations between student vision, eyeglass usage, and mental health after controlling for student characteristics and math tests. Third, we divided our sample into a high academic performance group, defined by standardized math tests scores greater than or equal to the median, and a low academic performance group, defined by standardized math test scores lower than the median. We then ran multivariate OLS regressions to explore the associations between student eyeglasses usage and mental health within the high and low academic performance groups, controlling for student characteristics and math scores. Fourth, we ran multivariate OLS regressions to examine associations between student eyeglasses usage and educational aspiration among the high/low academic performance groups, respectively, after controlling for student characteristics and math scores. All statistical analyses were performed using Stata 17.0 (Stata Corp LLC, Texas, USA).

4.3 Results

Table 4.1 provides summary statistics for the sampled students. Out of 19,425 students, 37.2% (N = 7,221) had poor vision, 56.6% (N = 4,088) of whom did not wear glasses. Additionally, 48.6% (N = 9,440) were female, 64.9% (N = 12,606) belonged to the Hui ethnic minority group, 93.6% (N = 18,181) resided in rural areas, 6.8% (N = 1,320) were only children, and 34.8% (N = 6,760) had at least one parent who had migrated for work in the past six months and did not live at home. A 9.7% divorce rate among parents was observed. Fathers had a mean age of 41.3 years and an average of 7.08 years of schooling, while mothers had a lower mean age of 38.67 and fewer years of schooling, with a mean of 5.13 years. Among the children, 35.4% (N=6,876) reported boarding at school, and 11.3% (N = 2,195) attended extracurricular classes. As indicated in Table 4.1, there was a significantly higher prevalence of girls among students with visual impairment with/without glasses (62.5%/55.4%) compared to those with normal vision (42.7%) ($p < 0.001$).

Tables 4.2 presents summary statistics of student mental health, educational aspiration, and academic performance. On average, students obtained scores of 12.6, 6.4, and 6.1 on the Strengths and Difficulties Questionnaire (SDQ) for total scores, internalizing scores, and externalizing scores, respectively. Additionally, 9.0%, 15.1%, and 17.3% of students exhibited abnormal scores in SDQ total, internalizing, and externalizing categories, respectively. Students with visual impairment wearing glasses were 0.35 ($p < 0.001$), 0.27 ($p < 0.001$), 0.01 ($p = 0.006$), and 0.03 ($p < 0.001$) percentage points less likely to experience mental health problems in terms of their SDQ total score, externalizing score, SDQ prevalence of abnormal scores, internalizing abnormality, and externalizing abnormality than students with normal vision. Meanwhile, students with visual impairment without glasses were 0.27 ($p < 0.001$) and 0.23 ($p < 0.001$) percentage points more likely to have mental health problems in terms of their SDQ total score and externalizing score than students with normal vision.

The majority (N = 14,141, 72.8%) of students believed that they were very likely to take the high school entrance exam. Additionally, 73.4% of students expected to attend academic high school, 20.8% students expected to attend vocational high school (i.e., vocational education and training), and 5.8% of students expected to enter the labor market after graduation from junior high school. Students with visual impairment who wore glasses were 0.093 ($p < 0.001$) and 0.092 ($p < 0.001$) percentage points more likely to take the high school entrance exam and attend academic high school after graduation compared to students with normal vision. Surprisingly, students with visual impairment without glasses followed a similar pattern, with a higher likelihood of taking the high school entrance exam and attending academic high school after graduation compared to students with normal vision.

The large share of junior high school students who suffer from visual impairment and poor mental health status, as shown above, leads us to consider two important yet understudied questions: To what extent does visual impairment associate with mental health? What is the role of eyeglasses in mental health for students with visual impairment?

We began by examining the association between visual impairment and mental health in the overall sample. Table 4.3 panel A shows the results of our multivariate regression when student and family characteristics were held constant (as described earlier). We found significant associations between visual impairment and higher SDQ scores and, thus, poorer mental health. Specifically, we found that students with visual impairment scored 0.148 and 0.105 higher in their SDQ total score and internalizing score compared to students with normal vision. Second, we examined the associations between eyeglasses usage and mental health. The results of the analysis are presented in Table 4.5 Panel B. We found that the adverse impact of poor vision on mental health diminished when students with poor vision wore eyeglasses. Specifically, there were no significant differences in mental health between students with visual impairment wearing eyeglasses and those with normal vision. We further found that the negative impact of poor vision on mental health was more pronounced for students without eyeglasses. Specifically, students with visual impairment without eyeglasses scored 0.23 and 0.15 higher on their SDQ total score and internalizing score than students with normal vision. Third, we narrowed our sample to only those students with visual impairment to scrutinize the role of eyeglasses on mental health. The outcomes are presented in Table 4.5 Panel C, revealing a significant association between eyeglasses usage and improved mental health.

We further examined the heterogeneous treatment effect of relationships between eyeglasses usage and mental health across high/low academic performance groups. The results in Table 4.4 reveal that eyeglasses usage may have protected mental health for students with high academic performance, but not for the students with low academic performance. Specifically, results in Table 4.4 Panel A show that among high academic performance students who wore eyeglasses, they scored lower in all SDQ measures except for internalizing SDQ abnormality compared to students with similar academic performance with visual impairment who did not wear eyeglasses. The result in Table 4.4 Panel B revealed that for the low academic performance students, there were no significant relationships between eyeglasses usage and mental health.

In Table 4.5, we present the results for examining the role of eyeglasses in raising or lowering educational aspirations among high/low academic performance groups. The results in Table 4.5 Panel A show that students with high academic performance who wore glasses had significantly improved educational aspirations compared to students with similar academic performance who did not wear glasses. This relationship pattern persists in students with low levels of academic performance, with the improved effect even stronger. Specifically, Table 4.5 Panel B shows that students with low academic performance who wore eyeglasses were associated with higher educational aspirations, such as believing that they were very likely to take the high school entrance exam, very likely to attend an academic high school after graduation, and less likely to enter the labor market after junior high school graduation.

4.4 Discussion

We studied the multidimensional associations between visual impairment, eyeglasses use, and mental health among junior high school students in rural China. Using data from 19,425 students from 124 schools in the rural Ningxia region of northwest China, we found that 9.3% to 17.3% of these students were categorized as SDQ abnormal, and 37.2% had visual impairment. After studying associations, first, we found that students with visual impairment were more likely to have poorer mental health in terms of their SDQ total score and internalizing SDQ score. In addition, students with visual impairment who did not wear eyeglasses had significantly poorer mental health in terms of their SDQ total score and internalizing SDQ score compared to students with normal vision. Third, eyeglasses usage was significantly associated with improved mental health for the students with higher academic performance, but not for the students with lower academic performance. Finally, eyeglasses usage was significantly associated with higher educational aspirations for low academic performance students.

The rural students in our study exhibited a high prevalence of mental health problems and had elevated rates of uncorrected visual impairment. The mean SDQ total difficulties score in this study was 12.60. In comparison, more developed nations, such as the United States, reported a mean score of 7.10 (Bourdon et al., 2005), and in a more urban region like Wuhan, China, the mean score was 11.24 (Hu, Lu and Huang, 2014). In our study, 37.2% of the sampled students had poor vision, a rate higher than the global average of 10% (Yekta et al., 2022). Moreover, a significant majority (56.6%) of the rural students with poor vision did not use eyeglasses. This trend mirrors findings from a previous study in the rural Chaoshan region of China, where two-thirds of students with visual impairment did not use eyeglasses (Li et al., 2010). The notable disparity in eyeglasses usage between rural and urban areas is evident, as 75% of urban Chinese students have visual impairment, yet a substantial majority (65.9%) own and use eyeglasses (Ma et al., 2022; Wang et al., 2021; He et al., 2004).

The positive association that we found between visual impairment and student mental health problems supports the existing literature. Students in our study with visual impairment were significantly more likely to have poor mental health as measured by SDQ. These results agree with findings from a previous study conducted in rural China that sampled primary school students and found that visual impairment could have negative impacts on their academic achievement and mental health (Guan et al., 2018; Ma et al., 2018). In the competitive education system in rural China, poor vision may discourage students from academically achieving, leading to feelings of hopelessness or defeat (Suldo, Thalji and Ferron, 2011).

Leveraging the multidimensional framework of the SDQ, our study provided deeper insights into specific aspects of student mental health, specifically internalizing or externalizing problems as measured by the SDQ subscales, and their associations with visual impairment compared to prior research. Our results highlighted a particularly significant association between visual impairment and internalizing problems. This finding suggests that visually impaired students may encounter heightened emotional conflicts related to self-image and peer relationships, potentially leading to challenges in forming friendships or facing bullying in class (Pinquart and Pfeiffer, 2012; Sims, Celso and Lombardo, 2021; Miyauchi, 2020). Our results align with existing literature suggesting that students with vision impairment are more likely to be bully-victims rather than bullies (Guan et al., 2018; Horwood et al., 2005). Moreover, literature on student bullying has shown that internalizing problems typically occur among victims, while externalizing problems are more common among bullies (Kelly et al., 2015; Arseneault et al., 2008). Therefore, we believe that because students with visual impairment are more susceptible to being bullied, they have higher internalizing problems.

Moreover, although wearing eyeglasses can safely and easily correct visual impairment, we did not observe significant differences in mental health between visually impaired students who wore eyeglasses and students with normal vision. This lack of association could be attributed to the potential negative mental health impacts associated with wearing eyeglasses, such as peer bullying, which may outweigh positive benefits (Li et al., 2010; Yi et al., 2015; Holguin et al., 2006). However, our examination also revealed that students with visual impairment who did not wear eyeglasses exhibited significantly worse mental health than students with normal vision. Taken together, these findings suggest that eyeglasses are crucial for students with uncorrected visual impairment, and that their mental health may significantly decline if they do not wear eyeglasses. However, once visual impairment is corrected with eyeglasses, their mental health aligns with that of students with normal vision.

Heterogeneous effects impacting this association may also exist; for instance, specific groups of students could experience a strong positive impact from wearing eyeglasses and endure fewer negative effects, resulting in an overall positive net effect for these particular student groups (Guan et al., 2018). As the students in our study navigate a highly competitive education system, their academic performance becomes a pivotal factor influencing various aspects of their school experience, including the decision to wear eyeglasses. In our pursuit of understanding heterogeneous effects, we investigated the association between eyeglasses usage and mental health among students with high and low academic performance. After categorizing junior high school students based on their standardized math test scores, we observed a significant association between eyeglasses usage and improved mental health for students with higher math scores. Specifically, high academic achievers who wore eyeglasses scored lower in all SDQ measures except for internalizing abnormality. A plausible explanation for these findings lies in the intense academic competition within rural China's educational system, where students face substantial pressure to excel academically. In this context, students with higher academic proficiency may recognize that wearing eyeglasses enhances their learning capabilities, enabling them to achieve more and fostering confidence, thereby positively impacting their mental health (Suldo, Thalji and Ferron, 2011; Feng et al., 2022).

On the contrary, students with lower academic proficiency did not experience a significant change to their mental health after wearing eyeglasses. This result diverges from a prior study conducted in China, which identified a notable adverse impact on the mental health of visually impaired students with low academic performance after they started wearing eyeglasses (Guan et al., 2018). This finding was attributed to prevalent social norms in China. According to a 2013 study, wearing eyeglasses is often associated with being perceived as more studious. Consequently, when low-performing students started wearing eyeglasses, they faced bullying from both peers and teachers because they were not perceived as studious but were now wearing eyeglasses (Guan et al., 2018). The current study, conducted with 2019 data gathered in rural China, did not find a significant association between eyeglasses usage and poor mental health among students with low academic performance. This suggests a potential shift in social norms regarding eyeglasses usage in rural China, potentially indicating that knowledge about vision health has improved among rural residents.

Finally, we explored the relationship between eyeglasses usage and educational aspirations. First, we observed that wearing eyeglasses is associated with increased educational aspirations for both high and low academic performance students, possibly due to improved mental health boosting their confidence in learning (Almroth et al., 2018). Second, there is a heterogeneous effect, with the impact of eyeglasses on educational aspirations being over two times stronger for low academic performance students. The lower the academic performance, the more pronounced the positive effect of improved educational aspirations from wearing eyeglasses. Third, there is a significantly negative relationship between eyeglasses usage and the aspiration to work after graduation among low academic performance students, but not for high academic performance students. This may be explained by the persistence of high academic performance students in pursuing high school admission regardless of eyeglasses usage, whereas for low academic performance students in rural China, wearing eyeglasses significantly and strongly increases their confidence in attending academic high school after graduation, expanding their decision-making options and allowing them to choose not to join the labor market after graduation, a choice they might not have considered without the positive effects of eyeglasses.

4.5 Conclusion

This study has several strengths. First, our study advances the literature on visual impairment, eyeglasses usage, student mental health, and student educational aspiration in the context of a developing country. Second, our utilization of data collected in 2019, coupled with a substantial sample size of 19,425 junior high school students, enabled a thorough and updated exploration of the relationships between visual impairment, visual impairment corrected by wearing eyeglasses, and mental health, capturing relevant associations in rural China and allowing for robust conclusions. Third, using subscales of the SDQ, we were able to identify more nuanced and specific mental health impacts of visual impairment and eyeglasses usage. Several limitations should be acknowledged when considering our results. First, although the sample size was large, caution must be used in extrapolating the results across all rural areas of China or beyond. Second, it is advisable to conduct longitudinal research to establish causality between student mental health, visual impairment, and eyeglasses usage. Despite these limitations, our study yields crucial insights into the ramifications of uncorrected visual impairment on child mental health development and underscores the necessity for targeted research and interventions aimed at bolstering the mental health of children who with visual impairment could be easily solved by distributing eyeglasses with low price that live in low-income rural settings. Policymakers should pay closer attention to the potential human capital loss and social costs of rural students' vision impairment unadjusted borne by this already vulnerable subset of students. Urgent action seems to be called for to design interventions and support programs for uncorrected myopia of vulnerable students in low-income rural settings.

Table 4.1: Summary Statistics of Sample Students and Their Family Characteristics

Variable	Full sample	By visual impairment			Differences	
	N = 19,425	Normal vision N = 12,204	Visual impairment with glasses N = 3133	Visual impairment without glasses N = 4088	Visual impairment with glasses vs normal vision	Visual impairment without glasses vs normal vision
	(1)	(2)	(3)	(4)	(3)–(2)	(4)–(2)
	Mean (SD)				Coefficient (P-value)	
Girl (1 = yes)	0.486 (0.499)	0.427 (0.494)	0.625 (0.484)	0.554 (0.497)	0.197 (< 0.001)	0.127 (< 0.001)
Hui ethnicity (1 = yes)	0.649 (0.477)	0.678 (0.466)	0.510 (0.499)	0.670 (0.470)	– 0.169 (< 0.001)	– 0.009 (0.299)
Rural residence (1 = yes)	0.936 (0.243)	0.940 (0.237)	0.920 (0.270)	0.939 (0.237)	– 0.019 (< 0.001)	– 0.000 (0.964)
Only child (1 = yes)	0.068 (0.251)	0.069 (0.254)	0.073 (0.260)	0.059 (0.236)	0.003 (0.501)	– 0.010 (0.028)
Grade (1 = seven)	0.495 (0.499)	0.523 (0.499)	0.390 (0.488)	0.489 (0.499)	– 0.133 (< 0.001)	– 0.035 (< 0.001)
Boarding at school (1 = yes)	0.354 (0.478)	0.345 (0.475)	0.360 (0.480)	0.376 (0.484)	0.015 (0.127)	0.031 (< 0.001)
Attending extracurricular classes (1 = yes)	0.113 (0.317)	0.114 (0.318)	0.124 (0.330)	0.103 (0.305)	0.010 (0.113)	– 0.010 (0.069)
Parents divorced (1 = yes)	0.097 (0.296)	0.100 (0.300)	0.082 (0.275)	0.099 (0.299)	– 0.018 (< 0.001)	– 0.001 (0.857)
Either parent migrated (1 = yes)	0.348 (0.476)	0.341 (0.474)	0.371 (0.483)	0.353 (0.478)	0.029 (< 0.001)	0.012 (0.173)
Grandparent as primary caregiver (1 = yes)	0.118 (0.323)	0.117 (0.322)	0.112 (0.316)	0.124 (0.329)	– 0.005 (0.437)	0.006 (0.285)
Father's age (years)	41.343 (5.314)	41.343 (5.350)	41.231 (5.156)	41.431 (5.323)	– 0.112 (0.292)	0.088 (0.360)
Mother's age (years)	38.674 (5.030)	38.653 (5.074)	38.755 (4.901)	38.677 (4.995)	0.103 (0.309)	0.025 (0.786)
Father's education (years)	7.083 (3.202)	7.006 (3.221)	7.554 (2.977)	6.950 (3.282)	0.548 (< 0.001)	– 0.056 (0.339)
Mother's education (years)	5.137 (4.015)	5.091 (3.992)	5.623 (4.048)	4.901 (4.026)	0.532 (< 0.001)	– 0.190 (< 0.009)
Household assets (1 = bottom third of sample)	0.333 (0.471)	0.336 (0.472)	0.273 (0.445)	0.368 (0.482)	– 0.063 (< 0.001)	0.032 (< 0.001)

Table 4.2: Distribution of Mental Health, Education Aspiration, and Academic Performance of Sample Students

Variable	Full sample	By visual impairment			Differences	
	N = 19,425	Normal vision N = 12,204	Visual impairment with glasses N = 3133	Visual impairment without glasses N = 4088	Visual impairment with glasses vs normal vision	Visual impairment without glasses vs normal vision
	(1)	(2)	(3)	(4)	(3)–(2)	(4)–(2)
	Mean (SD)				Coefficient (P-value)	
Panel A: Student mental health						
Strengths and Difficulties Questionnaire (SDQ) total scores	12.605 (4.915)	12.604 (4.950)	12.253 (4.769)	12.876 (4.901)	− 0.351 (< 0.001)	0.272 (< 0.001)
Internalizing score	6.453 (2.980)	6.416 (2.983)	6.344 (2.944)	6.648 (2.991)	− 0.072 (0.225)	0.232 (< 0.001)
Externalizing score	6.151 (2.922)	6.188 (2.939)	5.909 (2.847)	6.227 (2.917)	− 0.279 (< 0.001)	0.040 (0.454)
SDQ prevalence of abnormal scores	0.093 (0.290)	0.094 (0.292)	0.078 (0.268)	0.102 (0.303)	− 0.016 (0.006)	0.009 (0.107)
Internalizing SDQ abnormality	0.151 (0.358)	0.150 (0.357)	0.138 (0.345)	0.161 (0.368)	− 0.012 (0.087)	0.011 (0.084)
Externalizing SDQ abnormality	0.173 (0.378)	0.178 (0.383)	0.146 (0.353)	0.178 (0.382)	− 0.032 (< 0.001)	− 0.001 (0.937)
Panel B: Student education aspiration						
Very likely to take high school entrance exam	0.728 (0.444)	0.709 (0.454)	0.802 (0.398)	0.727 (0.445)	0.093 (< 0.001)	0.018 (0.024)
Most likely to attend academic high school after junior high school	0.734 (0.442)	0.714 (0.451)	0.807 (0.394)	0.733 (0.442)	0.092 (< 0.001)	0.019 (0.020)
Most likely to attend vocation high school after graduation	0.208 (0.405)	0.221 (0.415)	0.155 (0.362)	0.208 (0.406)	− 0.065 (< 0.001)	− 0.013 (0.086)
Most likely to work after junior high school	0.058 (0.234)	0.063 (0.244)	0.037 (0.188)	0.057 (0.233)	− 0.027 (< 0.001)	− 0.006 (0.169)
Panel C: Student standardized test score						
Math test scores, (standardized)	0.006 (0.995)	− 0.037 (0.989)	0.261 (0.980)	− 0.054 (0.992)	0.299 (< 0.001)	− 0.017 (0.337)

Table 4.3: Associations Between Student Vision, Eyeglasses Usage and Mental Health

	SDQ scores			SDQ abnormality		
	SDQ Total score	Internalizing SDQ score	Externalizing SDQ score	SDQ total abnormality	Internalizing SDQ abnormality	Externalizing SDQ abnormality
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Full sample						
Normal vision as comparison						
Visual impairment (1 = yes)	0.148** (0.072)	0.105** (0.044)	0.044 (0.043)	0.004 (0.004)	0.002 (0.005)	0.002 (0.006)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	19,425	19,425	19,425	19,425	19,425	19,425
R^2	0.060	0.045	0.054	0.022	0.016	0.031
Panel B: Full sample						
Normal vision as comparison						
Visual impairment with eyeglasses (1 = yes)	0.034 (0.098)	0.035 (0.060)	− 0.001 (0.059)	− 0.002 (0.006)	− 0.003 (0.007)	− 0.002 (0.008)
Visual impairment without eyeglasses (1 = yes)	0.230*** (0.087)	0.155*** (0.053)	0.075 (0.052)	0.008 (0.005)	0.006 (0.006)	0.005 (0.007)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	19,425	19,425	19,425	19,425	19,425	19,425
R^2	0.075	0.052	0.051	0.027	0.019	0.038
Panel C: Among visual impairment students						
Visual impairment without eyeglasses as comparison						
Visual impairment with eyeglasses (1 = yes)	− 0.203* (0.116)	− 0.128* (0.072)	− 0.075 (0.069)	− 0.011 (0.007)	− 0.009 (0.009)	− 0.012 (0.009)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	7221	7221	7221	7221	7221	7221
R^2	0.061	0.045	0.054	0.022	0.018	0.026
Avg Dep Var	12.605	6.453	6.151	0.093	0.151	0.173

Table 4.4: Associations Between Student Eyeglasses Usage and Mental Health Among Students with Poor Vision, by academic performance

	SDQ scores			SDQ abnormality		
	SDQ Total score	Internalizing SDQ score	Externalizing SDQ score	SDQ total abnormality	Internalizing SDQ abnormality	Externalizing SDQ abnormality
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Standardized math test score above the median						
Visual impairment with eyeglasses (1 = yes)	− 0.436*** (0.158)	− 0.205** (0.098)	− 0.231** (0.095)	− 0.023*** (0.008)	− 0.017 (0.011)	− 0.028** (0.011)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	3696	3696	3696	3696	3696	3696
R^2	0.023	0.027	0.019	0.010	0.009	0.008
Panel B: Standardized math test score below the median						
Visual impairment with eyeglasses (1 = yes)	− 0.092 (0.172)	− 0.105 (0.106)	0.014 (0.102)	− 0.003 (0.012)	− 0.006 (0.014)	0.001 (0.014)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	3525	3525	3525	3525	3525	3525
R^2	0.025	0.027	0.026	0.017	0.013	0.018
Avg Dep Var	12.605	6.453	6.151	0.093	0.151	0.173

Table 4.5: Associations Between Eyeglasses Usage and Educational Aspirations Among Students with Poor Vision, by academic performance

	Very likely to take high school entrance exam	Most likely to attend academic high school after graduation	Most likely to work after graduation
	(1)	(2)	(4)
Panel A: Standardized math test score above the median			
Visual impairment with eyeglasses (1 = yes)	0.019* (0.012)	0.022* (0.012)	− 0.002 (0.005)
Control variables	Yes	Yes	Yes
No. of observations	3696	3696	3696
R^2	0.026	0.024	0.013
Panel B: Standardized math test score below the median			
Visual impairment with eyeglasses (1 = yes)	0.052*** (0.017)	0.050*** (0.017)	− 0.017* (0.009)
Controls variables	Yes	Yes	Yes
No. of observations	3525	3525	3525
R^2	0.020	0.017	0.024
Avg Dep Var	0.728	0.733	0.058

Bibliography

Abeberese, Ama Baafra, Todd J. Kumler, and Leigh L. Linden. 2014. "Improving Reading Skills by Encouraging Children to Read in School: A Randomized Evaluation of the Sa Aklat Sisikat Reading Program in the Philippines." *The Journal of Human Resources*, 49(3): 611–633. 7

Albouy, Valerie, and Laurent Lequien. 2009. "Does Compulsory Education Lower Mortality?" *Journal of Health Economics*, 28(1): 155–168. 69

Almroth, Melody C, Krisztina D László, Kyriaki Kosidou, and Maria R Galanti. 2018. "Association Between Adolescents' Academic Aspirations and Expectations and Mental Health: A One-year Follow-up Study." *The European Journal of Public Health*, 28(3): 504–509. 132

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association*, 103(484): 1481–1495. 74, 112, 113

Arseneault, Louise, Barry J Milne, Alan Taylor, Felicity Adams, Kira Delgado, Avshalom Caspi, and Terrie E Moffitt. 2008. "Being Bullied as an Environmentally Mediated Contributing Factor to Children's Internalizing Problems: a Study of Twins Discordant for Victimization." *Archives of Pediatrics & Adolescent Medicine*, 162(2): 145–150. 131

Avendano, Mauricio, Augustin de Coulon, and Vahé Nafilyan. 2020. "Does Longer Compulsory Schooling Affect Mental Health? Evidence From a British Reform." *Journal of Public Economics*, 183: 104137. 69

Banerjee, Abhijit V, Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence From Two Randomized Experiments in India." *The Quarterly Journal of Economics*, 122(3): 1235–1264. 6

Barrera-Osorio, Felipe, and Leigh L Linden. 2009. "The Use and Misuse of Computers in Education: Evidence from a Randomized Controlled Trial of a Language Arts Program." *Cambridge, MA: Abdul Latif Jameel Poverty Action Lab (JPAL)*, 20. 6

Barro, Robert J. 2001. "Human Capital and Growth." *American Economic Review*, 91(2): 12–17. 1

Barrow, Lisa, and Ofer Malamud. 2015. "Is College a Worthwhile Investment?" *Annual Review of Economics*, 7(1): 519–555. 65

Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1): 289–300. 18, 19, 32, 33, 34, 35, 36, 37, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 60, 61, 62, 63, 64

Bertrand, Marianne, and Sendhil Mullainathan. 2001. "Do People Mean What They Say? Implications for Subjective Survey Data." *The American Economic Review*, 91(2): 67–72. 91

Biederman, Joseph, Haitao Gao, Ann K. Rogers, and Thomas J. Spencer. 2006. "Comparison of Parent and Teacher Reports of Attention-deficit/hyperactivity disorder Symptoms From Two Placebo-controlled Studies of Atomoxetine in Children." *Biological Psychiatry*, 60(10): 1106–1110. 27

Bitsko, Rebecca H. 2022. "Mental Health Surveillance Among Children — United States, 2013–2019." *MMWR Supplements*, 71. 8

Bleakley, Hoyt. 2010. "Health, Human capital, and Development." *Annu. Rev. Econ.*, 2(1): 283–310. 1, 7

Borghans, Lex, Bart H. H. Golsteyn, James J. Heckman, and John Eric Humphries. 2016. "What Grades and Achievement Tests Measure." *Proceedings of the National Academy of Sciences of the United States of America*, 113(47): 13354–13359. 14

Borkum, Evan, Fang He, and Leigh L Linden. 2012. "The Effects of School Libraries on Language Skills: Evidence from a Randomized Controlled Trial in India." National Bureau of Economic Research. 7, 26

Bourdon, Karen H, Robert Goodman, Donald S Rae, Gloria Simpson, and Doreen S Koretz. 2005. "The Strengths and Difficulties Questionnaire: US normative data and psychometric properties." *Journal of the American Academy of Child & Adolescent Psychiatry*, 44(6): 557–564. 130

Bourne, Rupert RA, Seth R Flaxman, Tasanee Braithwaite, Maria V Cicinelli, Aditi Das, Jost B Jonas, Jill Keeffe, John H Kempen, Janet Leasher, Hans Limburg, et al. 2017. "Magnitude, Temporal trends, and Projections of the Global Prevalence of Blindness and Distance and Near Vision Impairment: a Systematic Review and Meta-analysis." *The Lancet Global Health*, 5(9): e888–e897. 122

Brunello, Giorgio, Daniele Fabbri, and Margherita Fort. 2013. "The Causal Effect of Education on Body Mass: Evidence from Europe." *Journal of Labor Economics*, 31(1): 195–223. 69

Buckles, Kasey, Andreas Hagemann, Ofer Malamud, Melinda Morrill, and Abigail Wozniak. 2016. "The Effect of College Education on Mortality." *Journal of Health Economics*, 50: 99–114. 69

Cai, Xiqian, Yi Lu, Jessica Pan, and Songfa Zhong. 2019. "Gender Gap under Pressure: Evidence from China's National College Entrance Examination." *The Review of Economics and Statistics*, 101(2): 249–263. 70

Card, David. 1999. "The Causal Effect of Education on Earnings." *Handbook of Labor Economics*, 3: 1801–1863. 1

Carnoy, Martin, Prashant Loyalka, Maria Dobryakova, Rafiq Dossani, Isak Froumin, Katherine Kuhns, Jandhyala Tilak, and Rong Wang. 2013. *University Expansion in a Changing Global Economy: Triumph of the BRICs?* Stanford University Press. 65, 71

Chen, Yan, and Onur Kesten. 2016. "Chinese College Admissions and School Choice Reforms: A Theoretical Analysis." *Journal of Political Economy*, 125(1): 99–139. 70

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *The Quarterly Journal of Economics*, 129(4): 1553–1623. 1

Che, Yi, and Lei Zhang. 2018. "Human Capital, Technology Adoption and Firm Performance: Impacts of China's Higher Education Expansion in the Late 1990s." *The Economic Journal*, 128(614): 2282–2320. 68

China Statistics Press. 2018. "China National Bureau of Statistics 2018 [in Chinese GDP]." 124

Clark, Andrew E. 2018. "Four Decades of the Economics of Happiness: Where Next?" *Review of Income and Wealth*, 64(2): 245–269. 68

Clark, Andrew E., and Tom Lee. 2021. "Early-life Correlates of Later-life Well-being: Evidence From the Wisconsin Longitudinal Study." *Journal of Economic Behavior & Organization*, 181: 360–368. 91

Clark, Damon, and Heather Royer. 2013. "The Effect of Education on Adult Mortality and Health: Evidence from Britain." *American Economic Review*, 103(6): 2087–2120. 69

Congdon, Nathan, Yunfei Wang, Yue Song, Kai Choi, Mingzhi Zhang, Zhongxia Zhou, Zhenling Xie, Liping Li, Xueyu Liu, Abhishek Sharma, et al. 2008. "Visual Disability, Visual Function, and Myopia Among Rural Chinese Secondary School Children: the Xichang Pediatric Refractive Error Study (X-PRES)—Report 1." *Investigative Ophthalmology & Visual Science*, 49(7): 2888–2894. 122

Cordell, Antoinette S, and Terry Cannon. 1985. "Gifted Kids Can't Always Spell." *Academic Therapy*, 21(2): 143–152. 8, 11

Crespo, Laura, Borja López-Noval, and Pedro Mira. 2014. "Compulsory Schooling, Education, Depression and Memory: New Evidence from SHARELIFE." *Economics of Education Review*, 43: 36–46. 69

Currie, Janet, and Mark Stabile. 2006. "Child Mental Health and Human Capital Accumulation: The Case of ADHD." *Journal of Health Economics*, 25(6): 1094–1118. 1, 7, 8, 11

Cutler, David M, and Adriana Lleras-Muney. 2006. "Education and Health: Evaluating Theories and Evidence." National Bureau of Economic Research. 82

Cutler, David M, and Adriana Lleras-Muney. 2012. "Education and Health: Insights From International Comparisons." National Bureau of Economic Research. 82

Dahmann, Sarah C., and Daniel D. Schnitzlein. 2019. "No Evidence for a Protective Effect of Education on Mental Health." *Social Science & Medicine*, 241: 112584. 69

Daley, David, and James Birchwood. 2010. "ADHD and Academic Performance: Why Does ADHD Impact on Academic Performance and What Can Be Done to Support ADHD Children in the Classroom?" *Child: Care, Health and Development*, 36(4): 455–464. 7, 8, 11

Danielson, Melissa L, Rebecca H Bitsko, Reem M Ghandour, Joseph R Holbrook, Michael D Kogan, and Stephen J Blumberg. 2018. "Prevalence of Parent-reported ADHD Diagnosis and Associated Treatment Among US Children and Adolescents, 2016." *Journal of Clinical Child & Adolescent Psychology*, 47(2): 199–212. 8

Davey, Gareth, Chuan De Lian, and Louise Higgins. 2007. "The University Entrance Examination System in China." *Journal of Further and Higher Education*, 31(4): 385–396. 70

Deci, Edward L, and Richard M Ryan. 2004. *Handbook of Self-determination Research.* University Rochester Press. 21, 24

Deci, Edward L, and Richard M Ryan. 2013. *Intrinsic Motivation and Self-determination in Human Behavior.* Springer Science & Business Media. 21, 24

De Walque, Damien. 2007. "Does Education Affect Smoking Behaviors?: Evidence Using the Vietnam Draft as an Instrument for College Education." *Journal of Health Economics*, 26(5): 877–895. 69

Dorman, Jeffrey P. 2001. "Associations Between Classroom Environment and Academic Efficacy." *Learning Environments Research*, 4: 243–257. 10

Dudovitz, Rebecca N, Nilufar Izadpanah, Paul J Chung, and Wendelin Slusser. 2016. "Parent, Teacher, and Student Perspectives on How Corrective Lenses Improve Child Wellbeing and School Function." *Maternal and child health journal*, 20: 974–983. 122

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence From an Unusual Policy Experiment." *American Economic Review*, 91(4): 795–813. 66, 77, 90

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2015. "School Governance, Teacher Incentives, and Pupil–teacher Ratios: Experimental Evidence from Kenyan Primary Schools." *Journal of Public Economics*, 123: 92–110. 7

Duflo, Esther, Rema Hanna, and Stephen P Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review*, 102(4): 1241–1278. 6

Du, Kang, Jiaqi Zhu, Hongyu Guan, Yunyun Zhang, Huan Wang, Decai Wang, and Yaojiang Shi. 2023. "Factors Associated with the Spectacle Wear Compliance Among Primary School Students with Refractive Error in Rural China." *Ophthalmic Epidemiology*, 30(1): 17–26. 123

DuPaul, George J, and Gary Stoner. 2014. *ADHD in the Schools: Assessment and Intervention Strategies.* Guilford Publications. 11

Dursun, Bahadir, and Resul Cesur. 2016. "Transforming Lives: the Impact of Compulsory Schooling on Hope and Happiness." *Journal of Population Economics*, 29(3): 911–956. 69

Dörnyei, Zoltán, and Christine Muir. 2019. "Creating a Motivating Classroom Environment." *Second handbook of English language teaching*, 719–736. 11

Erskine, Holly E., Rosana E. Norman, Alize J. Ferrari, Gary C. K. Chan, William E. Copeland, Harvey A. Whiteford, and James G. Scott. 2016. "Long-Term Outcomes of Attention-Deficit/Hyperactivity Disorder and Conduct Disorder: ASystematic Review and Meta-Analysis." *Journal of the American Academy of Child and Adolescent Psychiatry*, 55(10): 841–850. 8, 11

Essau, Cecilia A, Patrick WL Leung, Judith Conradt, Halina Cheng, and Tiffany Wong. 2008. "Anxiety Symptoms in Chinese and German Adolescents: Their Relationship with Early Learning Experiences, Perfectionism, and Learning Motivation." *Depression and anxiety*, 25(9): 801–810. 123

Fan, Elliott, Xin Meng, Zhichao Wei, and Guochang Zhao. 2018. "Rates of Return to Four-Year University Education: An Application of Regression Discontinuity Design." *The Scandinavian Journal of Economics*, 120(4): 1011–1042. 70

Faraone, Stephen V., Joseph Biederman, and Brenda Zimmerman. 2005. "Correspondence of Parent and Teacher Reports in Medication Trials." *European Child & Adolescent Psychiatry*, 14(1): 20–27. 27

Fazel, Mina, Kimberly Hoagwood, Sharon Stephan, and Tamsin Ford. 2014*a*. "Mental Health Interventions in Schools in High-income Countries." *The Lancet Psychiatry*, 1(5): 377–387. 1, 7, 8

Fazel, Mina, Vikram Patel, Saji Thomas, and Wietse Tol. 2014*b*. "Mental Health Interventions in Schools in Low-income and Middle-Income Countries." *The Lancet Psychiatry*, 1(5): 388–398. 8

Feng, Shuaizhang, and Xiaoyu Xia. 2022. "Heterogeneous Firm Responses to Increases in High-skilled Workers: Evidence from China's College Enrollment Expansion." *China Economic Review*, 73: 101791. 68, 72, 77

Feng, Tianli, Xiyuan Jia, Lucy Pappas, Xiaojun Zheng, Teresa Shao, Letao Sun, Charlie Weisberg, Madeline Lu Li, Scott Rozelle, and Yue Ma. 2022. "Academic Performance and the Link with Depressive Symptoms among Rural Han and Minority Chinese Adolescents." *International Journal of Environmental Research and Public Health*, 19(10): 6026. 132

Feng, Yanxiang, Xinyang Tan, and Ruixin Wang. 2022. "The Value of Higher Education to Entrepreneurial Performance: Evidence from Higher Education Expansion in China." *China Economic Review*, 73: 101789. 68

Ferreyra, María Marta, Ciro Avitabile, Javier Botero Álvarez, Francisco Haimovich Paz, and Sergio Urzúa. 2017. *At a Crossroads: Higher Education in Latin America and the Caribbean.* World Bank Publications. 65

Fletcher, Jason, and Barbara Wolfe. 2008. "Child Mental Health and Human Capital Accumulation: the Case of ADHD Revisited." *Journal of Health Economics*, 27(3): 794–800. 8, 11

Fraser, Barry. 2012. *Classroom Environment.* Routledge. 10

Frazier, Thomas W, Eric A Youngstrom, Joseph J Glutting, and Marley W Watkins. 2007. "ADHD and Achievement: Meta-analysis of the Child, Adolescent, and Adult Literatures and a Concomitant Study with College Students." *Journal of learning disabilities*, 40(1): 49–65. 8

Fu, Hongqiao, Run Ge, Jialin Huang, and Xinzheng Shi. 2022. "The Effect of Education on Health and Health Behaviors: Evidence from the College Enrollment Expansion in China." *China Economic Review*, 72: 101768. 68

Ganimian, Alejandro J, and Richard J Murnane. 2016. "Improving Education in Developing Countries: Lessons from Rigorous Impact Evaluations." *Review of Educational Research*, 86(3): 719–755. 6

Gao, Qiufeng, Huan Wang, Di Mo, Yaojiang Shi, Kaleigh Kenny, and Scott Rozelle. 2018. "Can Reading Programs Improve Reading Skills and Academic Performance in Rural China?" *China Economic Review*, 52: 111–125. 11, 26

Gao, Qiufeng, Huan Wang, Fang Chang, Hongmei Yi, and Yaojiang Shi. 2021. "Reading Achievement in China's Rural Primary Schools: a Study of Three Provinces." *Educational Studies*, 47(3). 11

George, Anne L, and Cheryll Duquette. 2006. "The Psychosocial Experiences of a Student with Low Vision." *Journal of Visual Impairment & Blindness*, 100(3): 152–163. 122

Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 2009. "Many Children Left Behind? Textbooks and Test Scores in Kenya." *American Economic Journal: Applied Economics*, 1(1): 112–135. 6, 7

Glewwe, Paul, Michael Kremer, Sylvie Moulin, and Eric Zitzewitz. 2004. "Retrospective vs. prospective analyses of school inputs: the case of flip charts in Kenya." *Journal of Development Economics*, 74(1): 251–268. 7

Goodman, Robert. 1997. "The Strengths and Difficulties Questionnaire: a Research Note." *Journal of Child Psychology and Psychiatry*, 38(5): 581–586. 125

Grossman, Michael. 2006. "Education and Nonmarket Outcomes." *Handbook of the Economics of Education*, 1: 577–633. 1, 65

Guan, Hongyu, Huan Wang, Kang Du, Jin Zhao, Matthew Boswell, Yaojiang Shi, and Yiwei Qian. 2018. "The Effect of Providing Free Eyeglasses on Children's Mental Health Outcomes in China: a Cluster-randomized Controlled Trial." *International Journal of Environmental Research and Public Health*, 15(12): 2749. 122, 130, 131, 132

Haidar, Suzan A, NK De Vries, Mirey Karavetian, and Rola El-Rassi. 2018. "Stress, Anxiety, and Weight Gain Among University and College Students: a Systematic Review." *Journal of the Academy of Nutrition and Dietetics*, 118(2): 261–274. 66

Hanushek, Eric A, and Ludger Woessmann. 2008. "The Role of Cognitive Skills in Economic Development." *Journal of Economic Literature*, 46(3): 607–668. 1

Hautala, Helena, and Hannu Lehti. 2024. "Children's Perceived Economic Disadvantage and Social Relationships in School: Family Relationships as a Mediating Factor." *Child Indicators Research*, 1–24. 87

He, Mingguang, Junwen Zeng, Yizhi Liu, Jingjing Xu, Gopal P Pokharel, and Leon B Ellwein. 2004. "Refractive Error and Visual Impairment in Urban Children in Southern China." *Investigative Ophthalmology & Visual Science*, 45(3): 793–799. 130

Hill, Kent, Dennis Hoffman, and Tom R Rex. 2005. "The Value of Higher Education: Individual and Societal Benefits." *Arizona State University*. 65

Holguin, Aaron M Castanon, Nathan Congdon, Nita Patel, Amy Ratcliffe, Paul Esteso, Silvia Toledo Flores, Donna Gilbert, Marco A Pereyra Rito, and Beatriz Munoz. 2006. "Factors Associated with Spectacle-wear Compliance in School-aged Mexican Children." *Investigative Ophthalmology & Visual Science*, 47(3): 925–928. 131

Hong, Lin, Wang Yufeng, Kingsley Agho, and Jennifer Jacobs. 2011. "Preventing Behavior Problems among Elementary Schoolchildren: Impact of a Universal School-based Program in China." *The Journal of School Health*, 81(5): 273–280. 8

Horwood, Jeremy, Andrea Waylen, David Herrick, Cathy Williams, and Dieter Wolke. 2005. "Common Visual Defects and Peer Victimization in Children." *Investigative Ophthalmology & Visual Science*, 46(4): 1177–1181. 131

Hout, Michael. 2012. "Social and Economic Returns to College Education in the United States." *Annual Review of Sociology*, 38(1): 379–400. 65

Hovi, Matti, and Jani-Petri Laamanen. 2021. "Income, Aspirations and Subjective Well-being: International Evidence." *Journal of Economic Behavior & Organization*, 185: 287–302. 81, 91

Huang, Liqing, and Zhanjun Xing. 2019. "A Nationwide Study on Subjective Well-being Scale for Chinese Citizens (SWBS-cc)." *International Journal of Applied Psychology*, 9(5): 117–127. 73, 112

Hu, Hongwei, Shuang Lu, and Chien-Chung Huang. 2014. "The Psychological and Behavioral Outcomes of Migrant and Left-Behind Children in China." *Children and Youth Services Review*, 46: 1–10. 130

Jiang, Qi, Xinshu She, Sarah-Eve Dill, Sean Sylvia, Manpreet Kaur Singh, Huan Wang, Matthew Boswell, and Scott Rozelle. 2022. "Depressive and Anxiety Symptoms among Children and Adolescents in Rural China: A Large-Scale Epidemiological Study." *International Journal of Environmental Research and Public Health*, 19(9). 8, 11

Jiang, Wei, Yi Lu, and Huihua Xie. 2020. "Education and Mental Health: Evidence and Mechanisms." *Journal of Economic Behavior & Organization*, 180: 407–437. 69

Jia, Ruixue, and Hongbin Li. 2021. "Just Above the Exam Cutoff Score: Elite College Admission and Wages in China." *Journal of Public Economics*, 196: 104371. 65, 66, 70, 86

Jordans, Mark JD, Wietse A Tol, Dessy Susanty, Prudence Ntamatumba, Nagendra P Luitel, Ivan H Komproe, and Joop TVM de Jong. 2013. "Implementation of a Mental Health Care Package for Children in Areas of Armed Conflict: a Case Study from Burundi, Indonesia, Nepal, Sri Lanka, and Sudan." *PLoS Medicine*, 10(1): e1001371. 8

Kahneman, Daniel, and Angus Deaton. 2010. "High Income Improves Evaluation of Life But Not Emotional Well-being." *Proceedings of the National Academy of Sciences*, 107(38): 16489–16493. 81, 91

Karunamuni, Nandini, Ikuyo Imayama, and Dharshini Goonetilleke. 2021. "Pathways to Well-being: Untangling the Causal Relationships Among Biopsychosocial Variables." *Social Science & Medicine*, 272: 112846. 1, 7

Keane, Michael P, and Timothy Neal. 2024. "A Practical Guide to Weak Instruments." *Annual Review of Economics*, 16. 82

Kelly, Erin V, Nicola C Newton, Lexine A Stapinski, Tim Slade, Emma L Barrett, Patricia J Conrod, and Maree Teesson. 2015. "Suicidality, Internalizing Problems and Externalizing Problems Among Adolescent Bullies, Victims and Bully-victims." *Preventive Medicine*, 73: 100–105. 131

Kemptner, Daniel, Hendrik Jürges, and Steffen Reinhold. 2011. "Changes in Compulsory Schooling and the Causal Effect of Education on Health: Evidence from Germany." *Journal of Health Economics*, 30(2): 340–354. 69

Kieling, Christian, Helen Baker-Henningham, Myron Belfer, Gabriella Conti, Ilgi Ertem, Olayinka Omigbodun, Luis Augusto Rohde, Shoba Srinath, Nurper Ulkuer, and Atif Rahman. 2011. "Child and Adolescent Mental Health Worldwide: Evidence for Action." *The Lancet*, 378(9801): 1515–1525. 1, 7, 8

Kimenyi, Mwangi S. 2011. "Contribution of Higher Education to Economic Development: A Survey of International Evidence." *Journal of African Economies*, 20(suppl_3): iii14–iii49. 65

Kinderman, Peter, Sara Tai, Eleanor Pontin, Matthias Schwannauer, Ian Jarman, and Paulo Lisboa. 2015. "Causal and Mediating Factors for Anxiety, Depression and Well-being." *The British Journal of Psychiatry*, 206(6): 456–460. 1, 7

Knight, John, Quheng Deng, and Shi Li. 2017. "China's Expansion of Higher Education: The Labour Market Consequences of a Supply Shock." *China Economic Review*, 43: 127–141. 68

Kondirolli, Fjolla, and Naveen Sunder. 2022. "Mental Health Effects of Education." *Health Economics*, 31: 22–39. 69

Kremer, Michael, Conner Brannen, and Rachel Glennerster. 2013. "The Challenge of Education and Learning in the Developing World." *Science*, 340(6130): 297–300. 1, 6

Lai, Kelly YC, Ernest SL Luk, Patrick WL Leung, Ann SY Wong, Lawrence Law, and Karen Ho. 2010. "Validation of the Chinese Version of the Strengths and Difficulties Questionnaire in Hong Kong." *Social Psychiatry and Psychiatric Epidemiology*, 45: 1179–1186. 125

Lee, Albert. 2009. "Health-promoting Schools: Evidence for a Holistic Approach to Promoting Health and Improving Health Literacy." *Applied Health Economics and Health Policy*, 7(1): 11–17. 8

Leppink, Eric W, Brian L Odlaug, Katherine Lust, Gary Christenson, and Jon E Grant. 2016. "The Young and the Stressed: Stress, Impulse Control, and Health in College Students." *The Journal of Nervous and Mental Disease*, 204(12): 931–938. 66

Li, Dongfeng, Ving Fai Chan, Gianni Virgili, Prabhath Piyasena, Habtamu Negash, Noelle Whitestone, Sara O'Connor, Baixiang Xiao, Mike Clarke, David H Cherwek, et al. 2022. "Impact of Vision Impairment and Ocular Morbidity and Their Treatment on Depression and Anxiety in Children: a Systematic Review." *Ophthalmology*, 129(10): 1152–1170. 122

Li, Hongbin, Yueyuan Ma, Lingsheng Meng, Xue Qiao, and Xinzheng Shi. 2017. "Skill Complementarities and Returns to Higher Education: Evidence from College Enrollment Expansion in China." *China Economic Review*, 46: 10–26. 68

Li, Jiayu, Jing Li, Ruixia Jia, Yingquan Wang, Sheng Qian, and Yong Xu. 2020. "Mental Health Problems and Associated School Interpersonal Relationships Among Adolescents in China: a Cross-sectional Study." *Child and Adolescent Psychiatry and Mental Health*, 14: 1–10. 123

Li, Liping, Jasmine Lam, Yaogui Lu, Yanru Ye, Dennis SC Lam, Yang Gao, Abhishek Sharma, Mingzhi Zhang, Sian Griffiths, and Nathan Congdon. 2010. "Attitudes of Students, Parents, and Teachers Toward Glasses Use in Rural China." *Archives of Ophthalmology*, 128(6): 759–765. 122, 123, 130, 131

Li, Shi, John Whalley, and Chunbing Xing. 2014. "China's Higher Education Expansion and Unemployment of College Graduates." *China Economic Review*, 30: 567–582. 68, 73

Lleras-Muney, Adriana. 2005. "The Relationship Between Education and Adult Mortality in the United States." *The Review of Economic Studies*, 72(1): 189–221. 69

Lochner, Lance, and Enrico Moretti. 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." *American Economic Review*, 94(1): 155–189. 1

Lola, Hirbaye Mokona, Habte Belete, Abebaw Gebeyehu, Aemro Zerihun, Solomon Yimer, and Kassech Leta. 2019. "Attention Deficit Hyperactivity Disorder (ADHD) Among Children Aged 6 to 17 Years Old Living in Girja District, Rural Ethiopia." *Behavioural Neurology*, 2019(1): 1753580. 8

Loyalka, Prashant Kumar. 2009. *Three Essays on Chinese Higher Education After Expansion and Reform: Sorting, Financial Aid, and College Selectivity.* Stanford University. 70

Luo, Dongmei, Xiaojin Yan, Rongbin Xu, Jingshu Zhang, Xinran Shi, Jun Ma, Yi Song, George C. Patton, and Susan M. Sawyer. 2020. "Chinese Trends in Adolescent Marriage and Fertility between 1990 and 2015: A Systematic Synthesis of National and Subnational Population Data." *The Lancet Global Health*, 8(7): e954–e964. 65

MacAulay, Dolina J. 1990. "Classroom Environment: a Literature Review." *Educational Psychology*, 10(3): 239–253. 8, 10

Mannuzza, S., and R. G. Klein. 2000. "Long-term Prognosis in Attention-deficit/hyperactivity disorder." *Child and Adolescent Psychiatric Clinics of North America*, 9(3): 711–726. 8

Maul, Eugenio, Silviana Barroso, Sergio R Munoz, Robert D Sperduto, and Leon B Ellwein. 2000. "Refractive Error Study in Children: Results from La Florida, Chile." *American Journal of Ophthalmology*, 129(4): 445–454. 122

Ma, Xiao. 2024. "College Expansion, Trade, and Innovation: Evidence from China." *International Economic Review*, 65(1): 315–351. 68, 77, 123

Ma, Xiaochen, Nathan Congdon, Hongmei Yi, Zhongqiang Zhou, Xiaopeng Pang, Mirjam E Meltzer, Yaojiang Shi, Mingguang He, Yizhi Liu, and Scott Rozelle. 2015. "Safety of Spectacles for Children's Vision: a Cluster-randomized Controlled Trial." *American Journal of Ophthalmology*, 160(5): 897–904. 122

Ma, Xiaochen, Zhongqiang Zhou, Hongmei Yi, Xiaopeng Pang, Yaojiang Shi, Qianyun Chen, Mirjam E Meltzer, Saskia Le Cessie, Mingguang He, Scott Rozelle, et al. 2014. "Effect of Providing Free Glasses on Children's Educational Outcomes in China: Cluster Randomized Controlled Trial." *BMJ*, 349. 122, 123

Ma, Yue, Nathan Congdon, Yaojiang Shi, Ruth Hogg, Alexis Medina, Matthew Boswell, Scott Rozelle, and Mony Iyer. 2018. "Effect of a Local Vision Care Center on Eyeglasses Use and School Performance in Rural China: a Cluster Randomized Clinical Trial." *JAMA Ophthalmology*, 136(7): 731–737. 122, 130

Ma, Yue, Xinwu Zhang, Fei He, Xiaochen Ma, Hongmei Yi, Nathan Rose, Alexis Medina, Scott Rozelle, and Nathan Congdon. 2022. "Visual Impairment in Rural and Migrant Chinese School-going Children: Prevalence, Severity, Correction and Associations." *British Journal of Ophthalmology*, 106(2): 275–280. 122, 130

McTaggart, Jennifer A. 2009. "The Impact of Motivational Variables in a Reading Remediation Program using a Self-determination Theory Framework." PhD diss. University of Guelph. 3, 8

Miyauchi, Hisae. 2020. "A Systematic Review on Inclusive Education of Students with Visual Impairment." *Education Sciences*, 10(11): 346. 131

Morsink, Sarah, Edmund Sonuga-Barke, Gabry Mies, Nathalie Glorie, Jurgen Lemiere, Saskia Van der Oord, and Marina Danckaerts. 2017. "What Motivates Individuals with ADHD? A Qualitative Analysis from the Adolescent's Point of View." *European Child & Adolescent Psychiatry*, 26(8): 923–932. 8

Morsink, Sarah, Saskia Van der Oord, Inge Antrop, Marina Danckaerts, and Anouk Scheres. 2022. "Studying Motivation in ADHD: The Role of Internal Motives and the Relevance of Self Determination Theory." *Journal of Attention Disorders*, 26(8): 1139–1158. 3, 8, 11, 24

Mullis, Ina VS, Michael O Martin, Pierre Foy, and Kathleen T Drucker. 2012. *PIRLS 2011 international results in reading.* ERIC. 14, 22

Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy*, 119(1): 39–77. 6

Niemiec, Christopher P, and Richard M Ryan. 2009. "Autonomy, Competence, and Relatedness in the Classroom: Applying Self-determination Theory to Educational Practice." *Theory and Research in Education*, 7(2): 133–144. 3, 8

OECD. 2013. *OECD Guidelines on Measuring Subjective Well-being.* OECD. 81

Oreopoulos, Philip. 2006. "Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter." *American Economic Review*, 96(1): 152–175. 69

Oreopoulos, Philip, and Kjell G Salvanes. 2011. "Priceless: The Nonpecuniary Benefits of Schooling." *Journal of Economic Perspectives*, 25(1): 159–184. 1, 65

Pang, Xiaodong, Huan Wang, Sarah-Eve Dill, Matthew Boswell, Xiaopeng Pang, Manpreet Singh, and Scott Rozelle. 2021. "Attention Deficit Hyperactivity Disorder (ADHD) Among Elementary Students in Rural China: Prevalence, Correlates, and Consequences." *Journal of Affective Disorders*, 293: 484–491. 8, 11, 14

Patel, Vikram, Alan J. Flisher, Anula Nikapota, and Savita Malhotra. 2008. "Promoting Child and Adolescent Mental Health in Low and Middle Income Countries." *Journal of Child Psychology and Psychiatry*, 49(3). 8

Pinquart, Martin, and Jens P Pfeiffer. 2012. "Psychological Adjustment in Adolescents with Vision Impairment." *International Journal of Disability, Development and Education*, 59(2): 145–155. 122, 131

Pradhan, Menno, Daniel Suryadarma, Amanda Beatty, Maisy Wong, Arya Gaduh, Armida Alisjahbana, and Rima Prama Artha. 2014. "Improving Educational Quality Through Enhancing Community Participation: Results from a Randomized Field Experiment in Indonesia." *American Economic Journal: Applied Economics*, 6(2): 105–126. 6

Ralejoe, Malehlanye. 2021. "A Study to Understand the Inclusion of Learners With and Without Visual Impairment in a Secondary School in Lesotho." *South African Journal of Education*, 41(1): 1–12. 122

Raven, John C. 1941. "Standardization of Progressive Matrices, 1938." *British Journal of Medical Psychology*, 19(1): 137–150. 14

Resnikoff, Serge, Donatella Pascolini, Silvio P Mariotti, and Gopal P Pokharel. 2008. "Global Magnitude of Visual Impairment Caused by Uncorrected Refractive Errors in 2004." *Bulletin of the World Health Organization*, 86(1): 63–70. 123

Ribeiro, Ícaro JS, Rafael Pereira, Ivna V Freire, Bruno G de Oliveira, Cezar A Casotti, and Eduardo N Boery. 2018. "Stress and Quality of Life Among University Students: A Systematic Literature Review." *Health Professions Education*, 4(2): 70–77. 66

Robaei, Dana, Annette Kifley, Kathryn A Rose, and Paul Mitchell. 2006. "Refractive Error and Patterns of Spectacle Use in 12-year-old Australian Children." *Ophthalmology*, 113(9): 1567–1573. 122

Rogers, Maria, and Rosemary Tannock. 2018. "Are Classrooms Meeting the Basic Psychological Needs of Children with ADHD Symptoms? A Self-determination Theory Perspective." *Journal of Attention Disorders*, 22(14): 1354–1360. 8, 11, 24

Scheier, Michael F., Charles S. Carver, and Michael W. Bridges. 1994. "Distinguishing Optimism From Neuroticism (and Trait Anxiety, Self-mastery, and Self-esteem): A Reevaluation of the Life Orientation Test." *Journal of Personality and Social Psychology*, 67(6): 1063–1078. 73, 112

Schendel, Rebecca, and Tristan McCowan. 2016. "Expanding Higher Education Systems in Low-and Middle-income Countries: the Challenges of Equity and Quality." *Higher Education*, 72(4): 407–411. 65

Schofer, Evan, and John W Meyer. 2005. "The Worldwide Expansion of Higher Education in the Twentieth Century." *American Sociological Review*, 70(6): 898–920. 65

Sexton, Calvin, Heather Gelhorn, Jill Bell, and Peter Classi. 2011. "The Co-occurrence of Reading Disorder and ADHD." *Journal of Learning Disabilities*, 45: 538–64. 8

Sims, Stephanie Vetere, Brian Celso, and Ted Lombardo. 2021. "Emotional and Behavioral Assessment of Youths with Visual Impairments Utilizing the BASC-2." *Journal of Visual Impairment & Blindness*, 115(4): 310–318. 131

Sironi, Maria. 2012. "Education and Mental Health in Europe: School Attainment as a Means to Fight Depression." *International Journal of Mental Health*, 41(3): 79–105. 69

Smith, Zoe R., Joshua M. Langberg, Caroline N. Cusick, Cathrin D. Green, and Stephen P. Becker. 2020. "Academic Motivation Deficits in Adolescents with ADHD and Associations with Academic Functioning." *Journal of Abnormal Child Psychology*, 48(2): 237–249. 8

Snyder, C. R., Susie C. Sympson, Florence C. Ybasco, Tyrone F. Borders, Michael A. Babyak, and Raymond L. Higgins. 1996. "Development and Validation of the State Hope Scale." *Journal of Personality and Social Psychology*, 70(2): 321–335. 74, 113

Spitzer, Sonja. 2020. "Biases in Health Expectancies Due to Educational Differences in Survey Participation of Older Europeans: It's Worth Weighting For." *The European Journal of Health Economics*, 21(4): 573–605. 91

Spitzer, Sonja, and Daniela Weber. 2019. "Reporting Biases in Self-assessed Physical and Cognitive Health Status of Older Europeans." *Plos One*, 14(10): e0223526. 91

Suldo, Shannon, Amanda Thalji, and John Ferron. 2011. "Longitudinal Academic Outcomes Predicted by Early Adolescents' Subjective Well-being, Psychopathology, and Mental Health Status Yielded From a Dual Factor Model." *The Journal of Positive Psychology*, 6(1): 17–30. 130, 132

Sun, Jiandong, Michael P Dunne, Xiang-Yu Hou, and Ai-qiang Xu. 2011. "Educational Stress Scale for Adolescents: Development, Validity, and Reliability with Chinese Students." *Journal of Psychoeducational Assessment*, 29(6): 534–546. 123

Su, Yi Esther, Hong Wang, Yao-Guo Geng, Ling Sun, Ya-Song Du, Fang Fan, and Lin-Yan Su. 2015. "Parent Ratings of ADHD Symptoms in Chinese Urban Schoolchildren: Assessment With the Chinese ADHD Rating Scale-IV: Home Version." *Journal of Attention Disorders*, 19(12): 1022–1033. 9, 16, 32

the Ministry of Education of China. 2003. "Primary and Secondary School Library (Room) Regulations (Revised) (in Chinese)." 12

Thomas, Rae, Sharon Sanders, Jenny Doust, Elaine Beller, and Paul Glasziou. 2015. "Prevalence of Attention-deficit/hyperactivity disorder: a Systematic Review and Meta-analysis." *Pediatrics*, 135(4): e994–1001. 11

UNESCO. 2020. *Global Education Monitoring Report 2020: Inclusion and Education-All Means All.* UN. 1

Wang, Huan. 2012. "Thoughts on the Construction of Libraries in Primary Schools and Middle Schools in Zhejiang Province (in Chinese)." *Educational Equipment and Experiment*, 28(9): 53–55. 12

Wang, Huan, Brandon Barket, Sharon Du, Dimitris Friesen, Ezra Kohrman, Esther Tok, Baixiang Xiao, Wenyong Huang, Ving Fai Chan, Graeme MacKenzie, et al. 2021. "The Prevalence and Correlates of Vision Impairment and Glasses Ownership Among Ethnic Minority and Han Schoolchildren in Rural China." *PLoS One*, 16(8): e0256565. 123, 130

Wang, Huan, Chu Yang, Fei He, Yaojiang Shi, Qinghe Qu, Scott Rozelle, and James Chu. 2015. "Mental Health and Dropout Behavior: A Cross-sectional Study of Junior High Students in Northwest Rural China." 41: 1–12. 123

Wang, Huan, Hongyu Guan, Hongmei Yi, Emma Seevak, Reid Manheim, Matthew Boswell, Scott Rozelle, and Sarah Kotb. 2020. "Independent Reading in Rural China's Elementary Schools: A Mixed-Methods Analysis." *International Journal of Educational Development*, 78: 102241. 12

WHO. 2010. "Global Status Report on Noncommunicable Diseases 2010." World Health Organization. 1

WHO. 2022*a*. "Mental disorders." Geneva: World Health Organization. 1

WHO. 2022*b*. "World Mental Health Report: Transforming Mental Health for All." Geneva: World Health Organization. 1

WHO. 2023. "Vision Impairment and Blindness." Geneva: World Health Organization. 122

Wu, Yaowu, and Quan Zhao. 2010. "Higher Education Expansion and Employment of University Graduates." *Economic Research Journal*, 9: 93–108. 68

Yao, Shuqiao, Chenchen Zhang, Xiongzhao Zhu, Xiao Jing, Chad M McWhinnie, and John RZ Abela. 2009. "Measuring Adolescent Psychopathology: Psychometric Properties of the Self-report Strengths and Difficulties Questionnaire in a Sample of Chinese Adolescents." *Journal of Adolescent Health*, 45(1): 55–62. 125

Yekta, Abbasali, Elham Hooshmand, Mohammad Saatchi, Hadi Ostadimoghaddam, Amir Asharlous, Azadeh Taheri, and Mehdi Khabazkhoob. 2022. "Global Prevalence and Causes of Visual Impairment and Blindness in Children: a Systematic Review and Meta-analysis." *Journal of Current Ophthalmology*, 34(1): 1–15. 130

Yi, Hongmei, Di Mo, Huan Wang, Qiufeng Gao, Yaojiang Shi, Paiou Wu, Cody Abbey, and Scott Rozelle. 2019. "Do Resources Matter? Effects of an In-Class Library Project on Student Independent Reading Habits in Primary Schools in Rural China." *Reading Research Quarterly*, 54(3): 383–411. 7, 11, 22, 26

Yi, Hongmei, Linxiu Zhang, Xiaochen Ma, Nathan Congdon, Yaojiang Shi, Xiaopeng Pang, Junxia Zeng, Lei Wang, Matthew Boswell, and Scott Rozelle. 2015. "Poor Vision Among China's Rural Primary School Students: Prevalence, Correlates and Consequences." *China Economic Review*, 33: 247–262. 122, 131

Yu, Wenjing, Cody Abbey, Yiwei Qian, Huan Wang, Scott Rozelle, and Manpreet K Singh. 2022. "Behavioral Strengths and Difficulties and Their Associations with Academic Performance in Math Among Rural Youth in China." Vol. 10, 1642, MDPI. 125

Zhang, Mingzhi, Huan Lv, Yang Gao, Sian Griffiths, Abhishek Sharma, Dennis Lam, Liping Li, Yee Kit Tse, Xiaojian Liu, Daocheng Xu, et al. 2009. "Visual Morbidity Due to Inaccurate Spectacles Among School Children in Rural China: the See Well to Learn Well Project, Report 1." *Investigative Ophthalmology & Visual Science*, 50(5): 2011–2017. 122

Zhang, Xiulan. 2011. *China's Education Development and Policy, 1978-2008.* Vol. 9, Brill. 72