Does Having Advisor-Advisee Relationship with Content Teacher Impact Student Learning? Causal Evidence from Chinese Junior High Schools

Congli Zhang

College of Education, University of Oregon

Author Note

Congli (Claire) Zhang is currently a doctoral candidate at the University of Oregon, Eugene, OR.

I have no known conflict of interest to disclose.

Correspondence concerning this article should be addressed to Congli (Claire) Zhang, 102B Lokey Bldg, 5267 University of Oregon, Eugene, OR 97403. Email: congliz@uoregon.edu

Abstract

In this study, I present novel evidence of the causal impacts of advisor-advisee relationship with content teachers on student performance and subject self-concept. The advised-homeroom setting in China, where advisors also teach content subjects, creates a comparison condition between students being taught by advisors and those by non-advisor teachers. Importantly, the random assignment of advisors and students to homerooms allows me to isolate advisor-advisee relationship from confounding variables and unbiasedly estimate advisor-advisee relationship effects. Analyzing a nationally representative, longitudinal dataset, I found that advisor-advisee relationship improves student performance on math, self-concept in Chinese (language arts), and both performance on and self-concept in English (nationally mandated foreign language), all effect sizes measure up to approximately one standard deviation increase in teacher quality in teacher value-added literature (Hanushek & Rivkin, 2010). In all, advisor-advisee relationship has substantial causal impacts on multiple dimensions of student academic outcomes in Chinese junior high schools.

Key words: advisory program, teacher effects, student outcomes

Does Having Advisor-Advisee Relationship with Content Teacher Impact Student Learning? Causal Evidence from Chinese Junior High Schools

Introduction

Education policymakers and educators constantly search for evidence-based policy initiatives and school practice to improve student outcomes. An evidence-based junior high school movement in the U.S. in the 20th century was the implementation of teacher advisement programs, or advisory programs. Inspired by the strong associations between teacher-student relationship and student social and cognitive development (see Davis, 2003 for a literature review), many schools had developed some forms of advisory program, where a teacher-advisor is assigned to a small group of students to provide individualized support on students' academic and personal developments (Galassi et al., 1997; McClure et al., 2010). Some states and education agencies had further leveraged various policies to reenforce the teacher advisement implementation, for example, Florida passed legislation in the 80's to fund junior high school advisory programs (Galassi et al., 1997). Advisory programs had also been advocated by the National Association of Secondary School Principals (NASSP) and been integrated into multiple school reform efforts such as the Model Schools Project, IDEA's Individually Guided Education, and the Reform in Secondary Education Project in California (Jenkins, 1992).

The national trend has leveled out entering the 21th century but teacher advisement have been preserved as a complement to the school counseling system in many schools across the nation. Unlike the school counseling programs implemented by systematically trained professionals under national guidelines, however, advisory programs vary considerably from school to school. Subject to local policy, advisory programs may be designed to meet one or multiple student needs such as personal advocacy, group identity, development guidance, invigoration, academic performance, and general school business (Galassi et al., 1997), and teacher-advisors' responsibilities are difficult to be universally defined or clearly categorized. As a result, advisory program evaluation studies suffer from the substantial lack of uniformity across advisory programs, data limitation due to the small scopes of implementation, and methodological weaknesses in addressing omitted variable bias created by the prevalent sorting between advisors and students (Galassi et al., 1997). It is no surprise that advisory is an extremely understudied area and that within the scarce literature, findings are considerably inconclusive and mixed (Galassi et al., 1997; McClure et al., 2010).

This study is, to my knowledge, the first large-scale, quasi-experimental study that examines the causal impacts of teacher advisement on student academic outcomes. Of course, it is impossible, and I do not attempt, to evaluate the effects of advisory programs that are implemented differently from school to school, rather, I focus on only the existence of the advisor-advisee relationship and investigate, all else equal, whether it adds to student learning. This, at a minimum, requires not only a formally defined advisor-advisee relationship but also a comparison condition between students taught a subject by a traditional teacher and those taught the same subject by a teacher-advisor. This is not a typical school setting in the U.S. but fortunately common in many countries across the world including China. As detailed in the following Background section, China adopts an advised-homeroom school setting where students are formally grouped into homerooms with one content teacher serving as their advisor, therefore a natural comparison is formed between students who have advisor-advisee relationship with a subject teacher and those who do not (because their advisors teach other subjects).

More importantly, as detailed in the following Background section, most Chinese junior high schools (grades 7-9) have been implementing random assignment strategies upon students' entry to school since the 2006 Compulsory Education Law (henceforth referred to as the 2006 Law; see Appendix B1 for more details), which is documented in the first nationally representative, longitudinal data of junior high school students – China Education Panel Survey (CEPS; see Appendix A Data Description for details). Under random assignment, being taught a subject by advisors or non-advisors are decided exogenously rather than subject to self-selection of students and teachers. This allows me to methodologically isolate advisor-advisee relationship from other observed and unobserved confounding variables to answer a straightforward research question: whether and to what extent having an advisor-advisee relationship with content teacher impacts student academic outcomes?

Using the analytic sample drawn from CEPS, the national dataset of Chinese junior high school students in 2013-14 and 2014-15 and more importantly, leveraging the random assignment of teachers and students in the data, I found that advisor-advisee relationship has substantial impacts on student learning. Specifically, on average, having advisor-advisee relationship with teacher significantly improves student score in English (nationally mandated foreign language) and math by 0.155 standard deviation (SD) and 0.143 SD. Moreover, advisor-advisee relationship improves student self-concept in Chinese (language arts) and English by 0.203 SD and 0.187 SD, on average. These findings hold robust across three different model specifications. The placebo test using an alternative outcome variable, CEPS cognitive test score, yielded consistently insignificant results across three models and three subject samples.

In the following sections, I present the natural experiment and school setting background in China, the data and measures, the theoretical framework underlying teacher advisement, the methods and identification strategy that address the methodological challenges, the results of advisor effects, and a discussion of finding interpretations, study limitations, and policy implications.

Background

Natural Experiment Background in China

I conduct my research in China and identify my population of interest as Chinese public junior high school (grades 7-9) students based on a critical policy consideration: Chinese public junior high schools are under a national law that drives the implementation of random assignments of teachers and students. Specifically, in 2006, with great attention to education equality, the 2006 Law (Appendix B1) called off student tracking at all compulsory education levels (grades 1-9) and effectively eliminated national-, province-, and district-level academic exams below grade 9.

The 2006 Law has stimulated a trend of random assignment of teachers and students across the nation, which has been documented in the first nationally representative educational survey of junior high school students, China Education Panel Survey (CEPS): 83% of the randomly sampled schools across the nation reported that they randomly assigned teachers and students upon students' entry to junior high school. Under this national trend, a common assignment approach has been that, supervised by local education departments, schools create either random or stratified homerooms of students upon students' entry to school and the same amount of teacher groups (note that teachers are often assigned to multiple groups depending on their workload, for example, a math teacher is typically assigned to two homerooms because two classes per day, five days per week is the full time equivalent workload for a junior high school math teacher), then randomly assign teacher groups to homerooms. This random assignment is crucial to my identification strategy and will be discussed in further details in the Methods section.

Advised-Homeroom Setting

The prevalence and sustainability of this national trend are underlined by a structural support: China's advised-homeroom school setting. Unlike in the U.S. where each student has their own schedule and attends different classrooms every day, Chinese students are grouped into homerooms, put on a shared homeroom schedule, and assigned a group of subject teachers who rotate to the homeroom to teach – one of the teachers serves as the homeroom's advisor, or teacher-advisor, see Galassi et al. (1997) for more interchangeably used terms. Throughout all years in which they attend the same school, students typically remain grouped with the original homeroom cohorts. Their core content teachers (especially the homeroom advisor) are encouraged to follow the homerooms rising to higher grades.

Together, homeroom students, their teacher group, and their homeroom advisor form an ecosystem that is fundamental to Chinese school leaders' "homeroom accountability" leadership strategy. Specifically, Chinese junior high schools are heavily academic-oriented due the fact that upon graduation, students will take a high school admission exam administered by county-level education department then be placed in different tracks: those who pass certain cutoffs (approximately half nationwide) will be admitted to general high schools that are traditional pathways to academic-focused higher education, while others will be assigned to vocational or alternate schools that prepare them for the job market. To prepare students for this high-stakes exam, schools formally test students at least twice per semester on core content subjects and evaluate teachers based on their homerooms' performance. The average score of a homeroom is the most convenient and popular assessed value for a teacher and the evaluation results are often tied to human resources decisions.

Importantly, this advised-homeroom school setting, shared by China and many other countries such as France, Germany, India, Russia, Japan, South Korea, and other eastern countries, allows for the previous discussed random assignment of teachers and students to be feasible and manageable in terms of not only school administrative practice but also government policy regulation. For example, local education departments typically review their local public schools every year to check whether there are violations of the 2006 Law in school practice. Their strategies vary but many require schools to submit a copy of their original homeroom rosters for the purpose of documentation, conducting students/parents survey, and referencing (in times when parents complain about unlawful student tracking or kids being discriminated against during homeroom assignment). These policy regulations greatly reenforce the validity of random assignment and in turn help it become a national trend – an educational norm accepted by students, parents, and educators across the nation.

Homeroom Teacher-Advisors

Underlying the existence of teacher-advisors, an important policy context is that school counselor is not a professional position in China. In compensating for this policy void, every homeroom in the country has a teacher-advisor in position to implement a comprehensive advisory program that generally integrates four core components: moral education, student discipline, student development, and mentoring, according to the Ministry of Education's regulations of teacher-advisors in 2009 (henceforth referred to as 2009 Regulation; See Appendix B2 for more details). In practice, Chinese teacher-advisors take on various responsibilities that are probably performed by homeroom teachers, advisory teachers, and school counselors in the U.S.: 1) for the homeroom teacher side of duties, they take on logistic tasks involving their homeroom advisees, such as monitoring attendance, making announcements, distributing textbooks, and organizing students when they participate school or community engagement activities as a team; 2) as advisory teachers, they meet with homeroom students in regular advisory period once or more times a week, working on school-related topics they consider beneficial to students' academic, behavioral, and social-emotional growth; 3) they also take on school counselor's duties to provide individual guidance to advisees: they know well about each advisee, build academic and activity profile to monitor their progress, help them navigate school life, cooperate with parents, and prepare them to be academically and social-emotionally ready for next grade, college, and/or career. Acting on these responsibilities and utilizing the weekly advisory periods as critical channels of social-emotional education, teacher-advisors play a central role in students' school life and form much closer relationship with advised homeroom students than traditional teachers do.

Furthermore, although folk knowledge suggests that all subject teachers are expected to serve on this role when they are needed and their personal situation allows them to, the majority of teacher-advisors are teaching core content subjects that are required for all grade levels throughout junior high school so that it will be convenient for them to follow their advisees rising to higher grades and retain this unique advisor-advisee relationship.

Student Academic Performance

My findings will not be meaningful to policy and practice if the outcome variables are not valid measures of student performance. An issue at first sight is that, unlike in the US where student achievement tests are conducted at state level, in China, none of the national-, province-, county-, and district-level tests below grade 9 exists and each junior high school conducts their own tests to assess student performance. However, these test scores are in fact valid measures of student learning due to two major reasons: 1) students are educated on the same grade-level knowledge and skills regardless their school and location since all compulsory education schools follow a national curriculum and most use the same PEP (People's Education Press) textbooks; 2) being the key assessments in a school's homeroom accountability system, school-administered tests are designed to be fair evaluations of teaching and learning progress therefore often have various approaches in position to achieve high level validity, e.g., minimize test items not directly from current syllabus, avoid cheating or any types of manipulation of test score, include various types of items beyond multiple-choice questions to capture multiple dimensions of students' content knowledge and skills, just name a few.

Of course, this high degree of validity will be compromised across schools due to the large between-school variance: e.g., difficulty and quantity of test items are different (schools do not share test sheets) and scoring strategies vary (e.g., CEPS data indicates that schools were using a cap score of 100, 120, 130, or 150) from school to school. To address these issues, I will standardize student raw score to have mean zero and unit standard deviation within each school and use only within-school variation in student score to estimate teacher effects.

Beyond exam score, I also include student's subject-specific self-concept, a measure that is rarely examined as an outcome variable in teacher effect literature. Self-concept is generally defined as "individuals' general perceptions of themselves in given domains of functioning" (Möller et al., 2009) and, from a social-cognitive perspective, is a critical variable in explaining student performance behavior (Marsh, 1986). I include it as an academic outcome variable based on two major considerations. First, with research showing the substantial correlations between student achievement and corresponding self-concept (Marsh et al., 2001; Möller et al., 2020), self-concept can serve as a robustness check to score outcome. More importantly, self-concept has its own research value in capturing the motivational dimensions of student learning as it fuels into performance, subject interest, educational decisions, and longer-term academic outcomes.

Theoretical Framework of Teacher Advisement

One theoretical ground for teacher advisement is the extended attachment theory (Birch & Ladd, 1997; Pianta, 1999), which posits teachers as potential attachment figures to students at school (Rhodes et al., 2006). The central idea of original attachment theory is that children's emotional safety toward their mother allows them to explore environment and develop social and cognitive competencies (Bowlby, 1969). Extending this mechanism into a school setting, if a similar emotional bond is established between teachers and students, students will build confidence and motivation, become more engaged in learning activities, and actively develop academic skills (Birch & Ladd, 1997; Roorda et al., 2011; Pianta et al, 2012). Extended attachment theory is supported in the literature, for instance, Roorda et al (2011) conducted a meta-analysis of 92 articles from five regions and nations and found that the affective relationship between teachers and students was positively associated with students' school engagement and achievement.

Social-emotional learning literature also sheds light on one of the mechanisms through which advisors impact student learning. Explicitly or implicitly, all types of advisory programs have some components to foster a positive school climate and help students become social-emotionally competent (Elias et al., 1997). Contemporary research suggests that students' emotional attachment to school and engagement in classroom are critical components that influence student performance (Becker & Luthar, 2002; Hoffman, 2009).

The third, intuitive explanation of advisor effects is that, through direct advising, advisory teachers can help students stay motivated, engage in intellectual activities, and acquire and refine cognitive skills, which results in improved student learning outcomes. Although empirical evidence on direct advising is extremely limited, small-scale experimental studies do show that direct supports from teachers (e.g., setting academic goals, developing learning strategies, progress reviews, and positive feedback) significantly improve student academic outcomes such as GPA at school (Murray & Malmgren, 2005).

These three theoretical frameworks all shed light on the underlying mechanisms through which advisor-advisee relationship between Chinese teachers and students may impact student learning. First, since advisors build strong connections with advisees through advisory programming (Shi & Leuwerke, 2010), it is natural for advisees to be emotionally attached to their advisors. According to extended attachment theory, affective advisor-advisee relationship helps advisees to be more cognitively engaged in advisors' classroom and strongly motivated to acquire knowledge and content matter skills, which leads to improved performances. Second, from a social-emotional learning perspective, since a great proportion of advisory is focused on behavioral discipline and social-emotional competencies, advisees are used to physically behaving and emotionally engaged in front of their advisors, which not only benefits their own learning but also collectively creates a healthy environment for the whole homeroom. Lastly, advisors may use direct advising as an approach to help less motivated, low performing, and disengaged students, which also adds to positive academic performance.

Data and Measures

I draw analytic sample from China Education Panel Survey (CEPS), China's first nationally representative, longitudinal survey of middle-school students and take advantage of its two waves of data. Starting in school year 2013-14, CEPS team implemented a stratified, multi-stage sampling scheme to randomly select 112 middle schools from across the country. Administrators from each randomly selected school were surveyed. Within each school, the sampling scheme then selected two 7th grade and two 9th grade homerooms to survey. Within each homeroom, all students, parents, teacher-advisors, and content teachers in three core subjects (Chinese, English, and math) were surveyed. In school year 2014-15, most (N = 9,449, 91.93%) of the initial 7th grade cohort were successfully followed up in 8th grade, and these students will be the primary focus of my analysis. See Appendix A Data Description for more information about this data.

The two-wave CEPS data contains not only longitudinal information on a rich set of student-, family-, teacher-, and school-level variables but also whether the school randomly assigns teachers and students. Specifically, in the baseline survey, administrators were explicitly asked whether the school had randomly assigned teachers and students upon students' entry to middle school (begore 7th grade began) and 83 percent (N=93) schools identified as yes. This variable, coupled with the national random assignment trend stimulated by the 2006 Law, has been leveraged by researchers to overcome selection bias in estimating student outcomes and add causal evidence to a variety of educational research areas such as teacher-student identity match (Eble & Hu, 2020; Gong et al., 2018), peer effects (Xu et al., 2022), after-school tutoring (Sun et al., 2020), and urban-rural gap (Zhao et el., 2017). I will further show detailed evidence of this random assignment in each of my three articles.

Analytic Sample

The validity of random teacher-student assignments is central to my identification strategy. However, CEPS data was not collected from a randomized controlled trial where researchers had full control of the teacher-student assignment, instead, the assignments fell under the purview of local school administrators and the data was self-report in nature. Acknowledging this data limitation, I implement careful restriction criteria to obtain an analytic sample where students and teachers were mostly likely randomly assigned.

Beforehand, I theorize three major contaminants of random assignment: (A) some non-public schools or under-resourced public schools still sorted students and teachers under the radar; in schools truthfully implemented random assignment, after assignment, (B) some parents lobbied their children to be placed in the homerooms with their desired teachers, and (C) under the pressure of homeroom accountability, some teachers removed lower-achieving students from their homeroom (to other homerooms or another school) or schools used homeroom reassignment as some sort of policy intervention. The most recent study using CEPS data (Xu et al., 2022) delt with contaminants B and C by using only baseline data on the initial 7th graders in the 93 schools who reported random assignment, based on the rationale that 7th grade is the time when parents and teachers have the least knowledge about student academic ability therefore the least likely to sort students. I do not think this is strategy is sufficient because CEPS baseline data was collected after the mid-semester test, i.e., 2-3 months after initial assignment, which leaves enough time for student sorting if the school indeed allowed it to happen. More importantly, CEPS' valuable asset, the two-wave longitudinal data, allows for the inclusion of prior scores in same and other subjects as the most important covariates to mitigate measurement error (Lockwood & McCaffrey, 2014) and reduce estimation bias (Chetty et al., 2014a) – throwing it away is not a wise methodological decision.

I approach these three contaminants of random assignment in a different way and justify my steps of sample restriction in the following. First, I limit sample schools to 85 schools that were public schools (partially addressing contaminant A) and self-reported to have randomly assigned teachers to students before 7^{th} grade began. I then move on to address student sorting between baseline and Wave 2. Note that more than 80 percent homerooms had at least some change in their membership between two waves but most of these changes were driven by students moving in or out of school, indicated by 830 (8.07%) students unable to follow up and 471 (4.75%) newcomers in wave 2 data (Appendix A). For identification purposes, I am relatively unconcerned about this across-school sorting because in all the models I fit, I control for school fixed effects to absorb any time-invariant factors driving students to sort in or out of school. In contrast, I am concerned about within-school sorting (contaminants B and C), which will introduce considerable bias into the estimates of teacher effects. In dealing with this issue, I identify 22 schools that had at least one student change homeroom ID (but remain in the same school) between two waves and exclude them from my sample. I am left with 63 schools with two-wave data, which I use in my primary analyses throughout all three articles.

In Appendix C Table C1, I compare schools in my analytic sample (N=63) with the remaining schools (N=49) based upon observed descriptive statistics and show that these two groups of schools are indeed systematically different: my sample schools are more likely from coastal and urban area, serve a better educated population, and have smaller class sizes. This comparison suggests that excluding the 49 schools indeed helps address contaminant A. I believe sacrificing some degree of external validity in exchange for a much stronger internal validity is a sound decision and am more confident about the random assignment in my analytic sample.

It is important to note that, in the following Methods section, I will formally conduct a set of covariates balance checks to empirically evaluate the validity of random assignment based on baseline performance and observed characteristics.

Key Variables

Predictor Variable. The predictor variable is a dichotomous variable coded one for students whose teacher is also their homeroom advisor and zero otherwise.

Outcome Variables. In each of the three subjects (Chinese, English, and math), student academic outcome is measured by two variables, both of which contain unique information on student learning. First, I use student's subject score on school-administered mid-fall semester exam (obtained from their school records). The second outcome variable is subject-specific self-concept. I use student's (reversed) response to a 4-point Likert-scale survey item asking whether the subject is difficult as a proxy for self-concept. I reverse code the variable to represent four levels of self-concept: zero (very low), one (low), two (high), and three (very high). Overall, students report higher Chinese self-concept (70% reporting high or very high) than English and math self-concepts (50 and 52% reporting high or very high). Note that both outcome variables are standardized to be mean zero and unit variance within each school.

Covariates. I draw from baseline data three groups of covariates at student-, homeroom-, and teacher-level to improve estimation precision. The student-level covariates include student baseline standardized scores on three subjects (administered by the school) and cognitive test (administered by CEPS team), student gender, age, single child status, rural residency, migrant-worker family status, mother and father total years of education, and family wealth. The homeroom-level covariates include homeroom size and the homeroom means (leave-one-out mean, i.e., excluding self for each observation) of student characteristics. The teacher-level covariates include teacher gender and whether the teacher also serves as a teacher-advisor.

Within each of the analytic sample schools (N = 63), I match students with core content teachers to obtain separate samples for Chinese, English, and math subjects then examine the missingness. The predictor variable (advisor-advisee relationship indicator) is missing at 0.92%, 0%, and 0% for Chinese, English, and math sample, respectively. Across three subjects, on outcome variables (score and self-concept), the range of missing rate is 1.08%-1.19% and 0.48%-0.54%. On student and teacher covariates, the missing rate is all below 2% except for three variables: student age, teacher age, and teaching experience, the highest missing rate is 2.42%, 2.25%, and 4.78%, respectively. Because of the relatively large sample size and small missingness, I assume these missing is completely at random and drop all observations that have any missing value on predictor and outcome variables, then replace missing values on other variables with homeroom mean (for student covariates) or school mean (for teacher covariates). As shown in Table 1, Chinese, English, and math samples result in 5,080, 5,119, and 5,127 students, respectively. In the three samples, Chinese, English, and math, 28.11%, 27.58%, and 31.32% of the students were taught by their advisors, respectively.

Methods

Identification Strategy

China's unique advised-homeroom school setting creates a natural comparison condition between teachers who only teach a content subject and those who not only teach the same subject but also advise students, which provides me an opportunity to parse out the variations in student outcomes that are attributable to advisor-advisee relationship. My identification strategy relies on the following two assumptions.

Assumption 1: Balanced Teacher Groups

The first assumption is that the assignment of teachers to homerooms was random, which is nearly impossible to test due to the fact that only two homerooms were selected within each school. Acknowledging this limitation, I relax this assumption to that teacher-advisors and non-advisor teachers are not systematically different on variables that are associated with student academic outcomes, otherwise, one may argue that teacher-advisors are simply better teachers and the differences in student outcomes are not attributable to the advisor-advisee relationship. This assumption is likely warranted based on two main reasons. Firstly, the 2009 Regulation only requires advisors to have mentoring and communication skills (Appendix B2) – which are not pedagogical nor content specific, and folk knowledge among educators suggests that advisor appointment decision is finalized by teachers themselves and often hinges on their availability (e.g., health and family conditions) and willingness to take on such a committed role.

More importantly, a baseline characteristics balance check (Table 2) shows that, in my analytic sample, there is no systematic difference between 104 advisors and 251 non-advisor teachers in terms of pre-existing characteristics such as gender, age, teaching experience, education attainment, professional rank in the credential system and subject area. A robustness check based on full CEPS data also confirms this pattern (Appendix C Table C2). Although this analysis only demonstrates that these two teacher groups do not differ based on observables and cannot rule out the possibility that the two groups may differ in unobserved ways, it makes this difference less plausible.

Assumption 2: Balanced Student Groups

The second, and more critical assumption is that the assignment of students to homerooms was random so that students taught by advisors and non-advisor teachers were not systematically different before treatment. Indeed, the random assignment of students and teachers is not only enforced by the 2006 Law and reported by the surveyed schools (discussed in Background and Data sections), but also confirmed in the data. I conduct a series of student covariates balance check by regressing the predictor variable against baseline student covariates while controlling for school fixed effects and clustering standard errors at school level and report the results in Table 3. I show evidence that, both individually and jointly, four baseline scores (in three subject and CEPS cognitive test) and other socioeconomic variables did not predict whether the student was assigned for their subject-matter class to a teacher-advisor or not.

The only exception is that same-subject baseline score is significantly correlated with being taught by advisor, i.e., students having higher baseline math score were significantly more likely taught by advisors who taught math, same as to Chinese and English. Taking together the fact that baseline scores measure student performance on mid-fall semester exams and by that time, students had been in advisors' classroom for half a semester (2-3 months), one may argue that this indicates the possibility of students with either better learning ability or stronger motivation of learning a subject, say, math, sorting to math teacher-advisor during that first half semester. While I cannot rule out this possibility, noticing the small magnitude of the coefficients and the insignificant coefficients of other variables that are also correlated with academic ability and motivation, I argue that these coefficients are not evidence of sorting. Rather, they are highly more likely capturing the potential effects of being taught by advisor for a short period of time (although unfortunately, without prior ability controls and careful modeling, these coefficients cannot be directly interpreted as teacher-advisor effects).

In all, I argue that the two critical assumptions to my identification strategy are largely met: teacher-advisors were not systematically different from their colleagues and students being assigned to teacher-advisors were not systematically different from their school peers in 2013-14. As a result, the only explanation to any observed differences in student outcomes in 2014-15 is the effects of exogenously determined advisor-advisee relationship.

Estimation Strategy

My strategy for estimating the impacts of teacher characteristics has three major components. First, I control for school fixed effects, which means the coefficients are identified based off the variation in teacher characteristics within each school and effectively eliminate any biases associated with student across-school sorting. Taking together the fact that my samples were drawn under a no within-school sorting criterion, I am most confident about the validity of the random assignment of students and teachers in my analytic samples. Second, I control for a rich set of student-, homeroom-, and teacher-level covariates from baseline data to improve estimation precision and, importantly, I include a cubic function of baseline score in all three subjects and CEPS cognitive test in all the models I fit to capture varying functional forms of student prior learning ability as well as school and family inputs (Blazar & Kraft, 2017; Chetty et al., 2014a; Kane et al., 2008; Kraft, 2019). Third, I cluster standard error at the school level to account for the within-school correlations among residuals.

Model Specification

Based on my identification and estimation strategies, I recover the causal impact of advisor-advisee relationship on student outcomes by estimating a value-added ordinary least squares (OLS) regression model below:

$$OUT_{ijst} = \beta_0 + \beta_1(ADVISOR_{jt}) + \beta_2 X_{it-1} + \beta_3 H_{jt-1} + \beta_4 T_{jt-1} + \theta_s + \varepsilon_{ijst}$$

where i, j, s, t denote student, teacher, school, year; OUT_{ijst} is student i's academic performance or self-concept in year t; $ADVISOR_{jt}$ is teacher j's teacher-advisor status; X_{it-1}, T_{jt-1}, and H_{jt-1} are baseline student-, teacher-, and homeroom-level covariates; θ_s is the school fixed effects of school s; and ε_{ijst} is the student-level idiosyncratic error term. The coefficient of interest is β_1, which is the estimated causal effect of advisor-advisee relationship on student outcome.

Results

The estimated effects of having advisor-advisee relationship on student subject scores are presented in Table 4. To check for the robustness of the results, I fit three models to each of the three subject-specific samples: all three models control for student-level covariates including the cubic functions of baseline academic and CEPS cognitive scores, as well as student characteristics; the second model adds baseline teacher characteristics; and the third one further adds baseline homeroom characteristics. Within each subject, my findings were consistent across all three models. Specifically, on average, compared to their school peers, students having an advisor-advisee relationship with teachers scored significantly higher in English by 0.101-0.155 SD and math by 0.121-0.143 SD. No effect was detected in Chinese.

The estimated advisor effects on students' self-concept in Chinese, English, and math are presented in Table 5. The same three models are used to separately estimate each subject and the results are again robust across all models: compared to their school peers, students having an advisor-advisee relationship with teachers significantly reported higher self-concept in Chinese (0.219-0.203 SD) and English (0.169-0.187 SD), but not math.

I conduct a set of placebo tests by using the same three model specifications to estimate advisor-advisee relationship effects on student score in CEPS cognitive test. According to CEPS, the cognitive test was designed to measure student age-appropriate cognitive skills and should not be used to evaluate school learning outcomes. I present results in Table 6 and demonstrate that student cognitive score is indeed unresponsive to the advisor-advisee relationship. This suggests that my three model specifications do not return significant associations between predictor and outcome variables when there is no relationship.

Discussions

Teacher is the most important measured aspect of schools in determining student achievement (Hanushek, 2011) and an extensive literature body has used teacher's value-added to student achievement as a proxy of teacher quality to recover substantial variations among teachers (e.g., Aaronson et al., 2007; Chetty et al., 2014a, 2014b; Kane et al., 2008; Rockoff, 2004). Unfortunately, no such rigorous study has examined the role of advisory teachers, a role that is oftentimes not a standing alone job position but is assigned to teachers in K-12 schools all over the world.

Grounded in extended attachment theory, social-emotional learning literature, and intuitively theorized outcomes of direct advising, advisory programs are sound in theory; in practice, implementing an advisory program requires significant inputs from various parties including local government, school, teachers, and students; both lead to the critical interests of policymakers and educators in learning about whether teacher-advisors contribute meaningfully to student school outcomes. Responding to this call, my study adds novel evidence of the causal effects of teacher advisement by taking advantage of the advised-homeroom school setting and the random homeroom assignment in China, where teachers who only teach serve as ideal counterfactuals to their colleagues who simultaneously teach and advise a homeroom of students. I carefully draw analytic sample from a nationally representative dataset and provide considerable evidence on the validity of the random assignment of teachers and students in the data.

I found that advisor-advisee relationship significantly improves student English and math score by 0.155 and 0.143 SD, improves student Chinese and English self-concept by 0.203 and 0.187 SD, which are notable effect sizes compared to the magnitude of teacher effects in value-added literature: for example, Hanushek and Rivkin (2010) reviewed ten rigorous value-added studies leveraging within-school estimation and found that one SD increase in teacher quality improves student reading and math score by 0.13 and 0.17 SD. In other words, my findings can be largely interpreted as: on average, adding an additional advisor role to a content teacher has the approximately same effects as improving teacher quality by one SD.

Furthermore, it is critical to note the auxiliary nature of advisor-advisee relationship effects: recall these teacher-advisors are dual role teachers and the effects captured in my study are generated through channels that parallel instruction: relationship building, social-emotional skills development, and direct advising and support. These findings have substantial policy implications for school leaders and educators who seek to redefine teacher's role in students' school life and find additional ways to impact students' learning throughout traditional classrooms.

I also add to the literature on the multidimensional nature of student academic outcomes and show evidence that score measures do not capture all the impacts educators can make on students' learning (Kraft, 2019). My findings suggest that student score and self-concept respond differently to advisor effects and no effect on the former does not necessarily indicate no effect on the latter. For example, consistent with existing findings in the literature (e.g., Chetty et al., 2014), Chinese (corresponding to English Language Arts in U.S.) scores are not responsive to advisor influences as math score does, but advisors can help students increase their self-concept in Chinese, which potentially has long-term or ripple effects on student learning.

In terms of limitations of this study, I first emphasize that a certain degree of external validity has been sacrificed during sample restriction process, where I only include 63 schools in my analytic samples to (to my best ability) ensure that sample students were randomly assigned teachers upon their entry to junior high school in 2013-14 and did not sort in or out of their teachers' class by 2014-15. This leads to systematic differences between schools included and excluded in my analyses: my sample schools appear to be more likely located in economically developed and urban areas and have significantly smaller class sizes. As a result, I stress that my findings are not generalizable to schools in disadvantaged areas that are more likely to fail to restrictively implement random assignments. Further study should look closer to the assignment of teachers and students in schools from remote areas and if possible, conduct researcher-designed experiments to test the robustness of my findings. I also note that due to the self-report nature of CEPS data, there are potentially large measurement errors embedded in the key variables. For instance, the subject confidence variable only captures students' response to a single survey item therefore is potentially not accurately capturing the latent construct of students' self-concept. Great caution is suggested in interpreting these results and generalizing these findings to common practice.

References

Aaronson, Barrow, L., & Sander, W. (2007). Teachers and Student Achievement in the Chicago Public High Schools. Journal of Labor Economics, 25(1), 95–135. https://doi.org/10.1086/508733

Becker, B. E., & Luthar, S. S. (2002). Social-emotional factors affecting achievement outcomes among disadvantaged students: Closing the achievement gap. Educational Psychologist, 37(4), 197–214.

Birch, S. H., & Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. Journal of school psychology, 35(1), 61-79.

Blazar, & Kraft, M. A. (2017). Teacher and Teaching Effects on Students' Attitudes and Behaviors. Educational Evaluation and Policy Analysis, 39(1), 146–170. https://doi.org/10.3102/0162373716670260

Chetty, Friedman, J. N., & Rockoff, J. E. (2014a). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. The American Economic Review, 104(9), 2593–2632. https://doi.org/10.1257/aer.104.9.2593

Chetty, Friedman, J. N., & Rockoff, J. E. (2014b). Measuring the impacts of teachers, [part] II: teacher value-added and student outcomes in adulthood. The American Economic Review, 104(9), 2633–2679. https://doi.org/10.1257/aer.104.9.2633

Davis, H. A. (2003). Conceptualizing the role and influence of student-teacher relationships on children's social and cognitive development. Educational Psychologist, 38(4), 207–234.

Eble, & Hu, F. (2020). Child beliefs, societal beliefs, and teacher-student identity match. Economics of Education Review, 77, 101994. https://doi.org/10.1016/j.econedurev.2020.101994

Elias, M. J., Zins, J. E., Weissberg, R. P., Frey, K. S., Greenberg, M. T, Haynes, N. M., et al. (1997). Promoting social and emotional learning: Guidelines for educators. Alexandria, VA: Association for Supervision and Curriculum Development

Galassi, J., Gulledge, S., & Cox, N. (1997). Middle School Advisories: Retrospect and Prospect. Review of Educational Research, 67(3), 301-338.

Gong, Lu, Y., & Song, H. (2018). The effect of teacher gender on students' academic and noncognitive outcomes. Journal of Labor Economics, 36(3), 743–778. https://doi.org/10.1086/696203

Hanushek. (2011). The economic value of higher teacher quality. Economics of Education Review, 30(3), 466–479. https://doi.org/10.1016/j.econedurev.2010.12.006

Hanushek, & Rivkin, S. G. (2010). Generalizations about Using Value-Added Measures of Teacher Quality. The American Economic Review, 100(2), 267–271. https://doi.org/10.1257/aer.100.2.267

Hoffman, D. (2009). Reflecting on Social Emotional Learning: A Critical Perspective on Trends in the United States. Review of Educational Research, 79(2), 533-556.

Jenkins. (1992). Advisement programs: a new look at an old practice. National Association of Secondary School Principals.

Kane, Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness?: Evidence from New York City. Economics of Education Review, 27(6), 615–631. https://doi.org/10.1016/j.econedurev.2007.05.005

Kraft, M. A. (2019). Teacher effects on complex cognitive skills and social-emotional competencies. Journal of Human Resources, 54(1), 1-36.

Marsh. (1986). Verbal and math self concepts : an internal external frame of reference model. American Educational Research Journal, 23(1), 129–149. https://doi.org/10.3102/00028312023001129

Marsh, Kong, C.-K., & Hau, K.-T. (2001). Extension of the Internal/External Frame of Reference Model of Self-Concept Formation: Importance of Native and Nonnative Languages for Chinese Students. Journal of Educational Psychology, 93(3), 543–553. https://doi.org/10.1037/0022-0663.93.3.543

McClure, L., Yonezawa, S., & Jones, M. (2010). Can school structures improve teacher-student relationships? The relationship between advisory programs, personalization and students' academic achievement. Education Policy Analysis Archives, 18, 17.

Möller, Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A Meta-Analytic Path Analysis of the Internal/External Frame of Reference Model of Academic Achievement and Academic Self-Concept. Review of Educational Research, 79(3), 1129–1167. https://doi.org/10.3102/0034654309337522

Murray, C., & Malmgren, K. (2005). Implementing a teacher–student relationship program in a high-poverty urban school: Effects on social, emotional, and academic adjustment and lessons learned. Journal of School Psychology, 43(2), 137-152.

Pianta, R. C. (1999). Enhancing relationships between children and teachers. American Psychological Association.

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), Handbook of research on student engagement (pp. 365–386). Springer Science.

Rockoff. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. The American Economic Review, 94(2), 247–252. https://doi.org/10.1257/0002828041302244

Roorda, D., Koomen, H., Spilt, J., & Oort, F. (2011). The Influence of Affective Teacher–Student Relationships on Students' School Engagement and Achievement: A Meta-Analytic Approach. Review of Educational Research, 81(4), 493-529.

Shi, Q., & Leuwerke, W. (2010). Examination of Chinese homeroom teachers' performance of professional school counselors' activities. Asia Pacific Education Review, 11(4), 515-523.

Xu, Zhang, Q., & Zhou, X. (2022). The Impact of Low-Ability Peers on Cognitive and Noncognitive Outcomes: Random Assignment Evidence on the Effects and Operating Channels. The Journal of Human Resources, 57(2), 555–596. https://doi.org/10.3368/jhr.57.2.0718-9637R2

Zhao, G., Ye, J., Li, Z., & Xue, S. (2017). How and why do Chinese urban students outperform their rural counterparts? China Economic Review, 45, 103-123.Educational Evaluation and Policy Analysis, 37(2), 224-242.

Tables and Figures

Table 1. Analytic sample summary statistics

Key Variables	Chinese Sample	English Sample	Math Sample
	N = 5,080	N = 5,119	N = 5,127
Predictor Variable			
Advisor-advisee relationship	28.11%	27.58%	31.32%
Outcome Variables			
Score	0.00 (0.99)	0.00 (0.99)	0.00 (0.99)
Self-concept	0.00 (0.99)	0.00 (0.99)	0.00 (0.99)

Notes: Cells report mean and standard deviation for continuous variables and percentage of each category for categorical variables.

Table 2. Baseline covariates balance check between advisors and non-advisor teachers

Teacher Characteristics	Teacher-Advisor N = 104	Non-Advisor Teacher N = 251	*p*-value
Female teacher	74.04%	76.10%	0.7
Teacher age	39 (7)	39 (7)	>0.9
Teaching experience (years)	16 (8)	16 (8)	>0.9
Highest education			0.9
Associate degree	8.65%	7.97%	
Bachelor degree	88.46%	87.65%	
Graduate degree	2.88%	4.38%	
Professional rank			0.3
Novice teacher	0.96%	4.38%	
Intermediate teacher	31.73%	25.90%	
Advanced teacher	43.27%	47.01%	
Senior teacher	24.04%	22.71%	
Subject area			0.7
Chinese	30.77%	34.26%	
English	32.69%	33.07%	
Math	36.54%	32.67%	

Notes: Cells report mean and standard deviation for continuous variables and percentage of each category for categorical variables. The *p*-statistic was obtained from a) Pearson's Chi-squared test for gender, education, professional rank, and subject area, and b) Wilcoxon rank sum test for teacher age and experience.

Table 3. Baseline covariates balance check between students taught by advisors and non-advisors

	Advisor-Advisee Relationship		
	Chinese Sample	English Sample	Math Sample
Baseline Chinese	0.027*	-0.013	-0.014
	(0.011)	(0.012)	(0.013)
Baseline English	-0.018	0.045**	-0.026
	(0.015)	(0.015)	(0.017)
Baseline Math	-0.024	-0.023	0.039*
	(0.013)	(0.014)	(0.016)
Baseline cognitive	-0.004	0.007	0.018
	(0.008)	(0.016)	(0.016)
Female	-0.006	-0.006	0.011
	(0.008)	(0.007)	(0.009)
Age	0.007	-0.015	0.022
	(0.011)	(0.012)	(0.016)
Rural residency	-0.011	-0.005	-0.001
	(0.013)	(0.011)	(0.014)
Only child	-0.010	-0.002	-0.009
	(0.009)	(0.012)	(0.012)
Migrant family	-0.007	-0.009	0.015
	(0.012)	(0.011)	(0.013)
Mother education (years)	0.004	-0.004	0.001
	(0.002)	(0.002)	(0.002)
Father education (years)	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)
Family income	-0.009	0.010	0.014
	(0.009)	(0.010)	(0.011)
School FE	Yes	Yes	Yes
Clustered SE	Yes	Yes	Yes
F-Statistics	1.416 (*df* = 12; 62)	1.421 (*df* = 12; 62)	1.491 (*df* = 12; 62)
Observations	5,080	5,119	5,127
R^2	0.541	0.598	0.464

Notes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. Cells report coefficients and associated standard errors in parentheses. Each column reports results of a separate OLS regression where the predictor variable of interest, advisor-advisee relationship indicator, is regressed against baseline student score measures and characteristics. All models control for school fixed effects and cluster standard errors at school level.

Table 4. Estimates of the causal effects of advisor-advisee relationship on subject scores

	Subject Score								
	Chinese Sample			English Sample			Math Sample		
Advisor-advisee relationship	0.046 (0.048)	0.052 (0.051)	0.045 (0.058)	0.101** (0.032)	0.125** (0.037)	0.155*** (0.036)	0.121** (0.041)	0.130** (0.038)	0.143*** (0.039)
Student Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Homeroom Covariates	No	No	Yes	No	No	Yes	No	No	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered SE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,080	5,080	5,080	5,119	5,119	5,119	5,127	5,127	5,127
R^2	0.596	0.596	0.599	0.702	0.703	0.704	0.607	0.609	0.610

Notes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. Cells report coefficients and associated standard errors in parentheses. Each column reports results of a separate OLS regression where the outcome variable of interest, subject score, is regressed against the predictor variable, advisor-advisee relationship indicator. Each sample is estimated three times: the first model controls for student baseline covariates, the second model adds teacher covariates, and the third further adds homeroom covariates. All models control for school fixed effects and cluster standard errors at school level.

Table 5. Estimates of the causal effects of advisor-advisee relationship on subject self-concept

	Subject Self-Concept								
	Chinese Sample			English Sample			Math Sample		
Advisor-advisee relationship	0.219** (0.067)	0.216** (0.064)	0.203* (0.076)	0.169** (0.054)	0.138* (0.058)	0.187*** (0.046)	0.062 (0.050)	0.062 (0.046)	0.016 (0.041)
Student Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Homeroom Covariates	No	No	Yes	No	No	Yes	No	No	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered SE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,080	5,080	5,080	5,119	5,119	5,119	5,127	5,127	5,127
R^2	0.081	0.082	0.085	0.259	0.259	0.263	0.237	0.241	0.245

Notes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. Cells report coefficients and associated standard errors in parentheses. Each column reports results of a separate OLS regression where the outcome variable of interest, subject self-concept, is regressed against the predictor variable, advisor-advisee relationship indicator. Each sample is estimated three times: the first model controls for student baseline covariates, the second model adds teacher covariates, and the third further adds homeroom covariates. All models control for school fixed effects and cluster standard errors at school level.

Table 6. Robustness check: the effects of advisor-advisee relationship on cognitive score

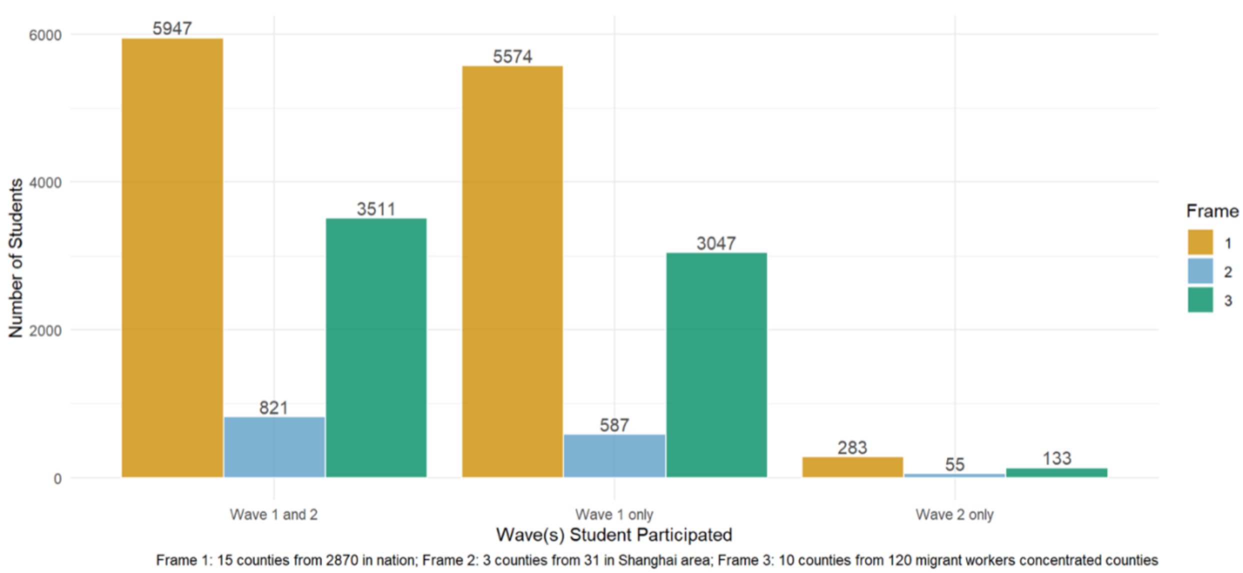
	CEPS Cognitive Test Score								
	Chinese Sample			English Sample			Math Sample		
Advisor-advisee relationship	-0.057 (0.047)	-0.052 (0.044)	-0.061 (0.044)	-0.052 (0.039)	-0.031 (0.037)	-0.058 (0.035)	0.053 (0.040)	0.048 (0.035)	0.066 (0.036)
Student Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Homeroom Covariates	No	No	Yes	No	No	Yes	No	No	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered SE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,080	5,080	5,080	5,119	5,119	5,119	5,127	5,127	5,127
R^2	0.579	0.581	0.584	0.578	0.579	0.582	0.577	0.580	0.583

Notes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. Cells report coefficients and associated standard errors in parentheses. Each column reports results of a separate OLS regression where the alternative outcome variable, CEPS cognitive test score, is regressed against the predictor variable, advisor-advisee relationship indicator. Each sample is estimated three times: the first model controls for student baseline covariates, the second model adds teacher covariates, and the third further adds homeroom covariates. All models control for school fixed effects and cluster standard errors at school level.

Appendix A. Data Description

The China Education Panel Survey (CEPS) started in school year 2013-2014 and employed a stratified, four-step random sampling procedure to draw a random sample of middle schools, teachers, and students from the nation. First, they randomly selected 28 school districts/counties with probability proportional to size (PPS) from three stratified sample frames, specifically, 15 from 2,870 districts/counties (frame 1) in the nation, 3 from 31 districts/counties in Shanghai area (frame 2), and 10 from 120 migrant labor concentrated districts/counties (frame 3). Second, within each district/county, they randomly selected four schools from all schools serving 7th and/or 9th grades with PPS. Third, within each school, they randomly selected two homerooms from 7th grade and another two from 9th grade. Fourth, within each homeroom, they included all students and administered separate surveys to students, parents, homeroom advisory teachers, classroom teachers for three core subjects (math, Chinese, and English), and school administrators.

Using this procedure, the CEPS team surveyed 10,279 7th grade and 9,568 9th grade students in school year 2013-14 and successfully followed up with 9,449 of the original 7th graders (follow-up rate 91.9%) along with 471 new students in school year 2014-15. Detailed numbers of students by wave and frame are visualized in the following bar chat. Note that the 9,449 students with two-wave data (the first three bars) will be the focus of my dissertation, see more discussion in the text.



Appendix B. National Education Policies in China

B1. Compulsory Education Law (2006)

The Compulsory Education Law[1] was amended and adopted at the 22nd Session of the 10th National People's Congress Standing Committee and issued as No. 52 Order of the President on June 29, 2006. Relevant to my research, the law highlighted that all school-age children and adolescents shall have equal right and the obligation to receive a 9-year compulsory education (Article 4) at the schools near their residency (Article 12). They shall go to school without taking any examination (Article 12). The county level governments and education departments shall promote the balanced development among schools and narrow down school quality gaps (Article 22). No education government may create key schools and non-key schools and no school may create key classes and non-key classes (Article 22). No school may expel students based on school management rules (Article 27). Legal liabilities are attached to the violations of these articles.

B2. Regulations of Advisory Teachers by Ministry of Education (2009)

The Ministry of Education issued the Regulations of Advisory Teachers[2] on August 12, 2009. Relevant to my research, the regulation specified advisory teacher's core responsibilities as moral education, student discipline, student development, and mentoring. The regulation emphasized that every homeroom in the country shall have an advisory teacher and the position is half-time equivalent. A homeroom's advisory teacher should teach the homeroom and should be ethical, psychologically healthy, caring, dedicated, and having strong communication ability and managerial skills.

1. See http://www.lawinfochina.com/Display.aspx?lib=law&Cgid=77520 for a translation of the Law.
2. No translation of this document was found on the internet. The Chinese version is here http://www.moe.gov.cn/srcsite/A06/s3325/200908/t20090812_81878.html.

Appendix C. Additional Tables and Figures

Table C1. Baseline covariates balance check between schools included and excluded from the analytic sample

School Characteristics	Included N = 63	Excluded N = 49	*p*-value
School district sampling frame			0.069
Sample frame 1	46.03%	63.27%	
Sample frame 2	15.87%	4.08%	
Sample frame 3	38.10%	32.65%	
School district location			0.03
East China	68.25%	51.02%	
Middle China	9.52%	28.57%	
West China	22.22%	20.41%	
School district administrative level			0.018
Municipality	28.57%	12.24%	
Urban area of provincial capital cities	20.63%	14.29%	
Urban area of prefecture-level cities	20.63%	14.29%	
County or county-level city	30.16%	59.18%	
District population average education (years)	9.88 (1.44)	9.27 (1.34)	0.024
School location			0.9
Center of the city/town	41.27%	32.65%	
Outskirts of the city/town	11.11%	10.20%	
Rural-urban fringe zone of the city/town	14.29%	16.33%	
Towns outside of the city/town	15.87%	20.41%	
Rural areas	17.46%	20.41%	
Proportion of rural residency students			0.003
Lower than 25%	33.33%	8.16%	
25% to 60%	30.16%	22.45%	
60% to 80%	15.87%	30.61%	
Higher than 80%	20.63%	38.78%	
Proportion of the local students			0.009
Lower than 50%	4.76%	14.29%	
50% to 70%	26.98%	12.24%	
70% to 90%	34.92%	18.37%	
higher than 90%	33.33%	55.10%	
Number of substitute teachers	1.38 (3.77)	4.27 (17.86)	0.5
Unknown	3	4	
Average homeroom size	48 (9)	52 (8)	0.011

Notes: Cells report mean and standard deviation for continuous variables and percentage of each category for dichotomous and categorical variables. The *p*-statistic was obtained from a) Wilcoxon rank sum test for district population average education, number of substitute teachers, and average homeroom size, and b) Pearson's Chi-squared test for all other characteristics.

Table C2. Baseline covariates balance check between advisors and non-advisor teachers (full CEPS data)

Teacher Characteristics	Teacher-advisor	Non-advisor teacher	*p*-value
	N = 235	N = 475	
Female teacher	66.38%	73.68%	0.043
Teacher age	38 (7)	39 (8)	0.11
Teaching experience (years)	15 (8)	17 (9)	0.12
Highest education			0.6
Associate degree	12.77%	12.63%	
Bachelor degree	84.68%	83.37%	
Graduate degree	2.55%	4.00%	
Professional rank			0.8
Novice teacher	8.09%	8.00%	
Intermediate teacher	30.21%	29.89%	
Advanced teacher	45.11%	42.32%	
Senior teacher	16.60%	19.79%	
Subject area			0.6
Chinese	36.17%	32.84%	
English	30.21%	33.47%	
Math	33.62%	33.68%	

Notes: Cells report mean and standard deviation for continuous variables and percentage of each category for dichotomous and categorical variables. The *p*-statistic was obtained from a) Pearson's Chi-squared test for gender, education, and professional rank, and b) Wilcoxon rank sum test for teacher age and experience.