ROCKWOOL FOUNDATION
BERLIN
INSTITUTE FOR THE ECONOMY AND THE FUTURE OF WORK

CReAM | Centre for Research & Analysis of Migration

DISCUSSION PAPER SERIES 20/24

Test Scores, Noncognitive Outcomes, and the Stereotyping of Non-Local Students

Qinyue Luo, Huihua Xie

OCTOBER 2024

ROCKWOOL Foundation Berlin
www.rfberlin.com

Centre for Research & Analysis of Migration
www.cream-migration.org

Test Scores, Noncognitive Outcomes, and the Stereotyping of Non-Local Students*

Qinyue Luo † Huihua Xie ‡

October 14, 2024

Abstract

This study investigates the impact of teachers' stereotyping of non-local students in terms of both academic performance and noncognitive outcomes using a random assignment of Chinese middle school students to teachers. We find that biased beliefs against non-local students, particularly among Chinese teachers, negatively affect non-local students by decreasing academic performance and increasing behavioral problems, with no significant effects on local students. Mechanism analysis suggests that these negative outcomes result from reduced teacher engagement with non-local parents, weaker classroom integration, and diminished self-confidence among non-local students. The negative effects are especially pronounced for non-local boys while non-local girls show resilience by increasing their efforts. These results highlight the critical role of teachers' stereotyping in shaping disparities in human capital development between local and non-local students.

Keywords: Teachers' Stereotypes, Non-Local Students, Test Scores, Mental Stress, Behaviour Problems Index.

JEL Classification: I24, J15, J24.

*We thank Thomas Cornelissen, Christian Dustmann, Zibin Huang, Gong Jie, Hyejin Ku, Haoming Liu, Mattsson Martin, Jessica Pan, Huiren Tan, Shamdasani Yogita and seminar participants of SIG Applied Micro at NUS, RFBerlin first Spring workshop, RFBerlin brown bag seminar, CReAM brown bag seminar at University College London, and 2024 Spring Labor Economics and Public Policy Workshop at Zhejiang University for their helpful comments. Any errors are our own.

†ROCKWOOL Foudation Berlin, E-mail: ql@rfberlin.com.

‡Zhejiang University. E-mail: huihuaxie@zju.edu.cn

1 Introduction

Individuals with migrant backgrounds often underperform in education and the labor market globally (OECD, 2023). This issue is critical as many countries have seen significant internal and external migration, including families with children (Carlana, La Ferrara, & Pinotti, 2022). For example, in 2018, 23.8% of students in OECD countries had immigrant backgrounds (Cerna, Brussino, & Mezzanotte, 2021). Similarly, China has experienced massive rural-to-urban migration since its economic reforms in the late 1970s, with the urban population rising from 20.9% in 1982 to 63.9% in 2020 (NBSC, 2021a). By 2020, 71.09 million children in China had migrated to cities with their parents, making up 23.9% of the total child population (NBSC, UNICEF, & UNFPA, 2023). These shifts have significantly altered school demographics and introduced challenges such as ensuring skill progression and reducing inequality between local and non-local students. The gap in academic and noncognitive skills between these groups tends to widen in school settings (Fryer Jr & Levitt, 2004; Cunha & Heckman, 2007; Carlana, La Ferrara, & Pinotti, 2022), contributing to disparities in labor market outcomes.[1]

One key factor that may contribute to exacerbating this gap relates to potential teacher biases. Stereotypes are widely held but fixed and oversimplified representations of group characteristics, leading to judgment errors and discrimination against certain groups (Bordalo et al., 2016). Such discrimination can discourage individuals, lower self-confidence, reduce effort, and hinder potential achievement (Coffman, 2014; Reuben, Sapienza, & Zingales, 2014; Glover, Pallais, & Pariente, 2017; Bordalo et al., 2019; Carlana, 2019).[2] Students are particularly susceptible to teachers' stereotypes for two reasons. First, teacher effects are substantial — students' achievements and long-term outcomes are significantly affected by teachers' characteristics, behaviours, expectations, beliefs, and attitudes (Chetty, Friedman, & Rockoff, 2014a, 2014b). Second, childhood is a key period for children's development of a self-concept. Thus, teachers' erroneous expectations in this period may create students' self-fulfilling prophecies, according to which prior stereotypes are self-confirming in equilibrium (Spencer, Steele, & Quinn, 1999; Ertl, Luttenberger, & Paechter, 2017; Carlana, 2019; Papageorge, Gershenson, & Kang, 2020). There is limited credible evidence on how teachers' stereotyping of non-local students affects their academic and personal development, mainly due to challenges in measuring these stereotypes and testing their causal effects.

In this paper, we examine the impact of teachers' negative stereotyping of non-local students in middle school on students' academic achievement, measured by standardized test scores in Chinese,

[1] According to the International Labour Organization (ILO), in 2020, migrants in high-income countries earned nearly 13% less than local workers. In China, the wage gap between migrant workers and urban residents was even more pronounced, with migrant workers earning approximately 30% (NBSC, 2021b).

[2] Stereotype threat theory suggests that individuals may conform to negative stereotypes about their group, increasing cognitive load and reducing academic performance (Steele & Aronson, 1995).

math, and English, and noncognitive outcomes, including mental stress and the Behavior Problems Index (BPI). "Non-local" refers to children without a local household registration ("*hukou*" in Chinese), which issued by the Chinese government to Chinese citizens to indicate their specific area of origin; it also restricts access to local benefits such as education, healthcare and the right to purchase housing. [3] Our analysis focuses on Chinese middle schools, offering a natural experiment characterized by three features: (1) a substantial portion of migrant students is present in the school population, [4] (2) random assignment of students to classrooms (and teachers) within schools, [5] and (3) detailed data on students, teachers, and parents. [6]

We use data from the 2014 China Education Panel Survey (CEPS), a nationally representative survey encompassing approximately 20,000 7th- and 9th-grade students and their teachers across 112 schools, with about 20% of students being non-local. To assess teachers' stereotypes regarding non-local students, we use items that ask teachers to compare classes with only local students to those where more than one-third are non-local, measuring implicit attitudes based on differences in evaluations of teaching effects, class discipline, and student relationships.[7] Section 3.1 details the construction of stereotype measures. A key strength of our dataset is its rich information on teacher-student-parent interactions, allowing us to uncover mechanisms by which teachers' negative stereotyping of non-local students affects student outcomes.

We estimate the effects of teachers' non-local stereotyping by taking advantage of the random assignment of teachers and students to classes. Our analysis adheres to the same sample restrictions as previous studies utilizing CEPS data (Hu, 2018; Gong, Lu, & Song, 2018; Eble & Hu, 2019; Huang, 2020; Gong, Lu, & Song, 2021) and focuses on schools meeting three criteria for random assignment: (1) random class formation confirmed by the school principal, (2) no changes to class

[3]In China's compulsory education period including primary and junior middle schools, teachers are typically well-informed about the basic background of each student. This information generally includes the student's "*hukou*", home address, and parents' educational and occupational background. Schools gather this information through several methods. Some collect it during enrollment and distribute it directly to teachers. Some issue questionnaires to students or their parents at the beginning of academic year. Others obtain it during the first parent-teacher meeting through direct communication with parents. Unlike university instructors, these teachers handle both teaching and administrative duties, giving them a detailed understanding of each student.

[4]According to NBSC, UNICEF, and UNFPA (2017), the number of migrant children reached 34.26 million in 2015, representing 13% of the total child population in China. In other words, about 1 in every 8 children in China traveled with their migrant parents. By 2020, this number had increased to 71.09 million, accounting for 24% of China's total child population. This indicates that nearly 1 in every 4 children in China were migrant children (NBSC, UNICEF, & UNFPA, 2023).

[5]We acknowledge that many Chinese schools still use criteria such as test scores for class assignments, and some schools may violate claimed random assignment policies. However, our analysis is restricted to schools that strictly adhere to three criteria of random assignment, as utilized in previous studies relying on CEPS data (Hu, 2018; Gong, Lu, & Song, 2018; Eble & Hu, 2019; Huang, 2020; Gong, Lu, & Song, 2021) We also conduct several balancing tests to confirm the validity of the randomization, with details provided in Section 4.2.

[6]Regarding teacher information, we have data on their attitudes towards local and non-local students, teaching practices, and education and job characteristics. However, we lack detailed demographic information, knowing only their gender, age, and marital status, but not their local or non-local status. This limits our ability to explore potential local vs. non-local matches between teachers and students.

[7]Related studies have used similar self-reported beliefs on Likert scales to capture stereotypes; see Alan, Ertac, and Mumcu (2018), Avitzour et al. (2020), and Rakshit and Sahoo (2023).

composition for the whole middle school period, and (3) no assignment based on test scores ensured by all head teachers in the grade. These criteria ensure the random matching of students and teachers, supporting the exogeneity assumption behind the teacher stereotyping measure and reducing concerns about student self-selection into classes or specific teachers. Additionally, balancing tests confirm that students with varying baseline characteristics are not systematically assigned to teachers with different levels of bias. To address the potential issue of reverse causality—where teachers' stereotyping of non-local students may be shaped by student performance—we demonstrate that the proportion of local students in a class is balanced against teachers' biases. Furthermore, we find no correlation between teachers' stereotyping and the baseline performance of local versus non-local students, confirming that biases are not driven by prior performance.

We obtain four sets of results. First, our analysis reveals that stereotyping of core-subject teachers (i.e., Chinese, math and English teachers) negatively affects non-local students' academic performance, with no significant impact on local students. Specifically, a one standard deviation increase in teachers' negative stereotyping reduces non-local students' test score index by approximately 0.04 standard deviations. We find no significant effects on mental stress or Behavior Problems Index (BPI) scores for either group.

Second, we examine the heterogeneity of these effects across teaching subjects, the dual role of Chinese teachers, student gender, and the length of time non-local students have resided in the area. Our results suggest that the impact of teachers' stereotyping is most pronounced among Chinese teachers compared to math and English teachers, where an increase of one standard deviation in Chinese teachers' stereotyping reduces non-local students' test scores by 0.1, 0.05, and 0.1 standard deviations in Chinese, math, and English, respectively.[8] In contrast, stereotyping by math and English teachers shows no significant effects on academic performance. Additionally, Chinese teachers' stereotyping negatively affects the BPI scores of non-local students. Notably, when Chinese teachers also serve as head teachers, their stereotypes not only impact non-local students' test scores but also intensify mental stress for both local and non-local students. This dual-role Chinese teachers' stereotyping decreases non-local students' academic performance but slightly improves local students' scores, thereby widening the achievement gap between the two groups.

We also find that non-local boys are more negatively affected by Chinese teachers' stereotypes than non-local girls. Increased stereotyping leads to greater reductions in test scores and an increase in behavioral issues such as tardiness and absenteeism for boys, whereas non-local girls show no significant academic or behavioral response to teacher biases. This finding aligns with prior studies

[8]Our findings align with previous research on the impact of teacher biases on student performance. For instance, Alan, Ertac, and Mumcu (2018) and Carlana (2019) reported that teachers' gender stereotypes negatively affect girls in terms of their academic performance but have negligible effects on boys.

suggesting that boys are more susceptible to environmental disruptions (Kling, Ludwig, & Katz, 2005; Bertrand & Pan, 2013; Deming et al., 2014; Dustmann, Ku, & Kwak, 2018; Autor et al., 2019; Stuart, 2022; García, Heckman, & Ronda, 2023). Additionally, we observe that the negative impacts of teachers' stereotypes are more pronounced for non-local students who have lived in the local area for more than 8 years. For these students, a one standard deviation increase in teacher bias results in a 0.14 standard deviation decrease in test scores, while those residing in the area for fewer than eight years do not exhibit significant effects. This suggests that long-term non-local students, potentially from more disadvantaged backgrounds or facing greater competition for local resources, are particularly vulnerable to teacher bias. [9]

Finally, we investigate the mechanisms underlying these disproportionate effects of teachers' stereotyping on local and non-local students. Our analysis indicates that teachers' stereotypes shape their behavior toward students. Specifically, Chinese teachers with stronger biases against non-local students are less likely to engage with their parents compared to those of local students. Furthermore, biased teachers' stereotypes affect classroom dynamics and peer interactions, as non-local students perceive their classmates as less friendly and tend to form friendships mainly with other non-local students. This social isolation raises concerns among non-local parents regarding their children's peer relationships. Teachers' stereotypes also negatively impact non-local students' self-confidence and academic effort. In classrooms where Chinese teachers exhibit higher levels of bias, the self-confidence gap between local and non-local students increases by 0.07 standard deviations. Additionally, non-local students in these biased environments are discouraged from dedicating effort to their studies, spending more time on non-academic activities like entertainment.

This paper contributes to several streams of research. It adds to the emerging literature on the effects of teachers' beliefs and attitudes on student achievement and development. Prior studies have shown that teachers' stereotypical gender attitudes significantly influence the gender gap in STEM fields (Alan, Ertac, & Mumcu, 2018; Lavy & Sand, 2018; Carlana, 2019; Terrier, 2020; Rakshit & Sahoo, 2023). Regarding teachers' stereotyping of immigrants or ethnic minorities, correlations have been found between teachers' prejudiced attitudes and the achievement gap among ethnic groups (Van den Bergh et al., 2010). Researchers also investigate the role of student-teacher racial match in teachers' biased beliefs towards other ethnic groups (Gershenson et al., 2022). However, while many studies have examined the consequences of teachers' gender biases, less is known about the causal effects of explicit measures of teachers' stereotyping of non-local students or ethnic minorities on student achievement and performance. In this context, Alan et al. (2023), closely related to us,

[9] This trend corresponds with China's institutional landscape, in which context prolonged residence without obtaining local "*hukou*" status may be the result of disadvantaged backgrounds or migration to larger cities where numerous migrants compete for limited local "*hukou*" allocations. Given their heightened vulnerability due to unstable residential status and limited access to local resources, these students are particularly susceptible to the detrimental impact of teacher bias.

demonstrate that teachers with prejudicial attitudes towards specific ethnic groups contribute to socially and spatially segregated classrooms.

Our study makes several contributions. First, it reveals that teachers' biased views of non-local students negatively impact both academic performance and noncognitive outcomes. This aligns with Jackson (2018), who asserted that teacher effects on non-test outcomes are strong predictors of long-term student outcomes, even when correlations between teacher effects on test scores and non-test behaviors are weak. Thus, our contribution highlights the importance of teachers' stereotyping of non-local students on both academic and noncognitive outcomes. Second, while much of the existing studies focuses on refugees and immigrants in developed countries, we analyze stereotypes and discrimination within the context of internal migration in China, the world's largest developing country with significant migratory movements. Finally, although recent studies have explored potential mechanisms through which teacher stereotypes affect student outcomes, empirical evidence is limited due to data constraints. Previous investigations have focused primarily on students' beliefs, offering little exploration of the roles played by student effort, teacher behavior, parental responses, and the broader class environment. Notable exceptions include Alan, Ertac, and Mumcu (2018), who examined students' gender role beliefs, self-confidence, and mindsets, revealing that teachers' gender biases significantly influence students' gender role beliefs. Similarly, Carlana (2019) demonstrated that teacher gender stereotypes negatively impact girls' math performance, mediated by lower self-confidence. Alan et al. (2023) highlighted the importance of the class environment, showing that refugee students are often spatially segregated by biased teachers. Using rich data, our paper addresses this empirical gap by systematically investigating the roles of teachers, students, parents, and classmates in shaping the impact of teacher stereotypes on student outcomes within a dynamic environment. In doing so, we provide a comprehensive picture of how teachers' stereotypes influence students' academic and noncognitive development.

A second line of research examines the role of teachers in shaping student achievement and long-term outcomes (e.g., Chetty, Friedman, and Rockoff (2014a, 2014b)). Studies have shown that student-teacher matches based on gender, ethnicity, religion, and hometown ties significantly influence teachers' perceptions and, consequently, student outcomes (Dee, 2005; Fairlie, Hoffmann, & Oreopoulos, 2014; Feld, Salamanca, & Hamermesh, 2016; Lusher, Campbell, & Carrell, 2018; Lavy, Sand, & Shayo, 2018; Fisman et al., 2018). Our paper extends this exploration by focusing on teachers' beliefs and attitudes, rather than just traditional characteristics, such as gender or ethnicity. We contribute to the literature by directly measuring teachers' stereotyping of non-local students and establishing that these biases significantly affect students' cognitive and noncognitive outcomes, while also shaping the broader educational environment involving students, peers, and parents.

More broadly, our research contributes to the literature on the development of disadvantaged children. China's rapid economic growth and large-scale internal migration have created a unique socioeconomic landscape, which includes children left behind by migrating parents and those who migrate with their families to new areas. While the literature on left-behind children in developing countries has extensively explored the negative effects of parental absence on cognitive development (Zhang et al., 2014), health (Mu & De Brauw, 2015; Meng & Yamauchi, 2017), and economic preferences (Cadsby, Song, & Yang, 2020), less attention has been given to non-local students, another disadvantaged group (Yang, Wei, & Qin, 2016; Huang, 2020). Our paper addresses this gap by focusing on how teacher bias hampers non-local students' assimilation and development. By shining a light on the specific challenges faced by non-local students, our research offers a deeper understanding of the obstacles these students encounter in developing countries and societies undergoing rapid transformation.

The remainder of the paper is structured as follows. Section 2 provides background information on the household registration system and the middle school education system in China. Section 3 describes the datasets and variables. Section 4 explains the estimation strategy, the random matching of teachers and students. Section 5 presents the main results, focusing on the impact of teachers' stereotyping of non-local students on their outcomes. Section 6 explores potential channels through which teachers' stereotypes adversely affect the academic performance and noncognitive outcomes of non-local students. Finally, Section 7 concludes the paper.

2 Background and Setting

2.1 The "*hukou*" System and Non-Local Students in China

Since the beginning of the economic reform of China in the late 1970s, the country has experienced unprecedented economic growth as well as a substantial shift in the population from rural to urban areas. Data provided by the National Bureau of Statistics of China (2021) indicate that the fraction of the urban population increased from 20.9% in 1982 to 63.9% in 2020. As China aims to promote urbanization and industrialization even further, the trend of extensive internal migration is likely to persist. Nevertheless, the intricate dynamics associated with China's unique household registration system, which is known as the "*hukou*" system, often present obstacles for migrants. This system links an individual's identity with their place of birth, frequently resulting in resistance and discrimination during the migration process.[10] The "*hukou*" system has significant implications

[10]The "*hukou*" system, which was established in the 1950s to restrict mobility, assigns households to 'rural' or 'urban' categories and anchors them to a geographic location based on their origin. The governance of the "*hukou*"

with regard to an individual's access to social welfare and public services, and it thus plays a crucial role in determining individuals' eligibility for and the quality of benefits in areas such as education and health insurance. Consequently, migrant workers do not receive the same welfare entitlements as local workers.

Moreover, children of rural workers who migrate alongside their parents encounter unique challenges. Since they lack a local, nonagricultural "*hukou*", their access to public social infrastructure is restricted. For instance, urban students have far better educational opportunities than their migrant counterparts. First, non-local families must meet several conditions imposed by local authorities to enrol their children in local schools. These requirements often include the need to provide proof of permanent residence, such as uninterrupted residency within the school district for more than a year, and evidence of the parents' stable employment, which might require the parents to have worked and consistently paid social security in the local area for the same duration. Second, due to the central government's practice of allocating subsidies to public schools based on the enrolment of children with local "*hukou*" migrant children must often pay higher fees to attend these schools.[11]

Third, non-local students face significant obstacles with regard to standardized tests, such as the senior high school entrance exam or the national college entrance exam, within their current municipality. Often, these students are required to take these exams at their "*hukou*" locality.[12] As a result, the academic achievements of non-local students are not factored into the metrics used to assess the quality of schools and teachers. Finally, apart from these systemic issues, non-local students often face continuous disadvantages due to their lack of familiarity with the local dialect and cultural subtleties. These challenges can significantly impede their ability to communicate effectively with their classmates and teachers, thereby creating obstacles to their full integration and engagement in the school community.

system was stringent until the late 2010s, a situation which led to the emergence of significant challenges for individuals seeking to change their registered location. This situation was especially difficult for individuals seeking to relocate from economically disadvantaged regions to more prosperous urban centres.

[11] Since 2016, the government has implemented a policy that allows educational funds allocated to non-local students to follow them to their new schools after relocation. For comprehensive details regarding this policy, please visit `https://www.gov.cn/zhengce/content/2015-11/28/content_10357.htm`. However, according to our data, the educational funds allocated to non-local students are still retained in their place of origin, thus restricting their immediate access to or use of these resources. Our data also reveal a notable trend, such that, on average, non-local students contribute 36% more in fees to their local schools than do their local peers.

[12] In our dataset, only half of the non-local students' parents were confident in their children's ability to gain admission to local high schools. Additionally, 60% of these parents reported extra requirements for non-local students seeking entry into local high schools. Notably, only two-thirds of these parents had a concrete plan regarding their child's education continuation in local high school settings.

2.2 The Education System and Class Assignment

In 1986, the Compulsory Education Law in China established a nine-year mandatory education system, which includes six years of primary school and three years of junior middle school. The students included in our sample were in the junior middle school phase and were generally between the ages of 12 and 14. Upon entering middle school, each student is allocated to a "general class", after which the student remains in this cohort for the duration of their middle school experience.[13] Within a "general class", students attend academic courses and engage in extracurricular activities on a collective basis. On any given school day, students typically remain in their designated classroom, where teachers from various disciplines deliver instruction in line with the established curriculum on a rotating basis. A head teacher, who is often an instructor in one of the core subjects, i.e., Chinese, math, or English, oversees the "general class" with the goals of managing classroom behaviour, coordinating school events, monitoring individual students' progress, and fostering relationships among students. Additionally, the head teacher routinely conducts meetings with parents to discuss students' academic performance and conduct alongside other core-subject teachers.

Middle schools employ various strategies for class assignment. Some schools base class placement on students' residential areas, while others require entrance exams and allocate students to classes based on their test results. In its efforts to maintain equity and fairness in education, the Chinese government has prohibited ability- and background-based class selection in compulsory education since the late 2000s.[14] Consequently, an increasing number of schools employ random student assignment when forming classes. Two methods for randomizing class assignments are prevalent: the random scheme and the average assignment scheme.[15] The random scheme involves using a random number generator or asking parents to draw lots, thereby taking into account the class size and demographic diversity when balancing the class composition. The average assignment scheme stratifies students based on their entrance exam scores and then randomly selects an equal number of students from each tier to create a class. Once students are randomly assigned to classes, teachers are also randomly assigned, thereby ensuring that the collective learning and activity participation of the "general class" are maintained throughout the school day, as students remain in one classroom while various subject teachers travel come and go to deliver lessons.

[13] Fewer than 0.1 percent of the students in our sample changed classes within a school.

[14] See the Compulsory Education Law of the People's Republic of China revised in 2006 at `www.gov.cn/flfg/2006-06/30/content_323302.htm`.

[15] Hu (2018) documented several other randomization schemes. For details, see Hu (2018).

3 Data and Variables

Our study utilizes data drawn from the 2014 China Education Panel Survey (CEPS), an extensive and nationally representative survey that covers approximately 20,000 students from 112 middle schools in 28 counties throughout mainland China. The CEPS collects comprehensive data through diverse questionnaires targeted at students, parents, teachers, and school principals, thereby offering valuable insights into students' educational accomplishments, family backgrounds, school settings, and community dynamics. We narrowed our sample to schools that employ the random assignment of students to teachers according to the responses provided by school principals and teachers regarding class assignment methods. This process yielded a final sample of 8,955 students in 208 classrooms across 67 schools.

3.1 Teacher Characteristics and the Stereotyping of Non-Local Students

We assessed teachers' stereotyping of non-local students using a set of items that are designed to reflect teachers' subjective beliefs regarding and attitudes towards the presence of non-local students in their classes. Teachers were asked to evaluate two types of classes on a five-point Likert scale: classes consisted solely of local students and classes with more than one-third of non-local students. The teachers rated aspects such as teaching effectiveness, classroom discipline, and student relationships on a scale ranging from 'very bad' (1) to 'very good' (5). To measure stereotyping, we calculated the differences in teachers' evaluations between the two class types in terms of these dimensions. The standardized average of these differences represents teachers' stereotyping of non-local students, in which context higher values indicate higher levels of biased views against non-local students. The specific survey questions included in the teacher questionnaire that were used to construct the stereotyping index are listed in Table 1. [16]

[Insert Table 1 here]

Figure 1 illustrates the complete distribution of teachers' stereotyping across the three core subjects. These measures are normalized based on a mean of zero and a standard deviation of one, and higher values indicate more exclusionary attitudes towards non-local students. A notable degree of variation was observed in teachers' stereotyping towards non-local students, as demonstrated by the broad range of bias measures pertaining to both individual subjects and the different disciplines.

[16] This paper, along with several others (e.g., Alan, Ertac, and Mumcu (2018), Avitzour et al. (2020), Dhar, Jain, and Jayachandran (2022), and Rakshit and Sahoo (2023), etc.), employs survey questions to assess bias. An alternative approach, commonly used in recent studies (e.g., Carlana (2019) and Alan et al. (2023)) and in the social psychology literature, is the Implicit Association Test (IAT). It is important to note, however, that both explicit measures and implicit tests are widely utilized and are not interchangeable (Oswald et al. (2013) and Rakshit and Sahoo (2023)).

Specifically, the stereotyping measures ranged from -2.29 to 2.76 in Chinese, from -3.03 to 2.45 in math, and from -2.68 to 3.32 in English. This considerable degree of variation in stereotypes among teachers teaching the same subject enabled us to conduct an analysis to contrast students who were randomly assigned to teachers with different degrees of bias against non-local students in the context of a particular subject. Additionally, Figure 2 illustrates the potential influence of the current classroom composition on variations in teacher stereotypes. The figure reveals notable disparities in the stereotypes held by teachers who lead classes that are composed exclusively of local students and those whose classes feature a substantial proportion of non-local students. These findings not only corroborate the patterns identified in Figure 1 but also suggest that our measures of stereotypes reflect teachers' general beliefs and attitudes rather than their observations in a specific classroom during the current academic year.

[Insert Figure 1 & Figure 2 here]

Furthermore, we utilized teacher questionnaires to construct profiles that included participants' demographic information (gender, age, marital status) and professional attributes (education level, years of experience, tenure, pedagogical qualifications, professional rank, and prior teaching experience). Table 2 provides summary statistics pertaining to teacher demographics, thus revealing that the majority of teachers are female, especially in the subjects of Chinese and English. The average age of teachers is 37 years; in addition, most teachers have more than 15 years of teaching experience, hold tenured positions and have prior teaching experience in other schools. Overall, no obvious variation is teacher characteristics is observed across different subjects. On average, teachers in all three subjects rated classes featuring only local students higher than classes with more than one-third of non-local students in terms of teaching effectiveness, discipline, and student relations. While teachers in all three subjects expressed similar views regarding classes featuring only local students, Chinese teachers assigned lower ratings to classes featuring non-local students than math and English teachers across all three dimensions.

[Insert Table 2 here]

Table 3 presents estimates obtained by regressing each of the teachers' characteristics on teachers' stereotyping of non-local students while controlling for fixed effects at the block level (including classes in the same school and the same grade). The results highlight the potential predictive relationship between teachers' characteristics and stereotypes. As shown in the table, a variety of factors, including teachers' demographics, education attainment, pedagogical qualifications, teaching experience, professional rank, and status as head teachers, exhibit no meaningful statistical links to

teachers' stereotypical attitudes. However, we observe a notable exception to this claim: a discernible negative association is evident between the tenure status of teachers and their biased stereotypes regarding non-local students. One possible explanation for this relationship is that teachers who have not yet secured tenure may be more attuned to elements that they believe could impact their students' scholastic performance, including the presence of non-local students in their classrooms, since their own job evaluations and tenure assessments are often contingent upon their students' academic success. Nevertheless, this evidence highlights the importance of accounting for teacher characteristics when exploring the effects of teacher stereotypes on student outcomes.

[Insert Table 3 here]

3.2 Student Outcomes and Family Characteristics

We utilize student examination data provided by school administration offices in the CEPS to measure students' academic achievement. We focus on standardized test scores in three core subjects, Chinese, math, and English, due to their importance in middle school curricula and their significance with regard to high school admission exams. Within a grade, all teachers of the same subject employ a similar teaching plan and regularly discuss their teaching progress and adjust their teaching approach. Simultaneously, all students in the same grade take the same exams as stipulated by the school. Therefore, test scores in Chinese, math, and English consistently measure students' academic performance within the same grade in the same school.[17] We standardized these test scores across subjects, grades, and schools by conducting principal component analysis (PCA) to generate a composite test score index, which was also standardized to have a mean of zero and a standard deviation of one.

Additionally, we construct two indices to measure students' noncognitive outcomes. The mental stress index is derived from the self-reported frequency with which students experienced five emotions over the past week, which is scored on a scale ranging from 1 (never) to 5 (always). These emotions include feeling "blue", "depressed", "unhappy", "not enjoying life" and "sad". The BPI is based on students' responses regarding punctuality and class attendance. Specifically, students indicated their agreement with the statements "I am always late for class" and "I always skip classes", on a scale ranging from 1 (strongly agree) to 4 (strongly disagree). Both indices were constructed using PCA and normalized to a mean of zero and a standard deviation of one.

[17]Examinations entail a stringent and uniform grading process, as part of which the grader remains unaware of individual student names, class affiliations, and ID numbers. Grading responsibilities within the same grade at the same school are distributed among teachers, thus ensuring that a specific question is consistently assessed by the same teacher and maintaining standardized grading criteria within a particular grade level. Simultaneously, each teacher is responsible for grading test papers for all students within a particular grade, including not only students enrolled in her own classes but also those enrolled in other classes at the same grade level.

To account for the influence of student and family characteristics on cognitive and noncognitive outcomes, we include a comprehensive set of predetermined variables in our study. These variables include demographic information, family background, and measures of students' cognitive and noncognitive outcomes prior to middle school. We use these variables in balancing tests and include them as control variables in our analysis. Table 4 presents summary statistics pertaining to both local and non-local students; our main outcomes are addressed in Panel A, and predetermined characteristics are addressed in Panel B. Columns 7 and 8 report the raw mean and standard error of the difference between local and non-local students. The results indicate that while non-local students exhibit comparable levels of academic achievement and BPI scores to their local peers, they are slightly more likely to experience mental stress. On the other hand, local students generally exhibit superior baseline performance. Local students are less likely to have repeated a grade in primary school and to have self-reported higher scores with regard to expressing their opinions clearly, responding quickly, and learning new material quickly in grade 6. Additionally, local students tend to come from families that exhibit a higher socioeconomic status. They are less likely to possess a rural "*hukou*", and their parents tend to have obtained higher levels of education.

[Insert Table 4 here]

4 Empirical Strategy

4.1 Estimating the Equations

To investigate the effects of teachers' stereotyping on students' outcomes, we use the following regression model:

$$\begin{aligned} Y_{ijcb} = \beta_0 + \beta_1 Non-local_{ijcb} + \beta_2 Stereotype_{jcb} + \beta_3 Non-local_{ijcb} \times Stereotype_{jcb} \\ + \beta_4 X_{1,ijcb} + \lambda_b + \delta_j + \beta_5 X_{2,jcb} + \beta_6 X_{3,jcb} + \beta_7 Non-local_{ijcb} \times X_{1,ijcb} \\ + \beta_8 Non-local_{ijcb} \times X_{2,jcb} + \beta_9 Non-local_{ijcb} \times X_{3,jcb} + \epsilon_{ijcb}. \quad (1) \end{aligned}$$

Let Y_{ijcb} represent the academic achievement and noncognitive outcomes of student i taught by a teacher of subject j in class c of block b (i.e., the same grade in the same school). $Non-local_{ijcb}$ is a dummy variable that is equal to one if student i is a non-local student. $Stereotypes_{jcb}$ is a standardized score that reflects the extent of stereotypical beliefs held by the teacher of subject j in class c; larger values represent more negative beliefs regarding non-local students. To account for the

possibility that teachers' stereotypes may have differential effects on local and non-local students, we include the interaction term between the measure of teachers' stereotyping and the non-local student indicator ($Non-local_{ijcb} \times Stereotype_{jcb}$). Additional predetermined characteristics of the students, such as their demographic information, family background, and measures of baseline cognitive and noncognitive outcomes (as detailed in Table 4), are included in the vector $X_{1,ijcb}$.

Given that randomization takes place within the same grade and same school, we pool matched student-teacher-level data from the seventh and ninth grades into a larger sample to improve the precision of the estimation, and we include block-level fixed effects λ_b to capture the aggregate influence of institutional characteristics specific to each block b. To control for disparities among different subjects, we also include fixed effects associated with each subject δ_j. Our identification strategy therefore takes advantage of the variation in stereotypical bias across teachers and the random assignment of students to teachers who engage in stereotyping to different degrees within the same subject in a specific block.

Additionally, we include teachers' demographic factors (gender, age, marital status) $X_{2,jcb}$ and control variables related to teachers' education and professional attributes (educational attainment, years of experience, tenure, graduation from a pedagogical college or with a teaching major, professional title, prior teaching experience in other schools, and status as a head teacher) $X_{3,jcb}$. We also include all student- and teacher-level controls interacted with the binary indicator $Non-local_{ijcb}$ to adjust for the potentially varying effects of student and teacher characteristics on the outcomes obtained by local versus non-local students. We employ robust standard errors that are clustered at the block level, thus taking into account potential outcome correlations among students within the same block.

The main coefficients of interest in this study are β_2 and β_3. β_2 measures the impact of teacher stereotyping on the academic scores and noncognitive outcomes of local students, while β_3 indicates the differential effect of teacher stereotypical bias on local students versus non-local students. The combined effect of β_2 and β_3 reveals the influence of teacher stereotyping on the outcomes of non-local students. The assumption underlying this approach, which focuses on guaranteeing the unbiased estimates of β_2 and β_3, is that, conditional on the included controls, the indicator of student identity $Non-local_{ijcb}$ and the measure of teacher stereotyping $Stereotypes_{jcb}$ are independent of the error term. By taking advantage of the random allocation of students, implementing block-level fixed effects, and controlling for a comprehensive set of teacher- and student-level control variables interacted with students' non-local status, we posit that the estimates of β_2 and β_3 can be interpreted as causal effects.

One advantage of equation (1) is that it allows us not only to identify the impact of teacher

stereotyping on the gap between local and non-local students (β_3) but also to explore different scenarios that can lead to changes in this gap by examining the effects on both local students (β_2) and non-local students ($\beta_2 + \beta_3$) separately. For example, a negative β_3 implies that teacher stereotyping impedes the developmental outcomes of non-local students as compared to local students. This situation may be due to the weaker performance of non-local students ($\beta_2 + \beta_3 < 0$), the stronger performance of local students ($\beta_2 > 0$), or a combination of these effects.

However, despite the fact that equation (1) includes a rich set of teacher-level variables, the omission of teacher fixed effects may still lead to omitted variable bias if unobserved characteristics at the teacher or class level are correlated with teacher stereotyping. Therefore, to address this concern, we further include teacher fixed effects in equation (1) to account for unobserved teacher-, class-, and school-level factors by conducting a robustness check. Notably, by controlling for teacher fixed effects, we are unable to identify the effects of observable teacher characteristics, including $Stereotypes_{jcb}$, on local and non-local students separately. Instead, we identify only the effects on the outcome gap between local and non-local students and cannot distinguish among different scenarios leading to changes in this gap.

4.2 Random Class Assignment and Validity Tests

This study focuses on institutions that employ a random assignment method for class formation and aims to investigate the impacts of teachers' stereotypes regarding non-local students on students' academic and noncognitive outcomes. In line with previous research that has relied on CEPS data (Hu, 2018; Gong, Lu, & Song, 2018; Eble & Hu, 2019; Huang, 2020; Gong, Lu, & Song, 2021), we established specific sample criteria. In particular, schools were required to adhere to three key principles during class assignments. First, the principal must employ random assignment as the strategy for creating classes; second, classes that are formed at the onset of the seventh grade must remain intact through the eighth and ninth grades. Finally, lead teachers at each grade level must ensure that student placement is not influenced by test scores. Based on these criteria, we refined our analysis to an estimation sample consisting of 8,955 students across 208 classes and 67 schools, thus accounting for approximately 59.8% of the schools contained within the original CEPS sample. The random pairing of students with teachers, alongside the stability of class assignments over a three-year period, reinforces the exogeneity of teacher stereotyping in our study, thus mitigating potential concerns related to students' self-selection of classes or teachers. Although parents may attempt to choose classes based on teacher attributes, the implicit nature of teachers' stereotypes towards non-local students renders such classes difficult to detect and predict.

To validate the random assignment in our sample, we conducted a balance test. We regressed

students' predetermined characteristics (i.e., attributes established prior to class assignment) on teachers' stereotypes of non-local students while controlling for block (i.e., classes in the same school and the same grade) fixed effects. The baseline variables pertaining to students included demographic details (gender, age, minority status, local residency, and only child status), baseline academic performance (as determined by reference to cognitive outcomes in kindergarten and primary schools, including kindergarten attendance, grade repetition or acceleration in primary school, and academic ranking in the sixth grade), initial noncognitive assessments (i.e., the self-rated ability to articulate one's opinions clearly, respond promptly, and learn new concepts swiftly in the sixth grade), and parental factors (including the educational attainment of both mothers and fathers).

Our findings, which are detailed in Table 5, indicate that the correlation between students' baseline characteristics and the stereotyping tendencies exhibited by three core-subject teachers (which are presented in columns 1 to 3) is largely statistically and economically negligible. This finding suggests the absence of systematic bias in the assignment of students with varying baseline profiles to teachers who may exhibit stronger or weaker stereotypical views regarding non-local students. The only exceptions are the minor correlations observed between only child status and the stereotyping behaviours of math teachers, although the point estimates in this context are very small in magnitude. Furthermore, all the F-statistics are small and not statistically significant, thus indicating that we cannot reject the null hypothesis that the coefficients of students' baseline characteristics are jointly nonsignificant. Overall, the results of this balancing test imply that the randomization process ensures a comparable distribution of student characteristics among students who have been exposed to varying degrees of teacher stereotyping.

[Insert Table 5 here]

Unlike previous studies that have measured baseline cognitive outcomes directly (e.g., Alan, Ertac, and Mumcu (2018) included fluid IQ and Carlana (2019) included standardized test scores in grade 5 in the balance test), we are unable to employ this approach due to a lack of direct measures in our dataset. This situation may give rise to the concern that students' academic performance in kindergarten and primary school might be insufficient to measure their baseline outcomes before they enter middle school; thus, our balance test might not provide full supporting evidence regarding randomized teacher-student matching. To address this concern, we include all students' baseline characteristics as well as the interaction terms between those characteristics and the student non-local indicator in the main analysis to ensure that the balance among students who are exposed to teachers who exhibit different levels of stereotyping is maintained.

In addition to concerns regarding nonrandomness, another potential threat to our estimation

is related to the possibility that teachers' stereotyping of non-local students may be influenced by the differing academic performance exhibited by local and non-local students. It is possible that teachers may harbour more traditional views and believe that classes that are composed solely of local students perform better in terms of teaching effectiveness, class discipline, and student relationships than do classes featuring non-local students, especially if those teachers have previously encountered an exceptional class consisting entirely of local students. Here, we describe four balance tests that suggest that reverse causality is unlikely in our context.

First, we investigate whether the proportion of local students in a class is associated with teachers' stereotyping. Should teachers' stereotyping of non-local students persist, we would not expect to observe a significant link between the share of local students and teachers' stereotyping. Panel A in Table 6 presents the regression outcomes of teachers' stereotyping in relation to the proportion of local students: the results indicate no statistically significant correlations across all teachers. [18]

Second, we assess the potential influence of academic performance by comparing the average academic ranking of local students in grade 6 with that of non-local students within the same class. Regressing teachers' stereotyping of non-local students against this class-specific academic performance gap reveals no meaningful or statistically significant correlation, as indicated in Panel B in Table 6. These findings suggest that teachers' stereotyping is not influenced by the comparative academic performance of local and non-local students.

Third, we must consider that non-local students often face higher costs when enrolling in local schools than their local peers. We investigate whether this disparity in educational expenses within a class is related to teachers' biases against non-local students. Our regression analysis, which calculates the average fees paid by both local and non-local students and then determines the within-class difference, reveals no statistically significant correlation between teachers' stereotyping and the class-level variation in terms of total educational expenses (presented in Panel C in Table 6). This result implies that the randomization process in our study effectively ensures that teachers' biases are not shaped by existing differences in educational expenditures within a class.

Finally, the CEPS asked teachers to evaluate the academic performance of their class in comparison to other classes in the same grade at the beginning of their teaching term. We use the subject teachers' responses to this question as a proxy for the class-average performance at the start of the academic year in Chinese, math, and English. Panel D in Table 6 explores the potential correlation between teachers' stereotyping of non-local students and class-average baseline performance;

[18] In further analysis of the effect of teachers' stereotyping on student outcomes, we also explored whether there are heterogeneous effects based on the share of non-local students in the class. The results, presented in Table A1, show no significant differences in the impact of teachers' stereotypes across classes with varying proportions of non-local students.

however, no statistically significant correlation is observed. In line with the balance test results presented in Table 5, we are confident that teachers' stereotyping of non-local students is not driven by any potential differences between the performance of local students and that of non-local students. In Section 5.4, we conduct additional robustness checks to further mitigate any concerns regarding reverse causality.

[Insert Table 6 here]

5 Main Results

5.1 Teachers' Stereotyping of Non-Local Students and Student Outcomes

Based on the baseline estimation equation (1), Table 7 presents the influence of teacher stereotypes regarding non-local students on those students' academic achievement (Columns 1-3) and noncognitive attributes (mental stress in Columns 4-6 and BMI scores in Columns 7-9). For each outcome, we first demonstrate the estimated effects based solely on a binary indicator for non-local status, teachers' stereotypes of non-local students, the interaction between these terms, and block and subject fixed effects (as shown in Columns 1, 4, and 7). We then add predetermined student characteristics, teacher demographics, and corresponding interaction terms with the non-local student identifier (Columns 2, 5, and 8). Ultimately, teacher educational and occupational traits as well as their interactions with the non-local student binary variable are integrated into the preferred model (Columns 3, 6, and 9). The outcomes remain consistent, regardless of whether student or teacher controls are taken into account; thus, the analysis of effect sizes is conducted using the results with the full set of control variables.

Our findings suggest that while the coefficients pertaining to the measure of teachers' stereotyping are neither economically nor statistically significant, those pertaining to the interaction terms between teacher stereotypes and the non-local student indicator are negatively significant with respect to the test score index. This finding suggests that stereotypes held by teachers of core subjects have a more detrimental impact on the academic performance of non-local students than on local students, primarily manifesting as a negative effect on non-local students. The incorporation of additional controls at both the student and teacher levels does not substantially alter these estimated patterns. However, no statistically significant effects were observed on either local or non-local students' mental stress or BPI scores. In terms of economic magnitudes, estimates based on the full specifications (Column 3) indicate that an increase of one standard deviation in teacher stereotypes against non-local students decreases the test score index for non-local students by approximately

0.04 standard deviations with respect to local students.

[Insert Table 7 here]

5.2 Heterogeneous Effects by Subject, Teacher Role, Student Gender and Duration of Residency

We further explore the heterogeneity of these effects across teaching subjects, the dual role of Chinese teachers, student gender, and the residency duration of non-local students in the area. The preceding analysis focused on a matched sample of students and teachers in three core subjects, namely, Chinese, math, and English. This analysis now allows us to present the average effect of teachers across these subjects.

Teaching subject. In light of the inherent variations in content and teaching methods across these disciplines, we investigate whether the effects of teachers' stereotyping of non-local students vary by subject, as shown in Table 8. Here, we present separate estimates for test scores in Chinese, math, and English, and we investigate the unique impacts of the stereotypes adopted by teachers of each subject in Panels A to C.

Panel A demonstrates that the coefficients of the interaction between teacher stereotypes and non-local student status are significantly negative across test scores in all three core subjects. However, the effect size of the teacher stereotype measure alone is negligible and does not deviate from zero. This pattern suggests that Chinese teacher stereotypes have a stronger adverse influence on the academic performance of non-local students than on that of their local peers, to the extent that no discernible effects are observed with regard to the latter group. In contrast, Panels B and C indicate no significant or consistent effects of math and English teachers' stereotypes on the test scores of either local or non-local students. Additionally, we test the equality of the interaction term coefficients across Panels A to C, and the p values are reported in Panel D. Our analysis indicates significant variation in the effects of teachers' stereotypes among the three subjects, thus suggesting that the negative impacts on non-local students' academic results are predominantly due to the stereotypes adopted by Chinese teachers and may influence performance in other subjects as well.

Chinese teachers' stereotyping significantly impairs the academic performance of non-local students, especially in language-related subjects such as Chinese and English. Our estimates while controlling for all variables indicate that an increase of one standard deviation in such bias among Chinese teachers leads to a decrease of approximately 0.1 standard deviations in the Chinese and English test scores obtained by non-local students. The impact on students' math test scores is less

pronounced; namely, an increase of one standard deviation in Chinese teacher stereotypes results in a decrease of 0.05 standard deviations for non-local students. These findings, which indicate that Chinese teachers' stereotyping affects only non-local students, are in line with studies conducted by Alan, Ertac, and Mumcu (2018) and Carlana (2019), who reported that teachers' gender stereotypes negatively affect girls in terms of their academic performance but have negligible effects on boys. Conversely, Lavy and Sand (2018) revealed that teachers' gender bias can improve boys' test scores while decreasing girls' scores. The economic magnitude of our estimates is on par with the figures previously reported in research on teachers' gender stereotypes, despite variations across different educational contexts, stages, and stereotype measures. [19] [20] Specifically, Alan, Ertac, and Mumcu (2018) discovered that when primary school students are exposed to biased teachers for two to three years, an increase of one standard deviation in a teacher's biased gender role views leads to decreases of 0.12 and 0.06 standard deviations in girls' math and verbal test scores, respectively. In an investigation of the long-term effects of teacher grading bias, Lavy and Sand (2018) revealed that an increase of one standard deviation in the measure of a primary teacher's bias leads to an improvement of 0.09 standard deviations in boys' high school matriculation exam scores but a decrease of 0.06 standard deviations with regard to girls. In the middle school context, Carlana (2019) estimated that an increase of one standard deviation in math teachers' IAT scores is correlated with an increase of 0.03 standard deviations in the gender gap in math performance.

Regarding noncognitive outcomes, no significant effects of three subject teachers' stereotyping of non-local students on the mental health of local or non-local students are observed. However, Chinese teachers' stereotyping of non-local students contributes negatively to the BPI scores of non-local students. In particular, a positive and significant coefficient is observed with regard to the interaction of Chinese teachers' stereotypes with non-local student status, thus indicating a detrimental impact on non-local students. Conversely, math teachers' stereotypes are observed to be linked to a reduction in behaviour problems for both student groups, as indicated by the negative and significant coefficient shown in Panel B. English teachers' stereotypes, however, do not have significant effects on any group's behaviour.

[Insert Table 8 here]

Dual role of Chinese teachers. As discussed in Section 2, a head teacher, who is often also a core subject teacher, is responsible for broader class management beyond the level of academic achieve-

[19] Our research focuses on middle schools in China, while Alan, Ertac, and Mumcu (2018) examined primary schools in Turkey, Lavy and Sand (2018) investigated the stereotypes adopted by primary school teachers in Israel, and Carlana (2019) analysed middle schools in Italy.

[20] Both Alan, Ertac, and Mumcu (2018) and our study feature a series of survey questions used to measure teachers' stereotypes. In contrast, Lavy and Sand (2018) used teachers' subjective assessments, and Carlana (2019) employed Implicit Association Test (IAT) scores.

ment. Table 8 suggests that the primary effects on non-local students' cognitive and noncognitive outcomes are the result of Chinese teachers' stereotyping. Given the additional time that head teachers allocate to the task of managing class discipline and student relations, we investigate whether the influence of Chinese teachers' biased beliefs increases when those teachers simultaneously serve as head teachers. Table 9 presents the differential effects of teacher stereotyping on students' test scores and noncognitive outcomes based on the dual role of Chinese teachers. In this context, the term "dual roles" refers to teachers who have dual duties as both Chinese and head teachers, while the term "subject teacher" refers solely to Chinese teachers.

Our regression models, which include the same comprehensive set of controls as in the baseline analysis, show that the stereotypes adopted by individuals serving as both Chinese and head teachers positively influenced local students' test scores, whereas the stereotypes adopted by individuals functioning solely as Chinese teachers had no significant effect. With regard to non-local students, the sum of the teacher stereotype coefficient and the interaction term coefficient reveals that although stereotyping by Chinese teachers with dual duties is slightly less detrimental to non-local students' test scores, it exacerbates the achievement gap between local and non-local students. In terms of noncognitive outcomes, the stereotypes of Chinese teachers who do not play such dual roles have no significant effect on mental stress but do increase the likelihood of non-local students skipping or arriving late to class; however, they have no impact on local students' BPI scores. However, when Chinese teachers simultaneously serve as head teachers, their stereotypes regarding non-local students intensify the mental stress experienced by both local and non-local students, although no significant effects on students' behavioural problems are observed.

[Insert Table 9 here]

Student gender. Table 10 presents students' gender differences with regard to the influence of Chinese teachers' stereotypes towards non-local students. Our analysis reveals more notable and statistically significant disparities with respect to the effects on test scores and BPI scores for non-local boys than with regard to their female counterparts. Specifically, the test scores of non-local boys who are exposed to Chinese teachers who exhibit an increase of one standard deviation in stereotype measures are notably lower, i.e., by 0.17 standard deviations. Furthermore, these boys are more likely to exhibit tardiness and absenteeism, as evidenced by the fact that an increase of 0.08 standard deviations in their BPI scores corresponds to an increase of one standard deviation in teachers' stereotypes. On the other hand, non-local girls' test scores and BPI scores remain unaffected by such biases, thus suggesting that the detrimental effects of teachers' stereotypes predominantly affect non-local boys. As indicated in the bottom panel in Table 10, the gender differences in the effects on test scores are statistically significant at the 1 percent level. Although the influence on non-

local boys' BPI scores is noteworthy and exceeds the influence observed with respect to non-local girls, we urge prudence in interpreting this gender-based discrepancy given the absence of statistical significance in the differences in these effects on BPI scores between genders. This pattern of gender-specific impacts is consistent with the growing body of research indicating that boys are often more susceptible to adverse conditions within educational and familial settings than are girls (Kling, Ludwig, & Katz, 2005; Bertrand & Pan, 2013; Deming et al., 2014; Dustmann, Ku, & Kwak, 2018; Autor et al., 2019; Stuart, 2022; García, Heckman, & Ronda, 2023).

[Insert Table 10 here]

Duration of residency. Finally, we examine heterogeneity in terms of the length of time for which non-local students have lived in the local area. Previous research has highlighted age-related variations in the ways in which children adapt to environmental shifts, in which context younger children frequently obtain more benefits from relocation than do older teens (Chetty, Hendren, & Katz, 2016; Chyn, 2018; Gross & Baron, 2022). To explore this issue, we separate non-local students into two groups according to the amount of time they have spent in the local county (i.e., 8 years or less vs. more than 8 years) while retaining all local students in our analysis. The results presented in Table 11 suggest that among non-local students who have lived in the local community for an extended period, the negative impact of teachers' negative stereotypes on students' test scores is more severe. Specifically, an increase of one standard deviation in teachers' stereotypes leads to a decrease of 0.14 standard deviations in test scores for non-local students who have lived in the local county for more than eight years. However, for students who have lived in the local county for a shorter span i.e., less than eight years, we observe no statistically significant effects.[21] This discrepancy must be interpreted with caution given the lack of statistical significance.

[Insert Table 11 here]

5.3 Robustness Checks

We conducted a series of robustness checks with regard to our main findings, as detailed in Table 8. Initially, our identification using equation (1) depends on within-block (i.e., the same school and same grade) variations through the inclusion of block fixed effects. In the first robustness check, we

[21]This variation based on the duration of residency is in line with the institutional context of China. Non-local students who have lived in the local area for more than eight years without securing local "*hukou*" status often come from families that face greater socioeconomic challenges. The prerequisites for acquiring local "*hukou*" are closely linked with parental education, occupation, and income. Moreover, these students may have moved to larger cities, where intense competition for scarce local "*hukou*" slots is the norm. Both situations may exacerbate the negative consequences of teacher stereotypes on the cognitive outcomes of students.

replace the block fixed effects with class fixed effects to ensure that any omitted variables at the teacher or class level are taken into account. The regression results on student academic performance and noncognitive outcomes, differentiated by subject based on the stereotypes of Chinese, math, and English teachers, are presented in Panels A to C of Table A2. This robustness test confirms our initial findings, especially concerning the interaction between teacher stereotypes and non-local students.

Second, our study reveals that while math and English teachers' stereotypes do not seem to have disproportionate impacts on the cognitive and noncognitive outcomes of non-local students, the stereotypes adopted by Chinese teachers do have detrimental effects on non-local students' test scores and BPI scores. This situation may give rise to the concern that noncognitive outcomes could be intertwined with academic success and that the observed behavioural issues could be largely the result of academic performance. For example, students with lower academic achievements might be more prone to tardiness or absenteeism. To assess the impact of teachers' stereotyping on students' BPI scores directly, we further control for students' test scores and their interaction with the non-local student indicator in our analysis of noncognitive outcomes. The estimations, which are shown in Table A3, are largely in line with our baseline results, thus indicating that even after controlling for academic performance, the prejudiced beliefs of Chinese teachers continue to have more negative impacts on the noncognitive outcomes of non-local students than on their local peers.

Third, our dataset lacked test scores and noncognitive outcome data concerning certain students. This lack could introduce bias if teachers' stereotypes influence the likelihood of data omission. [22] In Table A4, we regress an attrition indicator for each outcome against teacher stereotypes and the non-local student dummy variable while controlling for block fixed effects. The coefficients for teacher stereotypes across all subjects are negligible and not statistically significant, thus suggesting that attrition rates do not vary alongside the level of teachers' stereotypes.

Fourth, importantly, our sample was recruited based on a stringent criterion: the school principal and all head teachers from the same grade confirmed that student class assignments were random. To determine whether our baseline estimates would be biased by the inclusion of nonrandomized classes, we replicated our baseline model, i.e., equation (1), with schools that did not meet our randomization criteria. Hence, we can infer the direction of bias and conclude whether our baseline findings are biased upwards or downwards even if we were to include nonrandomized classes in our sample. The findings, which are reported in Table A5, indicate a stark contrast to our original results. With respect to Chinese teachers, the interaction term coefficients pertaining to math test scores, English test scores, and BPI scores become statistically nonsignificant and decrease in size. Even the marginally significant impact on Chinese test scores is reduced by more than one-

[22] The attrition may be due to the fact that some students refused to answer questions or did not attend the test.

third. Consequently, if our sample were to include nonrandomized classes, our estimates of the adverse effects of Chinese teacher stereotypes would likely be underestimated, thus serving as a lower boundary for the actual impact.

Fifth, the concern may arise that students could be influenced predominantly by exceptionally progressive teachers, i.e., those who exhibit the lowest stereotype scores. The influence on student outcomes might be attributed to unmeasured aspects of these teachers' quality rather than to stereotyping per se. To address this possibility, we excluded teachers whose stereotype scores were below the 10th percentile in their respective subjects. Table A6 illustrates that the impact of Chinese teachers' stereotypes on non-local students remains robust, though with a slight reduction in precision regarding Chinese test scores.

Relatedly, we conducted a permutation test in which one teacher was removed from the analysis sample in each iteration. This approach allows us to assess the sensitivity of our main results to the influence of individual teachers. As illustrated in Figure 3, the distribution of estimated coefficients predominantly clusters around the baseline coefficient, indicated by the vertical red line. Thus, our findings are not driven by any single teacher.

[Insert Figure 3 here]

Seventh, we performed a permutation test by randomly attributing stereotypes to teachers. Specifically, we reran the baseline regression (equation (1)) one thousand times based on randomly paired teacher stereotypes towards non-local students, and we plotted the coefficients of the interaction term between teachers' stereotypes and non-local student indicators, as illustrated in Figure 4. Clearly, the estimated coefficient distributions predominantly cluster around zero. However, plots in the first column regarding Chinese teachers reveal that the baseline coefficient, which is represented by the reference line from Table 8, is significantly different from zero, thus highlighting the effects of Chinese teachers' stereotypes on non-local students' test scores and BPI scores. Notably, out of 1,000 permutations, only 4 for Chinese test scores and none for English test scores yield a coefficient smaller than that reported in Table 8.

[Insert Figure 4 here]

Finally, in our baseline results, the teacher stereotype measure is a continuous variable. We conducted a robustness check using an alternative binary variable to measure teacher stereotypes against non-locals. Recall that we assess teacher stereotypes based on their evaluation of two hypothetical classes—one composed entirely of local students and the other with more than one-third of

non-local students—across three dimensions: teaching effectiveness, class discipline, and student relations. The binary variable is defined as 1 if the teacher believes that classes with all local students outperform those with one-third of non-local students in at least two of these three dimensions. The results, reported in Table A7, show that while the magnitude of the coefficients changes due to the different stereotype measures, the overall pattern remains consistent with the baseline results. We continue to find that Chinese teachers' stereotypes negatively affect non-local students, whereas no significant effects are observed for math or English teachers.

5.4 Further Checks Regarding the Issue of Reverse Causality

One potential concern is that teachers' assessments of two types of classes are influenced by their observation of their current cohorts rather than representing their inherent stereotyping with regard to non-local students. Consider a scenario in which local students in a teacher's current class consistently outperform non-local students both academically and in terms of their noncognitive skills. If a teacher's beliefs and attitudes regarding the presence of non-local students in a class are shaped by these observations, our analysis may be driven by reverse causality. Nevertheless, we contend that reverse causality is unlikely for several reasons.

First, our construct for teachers' stereotyping of non-local students is based on students' nonacademic assessments, thus decoupling notions of stereotyping from academic achievement. Importantly, the aspects of teacher evaluations regarding the difference between the two class types — the effectiveness of instruction, classroom discipline, and student interactions — are not directly tied to academic performance. Second, the balance tests reported in Section 4.2 did not indicate any correlation between teachers' stereotypes and the proportion of non-local students in the class, the initial academic disparities between local and non-local students, or the teachers' initial academic evaluations upon first teaching the class. Third, our primary findings further decrease the possibility of reverse causality. These findings suggest that only the biased perceptions of Chinese teachers have a significant and consistent influence on the test scores of non-local students across all three core subjects. If teacher stereotypes were merely a reflection of the existing academic gap between local and non-local students, we would expect a significant impact across all subject teachers, such that each teacher would influence only their respective subject. However, our findings do not support this scenario. Additionally, if teacher beliefs were based on disparities between local and non-local students, we would expect only one student group to be affected. However, as shown in Table 8, we observe that math teachers' stereotypes shape noncognitive outcomes and that English teachers' stereotypes influence English test scores for both local and non-local students.

To address concerns regarding reverse causality more robustly, we perform four empirical tests.

First, we refine our measure of teachers' stereotyping by excluding the assessment of teaching effects, which could be linked to the class's academic performance. This refined measure is used to replicate the findings presented in Table 8, and the results, which are presented in Table A8, remain statistically consistent with an analogous effect size. Second, we conduct a falsification test, as part of which we substitute students' current outcomes with their academic rankings and self-reported noncognitive skills from primary school. If teachers' stereotypes were influenced by the present cohort, we would expect significant correlations with the interaction between teachers' stereotyping and non-local status. However, the results shown in Table A9 indicate no significant influence on primary school outcomes, particularly with regard to Chinese teachers' biases, which drive the main effects on non-local student cognitive outcomes. Third, we focus on a subset of teachers with more than one decade of experience; these teachers, who have interacted with numerous cohorts, are less likely to alter their perceptions of local and non-local students based on the performance of their current class. The pattern of results observed with regard to this experienced subgroup, as shown in Table A10, is in line with our baseline findings. Finally, in Table A11, we limit our sample to teachers who simultaneously instruct other classes, thereby reducing the likelihood that their stereotypes are informed solely by the current class's performance. This additional precaution against reverse causality also yields consistent results, as presented in Table 8.

6 Mechanism

Our findings indicate that non-local students' cognitive performance suffers more due to teachers' stereotyping than does that of their local counterparts, as demonstrated in Table 7. Crucially, Table 8 reveals that in our subject-specific heterogeneous analysis, the detrimental effects on non-local students are driven mainly by the stereotypes adopted by Chinese teachers. In addition to harming non-local students' performance in Chinese tests, these biased perceptions of stereotypes spill over to other core subjects and negatively influence non-local students' noncognitive outcomes, such as by increasing their rates of tardiness and class absenteeism. This section explores the potential channels shaping the differential impacts of teachers' stereotypes on the outcomes of local and non-local students. Given that these effects vary across subjects, we examine each subject separately to uncover the underlying mechanisms.

6.1 Teacher Behaviour

One direct channel is teacher behaviour. Previous studies conducted by Alan, Ertac, and Mumcu (2018) and Lavy and Sand (2018) analysed scenarios in which teachers' biased gender beliefs influence

student performance.[23] Similarly, a teacher who exhibits strong biases against non-local students might perceive the academic abilities of those students as inferior or view those students as disruptive to classroom discipline. Such a teacher may implement practices that mirror these prejudices. For instance, by differentiating between local and non-local students in the context of questioning and feedback or by disproportionately praising one group over another. Furthermore, teachers who exhibit ingrained stereotypes regarding non-local students might also interact differently with the parents of non-local and local students, thus potentially facilitating learning for one group more effectively than for the other.

We begin by investigating whether teachers' behaviours towards local and non-local students differ. Our comprehensive dataset includes student-reported interactions with their Chinese, math, and English teachers, particularly with regard to aspects such as classroom questioning and praise.[24] Additionally, we assess teachers' parental engagement by determining how often teachers contacted parents during the semester under investigation.[25] Table 12 presents the estimated effects of teachers' stereotypes on their behaviour; the results for Chinese, math, and English teachers are presented in Panels A, B, and C, respectively. While Chinese teachers' stereotyping does not significantly alter how frequently they question or praise local versus non-local students, these teachers are less likely to reach out to parents of non-local students if they exhibit stronger biases against this group. With respect to math and English teachers, we do not observe any significant impacts of their stereotypes on their interactions.

[Insert Table 12 here]

6.2 Parents' Responses

An indirect yet possible pathway through which teachers' biased views may influence students' cognitive and noncognitive outcomes is through parental response. Increasing evidence in the fields of economics, psychology, and sociology has suggested that parental preferences, attitudes, beliefs, and personality traits are transmitted to their children.[26]; accordingly, this intergenerational transmission may account for a portion of intergenerational correlations in terms of economic outcomes

[23]For instance, Alan, Ertac, and Mumcu (2018) highlighted interactions between teachers and students, grading discrimination (whether against or in favour of girls), and classroom practices reflecting gender norms. Lavy and Sand (2018) discussed teachers' conscious discriminatory behaviours and unintentional teaching styles or personal traits that may favour one group's learning over that of another group.

[24]In the questionnaires, students indicated their agreement with various statements on a scale ranging from 1 (strongly disagree) to 4 (strongly agree); example statements include "My Chinese/math/English teacher always asks me to answer questions in class" and "My Chinese/math/English teacher always praises me."

[25]Each variable included in the regression analysis is standardized to have a mean of 0 and a standard deviation of 1.

[26]See, for example, Duncan et al. (2005),Dohmen et al. (2012),Anger (2012),Grönqvist, Öckert, and Vlachos (2017) and Zumbuehl, Dohmen, and Pfann (2021).

or health.[27]. Consequently, if parents are aware of a teacher's stereotyping of non-local students, their reactions may influence their children, thus potentially affecting the children's outcomes. For example, parents of local students who perceive bias on the part of a teacher might discourage their children from interacting with non-local peers, while parents of non-local students might respond by reducing their engagement with teachers.

To explore this mechanism, we first examine whether parents are conscious of teachers' biases towards non-local students. Parents are surveyed regarding their perceptions of teachers' prejudices against non-local students and the frequency of their communication with teachers throughout the semester.[28] Table 13 presents parents' responses to teachers' bias and the corresponding effects on parental perceptions and engagement with teachers. Our analysis reveals no significant impacts of teacher stereotyping on the beliefs and behaviours of parents for either local or non-local children.

[Insert Table 13 here]

6.3 Class Environment and Interactions with Classmates

Another possible mechanism through which teachers' stereotypes may disproportionately affect local and non-local students is student interaction. The impacts of peer effects in educational settings have been the subject of significant research interest.[29] If teacher bias against non-local students is perceived by the class, such perceptions could lead to segregation between local and non-local students. Local students may engage in discriminatory behaviours against their non-local peers if such biases are conveyed by the teacher, while non-local students, who sense a less welcoming atmosphere, may withdraw from interactions with local students. This segregation could have detrimental effects on the cognitive and noncognitive outcomes of students.

To assess this channel, we investigate the impacts of teachers' stereotypes on interactions between local and non-local students. Students are asked to report on the friendliness of their classmates and to indicate whether any of their five closest friends are non-local.[30] Additionally, we consider whether parents have contacted teachers to discuss their child's friendships to represent a measure of social assimilation at school. Table 14 details the estimated effects of teacher bias on the integration of local and non-local students. In the context of Chinese teachers, the findings suggest that biased

[27]See Björklund and Salvanes (2011), Black, Devereux, and Salvanes (2011), Holmlund, Lindahl, and Plug (2011) and Lindahl et al. (2016).

[28]In the questionnaire, parents rated their perceptions on a scale ranging from 1 (not prejudiced at all) to 4 (very prejudiced). Responses were normalized to a mean of 0 and a standard deviation of 1 for the regression analysis.

[29]Peer effects have been studied in various contexts, such as cognitive performance (Sacerdote, 2001; Ding & Lehrer, 2007; Lavy, Silva, & Weinhardt, 2012; Garlick, 2018), cheating (Carrell, Malmstrom, & West, 2008), and choice of major (De Giorgi, Pellizzari, & Redaelli, 2010).

[30]In the questionnaire, students were asked to rate their level of agreement with the statement "Most of my classmates are friendly to me" on a scale ranging from 1 (strongly disagree) to 4 (strongly agree).

views may lead non-local students to feel less accepted by their peers, encourage friendships within their own group, and increase parental concern regarding their children's social circles. No significant effects are observed with respect to local students.[31]

[Insert Table 14 here]

6.4 Student Responses

Finally, we examine how students react to teachers' stereotyping of non-local students. Research in the field of social psychology has shown that the formation of children's academic self-concepts begins in early childhood and is significantly influenced by the stereotypes adopted by parents and teachers after elementary school ((Ertl, Luttenberger, & Paechter, 2017)). Carlana (2019) explained why self-confidence may be a critical determinant of girls' performance in math and suggested that students use the information conveyed by teachers' feedback to form their own perceptions of their ability. If this feedback includes gender stereotypes, the self-confidence of female students in their math skills may decrease. Similarly, if teachers' perceptions are tainted by stereotypes regarding non-local students, these students might undervalue their academic abilities, thus leading to diminished educational investment. Our comprehensive data allow us to investigate the effects of teachers' stereotypes on students' self-assessed academic standing in class,[32] i.e., an approximated measure of student self-confidence in their academic ability, as well as on student effort by examining the amount of time they dedicate to study, leisure, and other activities.[33]

Table 15 presents the estimated impact of teachers' stereotyping of non-local students on student responses. With regard to student self-assessed academic rank, the interaction term between Chinese teachers' stereotypes and non-local students is negative and statistically significant. This finding indicates that Chinese teachers' stereotypes adversely affect non-local students' self-confidence. Specifically, one standard deviation increase in the Chinese teachers' stereotyping leads to a 0.07 standard deviation increase in the self-confidence gap between local and non-local students. While

[31] Conversely, math teachers' biases appear to establish a positive classroom environment, thus encouraging both local and non-local students to befriend non-local peers. However, no significant impact on parental communication regarding their children's friendships is observed. English teachers' bias does not seem to have any observable effects.

[32] In the questionnaire, students were asked to evaluate their academic rank in the class on a scale ranging from 1 (near the bottom) to 5 (near the top). For regression analysis, we standardized this measure with a mean of 0 and a standard deviation of 1.

[33] Students were asked to report the average amount of time they spent weekly on various activities, such as schoolwork assigned by teachers, additional schoolwork assigned by parents or supplementary schools, supplementary classes, sports, reading, watching television, playing computer games, and completing household chores. We categorized the amount of time spent on teacher-assigned schoolwork as "schoolwork within school", the amount of time spent on parent/supplementary school assignments and supplementary classes as "schoolwork outside school" and the sum of the amounts of time spent on sports, television, and gaming as "playtime". The amounts of time spent reading and completing household chores were categorized as "other activities". In our regression analyses, we normalized each variable with a mean of 0 and a standard deviation of 1.

math teachers' stereotypes do not seem to impact either local or non-local students' self-perceptions of academic ability, English teachers' biased views regarding non-local students decrease the self-confidence of both groups.

We also investigate whether teachers' stereotypes influence students' efforts by examining the average amount of time spent on various activities. We find that while English teachers' stereotypes motivate non-local students to spend more time on schoolwork assigned by teachers and math teachers' stereotypes encourage non-local students to take supplementary courses or complete homework assigned by parents or supplementary schools, Chinese teachers' stereotypes discourage non-local students from devoting time to their studies and increase the amount of time they spend on entertainment. This pattern may partially explain why we find adverse effects on non-local students' academic achievements only with regard to Chinese teachers' stereotypes and observe no corresponding effects among math or English teachers.

[Insert Table 15 here]

6.5 Understanding Student Gender Differences in the Impact of Teachers' Stereotyping of Non-Local Students on Non-Local Students

The gender differences observed in the impact of teachers' biases on non-local students can be elucidated by examining various mediating factors, as extensively detailed in Section 6. Our analysis explores potential gender disparities in these mediating factors, including teacher behaviours, parental responses, the class environment, and student reactions. Table 8 exclusively documents the significant effects of Chinese teachers' stereotyping of non-local students on student outcomes, thus prompting further investigation into these dimensions with the goal of obtaining deeper insights.

Table A12 presents the heterogeneous effects of Chinese teachers' stereotypes on their classroom practices between student genders. While minimal effects on teachers' classroom behaviours are observed for both boys (panel A) and girls (panel B), an intriguing pattern emerges. The hypothesis that the estimated effects on the classroom questioning of non-local students are equal for girls and boys can be rejected at the 10 percent level (p=0.057). This finding suggests that Chinese teachers who exhibit more biased beliefs tend to pose more questions to non-local girls during class while engaging less with non-local boys. Moreover, Chinese teachers who exhibit higher stereotyping scores are more likely to contact the parents of both local and non-local girls; however, this trend is not observed with regard to boys' parents. Consequently, it is plausible that increased teacher-parent interaction may compensate for the perceived disadvantages of non-local girls in the classroom.

Gender-specific disparities in the effects of teachers' stereotypes on parental response and behaviour are examined in Table A13. The results indicate no strong evidence of gender differences in the treatment effects. No statistically significant impacts on parents' beliefs are observed among either boys or girls, and we cannot reject the hypothesis that the point estimates are equal for girls and boys. However, parents of non-local girls exhibit reduced contact with teachers, whereas no significant change is observed among parents of non-local boys. This divergence could be due to the increased teacher-initiated contact documented in Table A12.

Additionally, we examine gender differences in the impacts of teachers' stereotyping on the dynamic class environment and student interactions presented in Table A14. Notably, the adverse effects of Chinese teachers' stereotyping solely impact non-local boys, as evidenced by a reduced belief in their classmates' friendliness. Conversely, non-local girls experience no significant such impact. Notably, significant gender differences in perceptions of class friendliness are observed at the 5 percent level, thus indicating dissimilar effects for girls and boys. Furthermore, the results provide little evidence regarding differences between boys (panel A) and girls (panel B) in the remaining two outcomes. While decisions regarding befriending non-local students are not significantly influenced by teacher stereotypes, statistically significant impacts are observed with respect to parents' concerns about friendship issues for both genders.

As a final point regarding mediating factors, we examine the gender-specific effects of teachers' stereotypes on student responses. Table A15 presents the impacts on self-confidence level and time allocation for both boys (panel A) and girls (panel B). The results indicate a significant, negative impact on self-evaluated academic rank among non-local boys, which contrasts with the absence of such a significant impact among non-local girls. The point estimates differ statistically at the 10 percent level (p=0.09), thus suggesting a gender difference in terms of self-confidence. This finding is in line with the literature, which has suggested that girls are more resilient than boys with regard to changes in the school or family environment. For instance, Dustmann, Ku, and Kwak (2018) reported that boys' presence does not distract girls from academic pursuits, whereas girls' presence seems to negatively influence boys' performance. Chetty and Hendren (2018) and Chetty et al. (2016) revealed that boys' early adult employment rates are differentially sensitive to neighbourhood exposure, unlike those of girls.

Columns 2-5 reveal that in response to Chinese teachers' stereotyping, non-local girls allocate more time to schoolwork at school, while non-local boys spend more time on play and entertainment. Substantial differences in time allocation for schoolwork are observed at the 1 percent level. Although we cannot statistically reject the hypothesis regarding equal effects on entertainment time allocation, the point estimates suggest that the effect on non-local boys is larger and more pre-

cisely estimated than that on non-local girls. Non-local girls respond to teacher stereotyping by exhibiting increased effort and investment in their studies, while the study efforts of non-local boys are reduced. This pattern may verify the argument that girls exhibit heightened responsiveness to adverse environments, which encourages them to intensify their academic efforts. Previous studies have highlighted the fact that girls are more inclined to complete homework than boys and possess superior study habits (Jacob, 2002; Hastings, Kane, & Staiger, 2006; Frenette & Zeman, 2007; Deming et al., 2014).

In summary, our analysis of the heterogeneous effects of student gender on the mediating factors reveals distinct responses to Chinese teachers who exhibit high-level stereotypes regarding non-local students. Boys perceive a less amiable classroom environment, exhibit reduced self-confidence, and allocate less effort to their studies. Conversely, girls exhibit greater adaptability to environmental shifts as well as a stronger focus on academics and enhanced study efforts. These gender disparities in class environments and student responses may significantly contribute to attempts to understand why the negative impacts on academic achievements and the BPI scores are observed predominantly among non-local boys.

7 Conclusion

In this paper, we investigate the impact of middle school teachers' stereotyping of non-local students on their academic performance and noncognitive outcomes. Leveraging the random assignments of students to classes, we compare outcomes among students in the same school and grade who experience varying levels of teacher stereotyping. Our results show that teacher stereotypes significantly contribute to disparities in cognitive achievement and noncognitive behaviors between local and non-local students. Non-local students with teachers who hold strong stereotypical views perform worse on standardized tests and are more prone to absenteeism and tardiness. In contrast, local students' outcomes remain unaffected by these stereotypes. The negative effects are primarily driven by Chinese teachers, whose biases reduce non-local students' performance not only in Chinese but also spill over to other core subjects and have detrimental impacts on noncognitive outcomes.

Our analysis uses data from the China Education Panel Survey (CEPS), enabling us to control for a broad range of teacher and student characteristics relevant to the effects of teacher stereotyping on non-local students' outcomes. Using responses from teachers, parents, and students, we investigate the mechanisms through which Chinese teachers' stereotypes may shape student achievement. We find that teacher biases impact not only their behavior but also classroom dynamics and student interactions. Teachers with stronger biases are less likely to engage with non-local students' parents

and create a classroom environment that feels unwelcoming to non-local students, shaping their social relationships and self-perceptions. These biased signals undermine non-local students' self-confidence and reduce their efforts, potentially widening the long-term developmental gap between local and non-local students.

Given the importance of the school environment for shaping sequential choices and outcomes, our results suggest that teacher stereotyping of one particular group over another may foster underconfidence and poor performance among stigmatized students, further widening long-term developmental gaps between groups. These findings point to several policy interventions that could mitigate the effects of teacher bias. Educating teachers about their biases and training them to promote equality in the classroom are potential first steps. For example, Alesina et al. (2024) demonstrates that revealing teachers' implicit stereotypes before assigning grades reduces the grade gap between native and immigrant students. Alternatively, schools could match teachers and students more strategically—assigning local students to teachers with higher levels of stereotyping and non-local students to teachers with lower levels of bias. Enhancing non-local students' self-confidence could also help them resist the effects of stereotyping. Additional strategies might include recruiting more non-local teachers as relatable role models and keeping students' "*hukou*" status confidential. Future research could explore the effectiveness of these interventions in mitigating the impact of teacher stereotyping on non-local students.

References

Alan, Sule, Enes Duysak, Elif Kubilay, and Ipek Mumcu (2023). "Social Exclusion and Ethnic Segregation in Schools: The Role of Teachers' Ethnic Prejudice". *Review of Economics and Statistics* 105.5, pp. 1039–1054.

Alan, Sule, Seda Ertac, and Ipek Mumcu (2018). "Gender Stereotypes in the Classroom and Effects on Achievement". *Review of Economics and Statistics* 100.5, pp. 876–890.

Alesina, Alberto, Michela Carlana, Eliana La Ferrara, and Paolo Pinotti (2024). "Revealing Stereotypes: Evidence from Immigrants in Schools". *American Economic Review* 114.7, pp. 1916–1948.

Anger, Silke (2012). "The Intergenerational Transmission of Cognitive and Non-Cognitive Skills During Adolescence and Young Adulthood".

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman (2019). "Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes". *American Economic Journal: Applied Economics* 11.3, pp. 338–381.

Avitzour, Eliana, Adi Choen, Daphna Joel, and Victor Lavy (2020). *On the Origins of Gender-Biased Behavior: The Role of Explicit and Implicit Stereotypes*. Tech. rep. No. w27818. National Bureau of Economic Research.

Bertrand, Marianne and Jessica Pan (2013). "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior". *American Economic Journal: Applied Economics* 5.1, pp. 32–64.

Björklund, Anders and Kjell G Salvanes (2011). "Education and Family Background: Mechanisms and Policies". *Handbook of the Economics of Education*. Vol. 3. Elsevier, pp. 201–247.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes (2011). "Too Young to Leave the Nest? The Effects of School Starting Age". *The Review of Economics and Statistics* 93.2, pp. 455–467.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer (2016). "Stereotypes". *The Quarterly Journal of Economics* 131.4, pp. 1753–1794.

— (2019). "Beliefs about Gender". *American Economic Review* 109.3, pp. 739–73.

Cadsby, C Bram, Fei Song, and Xiaolan Yang (2020). "Are "Left-Behind" Children Really Left Behind? A Lab-in-Field Experiment Concerning the Impact of Rural/Urban Status and Parental Migration on Children's Other-Regarding Preferences". *Journal of Economic Behavior & Organization* 179, pp. 715–728.

Carlana, Michela (2019). "Implicit Stereotypes: Evidence from Teachers' Gender Bias". *The Quarterly Journal of Economics* 134.3, pp. 1163–1224.

Carlana, Michela, Eliana La Ferrara, and Paolo Pinotti (2022). "Goals and Gaps: Educational Careers of Immigrant Children". *Econometrica* 90.1, pp. 1–29.

Carrell, Scott E, Frederick V Malmstrom, and James E West (2008). "Peer Effects in Academic Cheating". *Journal of Human Resources* 43.1, pp. 173–207.

Cerna, Lucie, Ottavia Brussino, and Cecilia Mezzanotte (2021). "The Resilience of Students with an Immigrant Background: An Update with PISA 2018".

Chetty, Raj, John N Friedman, and Jonah E Rockoff (2014a). "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates". *American Economic Review* 104.9, pp. 2593–2632.

— (2014b). "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood". *American Economic Review* 104.9, pp. 2633–79.

Chetty, Raj and Nathaniel Hendren (2018). "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects". *The Quarterly Journal of Economics* 133.3, pp. 1107–1162.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz (2016). "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment". *American Economic Review* 106.4, pp. 855–902.

Chetty, Raj, Nathaniel Hendren, Frina Lin, Jeremy Majerovitz, and Benjamin Scuderi (2016). "Childhood Environment and Gender Gaps in Adulthood". *American Economic Review* 106.5, pp. 282–288.

Chyn, Eric (2018). "Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children". *American Economic Review* 108.10, pp. 3028–3056.

Coffman, Katherine Baldiga (2014). "Evidence on Self-Stereotyping and the Contribution of Ideas". *The Quarterly Journal of Economics* 129.4, pp. 1625–1660.

Cunha, Flavio and James Heckman (2007). "The Technology of Skill Formation". *American Economic Review* 97.2, pp. 31–47.

De Giorgi, Giacomo, Michele Pellizzari, and Silvia Redaelli (2010). "Identification of Social Interactions through Partially Overlapping Peer Groups". *American Economic Journal: Applied Economics* 2.2, pp. 241–75.

Dee, Thomas S (2005). "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?" *American Economic Review* 95.2, pp. 158–165.

Deming, David J, Justine S Hastings, Thomas J Kane, and Douglas O Staiger (2014). "School Choice, School Quality, and Postsecondary Attainment". *American Economic Review* 104.3, pp. 991–1013.

Dhar, Diva, Tarun Jain, and Seema Jayachandran (2022). "Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India". *American Economic Review* 112.3, pp. 899–927.

Ding, Weili and Steven F Lehrer (2007). "Do Peers Affect Student Achievement in China's Secondary Schools?" *The Review of Economics and Statistics* 89.2, pp. 300–312.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde (2012). "The Intergenerational Transmission of Risk and Trust Attitudes". *The Review of Economic Studies* 79.2, pp. 645–677.

Duncan, Greg, Ariel Kalil, Susan E Mayer, Robin Tepper, and Monique R Payne (2005). "The Apple Does not Fall Far from the Tree". *Unequal Chances: Family Background and Economic Success*, pp. 23–79.

Dustmann, Christian, Hyejin Ku, and Do Won Kwak (2018). "Why Are Single-sex Schools Successful?" *Labour Economics* 54, pp. 79–99.

Eble, Alex and Feng Hu (2019). "How Important Are Beliefs about Gender Differences in Math Ability? Transmission across Generations and Impacts on Child Outcomes". *CDEP-CGEG Working Paper* 53.

Ertl, Bernhard, Silke Luttenberger, and Manuela Paechter (2017). "The Impact of Gender Stereotypes on the Self-Concept of Female Students in STEM Subjects with an Under-Representation of Females". *Frontiers in Psychology* 8, p. 703.

Fairlie, Robert W, Florian Hoffmann, and Philip Oreopoulos (2014). "A Community College Instructor Like Me: Race and Ethnicity Interactions in the Classroom". *American Economic Review* 104.8, pp. 2567–91.

Feld, Jan, Nicolás Salamanca, and Daniel S Hamermesh (2016). "Endophilia or Exophobia: Beyond Discrimination". *The Economic Journal* 126.594, pp. 1503–1527.

Fisman, Raymond, Jing Shi, Yongxiang Wang, and Rong Xu (2018). "Social Ties and Favoritism in Chinese Science". *Journal of Political Economy* 126.3, pp. 1134–1171.

Frenette, Marc and Klarka Zeman (2007). *Why Are Most University Students Women?Evidence Based on Academic Performance, Study Habits and Parental Influences.*

Fryer Jr, Roland G and Steven D Levitt (2004). "Understanding the Black-White Test Score Gap in the First Two Years of School". *Review of Economics and Statistics* 86.2, pp. 447–464.

García, Jorge Luis, James J Heckman, and Victor Ronda (2023). "The Lasting Effects of Early-Childhood Education on Promoting the Skills and Social Mobility of Disadvantaged African Americans and Their Children". *Journal of Political Economy* 131.6, pp. 000–000.

Garlick, Robert (2018). "Academic Peer Effects with Different Group Assignment Policies: Residential Tracking versus Random Assignment". *American Economic Journal: Applied Economics* 10.3, pp. 345–69.

Gershenson, Seth, Cassandra MD Hart, Joshua Hyman, Constance A Lindsay, and Nicholas W Papageorge (2022). "The Long-Run Impacts of Same-Race Teachers". *American Economic Journal: Economic Policy* 14.4, pp. 300–342.

Glover, Dylan, Amanda Pallais, and William Pariente (2017). "Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores". *The Quarterly Journal of Economics* 132.3, pp. 1219–1260.

Gong, Jie, Yi Lu, and Hong Song (2018). "The Effect of Teacher Gender on Students' Academic and Noncognitive Outcomes". *Journal of Labor Economics* 36.3, pp. 743–778.

— (2021). "Gender Peer Effects on Students' Academic and Noncognitive Outcomes Evidence and Mechanisms". *Journal of Human Resources* 56.3, pp. 686–710.

Grönqvist, Erik, Björn Öckert, and Jonas Vlachos (2017). "The Intergenerational Transmission of Cognitive and Noncognitive Abilities". *Journal of Human Resources* 52.4, pp. 887–918.

Gross, Max and E Jason Baron (2022). "Temporary Stays and Persistent Gains: The Causal Effects of Foster Care". *American Economic Journal: Applied Economics* 14.2, pp. 170–199.

Hastings, Justine S, Thomas J Kane, and Douglas O Staiger (2006). "Gender and Performance: Evidence from School Assignment by Randomized Lottery". *American Economic Review* 96.2, pp. 232–236.

Holmlund, Helena, Mikael Lindahl, and Erik Plug (2011). "The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods". *Journal of Economic Literature* 49.3, pp. 615–51.

Hu, Feng (2018). "Migrant Peers in the Classroom: Is the Academic Performance of Local Students Negatively Affected?" *Journal of Comparative Economics* 46.2, pp. 582–597.

Huang, Zibin (2020). "Peer Effects of Migrant and Left-Behind Children: Evidence from Classroom Random Assignment in China". *Available at SSRN 3434160*.

Jackson, C Kirabo (2018). "What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes". *Journal of Political Economy* 126.5, pp. 2072–2107.

Jacob, Brian A (2002). "Where the Boys Aren't: Non-Cognitive Skills, Returns to School and the Gender Gap in Higher Education". *Economics of Education review* 21.6, pp. 589–598.

Kling, Jeffrey R, Jens Ludwig, and Lawrence F Katz (2005). "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment". *The Quarterly Journal of Economics* 120.1, pp. 87–130.

Lavy, Victor and Edith Sand (2018). "On the Origins of Gender Gaps in Human Capital: Short-and Long-Term Consequences of Teachers' Biases". *Journal of Public Economics* 167, pp. 263–279.

Lavy, Victor, Edith Sand, and Moses Shayo (2018). *The Surprisingly Small Effects of Religion-Based Discrimination in Education*. Tech. rep. No. w24922. National Bureau of Economic Research.

Lavy, Victor, Olmo Silva, and Felix Weinhardt (2012). "The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools". *Journal of Labor Economics* 30.2, pp. 367–414.

Lindahl, Mikael, Evelina Lundberg, Mårten Palme, and Emilia Simeonova (2016). *Parental Influences on Health and Longevity: Lessons from a Large Sample of Adoptees*. Tech. rep. No. w21946. National Bureau of Economic Research.

Lusher, Lester, Doug Campbell, and Scott Carrell (2018). "TAs Like Me: Racial Interactions Between Graduate Teaching Assistants and Undergraduates". *Journal of Public Economics* 159, pp. 203–224.

Meng, Xin and Chikako Yamauchi (2017). "Children of Migrants: The Cumulative Impact of Parental Migration on Children's Education and Health Outcomes in China". *Demography* 54.5, pp. 1677–1714.

Mu, Ren and Alan De Brauw (2015). "Migration and Young Child Nutrition: Evidence from Rural China". *Journal of Population Economics* 28.3, pp. 631–657.

NBSC (2021a). *Communiqué of the Seventh National Population Census (No. 7)*. Tech. rep. National Bureau of Statistics of China.

— (2021b). *The 2020 Migrant Workers Monitoring and Survey Report*. Tech. rep. National Bureau of Statistics of China.

NBSC, UNICEF, and UNFPA (2017). *Population Status of Children in China in 2015*. Tech. rep. National Bureau of Statistics of China, the United Nations Children's Fund China, and the United Nations Population Fund China.

— (2023). *What the 2020 Census Can Tell Us About Children in China*. Tech. rep. National Bureau of Statistics of China, the United Nations Children's Fund China, and the United Nations Population Fund China.

OECD (2023). *Review Education Policies: Migration Background*. Tech. rep. OECD Publishing, Paris.

Oswald, Frederick L, Gregory Mitchell, Hart Blanton, James Jaccard, and Philip E Tetlock (2013). "Predicting Ethnic and Racial Discrimination: A Meta-Analysis of IAT Criterion Studies". *Journal of Personality and Social Psychology* 105.2, p. 171.

Papageorge, Nicholas W, Seth Gershenson, and Kyung Min Kang (2020). "Teacher Expectations Matter". *Review of Economics and Statistics* 102.2, pp. 234–251.

Rakshit, Sonali and Soham Sahoo (2023). "Biased Teachers and Gender Gap in Learning Outcomes: Evidence from India". *Journal of Development Economics* 161, p. 103041.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales (2014). "How Stereotypes Impair Women's Careers in Science". *Proceedings of the National Academy of Sciences* 111.12, pp. 4403–4408.

Sacerdote, Bruce (2001). "Peer Effects with Random Assignment: Results for Dartmouth Roommates". *The Quarterly Journal of Economics* 116.2, pp. 681–704.

Spencer, Steven J, Claude M Steele, and Diane M Quinn (1999). "Stereotype Threat and Women's Math Performance". *Journal of Experimental Social Psychology* 35.1, pp. 4–28.

Steele, Claude M and Joshua Aronson (1995). "Stereotype Threat and the Intellectual Test Performance of African Americans." *Journal of Personality and Social Psychology* 69.5, p. 797.

Stuart, Bryan A (2022). "The Long-Run Effects of Recessions on Education and Income". *American Economic Journal: Applied Economics* 14.1, pp. 42–74.

Terrier, Camille (2020). "Boys Lag Behind: How Teachers' Gender Biases Affect Student Achievement". *Economics of Education Review* 77, p. 101981.

Van den Bergh, Linda, Eddie Denessen, Lisette Hornstra, Marinus Voeten, and Rob W Holland (2010). "The Implicit Prejudiced Attitudes of Teachers: Relations to Teacher Expectations and the Ethnic Achievement Gap". *American Educational Research Journal* 47.2, pp. 497–527.

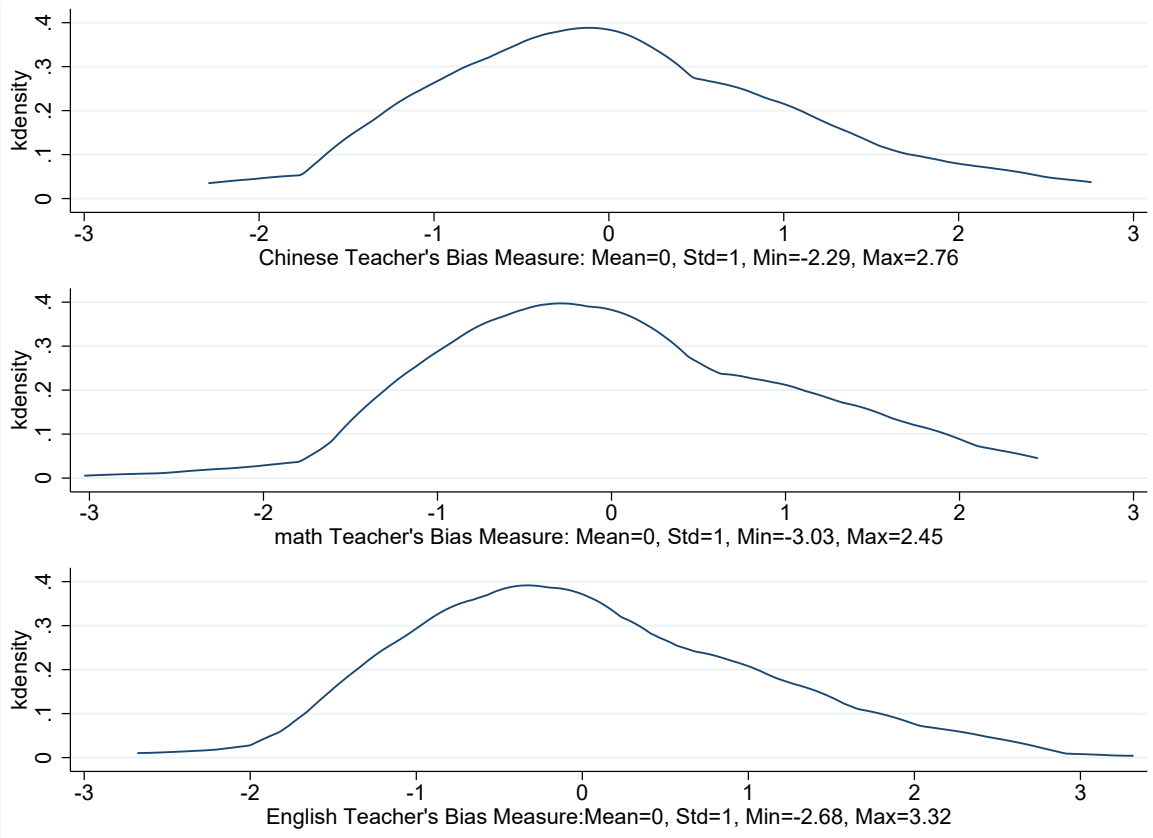
Yang, DP, JY Wei, and DP Qin (2016). "Annual Report on Education for China's Migrant Children". *Social Sciences Academic Press*.

Zhang, Hongliang, Jere R Behrman, C Simon Fan, Xiangdong Wei, and Junsen Zhang (2014). "Does Parental Absence Reduce Cognitive Achievements? Evidence from Rural China". *Journal of Development Economics* 111, pp. 181–195.

Zumbuehl, Maria, Thomas Dohmen, and Gerard Pfann (2021). "Parental Involvement and the Intergenerational Transmission of Economic Preferences, Attitudes and Personality Traits". *The Economic Journal* 131.638, pp. 2642–2670.

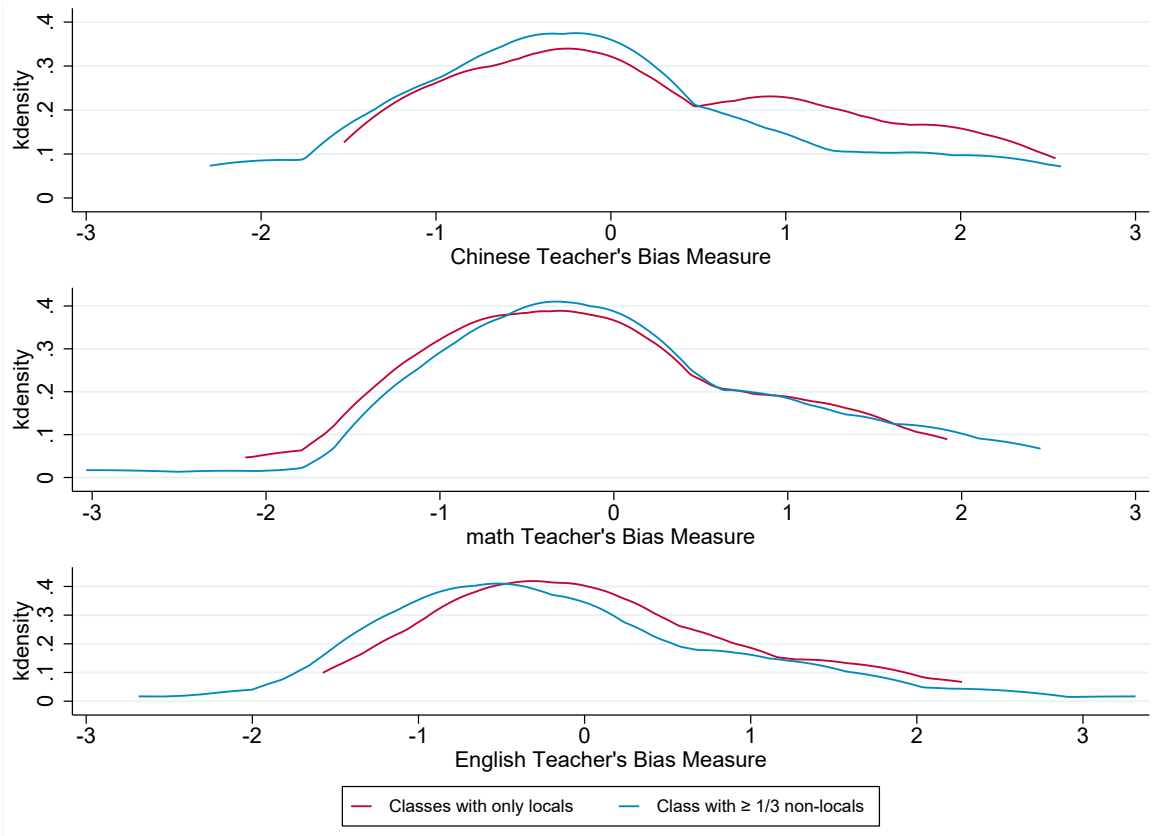
Figures and Tables

Figure 1: Distribution of Teachers' Stereotypes



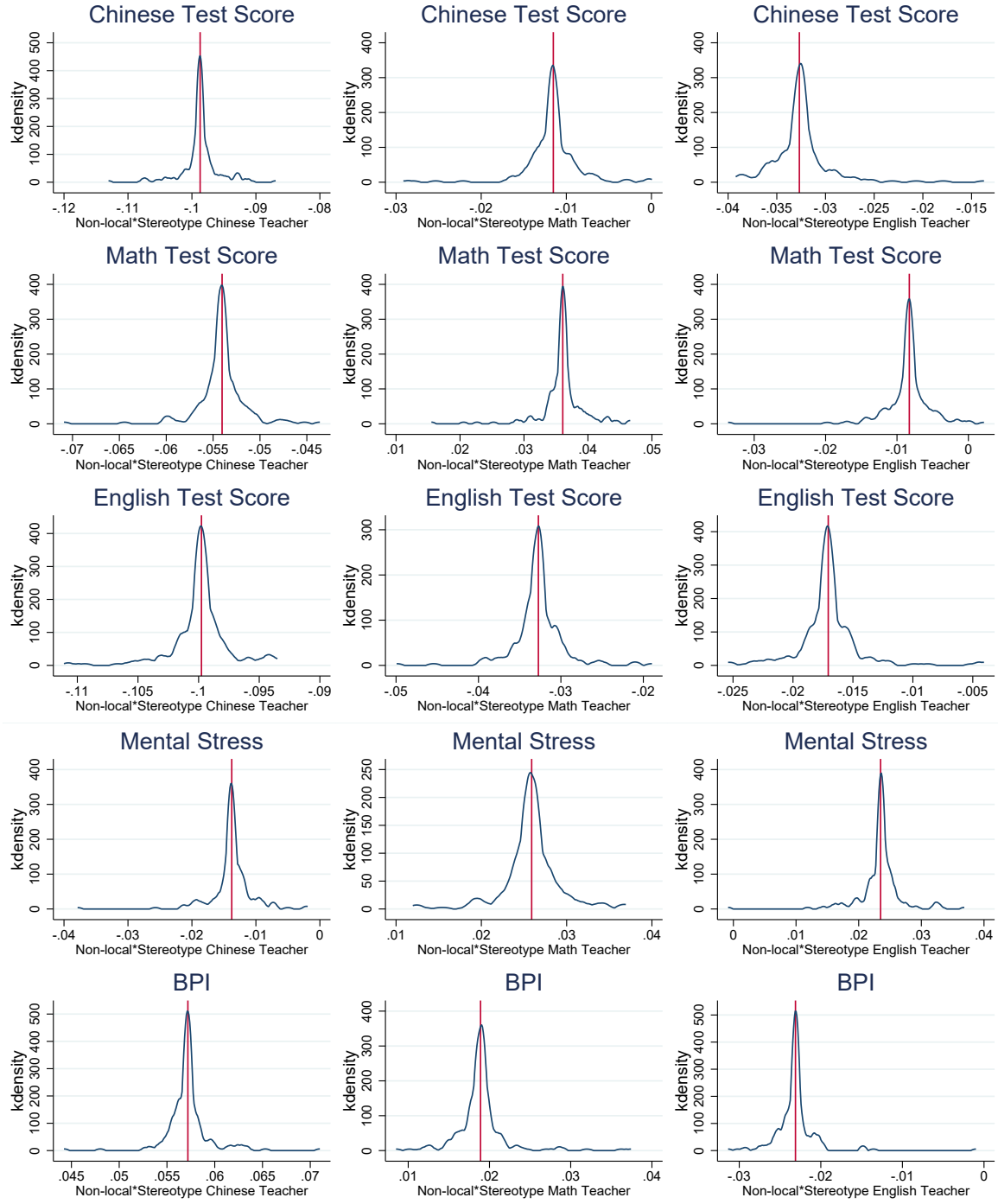
Note: This figure illustrates the distribution of standardized stereotype scores for Chinese, math, and English teachers. A higher stereotype score signifies that the teacher is more likely to believe that classes with only local students perform better in terms of teaching effectiveness, class discipline, and student relations compared to classes with over one-third of non-local students.

Figure 2: Distribution of Teachers' Stereotypes by Class Types



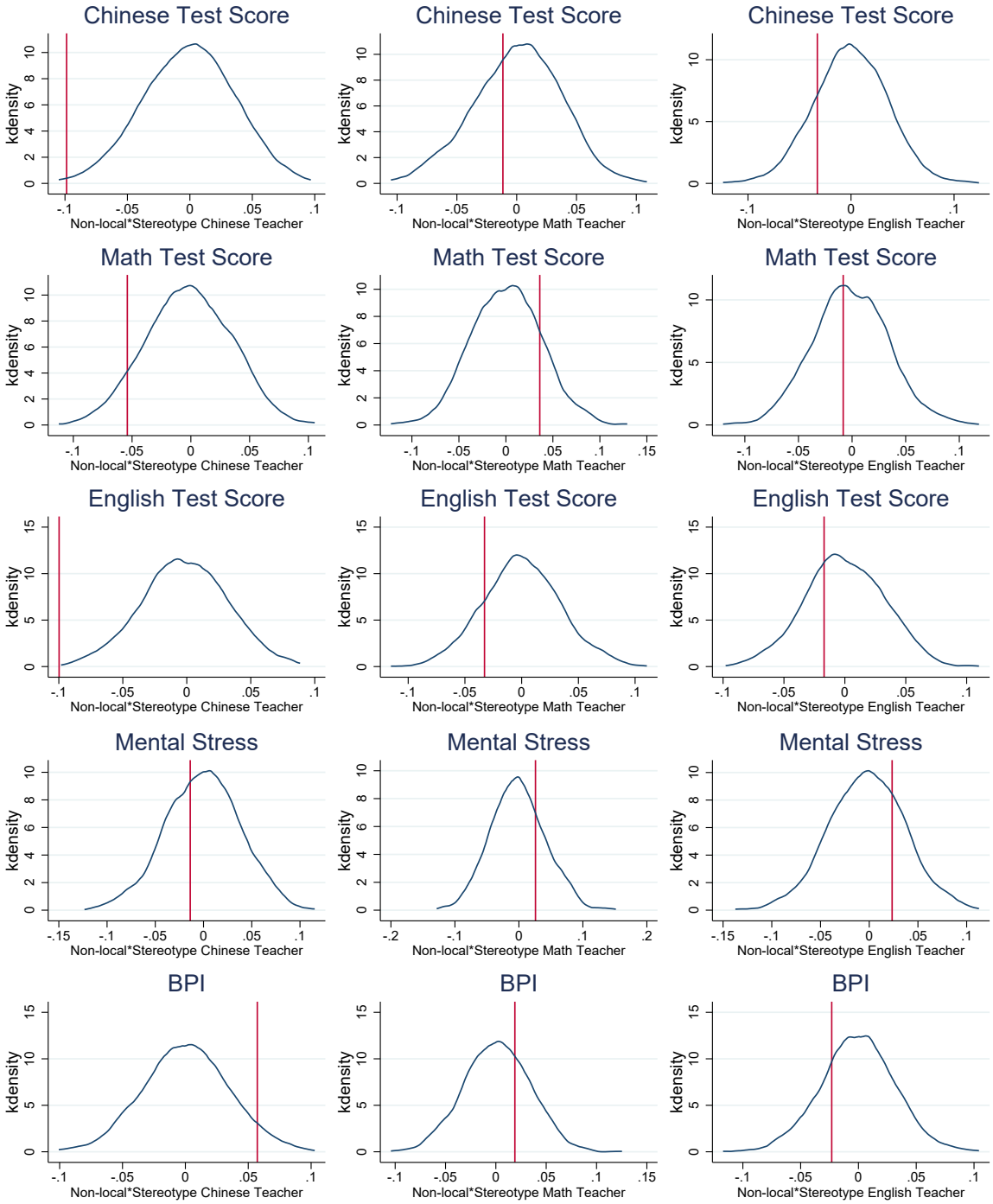
Note: This figure presents the distribution of standardized stereotype scores for Chinese, math, and English teachers by the type of classes they currently teach. A higher stereotype score signifies that the teacher is more likely to believe that classes with only local students perform better in terms of teaching effectiveness, class discipline, and student relations compared to classes with over one-third of non-local students. While teachers provided evaluations based on hypothetical class types, independent of their current class composition, this figure also examines the stereotype distribution based on the actual composition of the classes they teach—comparing those with only local students to those with over one-third of non-local students.

Figure 3: Permutation Tests, Drop One Teacher



Note: This figure shows the "Non-local*Stereotypes" results from a permutation test, where one teacher is removed from the analysis sample in each iteration. The vertical red line indicates the coefficient from the regressions in Table 8, with panel A corresponding to Chinese teachers, panel B to math teachers, and panel C to English teachers.

Figure 4: Permutation Tests, Randomly Assign Stereotypes to Teachers



Note: This figure shows the "Non-local*Stereotypes" results from a permutation test, which runs the main regression from equation (1) 1000 times by randomly assigning stereotype scores to Chinese, math, and English teachers. The vertical red line represents the coefficient obtained from the regressions in Table 8, with panel A corresponding to Chinese teachers, panel B to math teachers, and panel C to English teachers.

Table 1: Variable and Survey Question

Variable	Construction	Survey Question
Teacher's Stereotyping of Non-local students	Mean of standardized differences (1a-1b), (2a-2b), and (3a-3b).	(1a) What's your opinion of effect of teaching of classes composed of only local students? (1-Very bad; 5-Very good) (1b) What's your opinion of effect of teaching of classes with more than one third from non-local counties/districts? (1-Very bad; 5-Very good) (2a) What's your opinion of class discipline of classes composed of only local students? (1-Very bad; 5-Very good) (2b) What's your opinion of class discipline of classes with more than one third from non-local counties/districts? (1-Very bad; 5-Very good) (3a) What's your opinion of relations among students of classes composed of only local students? (1-Very bad; 5-Very good) (3b) What's your opinion of relations among students of classes with more than one third from non-local counties/districts? (1-Very bad; 5-Very good)
Test Score Index	Score of each subject is standarized by grade and school. The principal components analysis was performed to create the corresponding index outcomes which are then standardized with mean 0 and standard deviation of 1.	(1) Chinese test score. (0-100) (2) Math test score. (0-100) (3) English test score. (0-100)
Mental Stress Index		(1) Do you have the feelings below in the last seven days? Feeling blue (1-Never; 5-Always) (2) Do you have the feelings below in the last seven days? Depressed (1-Never; 5-Always) (3) Do you have the feelings below in the last seven days? Unhappy (1-Never; 5-Always) (4) Do you have the feelings below in the last seven days? Not enjoying life (1-Never; 5-Always) (5) Do you have the feelings below in the last seven days? Sad (1-Never; 5-Always)
Behavior Problem Index (BPI)		(1) I am always late for class. (1-Strongly disagree; 4-Strongly agree) (2) I always skip classes. (1-Strongly disagree; 4-Strongly agree)

Table 2: Summary Statistics of Teachers

	All Teacher			Chinese Teacher			Math Teacher			English Teacher		
	Count (1)	Mean (2)	SD (3)	Count (4)	Mean (5)	SD (6)	Count (7)	Mean (8)	SD (9)	Count (10)	Mean (11)	SD (12)
Panel A: Demographic												
Male	617	0.24	0.42	205	0.23	0.42	208	0.38	0.49	204	0.10	0.30
Age	609	37.19	7.89	203	37.66	7.90	206	37.82	7.77	200	36.08	7.94
Marital status (1-Single; 0-Married)	615	0.13	0.34	205	0.09	0.29	206	0.13	0.33	204	0.17	0.38
Panel B: Education and job characteristics												
Years of schooling	613	15.60	0.82	206	15.65	0.93	204	15.61	0.81	203	15.55	0.70
Pedagogical college/major (1-Pedagogical; 0-Other)	614	0.94	0.24	206	0.95	0.22	206	0.95	0.22	202	0.91	0.29
Years of experience	599	15.66	8.62	201	15.99	8.85	200	16.28	8.41	198	14.69	8.54
With tenure	605	0.87	0.33	202	0.88	0.32	202	0.90	0.30	201	0.84	0.37
With professional job title	614	0.94	0.24	204	0.96	0.21	206	0.93	0.25	204	0.93	0.26
With prior teaching tenure in other schools	612	0.56	0.50	205	0.56	0.50	204	0.53	0.50	203	0.60	0.49
Panel C: Stereotype												
Effect of teaching-all local student	589	3.89	0.72	196	3.86	0.74	197	3.86	0.73	196	3.95	0.71
Effect of teaching-some nonlocal student	565	3.25	0.90	189	3.17	0.95	191	3.32	0.86	185	3.25	0.89
Class discipline-all local student	587	3.95	0.75	195	3.91	0.78	197	3.95	0.72	195	3.98	0.75
Class discipline-some nonlocal student	565	3.43	0.92	189	3.34	1.02	191	3.48	0.83	185	3.48	0.90
Relations among students-all local student	587	4.08	0.64	195	4.07	0.66	197	4.09	0.63	195	4.08	0.64
Relations among students-some nonlocal student	566	3.58	0.86	190	3.55	0.95	191	3.63	0.78	185	3.56	0.86
Teacher's stereotyping of non-locals(Raw)	560	0.56	0.81	188	0.61	0.90	187	0.49	0.75	185	0.59	0.78

Table 3: Correlations between Teachhers' Characteristics and Stereotype

Dep. var: Stereotype					
Panel A: Independent variables (teachers' background and education)					
	Male	Age	Marital statu (Single)	College and above	Pedagogical college/major
	(1)	(2)	(3)	(4)	(5)
	0.017	-0.007	0.194	-0.000	0.270
	(0.119)	(0.007)	(0.134)	(0.056)	(0.188)
Observations	559	554	557	555	556
Panel B: Independent variables (teachers' quality and performance)					
	Years of experience	With tenure	Professional job title	Prior experience in other schools	Whether is a head teacher
	(1)	(2)	(3)	(4)	(5)
	-0.009	-0.484**	-0.149	-0.038	0.002
	(0.007)	(0.206)	(0.206)	(0.101)	(0.086)
Observations	545	550	557	554	560

Notes: This table reports OLS estimates of the correlation between teachers' stereotype score and their own characteristics. Standard errors (in parentheses) are robust and clustered at the block level. Subject fixed effects and block fixed effects are included in all regressions. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table 4: Summary Statistics of Students

	Local Students			Non-local Students			Difference	
	Count (1)	Mean (2)	SD (3)	Count (4)	Mean (5)	SD (6)	Mean (7)=(2)-(5)	SD (8)
Panel A: Outcome variable:								
Chinese Test Score (Mean=70, SD=10)	7,000	69.97	9.94	1,706	70.18	9.81	-0.209	(0.265)
Math Test Score (Mean=70, SD=10)	6,999	70.03	9.96	1,708	69.88	9.86	0.153	(0.267)
English Test Score (Mean=70, SD=10)	7,001	70.08	9.99	1,704	69.66	9.73	0.427	(0.264)
Feeling blue (1-Never; 5-Always)	6,976	1.97	1.05	1,734	2.02	1.09	-0.051*	(0.029)
Depressed (1-Never; 5-Always)	6,998	2.24	1.00	1,741	2.25	1.00	-0.010	(0.027)
Unhappy (1-Never; 5-Always)	6,984	2.27	1.05	1,745	2.32	1.06	-0.050*	(0.028)
Not enjoying life (1-Never; 5-Always)	6,968	1.73	1.06	1,732	1.84	1.13	-0.116***	(0.030)
Sad (1-Never; 5-Always)	6,988	2.01	1.03	1,739	2.06	1.04	-0.050*	(0.028)
Always late (1-Strongly disagree; 4-Strongly agree)	7,130	1.24	0.61	1,768	1.24	0.63	-0.001	(0.017)
Always skip classes (1-Strongly disagree; 4-Strongly agree)	7,125	1.09	0.43	1,767	1.09	0.41	0.003	(0.011)
Panel B: Predetermined variables:								
Student age	7,041	13.96	1.36	1,740	13.82	1.31	0.134***	(0.035)
Male student	7,177	0.51	0.50	1,778	0.54	0.50	-0.032**	(0.013)
Minority student	7,157	0.12	0.32	1,778	0.08	0.27	0.036***	(0.007)
Student rural residence	7,177	0.43	0.49	1,778	0.59	0.49	-0.164***	(0.013)
Only child in family	7,177	0.54	0.50	1,778	0.38	0.49	0.166***	(0.013)
Attend kindergarten	7,119	0.82	0.38	1,761	0.80	0.40	0.019*	(0.011)
Repeat grade in primary school	7,157	0.10	0.30	1,775	0.16	0.37	-0.059***	(0.009)
Skip grade in primary school	7,159	0.01	0.12	1,774	0.02	0.14	-0.005	(0.004)
Academic ranking in grade 6 (1-Strongly disagree; 4-Strongly agree)	6,500	15.58	11.92	1,596	15.69	11.28	-0.113	(0.319)
Express opinions clearly in grade 6 (1-Strongly disagree; 4-Strongly agree)	6,884	3.16	0.82	1,703	3.06	0.83	0.099***	(0.022)
Respond quickly in grade 6 (1-Strongly disagree; 4-Strongly agree)	6,887	3.04	0.79	1,706	2.97	0.80	0.074***	(0.022)
Learn new stuff quickly in grade 6 (1-Strongly disagree; 4-Strongly agree)	6,835	3.04	0.81	1,693	2.95	0.82	0.095***	(0.022)
Mother's years of schooling	7,160	10.26	3.71	1,772	9.47	3.43	0.785***	(0.093)
Father's years of schooling	7,160	10.88	3.39	1,772	10.30	3.10	0.578***	(0.084)

Notes: This table reports the summary statistics and the difference between the local and non-local students in outcomes and predetermined characteristics. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table 5: Balancing Test

	Chinese	Math	English
	Teacher's Stereotyping of Non-local Students		
	(1)	(2)	(3)
Student age	0.005	0.004	-0.003
	(0.012)	(0.011)	(0.011)
Male student	-0.014	0.012	0.005
	(0.012)	(0.011)	(0.011)
Minority student	0.031	-0.003	-0.005
	(0.041)	(0.029)	(0.038)
Student local residence	-0.061*	0.007	-0.033
	(0.033)	(0.030)	(0.026)
Student rural residence	0.010	0.003	-0.032
	(0.022)	(0.020)	(0.027)
Only child in family	0.057*	-0.050***	0.010
	(0.029)	(0.019)	(0.018)
Attend kindergarten	0.022	0.011	0.009
	(0.018)	(0.018)	(0.029)
Repeat grade in primary school	0.018	-0.026	-0.030
	(0.023)	(0.024)	(0.025)
Skip grade in primary school	0.009	-0.031	-0.135*
	(0.062)	(0.054)	(0.072)
Academic ranking in primary school	0.001	0.001	-0.001
	(0.001)	(0.001)	(0.001)
Express opinions clearly in primary school	0.004	-0.000	-0.006
	(0.014)	(0.013)	(0.013)
Respond quickly in primary school	0.010	0.005	-0.004
	(0.013)	(0.010)	(0.011)
Learn new stuff quickly in primary school	0.004	-0.001	-0.010
	(0.012)	(0.008)	(0.011)
Mother's education	0.004	-0.006*	0.002
	(0.004)	(0.003)	(0.003)
Father's education	0.001	0.004	0.001
	(0.004)	(0.004)	(0.004)
Test for joint significance:			
F-statistics	0.869	1.076	1.137
p-value	0.600	0.388	0.334
Observations	6,785	6,760	6,633

Notes: Each cell presents the coefficient for the listed student predetermined variables from the regression in which the dependent variable is the teacher's stereotyping of non-locals and the independent variables are all of the student predetermined characteristics. Standard errors (in parentheses) are robust and clustered at the block level. Block fixed effects are included in all regressions. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table 6: Random Tests

	Teacher's Stereotyping of Non-local Students											
	All			Chinese			Math			English		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Class composition												
Share of local students	-0.454			-0.055			0.336			-0.919		
	(0.453)			(1.064)			(0.710)			(0.789)		
Observations	534			179			179			176		
Panel B: Class-difference of Local and Non-local Student Predetermined Academic Performance												
Ranking in Grade 6	0.004			-0.011			0.017			0.011		
	(0.011)			(0.023)			(0.019)			(0.020)		
Observations	507			169			169			169		
Panel C: Class-difference of Local and Non-local Total Education Expense												
Ranking in Grade 6	-0.000			-0.000			0.000			-0.000		
	(0.000)			(0.000)			(0.000)			(0.000)		
Observations	481			163			159			159		
Panel D: Class-average Performance at the Beginning of Academic year												
Chinese Performance	0.001			0.068			-0.050			-0.063		
	(0.057)			(0.115)			(0.092)			(0.097)		
Math Performance		0.039			0.104			-0.043			-0.006	
		(0.052)			(0.092)			(0.094)			(0.107)	
English Performance			-0.038			-0.005			-0.046			-0.114
			(0.053)			(0.097)			(0.101)			(0.097)
Observations	557	560	549	188	188	183	186	187	182	183	185	184

Notes: This table reports OLS estimates of the correlation between teachers' stereotype score and class-level predetermined characteristics. Each cell presents the coefficient for the listed class characteristics from the regression in which the dependent variable is the teacher's stereotyping of non-locals and the independent variables are each class-level predetermined characteristics. For columns 4 to 12, the analyses are performed at the teacher level, whereas for columns 1 to 3, the regressions utilize a pooled sample of teachers from the same class. Block fixed effects are included in all regressions. Standard errors (in parentheses) are robust standard errors. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table 7: Effect of Teacher's Stereotyping of Non-local Students on Student Outcomes

	Test Score			Mental Stress			BPI		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stereotype	0.011	0.001	-0.003	0.003	0.003	0.005	-0.002	0.002	0.000
	(0.013)	(0.009)	(0.010)	(0.013)	(0.014)	(0.016)	(0.007)	(0.007)	(0.007)
Stereotype*Non-local	-0.058***	-0.046***	-0.042**	0.024	0.018	0.009	0.023	0.018	0.020
	(0.019)	(0.017)	(0.017)	(0.020)	(0.023)	(0.026)	(0.018)	(0.017)	(0.019)
Non-local	-0.031	0.261	-0.031	0.076**	0.460	0.904	0.010	0.790*	1.026*
	(0.044)	(0.387)	(0.481)	(0.034)	(0.465)	(0.593)	(0.030)	(0.442)	(0.576)
Subject FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Block FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student controls		Yes	Yes		Yes	Yes		Yes	Yes
Student controls*Non-local		Yes	Yes		Yes	Yes		Yes	Yes
Teacher basic controls		Yes	Yes		Yes	Yes		Yes	Yes
Teacher basic controls*Non-local		Yes	Yes		Yes	Yes		Yes	Yes
Teacher detailed controls			Yes			Yes			Yes
Teacher detailed controls*Non-local			Yes			Yes			Yes
Observations	23,490	19,376	18,498	23,369	19,537	18,637	24,015	19,872	18,958

Notes: Test score, mental stress and BPI are constructed according to Table 1. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table 8: Heterogeneous Effects on Student Outcomes-Subject

	Chinese Test Score (1)	Math Test Score (2)	English Test Score (3)	Mental Stress (4)	BPI (5)
Panel A: Chinese Teacher's Stereotype					
Stereotype of Chinese Teacher	0.012	-0.015	0.010	0.055	0.012
	(0.028)	(0.029)	(0.030)	(0.038)	(0.013)
Stereotype of Chinese Teacher*Non-local	-0.099***	-0.054**	-0.100***	-0.014	0.057**
	(0.034)	(0.027)	(0.025)	(0.042)	(0.027)
Non-local	-0.005	0.425	0.508	1.317	0.845
	(0.537)	(0.713)	(0.554)	(0.849)	(0.905)
Observations	6,198	6,197	6,196	6,239	6,345
Panel B: Math Teacher's Stereotype					
Stereotype of Math Teacher	0.063	0.002	0.027	-0.046	-0.075***
	(0.049)	(0.044)	(0.040)	(0.047)	(0.029)
Stereotype of Math Teacher*Non-local	-0.012	0.036	-0.033	0.026	0.019
	(0.032)	(0.040)	(0.034)	(0.040)	(0.038)
Non-local	-0.232	-0.718	0.278	1.670*	1.666*
	(0.858)	(1.052)	(0.814)	(1.003)	(0.874)
Observations	6,180	6,180	6,179	6,213	6,327
Panel C: English Teacher's Stereotype					
Stereotype of English Teacher	-0.019	-0.034	-0.039*	0.008	0.028
	(0.034)	(0.031)	(0.023)	(0.032)	(0.023)
Stereotype of English Teacher*Non-local	-0.033	-0.008	-0.017	0.023	-0.023
	(0.033)	(0.035)	(0.030)	(0.041)	(0.030)
Non-local	0.100	-0.770	1.058	0.316	0.468
	(0.748)	(0.791)	(0.824)	(1.020)	(0.885)
Observations	6,146	6,146	6,149	6,185	6,286
Test for difference in coefficients on Teacher Stereotype*Non-local across subjects:					
p-value	0.081	0.083	0.058	0.623	0.097

Notes: This table reports estimates of the heterogeneous impact of teachers' stereotyping against non-local students on test scores, mental stress and BPI by teacher's teaching subjects. All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table 9: Heterogeneous Effects on Student Outcomes-Dual Roles

	Subject Teacher			Dual Roles		
	Test Score	Mental Stress	BPI	Test Score	Mental Stress	BPI
	(1)	(2)	(3)	(4)	(5)	(6)
Stereotype	-0.020	0.073	0.009	0.138***	0.235***	-0.028
	(0.039)	(0.046)	(0.016)	(0.019)	(0.032)	(0.027)
Stereotype*Non-local	-0.097***	0.032	0.099***	-0.144***	-0.104	-0.034
	(0.036)	(0.048)	(0.033)	(0.047)	(0.101)	(0.053)
Non-local	0.271	1.899**	0.725	1.937	1.382	1.008
	(0.550)	(0.822)	(0.971)	(1.255)	(2.711)	(1.562)
Observations	4,393	4,426	4,490	1,795	1,813	1,855
Test for difference in coefficients on Teacher Stereotype*Non-local between subject and dual-role teachers:						
p-value	0.431	0.202	0.031			

Notes: This table reports estimates of the heterogeneous impact of teachers' stereotyping against non-local students on test scores, mental stress and BPI by teacher's roles. We focus on the Chinese teachers' sample as Table 9 only documents significant effects of Chinese teacher's stereotyping of non-locals on student outcomes. All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table 10: Heterogeneous Effects on Student Outcomes-Student Gender

	Male Students			Female Students		
	Test Score	Mental Stress	BPI	Test Score	Mental Stress	BPI
	(1)	(2)	(3)	(4)	(5)	(6)
Stereotype	0.019	0.068	-0.010	-0.005	0.053	0.028
	(0.034)	(0.046)	(0.017)	(0.032)	(0.043)	(0.022)
Stereotype*Non-local	-0.169***	0.040	0.084*	-0.012	-0.066	0.039
	(0.042)	(0.064)	(0.045)	(0.032)	(0.045)	(0.044)
Non-local	0.358	-0.128	0.911	0.646	2.646**	0.710
	(0.827)	(1.238)	(1.442)	(0.715)	(1.011)	(0.816)
Observations	3,119	3,170	3,221	3,069	3,069	3,124
Test for difference in coefficients on Teacher Stereotype*Non-local between male and female students:						
p-value	0.002	0.132	0.520			

Notes: This table reports estimates of the heterogeneous impact of teachers' stereotyping against non-local students on test scores, mental stress and BPI by students' gender. We focus on the Chinese teachers' sample as Table 9 only documents significant effects of Chinese teacher's stereotyping of non-locals on student outcomes. All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table 11: Heterogeneous Effects on Student Outcomes-Duration

	Non-locals stay $\geqslant$ 8 years			Non-locals stay $>$ 8 years		
	Test Score (1)	Mental Stress (2)	BPI (3)	Test Score (4)	Mental Stress (5)	BPI (6)
Stereotype	0.007	0.056	0.014	0.003	0.052	0.013
	(0.028)	(0.040)	(0.014)	(0.029)	(0.040)	(0.014)
Stereotype*Non-local	-0.050	0.013	0.049	-0.136***	-0.032	0.058
	(0.039)	(0.053)	(0.041)	(0.040)	(0.048)	(0.041)
Non-local	-1.045	3.097**	0.956	1.178*	-0.141	0.326
	(0.870)	(1.206)	(0.721)	(0.671)	(1.203)	(1.313)
Observations	5,575	5,607	5,703	5,586	5,629	5,720
Test for differences in coefficients of Teacher Stereotype*Non-local by duration of non-locals:						
p-value	0.121	0.370	0.881			

Notes: This table reports estimates of the heterogeneous impact of teachers' stereotyping against non-local students on test scores, mental stress and BPI by duration. All regressions include subject fixed effects, block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table 12: Mechanism: Teacher's Behavior

	Teacher's Class Questioning (1)	Teacher's Praise (2)	Frequency of Contacting Parent (3)
Panel A: Chinese Teacher's Stereotype			
Stereotype of Chinese Teacher	-0.004	-0.013	0.022
	(0.014)	(0.014)	(0.029)
Stereotype of Chinese Teacher*Non-local	-0.006	-0.007	-0.055*
	(0.017)	(0.017)	(0.030)
Non-local	0.405	0.827**	-0.871
	(0.370)	(0.328)	(0.670)
Observations	6,338	6,326	6,086
Panel B: Math Teacher's Stereotype			
Stereotype of Math Teacher	0.010	-0.007	-0.059
	(0.022)	(0.018)	(0.051)
Stereotype of Math Teacher*Non-local	0.005	0.013	-0.005
	(0.018)	(0.020)	(0.041)
Non-local	-0.174	0.631	0.002
	(0.450)	(0.515)	(1.010)
Observations	6,309	6,315	6,064
Panel C: English Teacher's Stereotype			
Stereotype of English Teacher	0.011	0.021	-0.000
	(0.011)	(0.016)	(0.031)
Stereotype of English Teacher*Non-local	-0.000	0.003	-0.002
	(0.015)	(0.018)	(0.031)
Non-local	1.164***	1.272**	-0.070
	(0.410)	(0.529)	(0.964)
Observations	6,275	6,278	6,004

Notes: All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table 13: Mechanism: Parent Response

	Parent Feels that		Frequency of
	Teachers Are Prejudiced against		Contacting
	Nonlocal Students	Nonlocal Parents	Teacher
	(1)	(2)	(3)
Panel A: Chinese Teacher's Stereotype			
Stereotype of Chinese Teacher	0.005	-0.004	0.002
	(0.021)	(0.016)	(0.028)
Stereotype of Chinese Teacher*Non-local	0.017	0.020	-0.028
	(0.040)	(0.036)	(0.031)
Non-local	1.508	0.749	0.026
	(1.163)	(0.961)	(0.799)
Observations	5,861	5,858	6,120
Panel B: Math Teacher's Stereotype			
Stereotype of Math Teacher	0.045	0.045	-0.053
	(0.036)	(0.027)	(0.035)
Stereotype of Math Teacher*Non-local	0.039	-0.038	-0.002
	(0.047)	(0.052)	(0.028)
Non-local	-0.615	0.424	-0.977
	(1.120)	(1.131)	(0.956)
Observations	5,824	5,821	6,104
Panel C: English Teacher's Stereotype			
Stereotype of English Teacher	0.006	0.016	0.006
	(0.024)	(0.023)	(0.022)
Stereotype of English Teacher*Non-local	0.001	-0.050	-0.052*
	(0.049)	(0.043)	(0.030)
Non-local	0.971	0.794	-0.046
	(1.277)	(1.030)	(0.715)
Observations	5,771	5,768	6,054

Notes: All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table 14: Mechanism: Class Environment and Interaction with Classmates

	Classmates Are Friendly to Me (1)	Having Nonlocal Friend (2)	Parents Contact Teachers for Friend Issues (3)
Panel A: Chinese Teacher's Stereotype			
Stereotype of Chinese Teacher	0.001	0.001	-0.004
	(0.006)	(0.013)	(0.009)
Stereotype of Chinese Teacher*Non-local	-0.019**	0.026*	0.042***
	(0.009)	(0.015)	(0.014)
Non-local	0.530**	0.952**	-0.709*
	(0.263)	(0.364)	(0.375)
Observations	6,322	5,202	6,087
Panel B: Math Teacher's Stereotype			
Stereotype of Math Teacher	0.015*	0.033**	-0.016
	(0.008)	(0.016)	(0.015)
Stereotype of Math Teacher*Non-local	0.002	-0.018	0.014
	(0.011)	(0.019)	(0.014)
Non-local	0.296	0.976**	-0.763**
	(0.272)	(0.475)	(0.365)
Observations	6,306	5,209	6,071
Panel C: English Teacher's Stereotype			
Stereotype of English Teacher	-0.007	-0.011	-0.012
	(0.006)	(0.020)	(0.008)
Stereotype of English Teacher*Non-local	-0.000	0.026	0.009
	(0.013)	(0.019)	(0.014)
Non-local	0.174	0.463	-0.174
	(0.301)	(0.473)	(0.471)
Observations	6,267	5,184	6,010

Notes: All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level.*** = p<0.01, ** = p<0.05, and * = p<0.1.

Table 15: Mechanism: Student Response

	Self-Evaluated	Hours Spent per Week on			
	Rank in Class	Schoolwork within School	Schoolwork outside School	Playing	Other Activities
	(1)	(2)	(3)	(4)	(5)
Panel A: Chinese Teacher's Stereotype					
Stereotype of Chinese Teacher	0.023	0.007	-0.007	0.024	0.002
	(0.017)	(0.026)	(0.020)	(0.026)	(0.022)
Stereotype of Chinese Teacher*Non-local	-0.069**	0.016	0.022	0.065**	-0.014
	(0.031)	(0.032)	(0.030)	(0.027)	(0.032)
Non-local	1.213	0.399	-0.015	0.249	0.242
	(0.835)	(0.565)	(0.679)	(0.583)	(0.688)
Observations	6,353	6,300	6,355	6,358	6,350
Panel B: Math Teacher's Stereotype					
Stereotype of Math Teacher	-0.006	-0.020	-0.029	-0.001	0.043
	(0.020)	(0.032)	(0.033)	(0.037)	(0.030)
Stereotype of Math Teacher*Non-local	0.036	0.050	0.113***	0.041	0.012
	(0.033)	(0.030)	(0.032)	(0.036)	(0.033)
Non-local	0.795	0.101	0.103	0.826	-0.061
	(0.758)	(1.059)	(0.944)	(0.750)	(0.807)
Observations	6,334	6,282	6,340	6,341	6,333
Panel C: English Teacher's Stereotype					
Stereotype of English Teacher	-0.041**	-0.000	-0.010	0.030	0.068**
	(0.019)	(0.030)	(0.028)	(0.028)	(0.029)
Stereotype of English Teacher*Non-local	0.037	0.075**	-0.007	0.013	0.030
	(0.034)	(0.034)	(0.028)	(0.033)	(0.033)
Non-local	1.939**	0.623	-0.009	-1.340	-0.268
	(0.829)	(0.813)	(0.813)	(1.134)	(0.953)
Observations	6,291	6,244	6,297	6,298	6,289

Notes: All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A1: Heterogeneous Effects on Student Outcomes-Class Composition

	More Non-Local Students			Less Non-Local Students		
	Test Score	Mental Stress	BPI	Test Score	Mental Stress	BPI
	(1)	(2)	(3)	(4)	(5)	(6)
Stereotype	-0.012	-0.006	-0.014	0.059	0.092**	0.029
	(0.042)	(0.060)	(0.028)	(0.047)	(0.038)	(0.020)
Stereotype*Non-local	-0.086**	-0.066	0.069	-0.093***	-0.006	0.060*
	(0.042)	(0.086)	(0.053)	(0.030)	(0.046)	(0.031)
Non-local	-0.060	5.556**	3.461**	0.816	0.939	-0.146
	(1.357)	(2.089)	(1.405)	(0.822)	(1.026)	(1.128)
Observations	3,197	3,224	3,283	2,991	3,015	3,062
Test for differences in coefficients of Teacher Stereotype*Non-local by class composition:						
p-value	0.876	0.527	0.881			

Notes: This table reports estimates of the heterogeneous impact of teachers' stereotyping against non-local students on test scores, mental stress and BPI by class composition. Classes with more non-local students are defined as those exceeding the median share, while classes with less non-local students are those below the median share. All regressions include subject fixed effects, block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A2: Effect of Teacher's Stereotyping of Non-local Students on Student Outcomes, Controlling for Classroom Fixed Effect

	Chinese Test Score (1)	Math Test Score (2)	English Test Score (3)	Mental Stress (4)	BPI (5)
Panel A: Chinese Teacher's Stereotype					
Stereotype of Chinese Teacher*Non-local	-0.098***	-0.062*	-0.108***	-0.008	0.058**
	(0.031)	(0.031)	(0.026)	(0.041)	(0.028)
Non-local	0.211	0.462	0.565	1.353	0.795
	(0.632)	(0.761)	(0.555)	(0.899)	(0.894)
Observations	6,198	6,197	6,196	6,239	6,345
Panel B: Math Teacher's Stereotype					
Stereotype of Math Teacher*Non-local	-0.028	0.031	-0.038	0.023	0.026
	(0.033)	(0.038)	(0.037)	(0.040)	(0.037)
Non-local	-0.573	-0.739	0.247	1.814*	1.407
	(0.803)	(0.981)	(0.807)	(1.081)	(0.902)
Observations	6,180	6,180	6,179	6,213	6,327
Panel C: English Teacher's Stereotype					
Stereotype of English Teacher*Non-local	-0.009	-0.021	-0.017	0.001	-0.033
	(0.033)	(0.034)	(0.030)	(0.046)	(0.032)
Non-local	0.091	-1.178	0.864	-0.086	-0.117
	(0.770)	(0.933)	(0.843)	(1.104)	(0.880)
Observations	6,146	6,146	6,149	6,185	6,286

Notes: All regressions include classroom fixed effects, student controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the classroom level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A3: Effect of Teacher's Stereotyping of Non-local Students on Noncognitive Outcomes, Controlling for Test Score

	Mental Stress (1)	BPI (2)
Panel A: Chinese Teacher's Stereotype		
Stereotype of Chinese Teacher	0.054	0.012
	(0.039)	(0.013)
Stereotype of Chinese Teacher*Non-local	-0.026	0.056**
	(0.042)	(0.028)
Non-local	1.722*	0.947
	(0.937)	(0.928)
Observations	6,074	6,172
Panel B: Math Teacher's Stereotype		
Stereotype of Math Teacher	-0.041	-0.077***
	(0.046)	(0.028)
Stereotype of Math Teacher*Non-local	0.024	0.022
	(0.040)	(0.040)
Non-local	1.969*	1.718*
	(0.999)	(0.940)
Observations	6,048	6,153
Panel C: English Teacher's Stereotype		
Stereotype of English Teacher	0.011	0.036
	(0.033)	(0.022)
Stereotype of English Teacher*Non-local	0.014	-0.030
	(0.042)	(0.030)
Non-local	0.517	0.549
	(1.078)	(0.908)
Observations	6,030	6,124

Notes: All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A4: Sample Attrition

	Chinese Test Score (1)	Math Test Score (2)	English Test Score (3)	Mental Stress (4)	BPI (5)
Panel A: Chinese Teacher's Stereotype					
Stereotype of Chinese Teacher	-0.002	-0.002	-0.002	-0.003	0.001
	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)
Observations	8,112	8,112	8,112	8,112	8,112
Panel B: Math Teacher's Stereotype					
Stereotype of Math Teacher	0.000	0.000	0.000	-0.000	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)
Observations	8,113	8,113	8,113	8,113	8,113
Panel C: English Teacher's Stereotype					
Stereotype of English Teacher	0.005	0.004	0.005	0.003	0.001
	(0.004)	(0.004)	(0.004)	(0.005)	(0.002)
Observations	7,968	7,968	7,968	7,968	7,968

Notes: Each cell represents a separate regression in which the dependent variable is a dummy variable indicating a missing value for the respective survey item, and the independent variable is teacher's stereotyping of non-locals and the non-local student indicator. Standard errors (in parentheses) are robust and clustered at the block level. Block fixed effects are included in all regressions. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A5: Effect of Teacher's Stereotyping of Non-local Students on Student Outcomes, Using Nonrandom Sample

	Chinese Test Score (1)	Math Test Score (2)	English Test Score (3)	Mental Stress (4)	BPI (5)
Panel A: Chinese Teacher's Stereotype					
Stereotype of Chinese Teacher	-0.032	-0.046	-0.010	0.044	-0.038
	(0.053)	(0.047)	(0.048)	(0.031)	(0.035)
Stereotype of Chinese Teacher*Non-local	-0.059	0.051	0.029	-0.041	-0.015
	(0.037)	(0.036)	(0.037)	(0.033)	(0.029)
Non-local	-0.435	-0.900	-0.661	-0.215	-0.737
	(0.742)	(0.875)	(0.851)	(0.839)	(0.642)
Observations	6,894	6,882	6,889	6,887	7,004
Panel B: Math Teacher's Stereotype					
Stereotype of Math Teacher	0.088**	0.035	0.084*	-0.009	-0.046*
	(0.036)	(0.037)	(0.045)	(0.027)	(0.027)
Stereotype of Math Teacher*Non-local	-0.010	-0.012	0.003	-0.028	-0.003
	(0.033)	(0.033)	(0.029)	(0.026)	(0.035)
Non-local	-2.686**	-2.335*	-1.999*	3.176***	-0.353
	(1.150)	(1.289)	(1.196)	(1.061)	(1.043)
Observations	6,634	6,624	6,629	6,644	6,735
Panel C: English Teacher's Stereotype					
Stereotype of English Teacher	0.132**	0.120**	0.143**	-0.004	-0.085**
	(0.056)	(0.048)	(0.057)	(0.031)	(0.036)
Stereotype of English Teacher*Non-local	-0.045	-0.033	-0.013	-0.043	0.001
	(0.043)	(0.043)	(0.040)	(0.038)	(0.034)
Non-local	-0.644	-2.228**	-0.001	0.716	-0.342
	(0.851)	(0.867)	(0.885)	(1.082)	(0.726)
Observations	7,035	7,022	7,027	7,018	7,131

Notes: The regression sample consists of blocks that fail our randomization restriction (i.e., principals or teachers report nonrandom assignment). All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A6: Effect of Teacher's Stereotyping of Non-local Students on Student Outcomes, Excluding Progressive Teachers

	Chinese Test Score (1)	Math Test Score (2)	English Test Score (3)	Mental Stress (4)	BPI (5)
Panel A: Chinese Teacher's Stereotype					
Stereotype of Chinese Teacher	-0.004	-0.055*	-0.028	0.040	-0.002
	(0.035)	(0.029)	(0.028)	(0.046)	(0.014)
Stereotype of Chinese Teacher*Non-local	-0.084*	-0.057**	-0.096***	-0.019	0.100***
	(0.043)	(0.028)	(0.034)	(0.057)	(0.033)
Non-local	0.025	0.025	0.351	1.264	0.375
	(0.513)	(0.834)	(0.622)	(0.921)	(1.098)
Observations	5,303	5,302	5,300	5,340	5,426
Panel B: Math Teacher's Stereotype					
Stereotype of Math Teacher	0.109*	0.050	0.101*	-0.034	-0.074***
	(0.063)	(0.068)	(0.056)	(0.068)	(0.027)
Stereotype of Math Teacher*Non-local	-0.010	-0.038	-0.014	0.015	0.020
	(0.042)	(0.047)	(0.037)	(0.059)	(0.057)
Non-local	-0.513	-1.653	-0.480	0.636	1.727
	(0.952)	(1.140)	(0.822)	(1.254)	(1.041)
Observations	4,904	4,905	4,903	4,932	5,027
Panel C: English Teacher's Stereotype					
Stereotype of English Teacher	0.070*	-0.066	-0.019	-0.005	0.016
	(0.041)	(0.042)	(0.034)	(0.042)	(0.037)
Stereotype of English Teacher*Non-local	-0.017	-0.012	-0.007	0.011	-0.040
	(0.049)	(0.044)	(0.048)	(0.057)	(0.038)
Non-local	-0.664	-1.161	1.016	-0.110	0.739
	(0.897)	(1.097)	(1.168)	(1.357)	(1.066)
Observations	4,947	4,948	4,948	4,970	5,045

Notes: All regressions exclude teachers with stereotype scores lower than the 10th percentile within each subject and include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A7: Effect of Teacher's Stereotyping of Non-local Students on Student Outcomes, Dummy Measurement of Teacher Stereotype

	Chinese Test Score (1)	Math Test Score (2)	English Test Score (3)	Mental Stress (4)	BPI (5)
Panel A: Chinese Teacher's Stereotype					
Stereotype of Chinese Teacher	0.087**	0.018	0.046	0.149**	-0.014
	(0.043)	(0.056)	(0.053)	(0.067)	(0.030)
Stereotype of Chinese Teacher*Non-local	-0.199***	-0.110*	-0.165***	-0.040	0.151**
	(0.068)	(0.064)	(0.054)	(0.092)	(0.061)
Non-local	-0.118	0.357	0.367	1.320	0.896
	(0.575)	(0.694)	(0.579)	(0.868)	(0.901)
Observations	6,198	6,197	6,196	6,239	6,345
Panel B: Math Teacher's Stereotype					
Stereotype of Math Teacher	0.088	0.034	0.046	-0.074	-0.144***
	(0.080)	(0.074)	(0.065)	(0.085)	(0.044)
Stereotype of Math Teacher*Non-local	-0.051	0.020	-0.060	0.008	0.071
	(0.070)	(0.083)	(0.069)	(0.089)	(0.071)
Non-local	-0.263	-0.757	0.307	1.707	1.714**
	(0.856)	(1.062)	(0.825)	(1.033)	(0.862)
Observations	6,180	6,180	6,179	6,213	6,327
Panel C: English Teacher's Stereotype					
Stereotype of English Teacher	0.050	-0.044	-0.033	-0.017	0.005
	(0.062)	(0.052)	(0.045)	(0.065)	(0.037)
Stereotype of English Teacher*Non-local	-0.013	0.011	0.030	-0.040	-0.049
	(0.072)	(0.068)	(0.060)	(0.074)	(0.059)
Non-local	0.209	-0.724	1.138	0.235	0.480
	(0.745)	(0.803)	(0.826)	(0.989)	(0.877)
Observations	6,146	6,146	6,149	6,185	6,286

Notes: The alternative dummy measurement of a teacher's stereotype against non-locals equals one if the teacher believes that classes with all local students outperform those with one-third non-local students in at least two of three areas: teaching effectiveness, class discipline, and student relations; 0 otherwise. Alternative measurement of teacher's stereptype against non-locals is constructed by (2a), (2b), (3a) and (3b) in Table 1. All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A8: Effect of Teacher's Stereotyping of Non-local Students on Student Outcomes, Alternative Measurement of Teacher Stereotype

	Chinese Test Score (1)	Math Test Score (2)	English Test Score (3)	Mental Stress (4)	BPI (5)
Panel A: Chinese Teacher's Stereotype					
Stereotype of Chinese Teacher	-0.004	-0.029	-0.002	0.044	0.013
	(0.029)	(0.028)	(0.029)	(0.037)	(0.012)
Stereotype of Chinese Teacher*Non-local	-0.098***	-0.068**	-0.103***	-0.007	0.059**
	(0.034)	(0.027)	(0.024)	(0.042)	(0.027)
Non-local	0.018	0.484	0.540	1.307	0.828
	(0.543)	(0.701)	(0.553)	(0.849)	(0.906)
Observations	6,198	6,197	6,196	6,239	6,345
Panel B: Math Teacher's Stereotype					
Stereotype of Math Teacher	0.084*	0.018	0.028	-0.040	-0.048*
	(0.048)	(0.041)	(0.037)	(0.038)	(0.026)
Stereotype of Math Teacher*Non-local	-0.027	0.017	-0.051	0.026	-0.010
	(0.034)	(0.040)	(0.032)	(0.040)	(0.042)
Non-local	-0.287	-0.736	0.275	1.698*	1.719*
	(0.848)	(1.053)	(0.809)	(1.010)	(0.885)
Observations	6,180	6,180	6,179	6,213	6,327
Panel C: English Teacher's Stereotype					
Stereotype of English Teacher	-0.012	-0.033	-0.033	0.013	0.035
	(0.033)	(0.032)	(0.024)	(0.032)	(0.023)
Stereotype of English Teacher*Non-local	-0.035	0.013	-0.017	0.025	-0.035
	(0.033)	(0.034)	(0.031)	(0.043)	(0.030)
Non-local	0.116	-0.696	1.091	0.314	0.425
	(0.752)	(0.795)	(0.828)	(1.023)	(0.876)
Observations	6,146	6,146	6,149	6,185	6,286

Notes: Alternative measurement of teacher's stereptype against non-locals is constructed by (2a), (2b), (3a) and (3b) in Table 1. All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A9: Effect of Teacher's Stereotyping of Non-local Students on Approximation of Baseline Outcomes

	Academic Ranking in Grade 6 (1)	Express Opinions Clearly in Grade 6 (2)	Respond Quickly in Grade 6 (3)	Learn New Stuff Quickly in Grade 6 (4)
Panel A: Chinese Teacher's Stereotype				
Stereotype of Chinese Teacher	0.029	0.026	0.024	0.038*
	(0.022)	(0.022)	(0.023)	(0.022)
Stereotype of Chinese Teacher*Non-local	-0.018	-0.053	-0.037	-0.046
	(0.025)	(0.038)	(0.031)	(0.034)
Non-local	-0.694	-0.526	0.121	0.593
	(0.699)	(0.647)	(0.611)	(0.677)
Observations	6,746	7,150	7,151	7,099
Panel B: Math Teacher's Stereotype				
Stereotype of Math Teacher	0.003	-0.008	-0.015	-0.015
	(0.032)	(0.040)	(0.028)	(0.030)
Stereotype of Math Teacher*Non-local	0.005	-0.016	0.009	-0.037
	(0.030)	(0.045)	(0.034)	(0.039)
Non-local	-0.139	0.451	-0.852	0.386
	(0.950)	(0.914)	(0.673)	(0.779)
Observations	6,730	7,147	7,150	7,095
Panel C: English Teacher's Stereotype				
Stereotype of English Teacher	-0.024	-0.008	-0.006	-0.005
	(0.020)	(0.033)	(0.025)	(0.026)
Stereotype of English Teacher*Non-local	-0.009	0.003	-0.008	-0.031
	(0.028)	(0.036)	(0.029)	(0.031)
Non-local	0.275	-0.267	-0.298	0.381
	(0.936)	(1.028)	(0.928)	(1.093)
Observations	6,676	7,088	7,090	7,038

Notes: All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A10: Effect of Experienced Teacher's Stereotyping of Non-local Students on Student Outcomes

	Chinese Test Score (1)	Math Test Score (2)	English Test Score (3)	Mental Stress (4)	BPI (5)
Panel A: Chinese Teacher's Stereotype					
Stereotype of Chinese Teacher	0.007	-0.033	-0.002	0.042	-0.002
	(0.045)	(0.047)	(0.034)	(0.037)	(0.016)
Stereotype of Chinese Teacher*Non-local	-0.121***	-0.082**	-0.115***	-0.044	0.050*
	(0.038)	(0.032)	(0.036)	(0.044)	(0.025)
Non-local	0.901	1.226*	0.103	3.256***	1.455**
	(0.592)	(0.629)	(0.777)	(1.127)	(0.659)
Observations	4,619	4,618	4,619	4,662	4,719
Panel B: Math Teacher's Stereotype					
Stereotype of Math Teacher	0.015	-0.006	0.010	-0.056	-0.061*
	(0.058)	(0.039)	(0.053)	(0.049)	(0.037)
Stereotype of Math Teacher*Non-local	-0.036	-0.007	-0.065*	0.042	0.012
	(0.036)	(0.040)	(0.038)	(0.046)	(0.037)
Non-local	-0.161	-0.458	0.528	1.883*	2.665***
	(0.953)	(1.115)	(0.877)	(1.107)	(0.942)
Observations	4,992	4,993	4,991	5,035	5,123
Panel C: English Teacher's Stereotype					
Stereotype of English Teacher	-0.002	-0.066*	-0.037	-0.015	0.029
	(0.030)	(0.035)	(0.027)	(0.035)	(0.029)
Stereotype of English Teacher*Non-local	-0.026	0.017	-0.027	0.062	0.013
	(0.043)	(0.042)	(0.036)	(0.059)	(0.029)
Non-local	1.584	0.380	1.637	0.010	1.538
	(1.259)	(1.229)	(1.254)	(1.780)	(1.171)
Observations	4,435	4,434	4,437	4,487	4,556

Notes: All regressions are restricted to a sample of teachers with more than ten-year teaching experience and include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A11: Effect of Teacher's Stereotyping of Non-local Students on Student Outcomes, for Teachers Teaching More Than One Class

	Chinese Test Score (1)	Math Test Score (2)	English Test Score (3)	Mental Stress (4)	BPI (5)
Panel A: Chinese Teacher's Stereotype					
Stereotype of Chinese Teacher	0.003	-0.025	-0.000	0.062	0.009
	(0.032)	(0.034)	(0.035)	(0.049)	(0.015)
Stereotype of Chinese Teacher*Non-local	-0.104***	-0.045	-0.088***	0.001	0.063**
	(0.037)	(0.030)	(0.025)	(0.048)	(0.030)
Non-local	0.283	0.507	0.699	1.222	0.800
	(0.506)	(0.739)	(0.573)	(0.860)	(0.945)
Observations	5,403	5,401	5,400	5,441	5,532
Panel B: Math Teacher's Stereotype					
Stereotype of Math Teacher	0.099*	0.017	0.039	-0.070	-0.073*
	(0.052)	(0.052)	(0.054)	(0.052)	(0.038)
Stereotype of Math Teacher*Non-local	-0.032	0.022	-0.040	0.047	0.020
	(0.035)	(0.043)	(0.037)	(0.042)	(0.037)
Non-local	-1.085	-0.866	-0.149	1.100	2.373**
	(0.891)	(1.132)	(0.851)	(1.080)	(0.911)
Observations	5,009	5,009	5,006	5,032	5,136
Panel C: English Teacher's Stereotype					
Stereotype of English Teacher	-0.010	-0.046	-0.052**	0.001	0.036
	(0.036)	(0.033)	(0.025)	(0.036)	(0.027)
Stereotype of English Teacher*Non-local	-0.030	-0.029	-0.023	0.027	-0.012
	(0.035)	(0.037)	(0.034)	(0.047)	(0.031)
Non-local	-0.119	-1.328	1.214	0.510	0.597
	(0.862)	(0.832)	(0.931)	(1.137)	(0.989)
Observations	5,090	5,090	5,092	5,118	5,204

Notes: All regressions focus on the sample of teachers teaching more than one class and include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, gender, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A12: Heterogeneous Effects by Gender: Teacher's Behavior

	Teacher's Class Questioning (1)	Teacher's Praise (2)	Frequency of Contacting Parent (3)
Panel A: Male Students			
Stereotype	-0.008	-0.022	-0.024
	(0.018)	(0.016)	(0.038)
Stereotype*Non-local	-0.024	-0.017	-0.021
	(0.022)	(0.023)	(0.037)
Non-local	0.452	1.492***	-3.023***
	(0.445)	(0.473)	(1.048)
Observations	3,215	3,209	3,055
Panel B: Female Students			
Stereotype	-0.005	-0.006	0.077**
	(0.017)	(0.018)	(0.032)
Stereotype*Non-local	0.025	0.003	-0.071
	(0.022)	(0.025)	(0.043)
Non-local	0.486	0.204	1.077
	(0.536)	(0.487)	(1.033)
Observations	3,123	3,117	3,031
Test for difference of coefficients on Stereotype*Non-local between boys and girls:			
p-value	0.057	0.537	0.363

Notes: This table reports estimates of the heterogeneous impact of teachers' stereotyping against non-local students on teacher's behavior by students' gender. We focus on the Chinese teachers' sample as Table 8 only documents significant effects of Chinese teacher's stereotyping of non-locals on student outcomes. All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A13: Heterogeneous Effects by Gender: Parent Response

	Parent Feels that		Frequency of
	Teachers Are Prejudiced against		Contacting
	Nonlocal Students	Nonlocal Parents	Teacher
	(1)	(2)	(3)
Panel A: Male Students			
Stereotype	0.000	-0.019	-0.018
	(0.028)	(0.025)	(0.039)
Stereotype*Non-local	0.047	0.035	0.019
	(0.045)	(0.041)	(0.038)
Non-local	1.448	0.191	-1.424
	(1.245)	(1.152)	(1.283)
Observations	2,943	2,940	3,068
Panel B: Female Students			
Stereotype	0.012	0.003	0.030
	(0.022)	(0.020)	(0.036)
Stereotype*Non-local	-0.004	0.040	-0.071*
	(0.052)	(0.044)	(0.041)
Non-local	2.461	1.631	0.923
	(1.757)	(1.319)	(1.100)
Observations	2,918	2,918	3,052
Test for difference of coefficients on Stereotype*Non-local between boys and girls:			
p-value	0.371	0.921	0.087

Notes: This table reports estimates of the heterogeneous impact of teachers' stereotyping against non-local students on parent response by students' gender. We focus on the Chinese teachers' sample as Table 8 only documents significant effects of Chinese teacher's stereotyping of non-locals on student outcomes. All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A14: Heterogeneous Effects by Gender: Class Environment

	Classmates Are Friendly to Me (1)	Having Nonlocal Friend (2)	Parents Contact Teachers for Friend Issues (3)
Panel A: Male Students			
Stereotype	0.006	0.009	-0.004
	(0.011)	(0.018)	(0.013)
Stereotype*Non-local	-0.042***	0.040	0.041**
	(0.016)	(0.027)	(0.020)
Non-local	0.929**	0.962*	-0.552
	(0.428)	(0.522)	(0.427)
Observations	3,205	2,621	3,053
Panel B: Female Students			
Stereotype	-0.004	-0.013	-0.005
	(0.011)	(0.017)	(0.012)
Stereotype *Non-local	0.008	0.015	0.046**
	(0.012)	(0.023)	(0.020)
Non-local	0.160	1.371**	-1.005*
	(0.387)	(0.583)	(0.536)
Observations	3,117	2,581	3,034
Test for difference of coefficients on Stereotype*Non-local between boys and girls:			
p-value	0.014	0.524	0.849

Notes: This table reports estimates of the heterogeneous impact of teachers' stereotyping against non-local students on class environment by students' gender. We focus on the Chinese teachers' sample as Table 8 only documents significant effects of Chinese teacher's stereotyping of non-locals on student outcomes. All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.

Table A15: Heterogeneous Effects by Gender: Student Response

	Self-Evaluated	Hours Spent per Week on			
	Rank in Class	Schoolwork within School	Schoolwork outside School	Playing	Other Activities
	(1)	(2)	(3)	(4)	(5)
Panel A: Male Students					
Stereotype	0.036	0.009	0.017	0.037	0.046
	(0.024)	(0.029)	(0.023)	(0.040)	(0.028)
Stereotype*Non-local	-0.123**	-0.049	0.006	0.112**	-0.017
	(0.051)	(0.038)	(0.046)	(0.046)	(0.055)
Non-local	1.559	0.515	-1.277	0.435	-0.236
	(1.171)	(0.984)	(0.999)	(0.939)	(1.093)
Observations	3,222	3,195	3,223	3,225	3,224
Panel B: Female Students					
Stereotype	0.016	0.008	-0.027	-0.001	-0.041
	(0.023)	(0.032)	(0.026)	(0.029)	(0.032)
Stereotype*Non-local	-0.017	0.092**	0.034	0.039	-0.006
	(0.035)	(0.046)	(0.038)	(0.040)	(0.051)
Non-local	0.857	0.996	0.854	0.149	-0.718
	(0.905)	(0.899)	(0.933)	(0.898)	(1.189)
Observations	3,131	3,105	3,132	3,133	3,126
Test for difference of coefficients on Stereotype*Non-local between boys and girls:					
p-value	0.090	0.003	0.651	0.241	0.893

Notes: This table reports estimates of the heterogeneous impact of teachers' stereotyping against non-local students on student response by students' gender. We focus on the Chinese teachers' sample as Table 8 only documents significant effects of Chinese teacher's stereotyping of non-locals on student outcomes. All regressions include block fixed effects, student controls, teacher basic controls, teacher detailed controls and interactions between the non-local student indicator and all student and teacher controls. Student controls include student age, baseline cognitive measurements, baseline noncognitive measurements, mother's education, father's education, and indicators showing minority, local residence, rural residence and only child in the family. Teacher basic controls include age, gender and marital status. Teacher detailed controls include years of schooling, experience, and indicators showing tenure status, with professional job title, whether the teacher graduated from a normal college, and whether the teacher had prior experience in other schools. Standard errors (in parentheses) are robust and clustered at the block level. *** = p<0.01, ** = p<0.05, and * = p<0.1.