I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17135

Can the Teaching Style Reduce Inequality in the Classroom? Evidence from a Quasi-Experiment

Lei Xu
Massimiliano Tani
Yu Zhu

JULY 2024

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17135

Can the Teaching Style Reduce Inequality in the Classroom? Evidence from a Quasi-Experiment

Lei Xu
Loughborough Business School, IZA and GLO

Massimiliano Tani
University of New South Wales, IZA and GLO

Yu Zhu
University of Dundee School of Business, IZA and GLO

JULY 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Can the Teaching Style Reduce Inequality in the Classroom? Evidence from a Quasi-Experiment*

We investigate the effects of 'lecture-based' (LBT) – i.e. individual work and rote learning - versus 'discussion-based' (DBT) – i.e. participative and focused on student-centred learning - teaching styles on the test scores and socio-economic inequality of middle-school students randomly assigned to classes using data from the China Education Panel Survey (CEPS) - a large-scale nationally representative survey. Estimates from Unconditional Quantile Regressions and decompositions based on the Recentered Influence Function suggest that LBT raises scores in mathematics, but the effect is non-linear, as students in the bottom and top quintiles are more likely to benefit from it. In contrast, LBT lowers scores in Chinese and English. LBT also has greater influence on socio-economically advantaged students, resulting in larger inequality within classrooms, especially between top and median students. These effects arise under various robustness checks, implying that: (i) teaching styles affect scores and classroom inequality, and (ii) they appear to be subject-specific. These results suggest that teaching styles can be used as a tool to influence students' academic performance as well as the socio-economic heterogeneity that they bring to their classrooms.

JEL Classification: I21, I24

Keywords: teaching style, achievement inequality, random class assignment, China

Corresponding author:
Yu Zhu
University of Dundee School of Business
3 Perth Road
Dundee, DD1 4HN
Great Britain

E-mail: yuzhu@dundee.ac.uk.

* We are grateful to comments and suggestions by participants of the 8th LEER Conference in Leuven, Belgium. We have also benefited from suggestions by Ian Walker.

1. Introduction

A common experience among college students and academics at the outset of their teaching career is exposure to teaching styles based on 'student-centred learning': namely a participatory environment where teachers skilfully present an authentic task through which relevant, memorable student learning is achieved (Brandes and Ginnis, 1996; Lea et al, 2003; Attard et al, 2010; Beaten et al, 2010). This *discussion-based teaching (DBT)* style differs from the more traditional *lecture-based teaching (LBT)* approach by teachers delivering information verbally to a largely passive student audience (Aitkin et al, 1981; Jarvis, 2006; Michel et al, 2009). Note that LBT and DBT are sometimes referred to as "traditional" and "modern" teaching styles respectively (see e.g. Schwerdt and Wuppermann 2011). However, these alternative terms could be controversial and misleading, as the DBT approach is sometimes referenced back to Socrates.

While there is a strong preference for DBT because of its longer-lasting effects on learning (e.g. Felder and Silverman, 1988; Doherty, 1995; McCarthy and Anderson, 2000), empirical evidence on whether the *LBT style* (Van Klaveren 2011; Schwerdt and Wuppermann 2011) or the *DBT style* (Freire 1998; Bientenbeck 2014; Lavy 2016; Hildalgo-Cabrillana and Lopez-Mayan 2018) is preferable is inconclusive. In addition, the analyses overwhelmingly focus on higher education, with much less evidence on the use and effectiveness of LBT versus DBT in earlier education like during mandatory schooling.

In this paper, we aim to fill this knowledge gap by contributing new evidence about the effects of LBT and DBT styles on the academic performance of students attending mandatory middle (lower secondary) school. In particular, we address the following questions:

1. do DBT and LBT style influence students' performance, if at all, as measured by their test scores?
2. do they uniformly influence students of various ability and socio-economic status?
3. do they produce similar effects across subjects such as mathematics and language?
4. can teaching styles be strategically used to reduce the inequality arising from students' heterogeneous socio-economic backgrounds and access to resources (e.g. teaching support, published material and technology, extra-curricular activities)?

We find that teaching styles have significant impacts on test scores, and these impacts vary by students' socio-economic status (SES) and prior achievement.

Our research aims to disentangle the effect of teaching styles from that of teacher characteristics on student test scores, complementing existing research, which has mainly focused on documenting the association between students' academic achievements and various teachers' characteristics, such as their quality, gender, experience and ability (Gong et al 2018; Huang et al 2021). For instance, Mansfield (2015) finds that teachers' allocation could account for 3% of the high school performance gap between the top and bottom deciles. Using administrative data in Chicago, Aaronson et al (2007) find significant relationship between teacher quality and student math scores. Kane et al (2011) find positive correlation between the observed measures of teaching effectiveness and achievement growth. This approach however implicitly assumes that it is teachers' characteristics (e.g. allocation, education, and gender) that generate the relevant observed outcome regardless of how the action of teaching is performed – an unlikely tenet.

By separating teachers' characteristics from teaching styles, our analysis also addresses a question of rising prominence for policy-makers: the growing socio-economic inequality in classrooms (as partly the result of the internationalisation of intermediate and higher education in several OECD countries), and what deliberate efforts can be implemented to make secondary and university education less ‘elitist’. ‘Between-school’ variation in student outcomes is prevalent in many countries due to selective admissions or residential segregation. Its alleviation through policy interventions is challenging due to the complex nature of the problem and the contrasting interests of the stakeholders. In contrast, understanding the heterogeneous impact of teaching styles within a classroom might hold a key for cost-effective interventions that can be managed within each school.

In trying to estimate the influence of different teaching styles, we face two formidable challenges in the identification of the relevant effects, which arise from (i) the non-random allocation of teachers, students, and schools (e.g. Clotfelter, 2010) and (ii) the endogenous choice that teachers make about their teaching style (Rivkin, et al, 2005; Schwerdt and Wuppermann 2011). Both contribute to bias the estimates obtained from regression analysis (endogeneity bias), which rely on the assumption of ‘exogeneity’: namely, that the independent variables X do not depend on the dependent variable Y.

Prior research has overcome those difficulties through a variety of estimation strategies, such as: (a) the use of student fixed-effects to cope with the selection between schools and students, as in Van Klaveren (2011), who finds a lack of relationship between time spent lecturing in front of the class and student performance across subjects; (b) the addition of teachers’ characteristics to better account for the selection of teaching styles, as in Bietenbeck (2014), who finds that teamwork and class discussions are strongly related to better achievements; and (c) the use of a random assignment design verified by balanced tests to deal with non-random within school selection, as in Hidalgo-Cabrillana and Lopez-Mayan (2018), and (d) a randomized field experiment that involves face-to-face talking between teachers and parents over two years, as in Islam (2019), who finds that test scores of students have raised by 0.26 and 0.38 standard deviations (SDs) respectively.

We address the two challenges through the choice of data and regression technique. With respect to data, we use the China Education Panel Survey (CEPS) - a large-scale nationally representative survey of middle school students with random class assignments. LBT has historically been the dominant teaching style in China’s schools (e.g. Fan and Ye, 2007; Tani, 2008; Sit, 2013). Teachers are the centre of the classroom, impart knowledge to students through structured lectures and frequent feedback on student assessments, with little room for discussions or student-led activities. This method was believed to be efficient in exam-focused schools which are resource constrained, in terms of class sizes and quality teachers. However, the limitations of the LBT, such as passive learning and its negative impact on student creativity, has increasingly raised concerns by educators, parents, and policymakers. Since the 1990s, alternative teaching styles have been promoted in both schools and higher education institutions to foster students’ all-round development in a shift from a subject-centred curriculum system to a more competence-based system (MOE 2014; Xie et al 2022).

One of CEPS’s features is information of whether students are randomly assigned into the sampled classrooms. This enables us to address streaming by students’ ability *within*-school while controlling for students’ and teachers’ self-selection into schools through school fixed-effects. The randomisation feature in CEPS has been extensively exploited by peer effect studies, which have verified the validity of the random assignment of students into classrooms (Gong et al 2018; Xu et al 2022). In using CEPS, we can hence effectively account for the sorting of students and teachers extending previous research on the impact of LBT (Schwerdt and Wuppermann 2011), to the extent that the variation in teaching styles might be regarded as plausibly exogenous in the presence of sufficient controls.

In addition, using CEPS has the advantage that its information includes all students in the sampled classrooms. This creates an ideal setting to examine the heterogenous impacts of teaching style across students of different socio-economic status and inequalities within classrooms. There has been compelling evidence that economically disadvantaged students generally underperform compared to advantaged students, as measured by SES index based on parental education and home possessions (Hanushek et al, 2022), or ethnic groups (Barber and Jones 2021). For instance, Hanushek et al (2022) show that the SES-achievement gap between the top and bottom quartiles is about 0.7 standard deviations in the U.S. in 2001, only one-fifth lower than the initial gap in 1961. As for ethnic groups, Hanushek and Rivkin (2006) find that most of the achievement gap between black and white students is driven by differences across schools rather than within schools. Fryer and Levitt (2004) attribute a considerable gap in test scores between black and white students in the U.S. to systematic differences in school quality attended by students. Following their paper, using a value-added (VA) model, Creel and Farell (2016) find that the gaps between black and white children is either constant or fading away as children grow older, depending on the measure and the family structure.

With reference to technique, we use the Unconditional Quantile Regression approach, and estimate the heterogenous impacts of a LBT style on students' academic achievements and the distribution of test scores in a classroom. Whereas most research on teaching effectiveness has been focusing on evaluating the average impacts of teacher's attributes, only a few studies have carefully examined how the impacts might vary by students' ability, prior attainment, or SES backgrounds. Using the 'STAR' project in the U.S., Konstantopoulos and Chung (2011) argue that students can benefit from effective teachers equally regardless of their gender and SES background. Antecol et al (2013) find evidence of heterogeneous effect of having a Teaching for America programme-trained teacher between Hispanic and black students.

To the best of our knowledge, our paper is the first to highlight the heterogenous impacts of teaching styles on student achievements and within-classrooms inequality, as proxied by prior attainment and SES, in the context of secondary education. The use of parental education as proxy for SES is a pragmatic choice, given that no index of multiple deprivation is available and education is the most important determinant of permanent income which is far superior than current income in explaining test score gaps (Fryer and Levitt 2006, Rothsten and Wozny 2013). Note that our definition of parental higher education status is based on whether holding at least a recognised "degree" from vocationally-oriented colleges or academically-oriented universities, which correspond to the minimum academic qualification for middle school teachers as mandated by the 1993 Teacher Law (Dai et al 2022).

We make several substantive contributions to the literature and policy debate on teaching styles. First, we show that teaching styles do influence academic performance in mandatory schooling. Second, such influence varies according to subjects: LBT has distinctive impacts on math and language subjects, having positive and negative impacts respectively compared with a teaching style that is either leaning towards DBT or indifferent between LBT and DBT. Third, bottom- and top-performing students are more likely to benefit from LBT. Fourth, teaching styles differently influence students belonging to different socio-economic groups: socio-economically advantaged students tend to benefit more from LBT on mathematics, LBT hence appears to aggravate within-classroom socio-economic inequality, as measured by the Gini index and the variance of test scores, as well as the upper-tail quantile ratios (such as the 90/50 ratio). Our finding thus shed new light in understanding the learning processes underpinning the forces at play in the delivery of education, driven by the complex interactions between students, their families and teachers.

The rest of the paper is organized as follows. Section 2 introduces the institutional background which provides the context for the empirical analysis. Section 3 presents data and analytical samples. Section 4 outlines the empirical model based on the Unconditional Quantile

Regression and the decomposition following the Recentered Influence Function (RIF) approach. Section 5 presents the empirical results and discusses the mediating channels. Section 6 concludes.

2. Institutional background

Following a comprehensive reform of education governance and management starting in 1985, China enacted its first Compulsory Education Law in 1986, mandating 6 years of primary school and 3 years of middle school education. Under the principle of "local responsibility and administration by levels", the central government delegated the responsibility for provision and financing of K-12 education to sub-provincial governments, to mobilise extra resources for the heavily under-funded basic education (Tsang, 1996). The increased competition and flexibility in the emerging quasi-marketized teacher labour market has apparently contributed to the phenomenal transition of Chinese educational system, helping to achieve near-universal 9-year compulsory education by 2005 and the massification of Higher Education by 2020 (Dai et al 2022). However, the decentralisation of basic education has not only led to a significant deterioration of the pre-existing urban-rural gap in school resources reflecting the restrictions imposed by the *hukou* (household registration) system since the late 1950s, but has also widened the gap by students' socio-economic background due to the growing income inequality and diminishing social mobility over time.

According to the theory of education production function, educational output is simply a function of key input factors from the school, including teacher quality, the family (parental financial and time investment among other things), as well as students' ability and prior educational attainment (Hanushek 2008). With large and increasing input gaps between urban and rural schools and between socio-economically advantaged and disadvantaged students, the theoretical prediction of an increased inequality in educational achievement is both unambiguous and well documented in statistics and academic research. For instance, using the China Family Panel Studies, Gruijters (2022) shows that whereas educational inequality in basic education has diminished as a result of the unprecedented education expansion in recent decades, the regional and SES disparities have persisted or even increased at senior high schools or above. On the other hand, Loyalka et al (2017) present compelling evidence showing that the largest source of unequal college access emerges from the transition between middle and high school. Even nowadays, no more than half of middle school graduates are enrolled in academic high schools each year through competitive selections based on the High School Entrance Exam (HSEE) scores.

Recent government policies have attempted to tackle the persistent educational inequality, especially in compulsory education by abolishing tuition fees, targeting government transfers to poor rural areas, raising of minimum qualification requirement for teachers and explicitly banning selective admissions by ability in primary and middle schools. Most of these policies are very costly, see e.g. the compelling evidence of selection by house price (Feng and Lu 2013; Huang et al 2020). However, there is hardly any evidence on how changing teaching styles, a change that is virtually free to implement, could impact on the average level of educational achievement, let alone its inequality. This is the very research gap we aim to address in this study.

While the evidence of the impacts of teaching style on student learning is thin in developing countries in general, there is virtually no study on the heterogeneity effect of teaching styles on the distribution of test results. The Chinese educational system has been traditionally associated with exam-oriented teacher-centred LBT, with students passively accepting knowledge most of the time (He 2021). While this approach might be good at strengthening students' grasp of basic knowledge, it hinders the cultivation of their creativity, compared to a student-centred DBT approach widely adopted in many OECD countries. To address the perceived deficiency of the traditional LBT, educators and policymakers in China have been actively promoting "*suzhi* (cultural quality) education" since the mid 1990s, through multiple rounds of curriculum reforms, to foster students' all-round development (Xie et al 2022). The "*suzhi* education" framework was further expanded in 2014 to "*hexin suyang* (core competencies)" by the Ministry of Education to shift the subject-centred curriculum system to a more competence-based system (MOE 2014; Xie et al 2022).

Despite the policy shift, it appears that the traditional LBT style is still commonly applied. Using a very conservative measure of LBT (with a hybrid mode grouped with DBT as the reference), we document that LBT is still the dominant teaching style for a significant portion of middle-school students in China on the basis of 2013-2014 CEPS data. We also document the notable heterogeneity between urban and rural areas, and between the more economically developed large cities (municipalities and provincial capitals) and the less developed "provincial backcountry" (referring to prefectural or county level cities, counties and townships).

<Table 1 Here>

Table 1 compares school-level differences between large cities and provincial backcountry in resources and teacher characteristics in CEPS data. Over 60% of the middle schools in the sample are located in provincial backcountry, while just under 40% are based in large cities. Consistent with decentralised funding mechanism for K-12 education in China, schools in the more developed large cities are much better resourced by all measures. In particular, per-student public spending for school in large cities is nearly twice as high as that for their provincial backcountry counterparts, despite all schools in the sample are publicly funded. Moreover, schools in larger cities also have higher quality teachers, with 95% having college degrees or above and 86% appointed with permanent contracts. This contrasts with 83% of teachers with college degrees and 66% with permanent contracts in provincial backcountries. On the other hand, the difference in teachers' years of experience is small, and if anything, slightly favours provincial backcountries.

Nearly all schools in large cities require college degree level qualifications and Teacher Certificates on newly recruited teachers, as mandated by the 1993 Teacher Law. In contrast, only 60% of schools located in provincial backcountry demand college degrees although the gap in Teacher Certificate is only 4 percentage point.

Figure 1 focuses on the contrast in student backgrounds between schools in large cities and schools in provincial backcountry. The mean standardized cognitive ability test scores for schools in large cities is 0.32 SD above that for their provincial backcountry counterparts. For comparison, the "urban advantage", between students attending schools in cities with more than 100k inhabitants, and their counterparts attending schools in villages and towns of up to 100k inhabitants, in the 2009 Programme for International Student Assessment (PISA), is 0.20 SD (or 20 PISA points) in the standardized tests for 15 year old students, who would correspond to Grade 9 in Chinese middle schools (OECD 2013). In contrast, the proportion of grade repeaters are 5 times as high for provincial backcountry compared to large cities. This pattern is in keeping with socio-economic gaps of similar magnitudes in favour of large cities as proxied by whether at least one parent holds a degree or whether the family receives income benefit from the government. Given the remarkable urban-rural development gap in China, the large discrepancies in students' and parents' backgrounds between well- and less-developed areas shown in Figure 1 is unsurprising.

<Figure 1 Here>

As conditioning on SES and parental investments, the variation in teacher quality and teaching styles might still contribute to inequality in secondary education, we provide further descriptive results to shed light on the large variations in terms of teacher quality and the adoption of LBT, by school types. Schools often seek to recruit teachers from local labour markets.

In Figure 2, Panel A shows there is a large variation in the implementation of LBT geographically at county-level, ranging from 10% to 40% in which teachers prefer LBT compared to other teaching styles. Panel B shows that the implementation of LBT is negatively correlated with principals' subjective ranking of their school while this pattern is less obvious based on the ranking of students' average cognitive skills, suggesting that the schools have various incentives to adopt LBT in Panel C. For instance, schools with higher cognitive scores are more likely to implement LBT in language courses. In contrast, no significant difference has been found in math courses.

We further break down the teachers' qualification and teaching style based on school types, as defined by city-tier in the administrative hierarchy and school location within-city in Table 2. The first panel shows that economically developed areas are more likely to have teachers with higher qualifications, while better educated teachers are not always located in city centre. In provincial and prefecture capitals, schools located at urban-rural junction have more college-educated teachers. In contrast, the adoption of LBT is more complicated. Schools in rural areas do not always have the strongest preference for LBT. This also applies to schools in urban areas. The detailed geographical differences in teachers' allocation and the implementation of LBT shed light on the complexity of the choice of teaching style at school level. Schools in rural areas might suffer from large class sizes and under-supply of teachers and hence might be more likely to implement LBT, while schools in urban areas more oriented towards may adopt a more efficient teaching style to better prepare their students for the high-stakes HSEE. The decision on teaching style is attributed to various factors, ranging from sorting of teachers in the local labour market to the individual preferences of teachers, students and their parents. Due to the complexity of the endogenous decision regarding teaching styles, we rely on the random assignment design in CEPS to address the endogenous selection between teachers and students.

<Figure 2 Here>

<Table 2 Here>

3. Data and sample:

The CEPS is a large-scale, nationally representative longitudinal survey carried out by the National Survey Research Centre (NSRC) at Renmin University of China. Data collection covers with two cohorts – the 7th and 9th graders in the baseline survey conducted in the academic year 2013-14 (https://ceps.ruc.edu.cn/index.php?r=index/index&hl=en). This is supplemented by a follow-up survey in 2014-15 of Grade 7 students in the baseline survey. The baseline survey contains 5 different questionnaires for the sampled students, parents, class headteachers and core subject teachers, and school principals respectively. The survey includes a standardized cognitive ability test for students in each grade respectively and an internet-based personality test for all sample students and collects transcripts of (mid-term) examinations in the three core subjects, i.e. mathematics, Chinese and English.

The CEPS follows a stratified, multistage sampling design with probability proportional to size (PPS), randomly selecting a school-based, nationally representative sample of approximately 20,000 students in 438 classrooms of 112 schools in 28 county-level administrative units in mainland China. In each relevant grade, all students from two randomly selected classes are included in the survey.

We construct two analytical samples. The Cross-sectional Sample is a harmonised pooled cross-section of all Grade 7 and grade 9 students from the 2012-2013 baseline data set, as well as Grade 8 students from the follow-up survey, used to examine the gaps between advantaged and disadvantaged students. The Longitudinal Sample is the two-wave panel of Grade 7 students in the baseline survey, which enables us to employ a value-added framework. To examine the effect of teaching styles on educational achievements, we select schools that have randomly assigned students into classes, according to reports by both the principals and all core subject teachers. The data set includes individual-level student's first mid-term test scores for three subjects, Chinese, Mathematics, and English taken in November in each academic year. Conventionally, students undertake two mid-term exams each year. The mid-term exams are important to students and teachers, as they are used to update beliefs about the performance of the students and to increase the effort if the outcome is under the expectation. Mid-term exams could also be used a performance indicator for teachers in their own appraisals and career promotions.

Measures of teaching styles

One of the difficulties in estimating the impact of teaching styles is the measurement of the implementation of different teaching styles. Previous research has typically used time spent in classes using different teaching styles (Van Klaveren 2011; Hidalgo-Cabrillana and Lopez-Mayan 2018) and students' responses of the extent to which a particular teaching style has been implemented (Bietenbeck, 2014).

We construct a measure of teaching style based on the three questions in the teachers' questionnaire about whether and how often they implement in the class a particular teaching style, namely LBT, DBT, and interactive. For simplicity, we group DBT and interactive teaching in the regression analysis, given the small differences that characterise the two relative to the LBT. Hence, we classify a teacher as using LBT more when the teacher reported using this lecture-based teaching style more often than the other two teaching styles. The comparison group consists of teachers who reported having DBT/interactive or having no differential frequencies in implementing the three teaching methods. By constructing the teaching style in this way, we develop a conservative binary treatment variable to identify if a teacher implements LBT.

We only use teachers' response in Grade 7 and assume that the teaching style is fixed between the mid-term of Grade 7 and Grade 8. There are too few teachers who have changed teaching styles in the sample period to allow a teacher fixed effect estimation. We are also aware of the potential measurement errors in our categorical measure of the teaching style. As we don't have responses from students, we cannot cross validate the measure of teaching style from students' perspective (Hidalgo-Cabrillana and Lopez-Mayan 2018). To alleviate the concerns, we examine the relationship between the choices of teaching style and teachers' and class's characteristics to show that the choice of teaching style is not correlated with teachers' characteristics and classroom environment.

It is worth emphasizing that the variation in teaching style is at the class-subject level. All students in each class are exposed to three teaching styles taught by three subject teachers. Students' academic achievement is measured by their subject-specific mid-term exam ranking across the grade in a school. Therefore, each student in a grade has three ranks measuring the relative performance of the three subjects, Chinese, Mathematics, and English. Since the test scores as absolute measures may introduce measurement errors when comparing students across schools, we use the Grade 8 decile rank as the dependent variable while controlling for the Grade 7 decile rank as an independent variable. Doing so enables us to capture the accumulative educational achievements.

Table 3 presents summary statistics of the Longitudinal Sample by teaching styles and by whether classes are randomly assigned. Figure A1 presents the relationships between the within-grade decile rank, average test scores, and cognitive scores in Grade 7. Students' standardised test scores range from 50 to 90, and the trends of test scores and cognitive scores are identical up to the median of the within-grade decile rank. As illustrated in the figure, the two trends only depart for students above the median.

Roughly speaking, LBT and DBT each account for a quarter of all teaching styles at the student level, with the remaining half classified as no preferences by the teacher, regardless of whether or not classes are randomly assigned. Conditional on random class assignment, LBT is associated with higher share of economically advantaged students, higher students' cognitive scores, larger class sizes, more teacher experience, and male teachers.

This pattern of correlations is confirmed (see Appendix Figure A2) using means of class pairs separately in the Longitudinal Sample, for randomly and non-randomly assigned classes. The balancing test, presented in Table A1, suggests that only class size is statistically significant. No other characteristics including teacher qualifications and ranks explain the variations in the frequency of the adoption of LBT in Column 1. This lends strong support to our binary classification of teaching styles in the empirical estimation. Although higher teacher ranks are correlated with the adoption of interactive teaching style in Column 3, we argue that the balancing test might not effectively identify the systematic unbalance in the data as idiosyncratic differences across classes may also drive significant results in a balancing test.

<Table 3 Here>

4. Empirical model

If teaching styles create uneven impacts on students with different backgrounds, such effect can be captured using the Firpo et al (2009, 2018) Recentred Influenced Function (RIF) approach to implement the unconditional quantile regression (UQR). One important merit of RIF is that it provides the flexibility to estimate a wide range of functional statistics, including inequality and decompositions. We employ the RIF regression to examine the impacts of teaching styles on the inequality within classrooms, taking advantage of the availability of information on classroom assignment to account for potential endogenous sorting into classrooms by students' prior attainment or ability.

Compared to Conditional Quantile Regressions (CQR), Firpo et al (2009) show that the RIF estimates the marginal effect of a change in covariate on the distributional change in a specific distribution statistic $v(F_Y)$.

$$RIF(y_i, v(F_Y)) = v(F_Y) + IF(y_i, v(F_Y)) \tag{1}$$

where the influence function (IF) of a quantile regression at the τ-th quantile ($0 \leq \tau \leq 1$) is defined as

$$IF(y_i; q_\tau, F_Y) = \frac{\tau - I(y_i \leq q_\tau)}{f_Y(q_\tau)} \tag{2}$$

where I(.) indicates an indicator function while $f_Y(.)$ and $F_Y(.)$ denote the probability density function and the cumulative density function, respectively.

The unconditional quantile treatment effects might be interpreted as the causal impact of teaching style on students with different backgrounds, if and only if the teaching style treatment is exogenous. The identification rests on the random assignment of students into classrooms, and the assumption that the choice of teaching style is independent from characteristics of students between the two classes in a school.

However, even after addressing the endogeneity of student assignment, interpreting the results as the direct impact of teaching style is probably a stretch because of the likely presence of omitted variables, as one's teaching style reflects potentially relevant unobservable teacher characteristics. Without a field experiment that randomly allocates teachers' teaching style and students with dedicated measures for teaching styles, it is problematic to identify the causal effect of a teacher's unique teaching style due to the association between teaching style and teacher's attributes. To our best knowledge, no data or experiment is available to answer such causal impacts.

Being aware of the limitations, we are cautious about the interpretation of causal effects of teaching styles of our results. Without a field experiment eliminating the association between teaching style and teacher's unobserved characteristics, we estimate heterogenous effects of teaching styles between advantaged and disadvantaged students to alleviate the concern based on an assumption that both types of students are affected equally by teachers' unobserved characteristics. More importantly, we investigate the inequality within classrooms, resulting from distinctive teaching styles.

To evaluate the impacts of teaching styles net of prior educational achievement, we estimate the UE based on a value-added framework. The outcome is the within-class decile rank of student *i* in school *s* of subject *k* at time *t*. As schools do not have uniform tests, the test scores across schools might not be comparable. In addition, we are interested in the relative ranking of students receiving different types of teaching styles. Given the average class size of approximately 50 in our sample, using within-class decile rather than percentile rank might be more robust. Therefore, the empirical model becomes:

$$Rank_{iskt} = \beta Teaching\ style_{ijk} + Rank_{iskt-1} + \boldsymbol{X}_i\theta + \boldsymbol{T_k}\delta + \varepsilon_{ij} \quad (3)$$

in which $Teaching\ style_{ijk}$ denotes the preferred teaching style of teacher *k*, implemented to student *i* in class *j*. The variable is a dummy variable and equals to one if a teacher has implemented LBT more frequently than DBT and the reference category of balanced teaching style.

$\boldsymbol{T_k}$ is a vector of teachers' characteristics, including gender, education, and experience. $\boldsymbol{X}_i$ denotes the students' and schools' characteristics, including school fixed effect, subject fixed effect, gender, *hukou* (household registration) status, cognitive skills, non-cognitive skills, parental interactions, etc. $Rank_{iskt-1}$ denotes the lagged ranking at Grade 7. β measures the impact of teaching style on the relative ranks in Grade 8 based on the original rank in a school in Grade 7. We employ UQR as our main empirical model because it systematically outperforms CQR because CQR fails to construct counter-factual scenario at quantiles if unobserved factors exist and differ across quantiles. Using UQR, we estimate the changes in the ranking distribution when one group of students receive LBT while the other group of students do not.

5. Results

We start in Section 5.1 with the estimation of the baseline Unconditional Quantile Regression using RIF, focusing on the unconditional effect of the locational shift of the share of LBT on different quantiles. This is followed by the corresponding unconditional partial effect (Bonacini et al 2021) estimation. We also estimate the effect of teaching styles on various measures of the within-class inequality in exam scores.

Section 5.2 further explores the potential channels that might drive the teaching style effects, as suggested by the recent literature. Specifically, we focus on the roles of allocating students to classes based on their ability, relative to random assignment, in each school, and the heterogeneous effects between more-selective and less-selective schools.

5.1. Baseline results

We find that distinctive impacts of LBT between mathematics and language subjects in Table 4. For both subjects and across all specifications, the lagged decile rank from Grade 7 has a strong and remarkably stable effect on the Grade 8 rank, at around 0.65. For Mathematics, increasing the share of LBT by 10 percentage points (relative to the reference category of balanced teaching style and DBT) would result in an upward shift in decile ranking by about 0.04 (i.e. 0.1x0.388), which is equivalent to 0.4 of one percentile rank, while the LBT has negative impacts on language rankings, resulting in a downward shift in decile ranking by about 0.02.

The positive and negative impacts of LBT on math and language subjects respectively are consistent with previous results found by Schwerdt and Wuppermann (2011) and Hidalgo-Cabrillana and Lopez-Mayan (2018). Based on the benchmark results, we further show that the impacts of teaching style on mathematics vary by student's backgrounds, as illustrated by the interaction terms. Students with higher cognitive skills or degree-educated parents are more likely to benefit from the LBT in mathematics while the corresponding interaction effects on language subjects are positive but statistically not significantly different from zero. This is consistent with Zhang and Xie (2016), which show compelling evidence that higher parental education and family income in China increases chances of the child receiving private tutoring, which in turn are predicative of higher verbal and math performance. The distinctive results motivate us to examine the heterogenous impacts of teaching style using quantile regressions.

<Table 4 Here>

<Figure 3 Here>

Figure 3 presents our main results, showing the distinctive impacts of teaching styles across quantiles of the unconditional Grade 8 distribution. While LBT is on average positively and negatively associated with mathematics and language subjects respectively, the impacts of teaching styles are heterogenous across students with different abilities. The bottom and top students in Mathematics are more likely to benefit from LBT while the students in between the two extremes of the distribution do not benefit from it. We found no previous result against which to benchmark what uncovered by our analysis. As a result, we interpret the finding as follows: that students with poorer academic background are perhaps less engaging and consequently they cannot benefit from a more interactive approach whereby teachers spend more time engaging (or trying to engage) with them. Those weaker students may benefit from LBT if teachers design their delivery to transfer knowledge with limited further engagement. Stronger students may spend more time on studying knowledge when LBT is implemented, having higher ranking in the exams. The heterogeneous impacts of LBT on language vary less by academic backgrounds and the impacts are negative among middle-ranked students.

We provide a range of robustness checks to validate the heterogeneous impacts of the teaching style. To mitigate the concern that teaching style might be itself a choice based on observable teacher characteristics, Figure A3 presents the UE with inverse probability weighting to adjust for differences in observed teachers' characteristics. The results are largely consistent with the benchmark model results reported in Figure 4.

<Figure A3 Here>

Following Bonacini et al (2021), in Figure A4, we show the 'unconditional partial effect' (UPE). Compared to UE that allows other covariates to change as a result of the shifts in the variable of interest, UPE estimates the marginal effect while explicitly holding other covariates constant. The UPE results are largely consistent with the UE and are statistically more significantly different from zero. This is not surprising in our sample of schools with random class assignments. Figure A5 shows the impacts of LBT on Chinese and English scores, respectively. Consistent with the combined results above, the impacts of LBT is negative for students within the middle range of the score distribution.

<Figure A4 and A5 Here>

5.2. Potential channels

To investigate the impacts of the endogeneity of teacher's choice on teaching style, in Figure 4 we also present the impacts of teaching styles based on non-random classrooms (typically ability-streamed) in which teachers are more likely to select an appropriate teaching method given students' backgrounds. Here, the patterns run contrary to the results found when the analysis is based on random allocation to classrooms.

The implementation of the LBT shows positive impacts for students across both mathematics and language subjects. This suggests that with ability streaming, teachers might strategically adopt teaching styles that can effectively cater to a broader spectrum of students, with a particular emphasis on the average student, as suggested by the relatively homogeneous impacts across the quantile distribution.

The sharp differences in the results between random and non-random classrooms may shed light on the differential impacts of the endogenous choice on the estimation of teaching styles, depending on whether students are randomly assigned to classrooms (by ability). Without accounting for the student sorting, it is likely to derive converse conclusions.

<Figure 4 Here>

To examine whether the impacts vary by the background of schools, we examine the heterogenous impacts based on the average cognitive skills of students across schools in Figure A6. After grouping schools into two groups, Panel A shows the distribution of cognitive scores between schools above and below the mean cognitive scores at school-level.

There exists a distinction between "more-selective" and "less-selective" schools, based on the average cognitive scores of Grade 7 students. This likely reflects residential segregation as explicit admission selection is prohibited in China during compulsory education. When examining the impact of LBT on math performance in Panel B, it becomes apparent that an "U" shape emerges within the category of "more-selective" schools across the student ability quantiles. Conversely, all but the top quantile students show improvements in math in "less-selective" schools. Similar results can be observed in the language domain, where the adverse effects of LBT are more pronounced among "less-selective" schools, as shown in Panel C.

These findings hold even when considering parental qualifications as a measure of school quality. Given the overlap in cognitive skills between the lower and higher ends of the distribution, as suggested in Panel A, it appears that the school's expectations affect students at the higher end.

<Figure A6 Here>

After showing that teaching styles have modest but statistically significant effects on educational achievement directly, which could vary along the unconditional quantile distribution, we now explore the within-class distributional impacts of teaching styles. These could be more relevant from a policy perspective as changes are more likely and easily implemented within rather than across schools. Taking advantage of the information on the entire population of students in the classroom in the CEPS, we examine the impacts of teaching style on the inequality within classrooms.

Intuitively, a mean-preserving innovation in teaching styles that reduces inequality could be still regarded as desirable by policymakers. One of the merits of RIF regression is the flexibility of the selection of the objective function. We estimate the impacts of teaching styles on five alternative measures of inequality in Table 5, with Mathematics and Language subjects pooled together in Panel A, but separately in Panels B and C respectively. The results suggest that the LBT style increases inequality, as measured by the Absolute Gini (A-Gini) index and variance of rankings when all subjects are pooled together.

The results further suggest that most of the impact on inequality is due to an inequality-enhancing effect at the top half of the unconditional quantile distribution, as measured by the ratio of the 90^{th} percentile to the 50^{th} percentile (i.e. the 90/50 ratio). The rankings of students taught by LBT are more dispersed within a school in Grade 8. More advantaged students may move to higher rankings compared to median students. This dominates the negative effect on inequality which is a statistically insignificant for the bottom half of the unconditional quantile distribution captured by the 50/10 ratio, such that the overall effect across the whole distribution proxied by the 90/10 ratio is still positive and statistically significant at the 5% level.

The effects of LBT on the corresponding inequality measures in Mathematics ranking are largely consistent with those for the overall effect, notwithstanding that the effects on the A-Gini and variance are no longer statistically significant, likely because of the small sample. The effects on the 90/50 and 90/10 ratios become even more pronounced and significant for Mathematics. Interestingly, the effect of LBT on languages is not significant, except for a negative effect on the 50/10 ratio.

<Table 5 Here>

6. Conclusions

We examine the heterogeneous impacts of teaching styles on students' academic achievements using a sample of schools with students randomly assigned to classes. Applying the Unconditional Quantile Regression approach based on the RIF regression on the value-added within-grade rank and allowing for school fixed-effects, we find that the lecture-based teaching (LBT) style is positively and negatively correlated with the test scores in mathematics and language subjects, respectively. LBT tends to benefit the bottom and the top students in the unconditional quantile distribution. More importantly, teaching styles are found to have more substantial impacts on various measures of within-class achievement inequality, with the LBT style significantly increasing the Gini index and variance, as well as the 90/50 percentile ratio.

Our results have important policy implications. While the LBT style might be favoured by exam-oriented schools under resource constraints, especially in larger and more homogeneous classes with strong ability tracking, the potential benefit could be outweighed by the significant aggravation in within-class achievement inequality compared to a more balanced teaching style.

While we contribute to the literature by using randomly assigned classes and school fixed-effects to address student and teacher sorting, there is a potential threat to the maintained identification assumption that the variation in teaching styles is plausibly exogenous, arising from for instance potential teacher responses to student endowments or parental pressure. In the absence of field experiments on teaching training or educational policy changes designed to influence teaching styles or rich longitudinal data with potential instruments for teaching styles, it is virtually impossible to fully address the association between teaching styles and teacher's unobserved characteristics and estimate the average treatment effect of different teaching styles. Another limitation is the framing of the choice of LBT style as dichotomous due to lack of variation in the survey. Nevertheless, our finding that teaching styles matters for student learning outcomes even when we control for school fixed-effect and the choice by each school as whether to stream students by ability is encouraging. Future research could use bespoke surveys with better measures of teaching styles (say with self or peer-reported time use associated with each style) in order to establish a more precise and robust causal relationship.

References:

Aaronson, D., Barrow, L. and Sander, W. 2007. Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, 25, 95-135.

Aitkin, M., Bennett, S. N., and Hesketh, J. 1981. Teaching styles and pupil progress: a re - analysis. *British Journal of Educational Psychology, 51*(2), 170-186.

Antecol, H., Eren, O. and Ozbeklik, S., 2013. The effect of Teach for America on the distribution of student achievement in primary school: Evidence from a randomized experiment. *Economics of Education Review*, 37, 113-125.

Attard, A., Di Iorio, E., Geven, K., and Santa, R. 2010. Student-Centred Learning: Toolkit for Students, Staff and Higher Education Institutions. *European Students' Union (NJ1)*.

Avvisati, F., Gurgand, M., Guyon, N. and Maurin, E., 2014. Getting parents involved: A field experiment in deprived schools. *Review of Economic Studies*, 81(1), 57-83.

Barber, M. and Jones, M.E., 2021. Inequalities in test scores between Indigenous and non-Indigenous youth in Canada. *Economics of Education Review*, *83*, 102139.

Baeten, M., Kyndt, E., Struyven, K., and Dochy, F. 2010. Using student-centred learning environments to stimulate deep approaches to learning: Factors encouraging or discouraging their effectiveness. *Educational Research Review*, *5*(3), 243-260.

Bietenbeck, J., 2014. Teaching practices and cognitive skills. *Labour Economics*, *30*, 143-153.

Bonacini, L., Gallo, G. and Scicchitano, S., 2021. Working from home and income inequality: risks of a 'new normal' with COVID-19. *Journal of Population Economics*, 34(1), 303-360.

Brandes, D., and Ginnis, P. 1996. *A guide to student-centred learning*. Nelson Thornes.

Clotfelter, C.T., Ladd, H.F. and Vigdor, J.L., 2010. Teacher credentials and student achievement in high school a cross-subject analysis with student fixed effects. *Journal of Human Resources*, 45(3), 655-681.

Creel, M. and Farell, M., 2016. On the production of cognitive achievement and gaps in test scores. *Oxford Bulletin of Economics and Statistics*, *78*(2), 228-247.

Dai, F., Xu, L., and Zhu, Y. 2022. Higher Education expansion and supply of teachers in China. *China Economic Review* 71, 101732.

Doherty, K. 1995. A quantitative analysis of three teaching styles. *Journal of Experiential Education*, *18*(1), 12-19.

Fan, W., and Ye, S. 2007. Teaching styles among Shanghai teachers in primary and secondary schools. *Educational Psychology*, *27*(2), 255-272.

Felder, R. M., and Silverman, L. K. 1988. Learning and teaching styles in engineering education. *Engineering education*, *78*(7), 674-681.

Feng, H., and Lu, M., 2013. School quality and housing prices: empirical evidence from a natural experiment in Shanghai, China. *Journal of Housing Economics* 22, 291–307.

Firpo, S.P., Fortin, N.M. and Lemieux, T. 2009. Unconditional Quantile Regressions. *Econometrica*, 77(3), 953-973.

Firpo, S. P., Fortin, N.M. and Lemieux, T. 2018. Decomposing wage distributions using Recentered Influence Function Regressions. *Econometrics*, 6(28). doi:10.3390/econometrics6020028.

Freire, P. 1998. *Pedagogy of freedom: Ethics, democracy, and civic courage*. New York: Rowman & Littlefield.

Fryer Jr, R.G. and Levitt, S.D., 2004. Understanding the black-white test score gap in the first two years of school. *Review of Economics and Statistics*, *86*(2), 447-464.

Fryer Jr, R.G. and Levitt, S.D., 2006. The Black-White Test Score Gap Through Third Grade. *American Law and Economics Review* 8(2), 249-81.

Gong, J., Lu, Y. and Song, H., 2018. The effect of teacher gender on students' academic and noncognitive outcomes. *Journal of Labor Economics*, 36(3), 743-778.

Gruijters, R.J., 2022. Trends in educational stratification during China's Great Transformation. *Oxford Review of Education*. 48(3), 320-340.

Hanushek, E.A. and Rivkin, S.G., 2006. School quality and the black-white achievement gap.

Hanushek, E.A. 2008. Education Production Functions. *The New Palgrave Dictionary of Economics* (2nd ed). Eds. Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan, 2008.

Hanushek, E.A, Light, J.D., Peterson, P.E., Talpey, L.M. and Woessmann, L., 2022. Long-run Trends in the U.S. SES-Achievement Gap. *Education Finance and Policy*, 17(4), 608-640.

He, Y., 2021. Analysis on the Comparison Between Chinese and Western Classroom Teaching. *Advances in Social Science, Education and Humanities Research 554.* Proceedings of the 7th International Conference on Humanities and Social Science Research (ICHSSR 2021).

Hidalgo-Cabrillana, A. and Lopez-Mayan, C., 2018. Teaching styles and achievement: Student and teacher perspectives. *Economics of Education Review*, 67, 184-206.

Huang, B., He, X., Xu, L., and Zhu, Y. 2020. Elite school designation and house prices: Quasi-experimental evidence from Beijing, China. *Journal of Housing Economics*, 50, 101730.

Huang, W., Li, T., Pan, Y. and Ren, J., 2021. Teacher characteristics and student performance: evidence from random teacher-student assignments in China. *IZA DP*, No. 14184.

Islam, A., 2019. Parent–teacher meetings and student outcomes: Evidence from a developing country. *European Economic Review*, 111, 273-304.

Jarvis, P. 2006. Teaching styles and teaching methods. *The theory and practice of teaching*, *2*, 28-38.

Kane, T.J., Taylor, E.S., Tyler, J.H. and Wooten, A.L., 2011. Identifying effective classroom practices using student achievement data. *Journal of Human Resources*, 46(3), 587-613.

Konstantopoulos, S. and Chung, V., 2011. Teacher effects on minority and disadvantaged students' grade 4 achievement. *The Journal of Educational Research*, 104(2), 73-86.

Lavy, V., 2016. What makes an effective teacher? Quasi-experimental evidence. *CESifo Economic Studies*, 62(1), 88-125.

Lea, S. J., Stephenson, D., and Troy, J. 2003. Higher education students' attitudes to student-centred learning: beyond 'educational bulimia'?. *Studies in Higher Education*, *28*(3), 321-334.

Loyalka, P., Chu, J., Wei, J., Johnson, N. and Reniker. J. 2017. Inequalities in the pathway to college in China: When do students from poor areas fall behind? *The China Quarterly*, 229, 172-194.

Mansfield, R.K., 2015. Teacher quality and student inequality. *Journal of Labor Economics*, 33(3), 751-788.

McCarthy, J. P., & Anderson, L. 2000. Active learning techniques versus traditional teaching styles: Two experiments from history and political science. *Innovative Higher Education*, *24*, 279-294.

Michel, N., Cater III, J. J., and Varela, O. 2009. Active versus passive teaching styles: An empirical study of student learning outcomes. *Human resource development quarterly*, *20*(4), 397-418.

MOE (Ministry of Education). 2014. *The Opinions of the Ministry of Education on Deepening of Curriculum Reform and the Implementation of the Fundamental Tasks of the Cultivating People of Moral Character* (*in Chinese*: 教育部关于全面深化课程改革落实立德树人根本任务的意见)." http://www.moe.gov.cn/srcsite/A26/jcj_kcjcgh/201404/t20140408_167226.html. Accessed 1 June 2023.

OECD, 2013. What makes urban schools different? *PISA in Focus* No. 28. OECD, Paris.

Opper, I.M., 2019. Does helping John help Sue? Evidence of spillovers in education. *American Economic Review*, *109*(3), 1080-1115.

Poropat, A. E. 2009. A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135(2), 322.

Richardson, M., Abraham, C., and Bond, R. 2012. Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychological Bulletin*, 138(2), 353.

Rios-Avila, F., 2020. Recentered Influence Functions (RIFs) in Stata: RIF regression and RIF decomposition. *The Stata Journal*, 20(1), 51-94.

Rivkin, S.G., Hanushek, E.A. and Kain, J.F., 2005. Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417-458.

Rothstein, J. and Wozny, N., 2013. Permanent income and the black-white test score gap. *Journal of Human Resources*, 48(3), 510-544.

Schwerdt, G. and Wuppermann, A.C., 2011. Is traditional teaching really all that bad? A within-student between-subject approach. *Economics of Education Review, 30*(2), 365-379.

Sit, H. H. W. 2013. Characteristics of Chinese students' learning styles. *International proceedings of economics development and research*, *62*, 36.

Tani, M., 2008, Raising the in-class participation of Asian students through a writing tool. *Higher Education Research & Development,* 27, 345-356.

Tsang, M. C. 1996. Financial reform of basic education in China. *Economics of Education Review*, 15, 423–444.

Van Klaveren, C., 2011. Lecturing style teaching and student performance. *Economics of Education Review*, 30(4), 729-739.

Xie, A., Li, J. and Ma, F., 2022. Understanding China's policy responses to PISA: using a *ti* and *yong* framework. *Comparative Education*, DOI: 10.1080/03050068.2022.2157014.

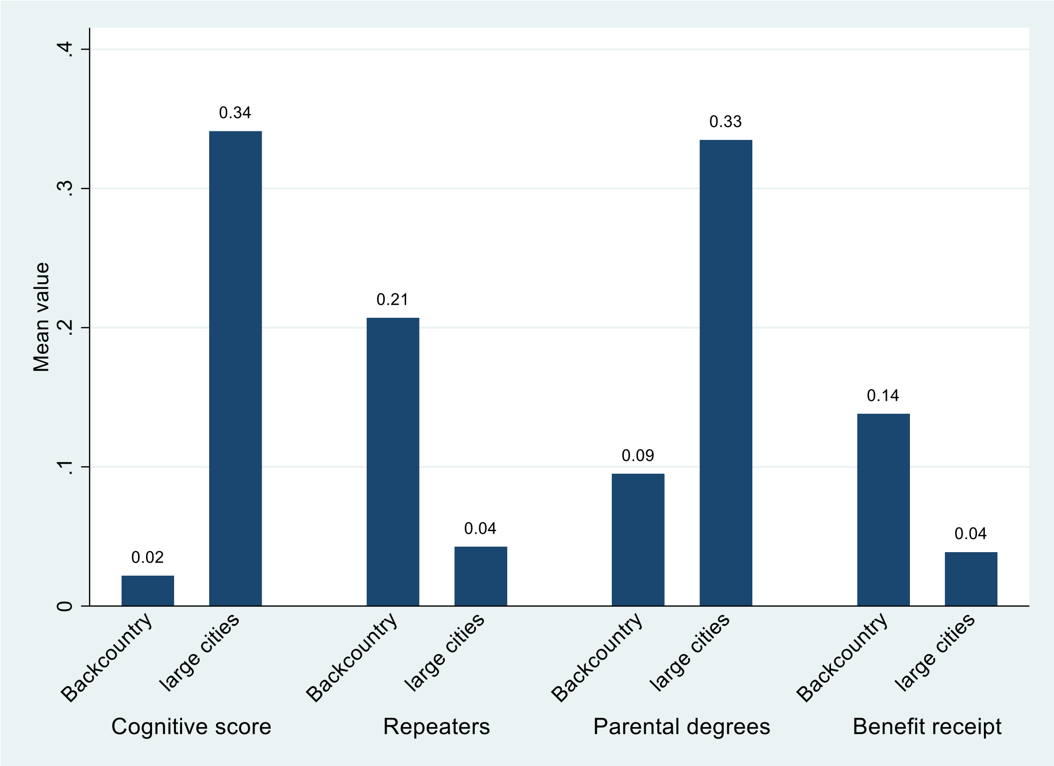
Xu, D., Zhang, Q. and Zhou, X., 2022. The Impact of Low-Ability Peers on Cognitive and Noncognitive Outcomes Random Assignment Evidence on the Effects and Operating Channels. *Journal of Human Resources*, *57*(2), 555-596.

Zhang, H., Qin, X. and Zhou, J., 2020. Do tiger moms raise superior kids? The impact of parenting style on adolescent human capital formation in China. *China Economic Review*, 63, 101537.

Zhang, Y. and Xie, X. 2016. Family Background, Private Tutoring, and Children's Educational Performance in Contemporary China, *Chinese Sociological Review*, 48:1, 64-82.

Figures

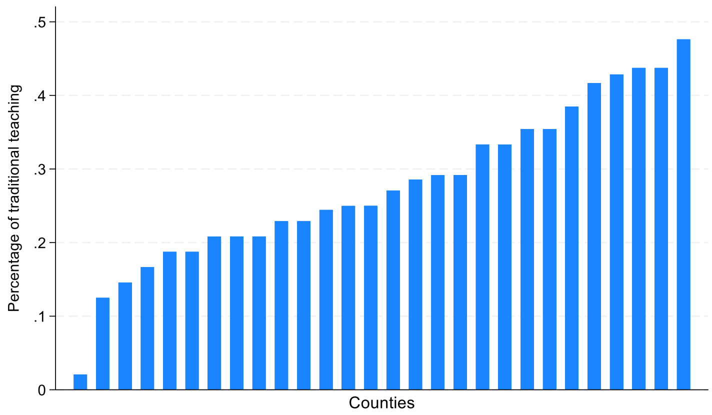
Figure 1. Students' backgrounds between provincial backcountry and large cities



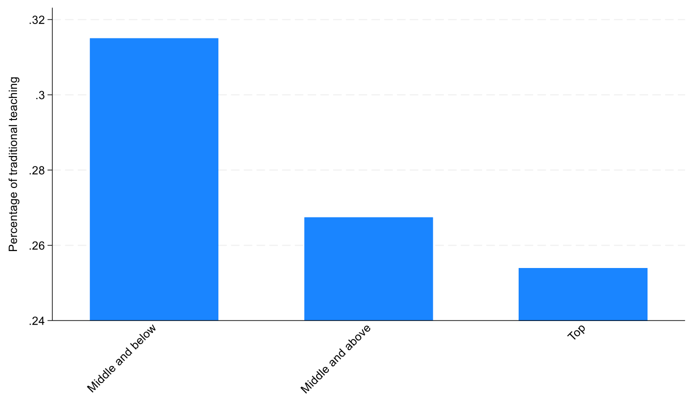
Note: Cross-sectional data. The Y-axis shows the fraction of students by cognitive ability, grade repetition, whether parents hold a college degree and income benefit receipt from the government.

Figure 2. The choice on the traditional teaching

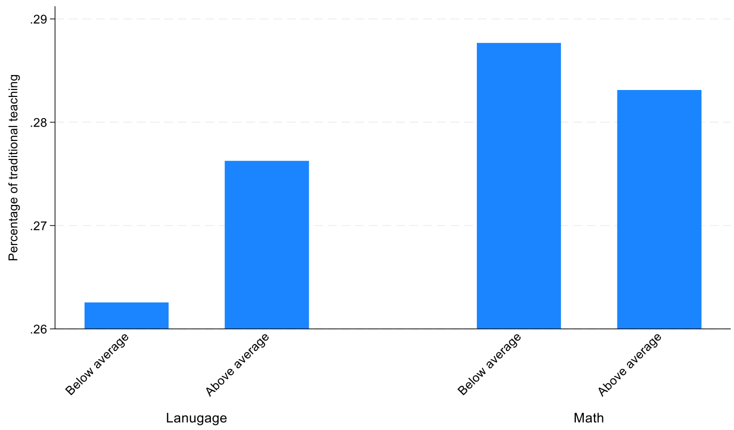
Panel A, By County



Panel B, By Within-County School Ranking (Based on Responses of Principals)

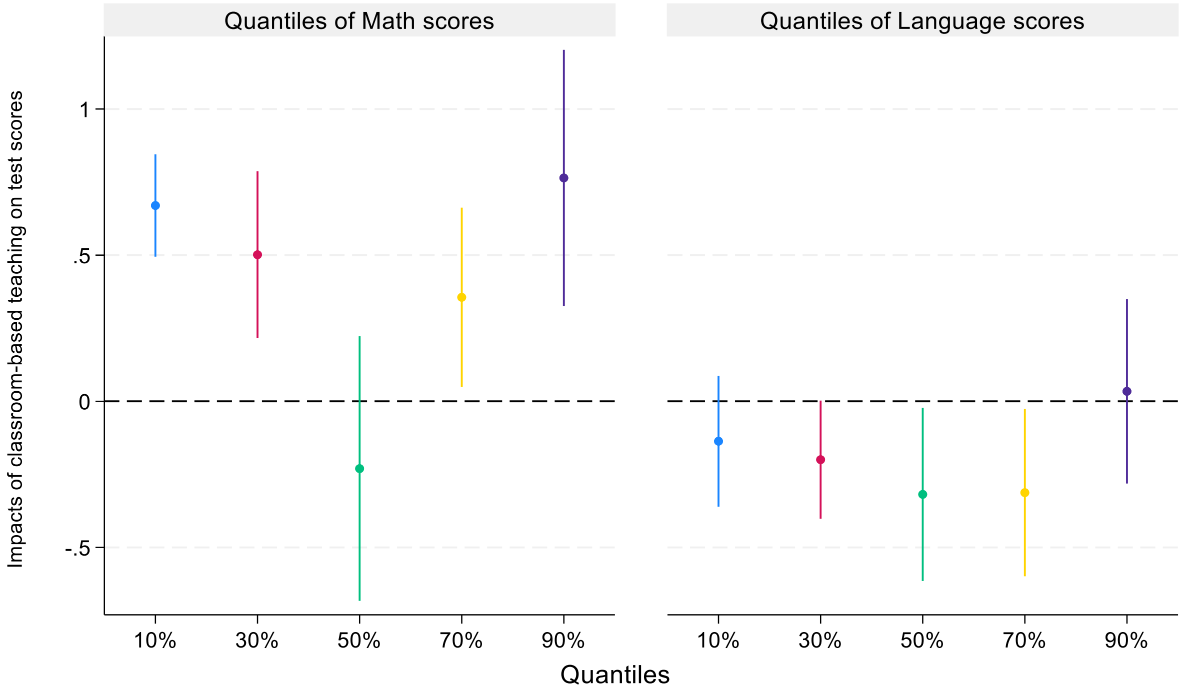


Panel C, By Average Cognitive Ability of Students



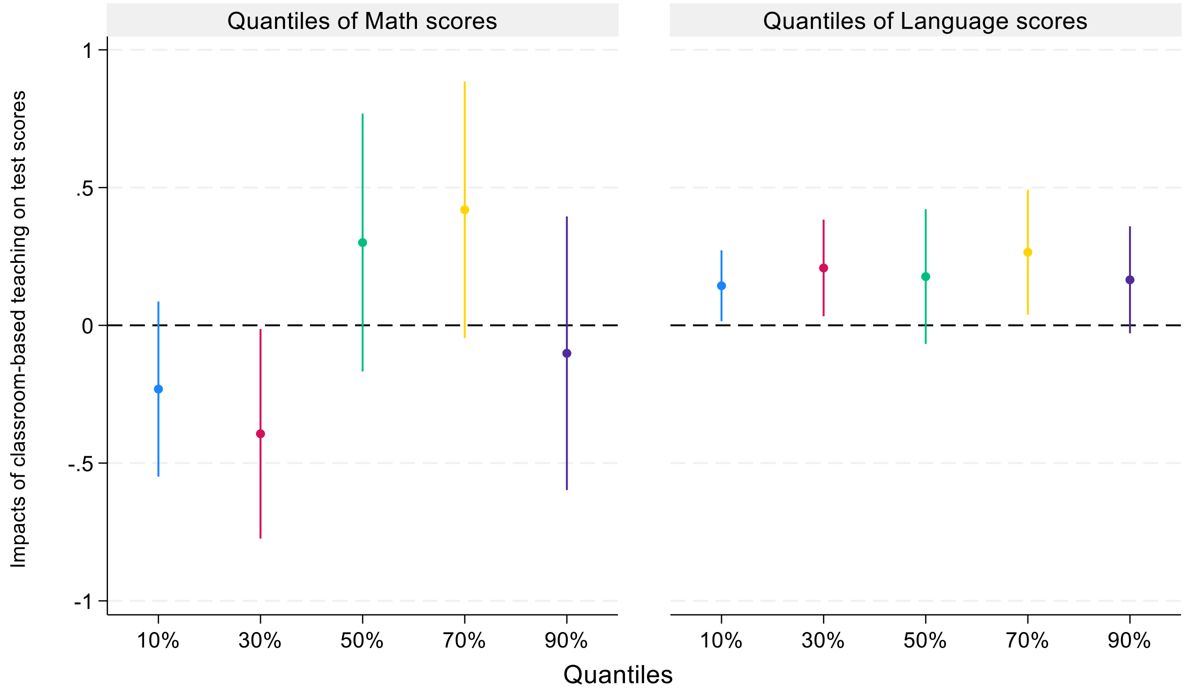
Note: Cross-sectional data. The sub-figures describe the percentage of teachers adopting the LBT style by counties, within-county school ranking (based on the responses of principals) and average cognitive ability of students, respectively.

Figure 3. Unconditional quantile regression using RIF



Note: Longitudinal Sample. Estimates of Unconditional Quantile Regression (UQR) are based on the RIF regression developed by Fortin et al (2009, 2018) and implemented by the Stata package developed by Rios-Avila (2020).

Figure 4. Unconditional quantile regression using RIF for non-random schools



Note: Non-random schools, based on teachers' and principal's responses. It includes different schools compared to random schools.

Tables

Table 1. Differences in school resources between urban and rural areas

	Provincial Backcountry	Large Cities	Total
School resources:			
Public spending per student (Yuan)	715.6	1585.5	1040.8
Computers in school	105.7	165.7	129.1
Books in library (1,000)	74.8	88.0	79.7
Teacher-level characteristics:			
Proportion of teachers having a degree	0.83	0.95	0.88
Proportion of teachers with permanent contract	0.66	0.86	0.74
Teacher experience (years)	16.4	15.4	16.0
School-level teacher requirements:			
College degree requirement (new teachers)	0.54	0.95	0.71
Teacher certificate requirement (new teachers)	0.94	0.98	0.96
Sample share (%)	61.5	38.5	100.0

Note: Cross-sectional data. The dummies for college degree requirement and teacher certificate requirement, as well as permanent contract (*Bianzhi*) share are based on the responses of principals, while teachers having a degree and years of experience are based on reports of subject teachers.

Table 2. The geographical differences of teaching styles and teacher's qualification

LBT(Traditional teaching)	Urban area	Urban-rural junction	Rural area	Total
Municipality	0.35	0.17	0.38	0.31
Province capital	0.21	0.27	0.27	0.23
Prefecture capital	0.20	0.29	0.35	0.27
County	0.37	0.27	0.25	0.28
Total	0.28	0.25	0.28	0.27

Teacher having a degree	Urban area	Urban-rural junction	Rural area	Total
Municipality	0.73	0.69	0.65	0.71
Province capital	0.58	0.67	0.58	0.60
Prefecture capital	0.47	0.52	0.43	0.47
County	0.4	0.25	0.22	0.26
Total	0.57	0.46	0.33	0.45

Note: Cross-sectional data. Each cell shows the percentage of teachers adopting traditional teaching or having a degree.

Table 3. Summary of variables

	Non-random schools			Random schools		
	LBT	DBT	No preference	LBT	DBT	No preference
Student characteristics:						
Cognitive scores	0.10	0.20	0.01	0.25	0.14	0.22
Grade 7 decile rank	5.35	5.56	5.39	5.42	5.46	5.53
Grade 8 decile rank	5.48	5.54	5.42	5.45	5.65	5.63
Repeaters	0.13	0.10	0.11	0.11	0.11	0.09
Male	0.53	0.50	0.52	0.51	0.51	0.51
Rural *hukou*	0.52	0.52	0.54	0.37	0.48	0.38
Parental degree	0.20	0.21	0.19	0.30	0.22	0.29
School/class characteristics:						
Class size	49.34	47.95	50.18	52.27	47.46	48.67
Female teacher	0.68	0.67	0.73	0.70	0.79	0.85
Average teacher education	5.37	5.32	5.43	5.46	5.21	5.47
Average teacher experience	14.29	18.02	14.96	17.91	15.64	16.25
N	1,795	2,024	3,966	2,504	2,628	5,064
%	23.1%	26.0%	50.9%	24.6%	25.8%	49.7%

Note: Longitudinal Sample. The ranks refer to students' decile ranks in the mid-term exams at Grade 7 and Grade 8 and the standardized cognitive ability tests. For instance, the highest decile, Decile 10, refers to students in the top 10% of the respective exam or test. The cognitive scores have been standardised. Repeaters measure the proportion of repeaters in a cell. Teacher's qualification is a categorical variable, including, diploma, degrees, or post-graduate degree. Teacher's experience is the years of working experience as a teacher.

Table 4. The impact of traditional teaching style

Dep = Grade 8 Decile Rank	(1) Math	(2) Language	(3) Math	(4) Language	(5) Math	(6) Language
LBT (Traditional teaching)	0.388^{***}	-0.173^{**}	0.289^{***}	-0.192^{**}	0.239	-0.237^{**}
	(0.10)	(0.08)	(0.11)	(0.08)	(0.15)	(0.10)
LBT X Cognitive score			0.323^{**}	0.053		
			(0.14)	(0.08)		
Parental degree					0.132	0.234^{***}
					(0.10)	(0.08)
LBT X Parental degree					0.374^{*}	0.220
					(0.20)	(0.16)
Rank at 7	0.652^{***}	0.672^{***}	0.651^{***}	0.671^{***}	0.649^{***}	0.667^{***}
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
N	2431	4485	2431	4485	2431	4485

Note: All regression results are based on the Longitudinal Sample which allows a VA model. The sample is based on the strict random class assignment where all teachers, including both principles and subject teachers, indicate that the classes are randomly assigned. Other control variables include subjects, gender, class size, *hukou*, teacher's characteristics, including gender, education, experience, teaching qualification, and school identifier. Standard errors are clustered at class-subject-level. The control variables are the same in the following regressions. ***, ** and * indicate statistical significance at 1%, 5% and 10% respectively.

Table 5. The effect of LBT on student achievement inequality

	(1) A-Gini index	(2) Variance	(3) 90/50 ratio	(4) 50/10 ratio	(5) 90/10 ratio
A): Math and Languages Pooled (N=6,916)					
LBT	0.045**	0.438**	0.391***	-0.144	0.247**
	(0.02)	(0.18)	(0.14)	(0.10)	(0.11)
B): Math Only (N = 2,431)					
LBT	0.049	0.463	0.995***	-0.230	0.764***
	(0.04)	(0.39)	(0.22)	(0.23)	(0.22)
C): Languages Only (N = 4,485)					
LBT	0.029	0.285	0.353	-0.319**	0.034
	(0.03)	(0.26)	(0.21)	(0.15)	(0.16)

Note: Compared to the conventional (relative) Gini index which shows the proportional differences between subgroups without accounting for the group mean, an absolute Gini (A-Gini) index shows the absolute magnitude of difference between subgroups of population. The absolute Gini index has been argued to be a better measure for inequality (Bandyopadhyay 2018). ***, ** and * indicate statistical significance at 1%, 5% and 10% respectively.

Appendix:

Figure A1 The relationships between rankings, test scores, and cognitive scores in Grade 7

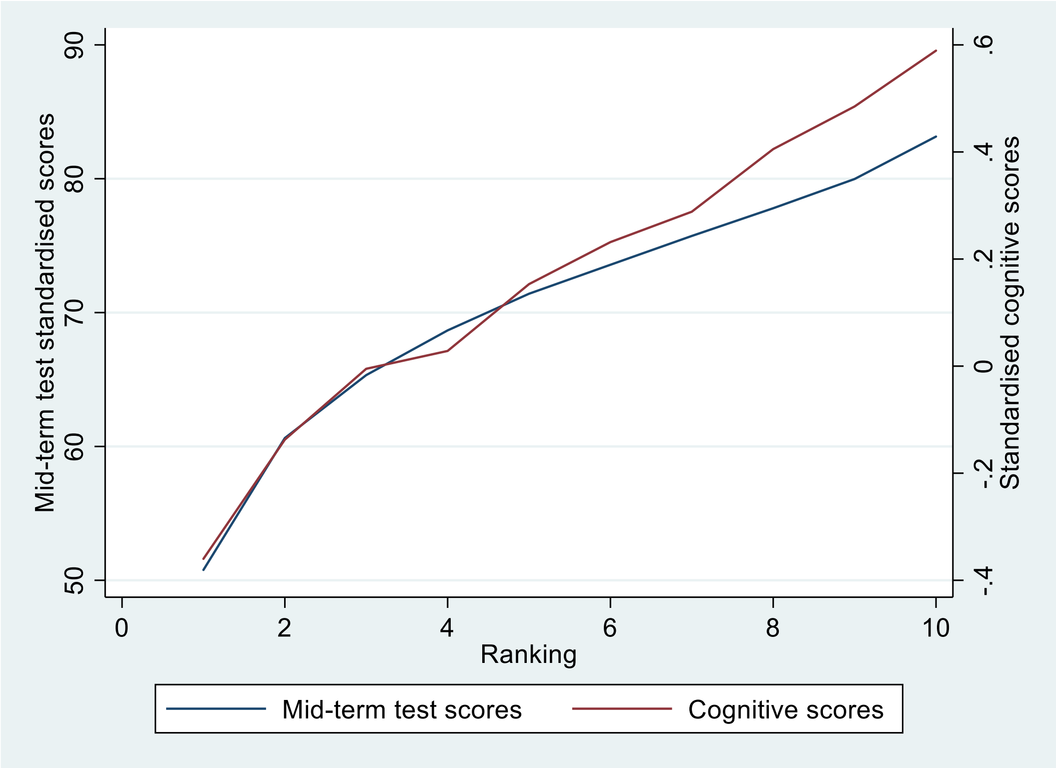
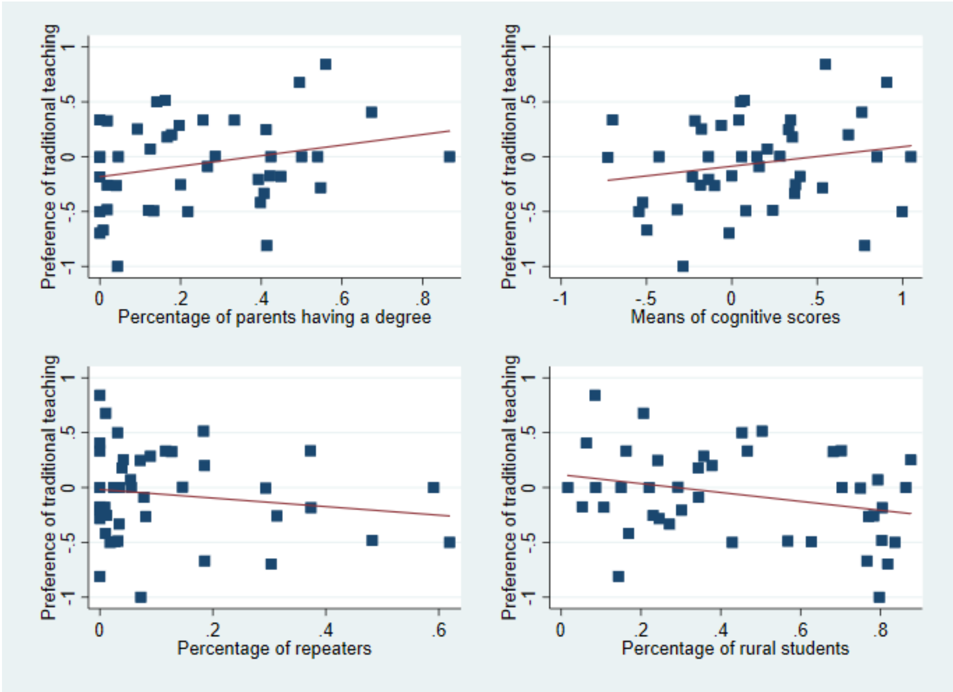
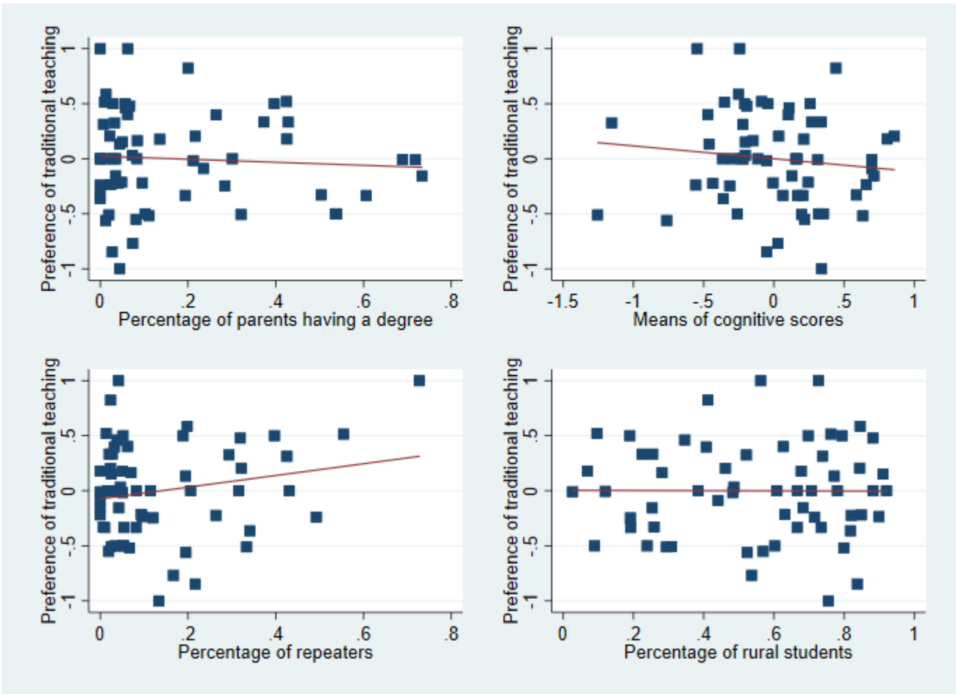


Figure A2. Correlation between teaching styles and average student characteristics across class pairs

Panel A. Random classes

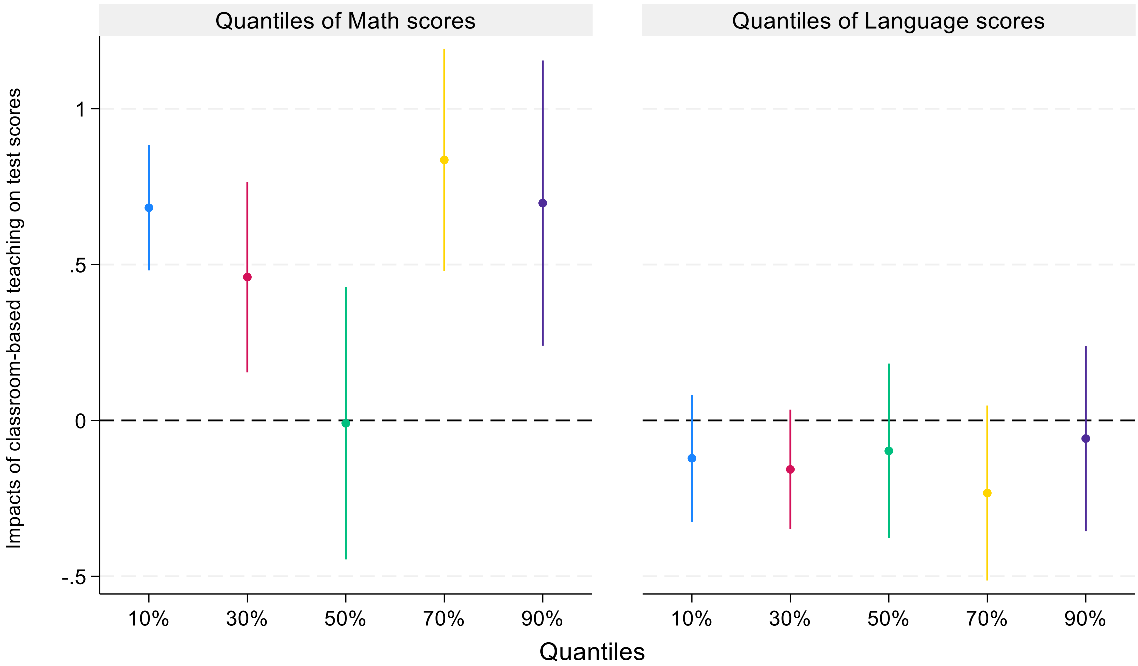


Panel B. Non-random classes



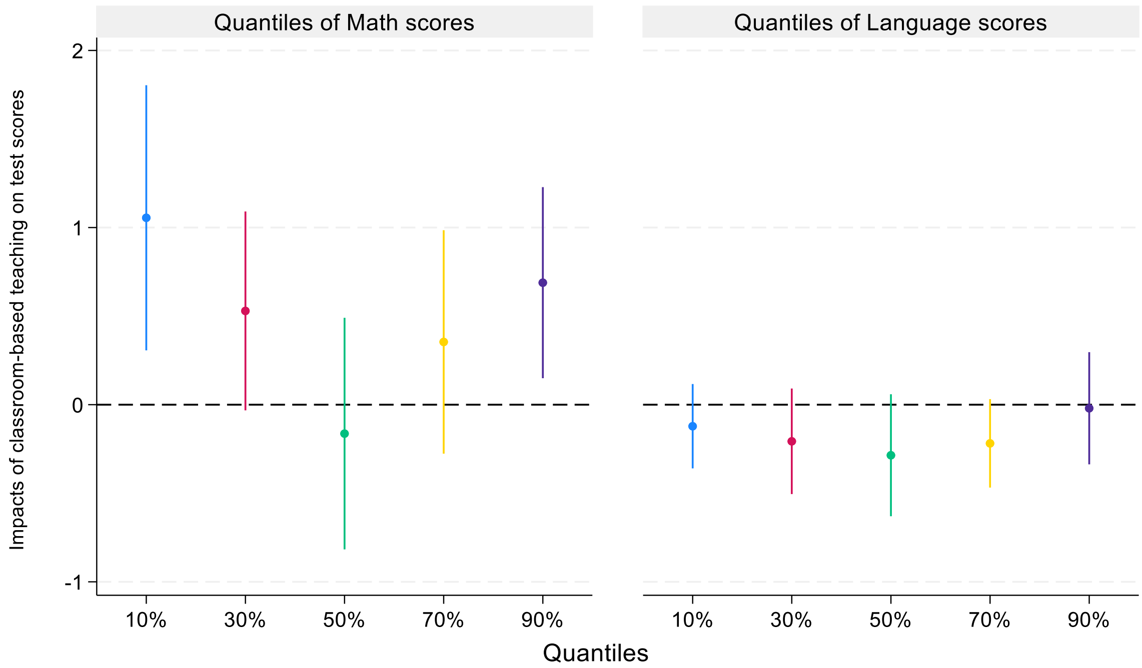
Note: Pair of classes based on the Cross-sectional Sample between schools randomly and non-randomly allocate students.

Figure A3. Unconditional quantile regression with IPW



Note: The variables used for weighting are teacher's characteristics, including, gender, education, experience, and teaching qualification. ***, ** and * indicate statistical significance at 1%, 5% and 10% respectively.

Figure A4. The partial effect of Unconditional quantile regression using RIF



Note: Following Bonacini et al (2021), the table includes the results of the partial effect of the unconditional quantile regression. ***, ** and * indicate statistical significance at 1%, 5% and 10% respectively.

Figure A5. Unconditional quantile regression using RIF for Chinese and English

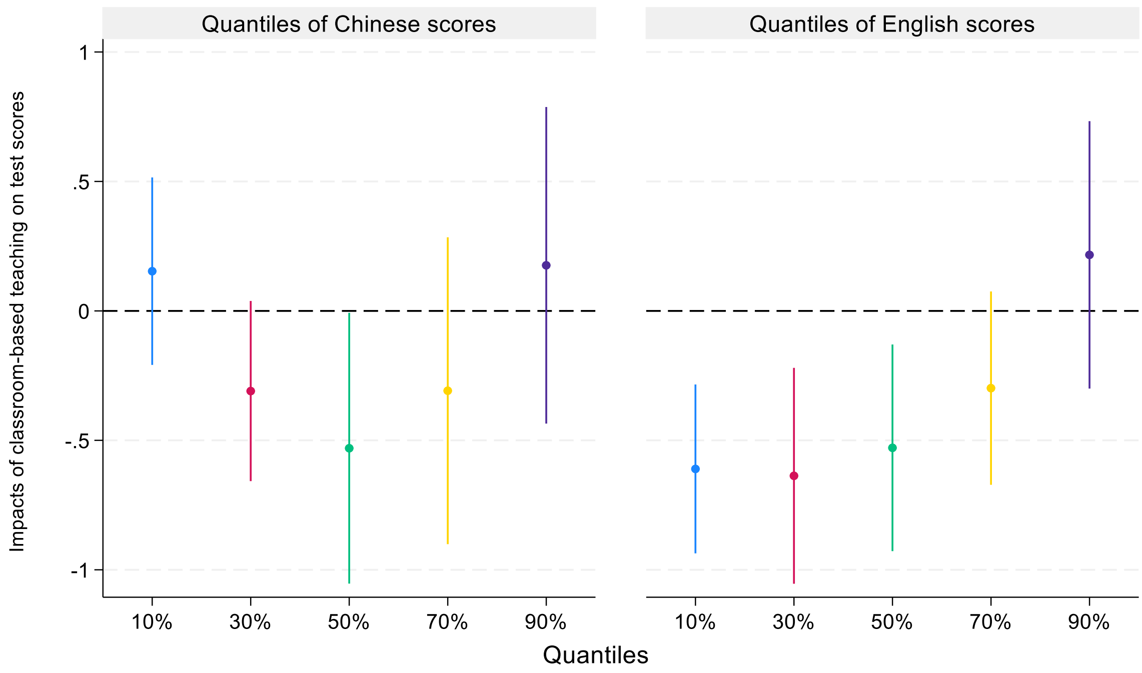
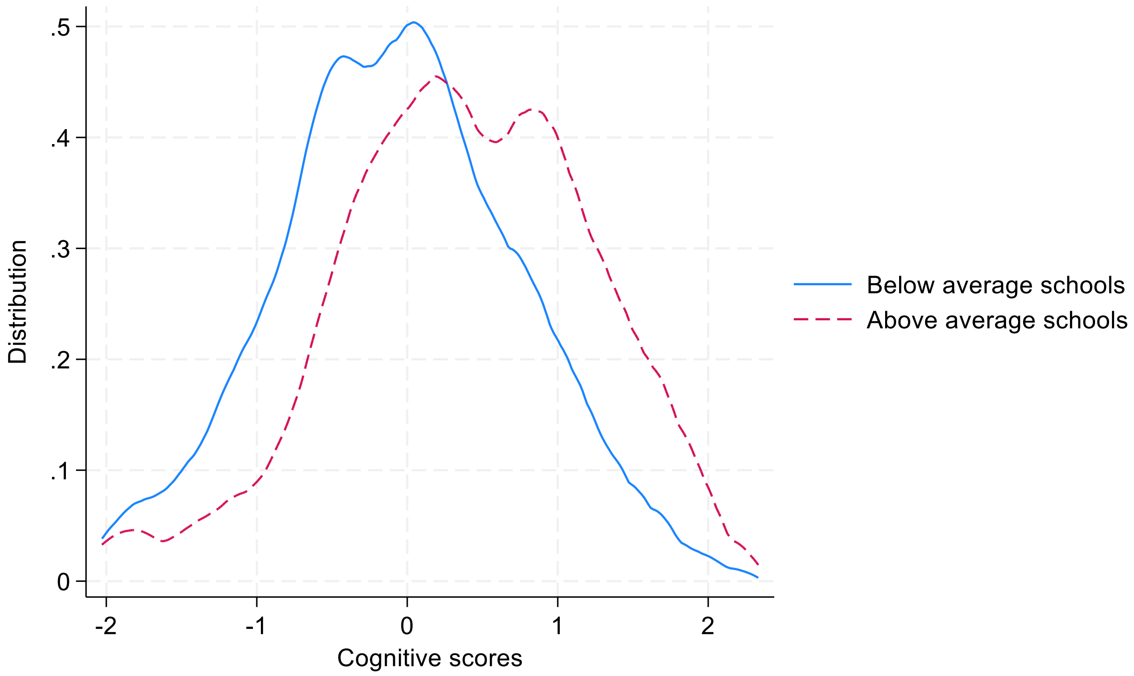
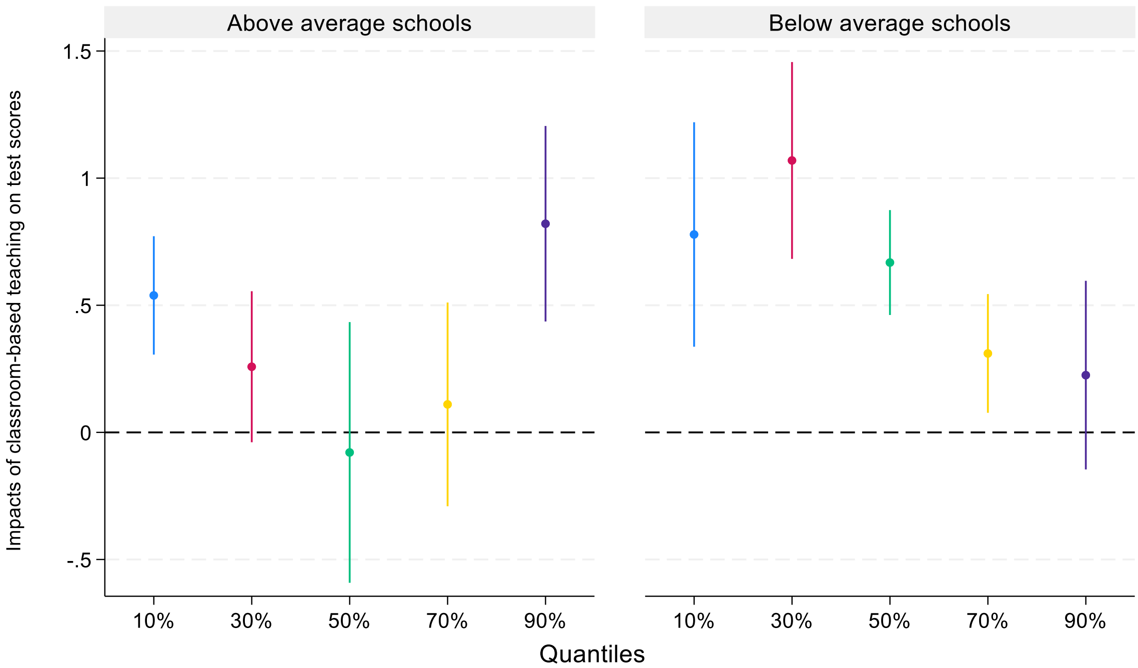


Figure A6. Heterogeneous impacts of teaching styles by average cognitive skills

Panel A, Distribution of cognitive scores between schools



Panel B, Math



Panel C, Languages

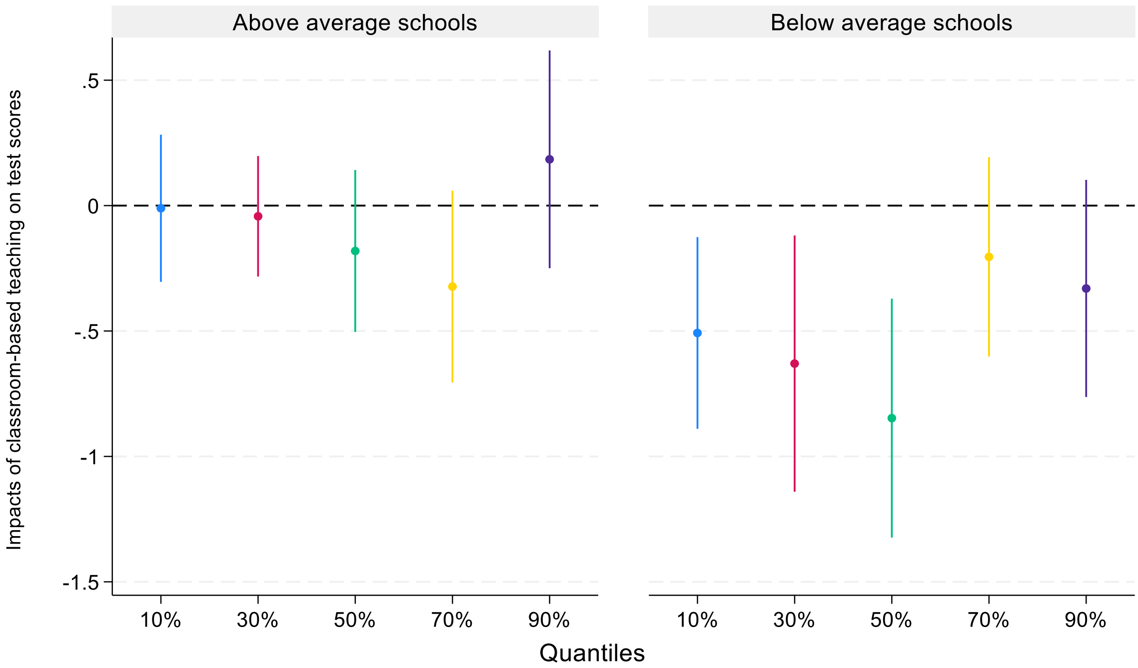


Table A1. Balancing test

	(1) Frequency of classroom-based teaching	(2) Frequency of discussion-based teaching	(3) Frequency of interactive teaching
Parental degree	-0.027	-0.051	-0.069
	(0.05)	(0.05)	(0.05)
Repeater	0.123*	-0.057	-0.077
	(0.07)	(0.09)	(0.06)
Cognitive score	-0.010	-0.011	-0.058
	(0.05)	(0.06)	(0.05)
Male	0.033	-0.020	-0.023
	(0.03)	(0.02)	(0.02)
Class size	0.156**	0.016	-0.027
	(0.07)	(0.09)	(0.07)
Rural	-0.010	0.087	0.029
	(0.06)	(0.05)	(0.06)
Endure	-0.001	-0.000	0.012
	(0.03)	(0.02)	(0.02)
Curiosity	0.010	0.008	0.022
	(0.03)	(0.02)	(0.03)
Male teacher	0.190	0.071	0.444
	(0.58)	(0.45)	(0.42)
Teacher experience	0.009	0.019	-0.037
	(0.04)	(0.03)	(0.03)
Degree (Adult Education)	-0.091	0.021	-0.863
	(1.18)	(0.65)	(0.78)
Undergraduate Degree	0.496	-0.134	-0.935
	(1.12)	(0.74)	(0.95)
Master's Degree	1.594	1.441	-1.827
	(1.80)	(2.18)	(2.19)
Second-level Teacher	0.568	-2.084	1.650*
	(0.86)	(1.62)	(0.90)
First-level Teacher	0.627	-2.376	2.224**
	(0.85)	(1.75)	(1.09)
Senior Teacher	1.041	-2.305	2.433**
	(1.08)	(1.84)	(1.08)
/			
cut1	2.517	-7.214*	-3.925
	(2.93)	(3.89)	(3.46)
cut2	4.746	-3.579	-1.676
	(2.99)	(3.78)	(3.44)
cut3	7.544**	-1.332	2.286
	(3.07)	(3.76)	(3.37)
cut4	11.793***	2.078	
	(3.12)	(3.79)	

Note: N=6,916. Ordered Logit estimates. The three dependent variables are the frequency of implementing the three teaching styles. The regressions also include school fixed effect. The standard error is clustered at the class-subject level. Omitted categories for teacher variables are female teacher, No Degree, and Third-Level Teacher (entry level). There is no teacher who has never used interactive teaching in our sample.