Economics of Education Review 78 (2020) 102034

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Can conditional grants attract better students? Evidence from Chinese teachers' colleges☆

Li Han*,[a], Jiaxin Xie[b]

[a] Division of Social Science, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong
[b] Institute for Economic and Social Research, Jinan University, Guangzhou, China

ARTICLE INFO

Keywords:
Teacher selection
Conditional transfer
Teacher quality
Financial aid
College major choice

ABSTRACT

This paper examines whether a conditional grant program for teacher trainees helped attract better students to teaching majors in Chinese universities. The Free Education for Future Teachers program implemented in top teachers' colleges provides tuition exemption and a generous stipend package to students in teaching majors under the condition that the recipients teach in their home provinces after graduation. By comparing score changes between teaching and nonteaching ("regular") majors and between program colleges and nonprogram teachers' colleges, we find that this grant program helps attract students scoring 0.4–0.5 percentiles higher in the score distribution to teaching majors. Further analysis shows that the program impact is stronger in places where more students are likely to be credit-constrained. Our results suggest that conditional grants targeting university-based teacher training programs are effective in improving the selection of teachers.

1. Introduction

Teachers make a difference. Evidence of how teacher quality impacts students' short- and long-term outcomes has been documented in both earlier literature (e.g., Hanushek, 2009; Hanushek & Rivkin, 2006) and more recent studies using richer data and more rigorous methodology (Chetty, Friedman, & Rockoff, 2014a; 2014b; Jackson, 2014; Jackson, Rockoff, & Staiger, 2014). Both empirical studies and practitioners' observations suggest that teacher effectiveness depends significantly on a number of personality traits or background characteristics prior to the teacher's entry into teaching or related training (Jacob, 2016; Corcoran and O'Flaherty, 2018). This consensus implies that improving the selection of teachers is crucial to increasing teacher effectiveness (Rothstein, 2010). As a primary route into teaching in many countries is through college- or university-based teacher training programs, various policies target applicants to such programs. An often-used policy tool is to provide financial aid, including loans, subsidies, and grants. For example, the British government has a nationwide program providing financial aid to teaching programs. Multiple state governments in the U.S. have similar programs. Such policy instruments are commonly used for hard-to-staff areas or schools. However, evidence remains scant on the effectiveness of these measures in improving the quality of prospective teachers.

The theoretical quality implications of such financial aid for teacher trainees are unclear. Career decisions depend on expected costs and benefits for many years in the future. Short-term financial aid may not significantly affect one's decision if one is not credit-constrained. Moreover, such aid is sometimes associated with commitments or conditions that limit the career choices of recipients, and such provisions are likely to dissuade high-aptitude applicants who are unwilling to make early commitments (Liu et al., 2011).

In this paper, we examine how a conditional grant program targeting teachers' colleges affects the quality of incoming teacher trainees in China. Rapid economic growth over three decades has created a vibrant private sector that increasingly attracts highly educated personnel away from the teaching profession. Moreover, entrenched economic disparities across regions make schools in low-income areas even harder to staff. To prevent the quality of the teaching force from declining, in 2007, the Chinese Ministry of Education (MOE) implemented the Free Education for Future Teachers (FEFT) program, a pilot conditional grant program in six top national teachers' colleges. All the tuition and fees of teaching majors in program colleges were waived, and in return, the respective students committed to teaching in their home provinces for ten years after graduation. The MOE has pressured provincial governments to implement similar programs in province-funded teachers' colleges since 2011. Members of the teaching profession have noted both pros and cons regarding the effectiveness of this program. While some believe that this program helps maintain a stable flow of teacher candidates, others are concerned that high-aptitude students would shun the program and that

☆ We thank Hong Kong Research Grant Council for providing funding support (GRF Project No. 645812).
* Corresponding author.
E-mail address: lihan@ust.hk (L. Han).

https://doi.org/10.1016/j.econedurev.2020.102034
Received 30 September 2019; Received in revised form 19 July 2020; Accepted 28 July 2020
Available online 19 August 2020

less motivated students would opt in for the sake of job security (Jia, Tao, & Yu, 2012; Wang, Yang, 2018). However, no rigorous study has evaluated the program.

To investigate the program's effect on the selection of teacher candidates, we draw on enrollment data from 2005 to 2009. The data contain information on both the number of students in each major enrolled from each province and the mean and maximum scores in the College Entrance Examinations (CEEs). As part of the planned economy's legacy, college majors in China are classified into very fine categories. A typical college has more than 200 majors. A quota of each major's enrollment is assigned to each of 31 provinces. As a result, only a few applicants from each province are admitted to each college major. Therefore, our data set is close to an individual-level data set and can be used to construct the distribution of CEE scores in each province for each year. To make it comparable across provinces and over time, we measure the quality of students admitted to each major using their percentiles in the distribution of CEE scores in their home province.

We first use a difference-in-differences (DID) approach to identify the program's effect on the quality of students admitted to teaching majors by exploiting the fact that the FEFT program is only implemented in teaching majors in the six program colleges from 2007 on and does not apply to regular majors. We estimate the impact of this grant program by comparing the change in the score percentiles of students in teaching majors before and after 2007 with the corresponding change in regular majors. To eliminate confounding effects of major-specific temporal trends, we further use other elite teachers' colleges as an additional comparison group and apply the difference-in-differences-in-differences (DDD) method to estimate the program effect.

Our main finding is that teaching majors indeed attract better students because of the FEFT program. There is no evidence that the program drew good students from regular majors in the program colleges or teaching majors in the nonprogram teachers' colleges in our sample. Evidence suggests that high-aptitude students drawn to the teaching majors in program colleges are likely to be those who would have attended regular colleges of similar ranks as the program colleges to study in majors similar to teaching majors in teachers' colleges but with no pedagogical training or teaching commitment.

Further analysis shows that the program effects are stronger for students in provinces with larger shares of economically disadvantaged students—those who are from rural areas, are female, and have more siblings. Taken together, these results suggest that a likely channel through which the FEFT program attracts high-aptitude students into teaching is to ease the credit constraints faced by those students.

Our study contributes to two strands of literature. First, there is a growing literature on ways of improving the selection of teachers (e.g., Rothstein, 2010, 2015; Jacob, 2016). As one of the primary routes to teaching, university-based teaching programs have recently attracted more attention from scholars and policymakers. For example, Backes, Goldhaber, Cade, Sullivan, and Dodson (2018) find that STEM teachers from UTeach, a university-based STEM teacher preparation program, are more effective in teaching secondary math and science. Our paper contributes to this literature by showing that such university-based programs can improve the selection of teachers. Second, a large body of literature has studied the effect of financial aid on college admission (e.g., Abraham & Clark, 2006; Kane, 2003; Long, 2003). Most studies focus on enrollment rates (Cornwell, Mustard, & Sridhard, 2006; Dynarski, 2000; Linsenmeier, Rosen, & Rouse, 2006; Monks, 2008). Some studies use standardized college admission exam scores (e.g., SAT, ACT), high school test scores or GPA as measures of quality. As the enrollment decision may depend on the quality of applicants, it is not always clear whether the quality effect of financial aid arises from the demand side or the supply side. In our setting, the enrollment quota is determined before the application process. Our result is best interpreted as evidence that financial aid is likely to change the composition of the applicants.

The remainder of this paper is organized as follows. Section 2 describes the institutional background of college admissions in China and the FEFT program. Section 3 describes our data and measures. Section 4 introduces the empirical strategy. We report and interpret the main results in Section 5 and explore in Section 6 the channels through which the policy effect arises. Section 7 concludes the paper.

2. Institutional background

The primary route into teaching in China entails pursuing teaching majors in teachers' colleges. The law stipulates that teacher candidates must have a teaching certificate granted by the government.[1] Graduates of teaching majors could automatically obtain this certificate in China before 2015, while students in other majors must take a series of certification trainings and exams. Schools usually prefer graduates of teaching majors for their rigorous pedagogical training and experience of serving as assistant teachers before graduation. A survey in Shandong Province in 2008 shows that 96.6% of high school teachers were graduates of teaching majors (Zhou, 2010). Most of the teachers from other sources are those who have similar levels of educational attainment and acquired the teaching certificate by taking exams.

A typical teachers' college offers majors in two tracks: teaching and regular.[2] Students in the teaching track are required to receive pedagogical training in addition to training in their fields, such as Chinese literature, English language, history, mathematics, physics, chemistry, and biology. Regular colleges also provide similar majors with the same field training but no pedagogical training.

Following *Regulations on the Admission Process of General Higher Education Institutions* stipulated by the MOE, students in both teaching and regular tracks are admitted through a centralized matching system that matches applicants with colleges and majors based on applicants' CEE scores and preferences.[3] The CEEs are annual exams administered separately by each province for applicants in the science cluster and the humanities cluster. Applicants have to choose their cluster at least one year before taking the CEEs. Before the CEEs, each college publishes a detailed enrollment plan. This plan specifies not only the quota for each major but also how the quota for each major is allocated to each cluster in each province. Some majors, e.g., economics and accounting, admit students from both science and humanities clusters. The enrollment plan is made by colleges and provincial admissions offices under the regulation and coordination of the MOE and published before college application. The allocation of the quota depends on many factors, including the capacity and the funding sources of the colleges, the relationship between colleges and provincial governments, the goals of the MOE, and so forth. Normally, the allocation of quotas is stable over time.

College applicants from each province list colleges and majors in order of preference on an application form designed by the provincial education bureau. One is only allowed to specify 3–9 colleges in each of the three tiers ordered by the level of academic prestige.[4] If one applies to teaching-track majors, the commitment is made before the admission process starts and is extremely difficult to change afterwards.

The first-tier colleges, the enrollment of which accounts for less than 10% of total college enrollment, are the most resourceful and prestigious institutions and include all the colleges affiliated with and funded by the central government and a small number of reputable provincially funded colleges. The second-tier colleges consist of the remaining public colleges funded by the provincial and/or prefectural governments, while the third-tier colleges are those funded by private organizations. Provincially funded colleges tend to allocate the lion's share of the admissions quota

[1] Source: The *Law of Teachers*, P.R. China.

[2] See *The List of Undergraduate Programs of General Higher Education Institutions*, an annual publication from the Ministry of Education.

[3] Only a very limited number of students could enter college without taking the CEEs during our sample period. We do not consider this case in our analysis.

[4] The number of colleges that students can choose in each tier ranges from 1 to 3, which is specified by the admissions office before application.

to the host provinces, whereas centrally funded colleges are obliged to recruit students from across all provinces due to equity concerns. Despite the alleged equity goal, all colleges disproportionately allocate more quotas to their host provinces (Li, 2017). The share of the enrollment quota allocated to host provinces during our sample period is 49% for provincially funded colleges and 30% for centrally funded colleges.

After the CEE scores are released, for each cluster in each college, the provincial admissions office ranks the scores of the applicants whose first choice is the same. The provincial cutoff score for each college is set such that the number of applicants above that cutoff is equal to the quota of the college assigned to that province. The applicants below the cutoff line are placed on the waiting list for their second-choice college. If the number of first-choice applicants is below the quota of a college, the office will turn to applicants who listed that college as their second choice but are not admitted by their first-choice colleges. Given the admission procedures, an applicant's admission result is determined by his/her CEE scores relative to his/her peers in the same cluster from the same home province.

Before the marketization reform in higher education began in the late 1990s, students in teaching-track majors enjoyed a tuition waiver and a monthly stipend at colleges and were posted through a mandatory allocation process upon graduation. After the marketization reform, the national government allowed teachers' colleges to charge tuitions and fees as other regular colleges do and dismantled the mandatory allocation of graduates. On the one hand, students had to search for teaching jobs by themselves. Without the advantage of low cost, teachers' colleges lost their attractiveness to high-aptitude applicants. On the other hand, it has become increasingly difficult for economically lagging regions to staff schools with qualified graduates of teachers' colleges.

To tackle the teacher-staffing problem, the national government implemented a conditional grant program in 2007 in six MOE-affiliated first-tier teachers' colleges: Beijing Normal University, Huadong Normal University, Dongbei Normal University, Huazhong Normal University, Shaanxi Normal University and Southwestern Normal University.[5] Students in teaching-track majors at program colleges are exempted from tuition and board during their four-year study and receive a monthly allowance of 400 *yuan*. Compared to students in regular majors, grant recipients save a minimum of 10,000 *yuan* (approximately 1500 in 2009 U.S. dollars) per year, which is close to the average annual income of a three-person rural household. Those grant recipients are required to teach in primary or middle schools in their home province for ten years after graduation, with the first two years being in rural schools. The mandatory durations of total and rural service changed to six years and one year, respectively, beginning in 2019. Those who do not follow the job assignment not only need to repay 150% of all the grants they have received but also have their names entered on a list of "discredited individuals" maintained and published by the government, which makes it difficult for them to secure loans or rent apartments.[6] Given this penalty, the default rate is low. Among the 2007 grant recipients, no more than five opted out of the teaching profession in most of the provinces.[7]

The conditional grant program implemented in 2007 is only a pilot experiment on a small scale. As Fig. 1 shows, the enrollment (including both teaching and regular tracks) in the six program colleges only accounts for approximately 1.6% of admissions in first-tier colleges and 20–30% of admissions in first-tier teachers' colleges between 2005 and 2009. Given that students who are eligible for first-tier colleges usually do not consider second-tier colleges as an option, we will focus on the comparison of program and nonprogram colleges in the first tier.

3. Data and measurement

Our data are mainly from two sources. The first data set is the admission data from 2005 to 2009 at a disaggregated level. The data source is *The Guide to the College Entrance Examinations*, a publication series authorized by the MOE to provide detailed information on previous years' examinations as a reference for college applicants since 2005. This collection contains data on incoming students from each of 31 provinces admitted to each major in each college every year, including the maximum and mean scores as well as the number of students admitted. Unfortunately, the information is not available for applicants from the following 6 province-years: Guangdong, Jiangsu, and Liaoning in 2005, Heilongjiang in 2007, and Jiangsu and Zhejiang in 2009. In addition, we exclude Tibet and Xinjiang from our analysis because these two provinces are subject to special policy treatment by the national government.[8] The median and mean admission counts from each province in each major of a college are 3 and 10, as shown in Table 1. Our data are close to individual-level data.

For the purpose of our study, we restrict our analysis to first-tier colleges. There are 36 teachers' colleges in the first tier, among which six are designated for the FEFT program. Our analysis focuses on these 36 teachers' colleges.

The second data set that we use contains provincial characteristics, including the main socioeconomic development indicators compiled from the *China Statistical Yearbooks* from 2005 to 2009 and measures of teachers' pay from the *China Fiscal Yearbooks*. In addition, we computed measures of demographic characteristics of senior high school students in the 2005 to 2009 graduating classes using 2005 mini-census data (covering 2% of the population). The demographic characteristics include the share of high school students who are rural and/or female, the average number of siblings that high school students have, and the average wage in the education sector. Table 1 illustrates very large regional disparities by students' home provinces.

Fig. 2 (a) illustrates a reallocation of enrollment slots from regular to teaching majors in program colleges: the enrollment of teaching majors increases steadily during the examined period and accelerates after the policy's adoption, while the enrollment of regular majors declines after the adoption of the policy in program colleges. In contrast, Fig. 2(b) shows that the trends of enrollment in both teaching and regular majors in nonprogram colleges are nearly parallel after 2007. Given that our panel of provinces is not balanced, the overall trend only reflects the enrollment in our sample. In 2007, five nonprogram teachers' colleges in our sample had abnormal expansions of enrollment due to local policy changes, which drive the spike in enrollment for the nonprogram colleges that year. As we will essentially compare teaching and regular majors within the same college for the same province, the general change in enrollment in a certain year will not affect our results as long as it has no asymmetric effects on the two tracks of majors.*Standardizing CEE scores.* Our primary measure of the academic quality of students is the CEE score. As the CEEs and the admission process are administered within individual provinces, we standardize the CEE scores by computing the corresponding percentiles of raw scores in the score distribution for each province-cluster-year, i.e., for each pool of applicants with identical CEE profiles. Our method of construction is as follows.

1. For each province-cluster-year, we know the total number of applicants[9](P) and the total enrollment in all universities (E), so we can calculate the enrollment rate ($\frac{E}{P}$).
2. We collected the mean raw score and the number of students from

[5] State Council [2007] No. 30.

[6] The program became more flexible after an amendment in 2017, which allows students to transfer to the regular track within the first year of college provided that they repay all the grants.

[7] Source: http://www.eol.cn/zt/201205/jiaoshishifanbk/, published on *Education Online* – a website affiliated with the MOE.

[8] The results remain qualitatively unchanged when (1) Xinjiang and Tibet are included or (2) Guangdong, Jiangsu, Zhejiang and Heilongjiang are excluded for all years.

[9] The numbers are obtained from news reports. We also cross-check the total number of applicants for each province-year using the *Yearbook of Education and Examinations*.

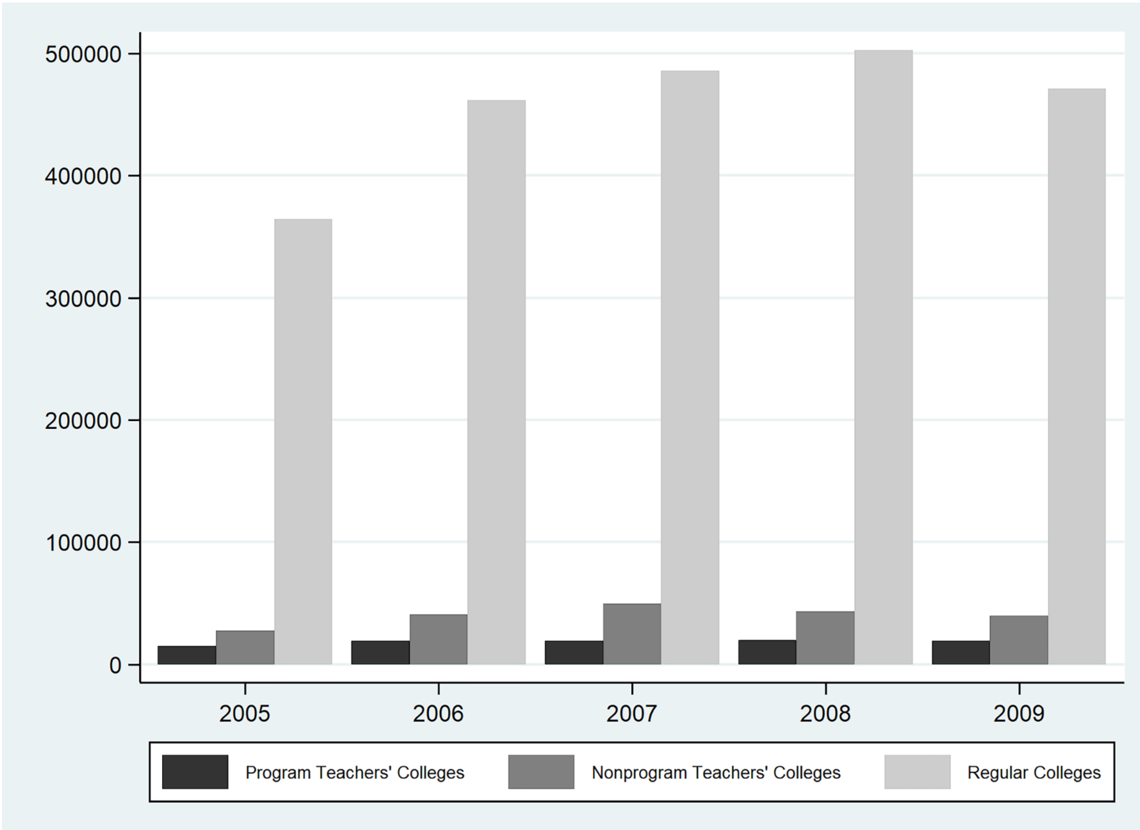

Fig. 1. Enrollment of first-tier colleges by college type, 2005–2009.

Table 1
Summary statistics.

Variable	Obs.	Median	Mean	Std. dev.	Min	Max
Panel A: CEE score and enrollment						
Standardized mean score	30,100	97.29	96.37	3.47	61.74	99.99
Standardized maximum score	30,100	98.15	97.24	2.93	65.88	100.00
Number of admitted students	30,100	3	9.79	25.49	1	628
Panel B: Share of high school students who are						
rural	30,100	0.69	0.62	0.17	0.16	0.79
rural female	30,100	0.30	0.29	0.09	0.08	0.45
Panel C: Provincial average number of siblings of						
high school students	30,100	1.11	1.02	0.42	0.08	1.92
rural high school students	30,100	1.37	1.32	0.48	0.17	2.91
rural female high school students	30,100	1.49	1.42	0.48	0.19	3.09
Panel D: Provincial development, 5-year average						
GDP per Capita	30,100	10579.2	14481.3	9222.6	4707.3	491600
Average wage of schoolteachers	30,100	12005	13294.7	5320.8	8278.7	32567.4
Average wage in the education sector	30,100	16573.6	19550.5	9180	12421.9	64713.1

each province admitted to each major in each university for all universities. Using this information, we construct a proxy score distribution for all universities. Then, we can determine the percentile of a specific score in this corresponding distribution (S).

3. Assuming that the upper tail of the score distribution of all applications in each province-cluster-year contains only those admitted to universities, we can calculate the percentile of a specific score in the entire distribution of applicants using the following equation:

$$Y = (1 - \frac{E}{P}) + s\% \cdot \frac{E}{P} \tag{1}$$

We will examine both the mean and the maximum scores. While the mean score can be used as a proxy for the average quality of the incoming students, it might be affected by changes in the enrollment quota. If the increase in the quota leads to a lower cutoff, the mean scores will decrease. Therefore, we use the maximum score as an alternative measure, which proxies for the quality of the top-performing student admitted to each major from a specific province and is not subject to the influence of the quota. Plots 3(a) and 3(b) in Fig. 3 exhibit the trends of the mean and the maximum scores, respectively, in program colleges. The mean and maximum percentiles of teaching and regular majors in program colleges diverge in 2007, when both scores of teaching majors exceed those of regular majors. In contrast, the scores of these two types of majors in nonprogram colleges follow one another closely throughout the sample period (Fig. 3(c) and (d)).

Table 2 reports the mean and maximum scores for teaching majors and regular majors in the sample teachers' colleges before and after 2007. Note that the scores for program colleges are 1.3–1.8 percentiles

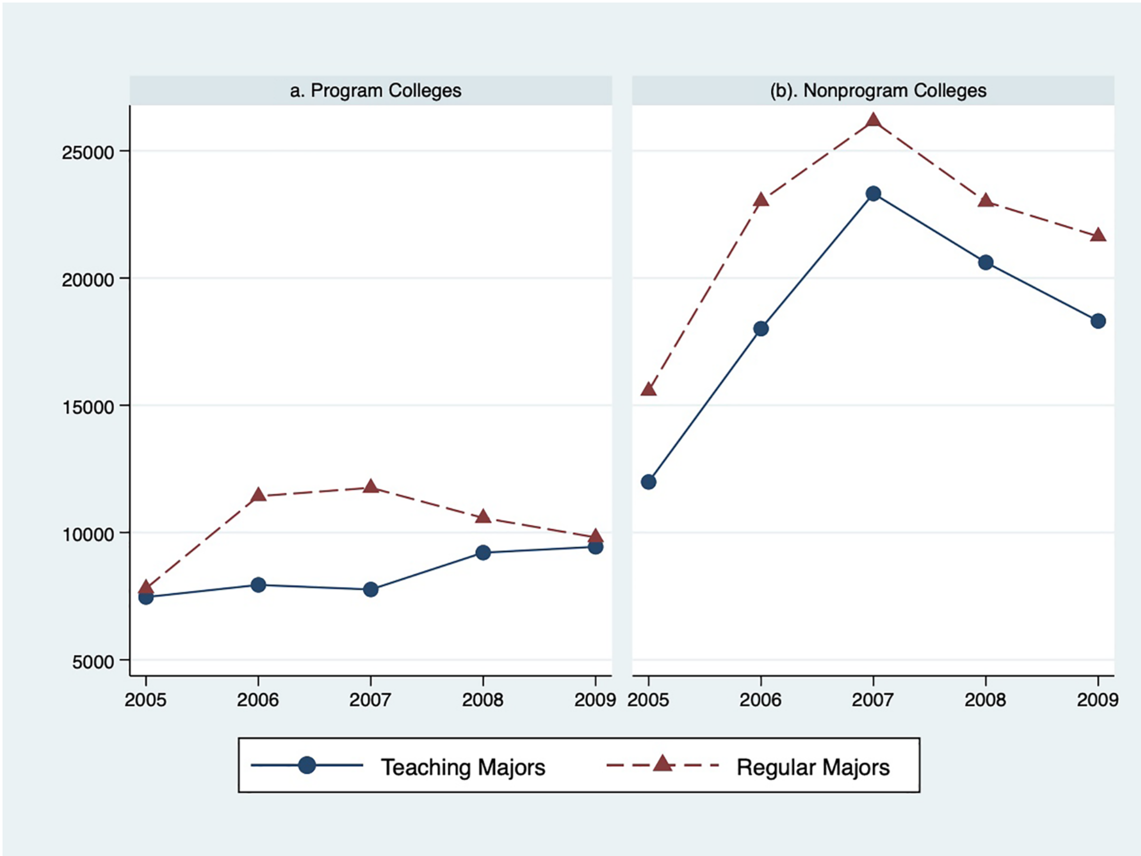

Fig. 2. Enrollment changes by major and college type, 2005–2009.

higher than those for nonprogram colleges. The mean scores in program colleges are approximately in the 97 percentile, while those in nonprogram colleges are in approximately the 95.2 percentile. We compare in Panel A the changes in standardized scores within the program colleges: for teaching majors, there is a 0.387 percentage point increase in the mean and 0.347 in the maximum after the introduction of the grant program, while for regular majors, decreases with magnitudes of 0.238 percentage points in the mean and 0.284 in the maximum are observed. Overall, relative to those of regular majors, the scores of teaching majors increase by 0.625 percentage points in the score distributions. In contrast, the same comparison for nonprogram colleges shows no significant difference in the score changes before and after 2007 between teaching and regular majors.

4. Identification strategy

To isolate the program effect, we first apply the difference-in-differences (DID) method to the sample of program colleges. As the enrollment size of the teaching majors in program colleges is small compared to the potential pool of qualified applicants, the program is not very likely to affect regular majors. We compare the score changes after program implementation in the teaching majors with those in the regular majors. The regression is specified as follows:

$$\begin{aligned} Y_{ijkst} &= \alpha_0 + \alpha_1 Teach_{ijkst} + \alpha_2 Teach_{ijkst} \times post \\ &\quad + \beta_1 quota_{ijkst} + \beta_2 quota^2_{ijkst} + \theta_I + \mu_{jkst} + \epsilon_{ijkst} \end{aligned} \tag{2}$$

where Y_{ijkst} is the percentile of the maximum or mean score of students from province s and cluster k admitted to major i in college j at time t, $Teach_{ijkst}$ is an indicator of teaching majors that takes value 1 if major i in college j admitting students from province s and cluster k in year t is a teaching major and 0 otherwise, $post$ is an indicator of the introduction of the grant program that takes value 1 before the year 2007 and 0 otherwise, and $quota_{ijkst}$ is the number of students admitted to major i from cluster k of province s in year t. We also control for subdiscipline fixed effects θ_I and cluster-province-college-year fixed effects μ_{jkst}. These fixed effects capture not only the unobserved characteristics of each pool of students taking the same exams but also the unobserved time-varying popularity of each college in different provinces. All the standard errors are clustered at the province level.

The assumption underlying the DID model above is that the scores of admitted students in teaching and regular majors would have followed the same trends had there been no grant program, i.e., the gap in the scores between teaching and regular majors would have remained the same over time had the program not been introduced. Under this assumption, α_2 in regression (2) captures the program effect on the scores of incoming students. One may question the validity of this assumption. For example, the willingness of students to join the teaching profession as opposed to other professions may be increasing or decreasing over time and is likely affected by changes in the labor market. To address this concern, we will examine the 32 nonprogram teachers' colleges in the first tier, which are comparable with the program colleges in terms of academic rankings and offer both teaching and regular majors.

We first test whether the teaching majors have a different trend in scores than the regular majors. We do so by comparing the score changes of the teaching majors in program colleges with those in the nonprogram teachers' colleges. The specification is similar to Eq. (2), except that the control group is the teaching majors in nonprogram colleges. If teaching majors in nonprogram colleges experience a similar increase in the scores, the DID estimate should not be significantly different from 0.

To further account for potential major-specific trends, we compare the differences in score changes of teaching and regular majors in

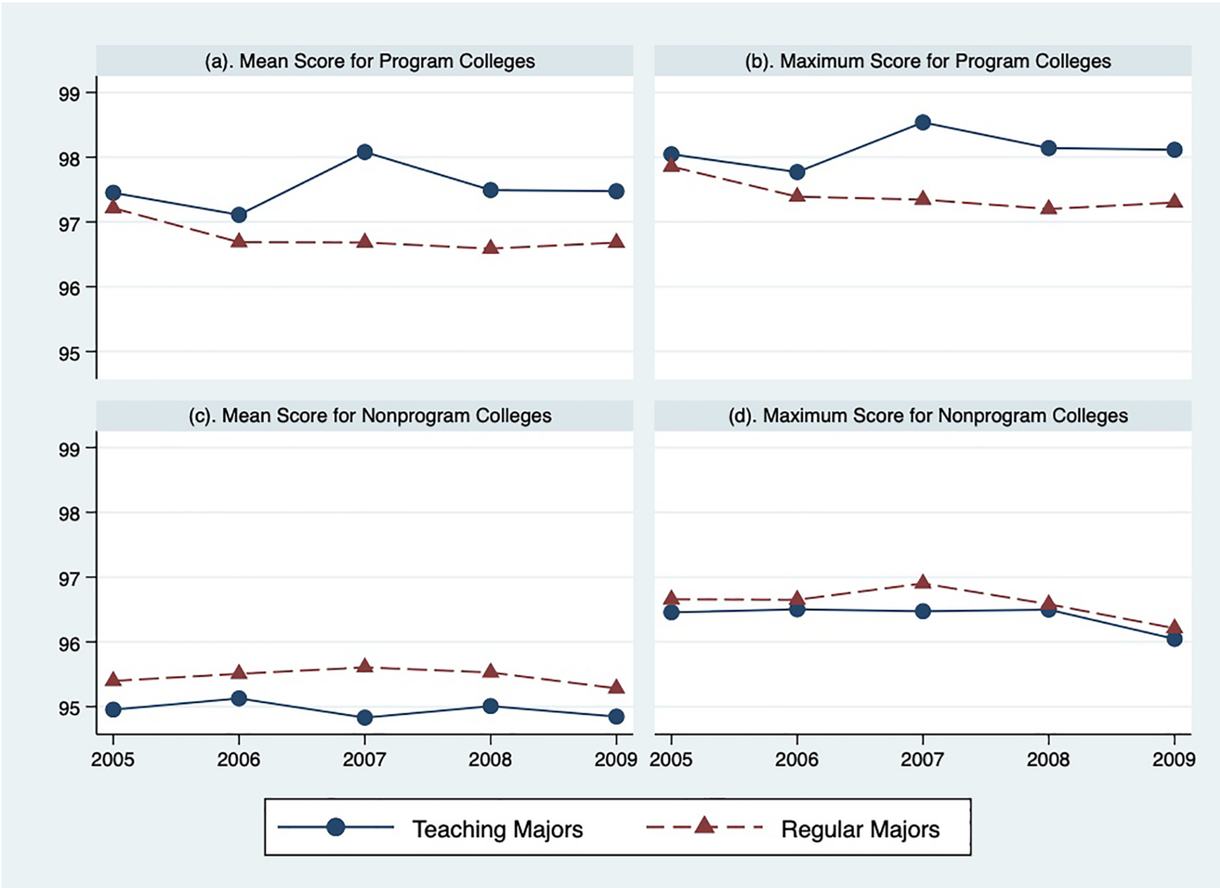

Fig. 3. Quality changes by major and college type, 2005–2009.

Table 2
Score percentiles by cohort, major and college.

	Standardized mean score		Standardized maximum score	
	Teaching major (1)	Regular major (2)	Teaching major (3)	Regular major (4)
Panel A: Program colleges				
Pre-policy	97.274	96.888	97.903	97.567
	(0.055)	(0.047)	(0.049)	(0.042)
Post-policy	97.661	96.650	98.250	97.282
	(0.038)	(0.036)	(0.032)	(0.021)
Difference	0.387**	−0.238**	0.347**	−0.284**
	(0.065)	(0.059)	(0.057)	(0.053)
Difference-in-differences	0.625**		0.632**	
	(0.090)		(0.080)	
Panel B: Nonprogram colleges				
Pre-policy	95.054	95.461	96.483	96.654
	(0.109)	(0.079)	(0.087)	(0.071)
Post-policy	94.899	95.461	96.332	96.539
	(0.082)	(0.052)	(0.063)	(0.055)
Difference	−0.154	−0.000	−0.151	−0.114
	(0.142)	(0.092)	(0.109)	(0.082)
Difference-in-differences	−0.154		−0.037	
	(0.162)		(0.134)	

Notes: Standard errors are shown in parentheses. Significance levels are indicated as follows: * $p < 0.05$, and ** $p < 0.001$. Constant terms are not reported.

program colleges with the differences in the respective majors in non-program colleges. We specify the following difference-in-differences-in-differences (DDD) model:

$$\begin{aligned} Y_{ijkst} = {} & \beta_0 + \beta_1 Teach_{ijkst} + \beta_2 Teach_{ijkst} \times post + \beta_3 Teach_{ijkst} \\ & \times ProgramCollege_j \\ & + \beta_4 post \times ProgramCollege_j \times + \beta_5 Teach_{ijkst} \times ProgramCollege_j \\ & \times post \\ & + \gamma_1 quota_{ijkst} + \gamma_2 quota^2_{ijkst} + \theta_I + \mu_{jkst} + \epsilon_{ijkst} \end{aligned} \tag{3}$$

where $ProgramCollege_j$ is an indicator of program colleges that takes value 1 if college j is a program college and 0 otherwise. All standard errors are clustered at the province level. The coefficient β_5 in Eq. (3) is thus the DDD estimate of the program effect.

5. Main results

5.1. Baseline results: DID estimates

Table 3 presents the DID estimates of Eq. (2). Panels A and B show the regression estimates obtained using the mean and maximum score percentiles, respectively, as the outcome variable. The estimates in column (1) show that compared to all regular majors, teaching majors exhibit a 0.306 percentage point increase in the mean score percentile because of the grant program, and the estimate is statistically significant at the 5% level. Furthermore, column (7) in Panel B shows that the program effect on the maximum score is 0.351 percentage points and that the estimate is statistically significant at the 1% level.

One may be concerned that the variety of regular majors is too broad for them to constitute a reasonable control group for teaching majors that only exist in seven of the total of 12 disciplines. To address this concern, we exclude regular majors in the other five disciplines from the control group and reestimate Eq. (2). The result is reported in

Table 3
DID effects on standardized scores in program colleges.

	All enrollment			Nonlocal enrollment		
	All majors	Similar disciplines	Similar majors	All majors	Similar disciplines	Similar majors
	Panel A: Dependent variable: standardized mean score					
	(1)	(2)	(3)	(4)	(5)	(6)
Teaching major (TM)	0.089	0.062	0.072	0.065	0.041	0.050
	(0.070)	(0.058)	(0.057)	(0.073)	(0.059)	(0.056)
Post*TM	0.306*	0.348**	0.374*	0.376**	0.413**	0.439**
	(0.116)	(0.120)	(0.148)	(0.101)	(0.104)	(0.111)
ln(*Admission*)	0.042	0.053	0.028	0.152	0.167	0.156
	(0.087)	(0.086)	(0.078)	(0.113)	(0.110)	(0.120)
ln(*Admission*)^2	−0.030	−0.035	−0.032	−0.089*	−0.094*	−0.094*
	(0.028)	(0.027)	(0.024)	(0.042)	(0.043)	(0.044)
No. of observations	18,950	16,964	11,989	17,187	15,404	11,002
R-squared	0.868	0.868	0.867	0.879	0.879	0.881
	Panel B: Dependent variable: standardized maximum score					
	(7)	(8)	(9)	(10)	(11)	(12)
Teaching major (TM)	0.018	−0.007	−0.005	0.004	−0.018	−0.009
	(0.075)	(0.065)	(0.067)	(0.077)	(0.064)	(0.064)
Post*TM	0.351**	0.399**	0.442**	0.379**	0.419**	0.450**
	(0.092)	(0.101)	(0.121)	(0.095)	(0.101)	(0.118)
ln(*Admission*)	0.572**	0.578**	0.553**	0.707**	0.723**	0.713**
	(0.114)	(0.106)	(0.101)	(0.182)	(0.177)	(0.172)
ln(*Admission*)^2	−0.061*	−0.067*	−0.062*	−0.132*	−0.140*	−0.136*
	(0.027)	(0.025)	(0.024)	(0.056)	(0.056)	(0.051)
No. of observations	18,939	16,951	11,977	17,176	15,391	10,990
R-squared	0.814	0.816	0.815	0.827	0.828	0.830

Notes: All controls and fixed effects as in Eq. (2). Standard errors are shown in parentheses and are clustered at the province level. Significance levels are indicated as follows: ** $p < 0.01$, and * $p < 0.05$. Constant terms are not reported.

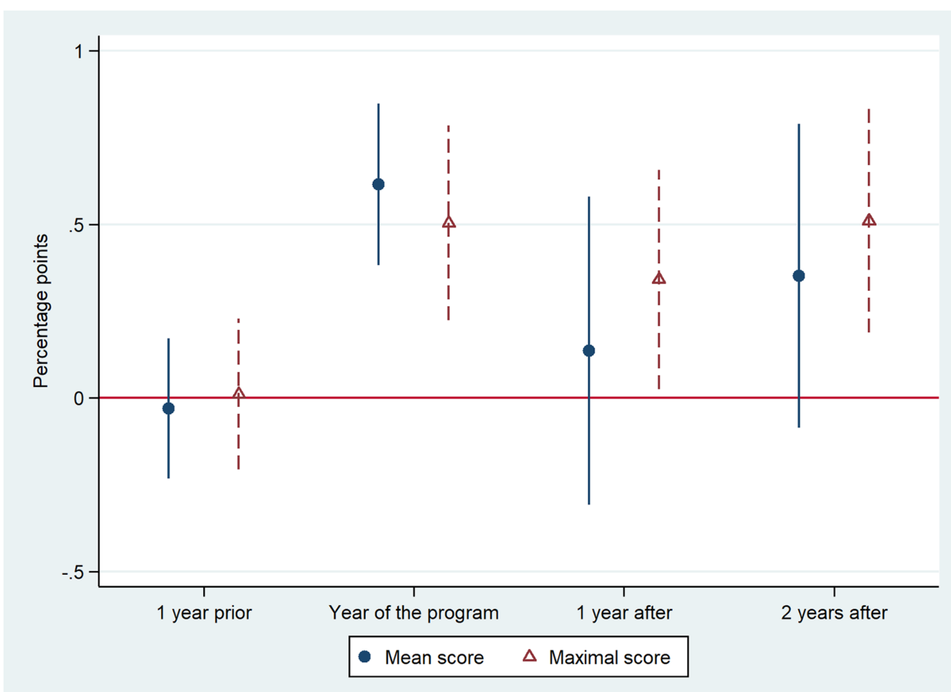

Fig. 4. DID trend, 2005–2009. Note: This figure displays the estimated coefficients of the event study model. The base year is two years prior to the program year, i.e. 2005.

column (2) of Table 3. The estimated program effect remains robust. We further refine the regression by restricting our sample to 14 of the 92 total subdisciplines with both regular and teaching majors. The estimated program effects become even stronger, and the magnitudes are slightly larger (column (3) of Table 3). These estimates show that compared to regular majors in the same subdisciplines within the same college, teaching majors attracted students ranked on average 0.374 percentage points higher after the FEFT program was introduced.

As mentioned in Section 2, colleges allocate more of the enrollment quota to their host provinces. One may be concerned that the program only benefits colleges' host provinces. We further examine whether the program effects mainly arise from colleges' host provinces. To this end, we exclude enrollment in the colleges' host provinces from the sample and replicate the results in columns (1)–(3) of Table 3. The new results, reported in columns (4)–(6) of Table 3, are consistent with those using the whole sample (columns (1)–(3)) and are of slightly greater magnitude. This finding suggests that the program does not only benefits colleges' host provinces.

5.2. Adjusting for the major-specific time trends

One threat to the DID approach is that the potential trends in scores would have differed between teaching and regular majors. To address this concern, we construct several checks.

We first examine whether the pre-existing trends in the scores differ between the teaching and regular majors in program colleges. To do so, we extend Eq. (2) by replacing the indicator "*post*" with a vector of indicators for each year. Fig. 4 shows the estimated coefficients and the 95% confidence intervals of the interactions between these indicators and the indicator for program colleges. This figure shows that the estimated program effects on the mean and maximum scores only appear in and after the year of program was implemented, which suggests that pre-existing trends are not different in these two tracks of majors.

Another concern is that teaching majors experience different trends than regular majors around the time when the program was implemented. If this is the case, the scores of incoming students in the teaching majors in nonprogram colleges move together with those in program colleges. To test for this alternative explanation, we compare the score changes of the teaching majors between program and nonprogram colleges using a specification similar to Eq. (2). The results, reported in columns (1) and (2) of Table 4, show that the increases in the mean and maximum scores of the teaching majors in the program colleges are 0.56 and 0.34 percentage points higher, respectively, than those in the nonprogram colleges. The magnitude of these estimates is similar to that in Table 4, suggesting that the estimated program effects in Table 3 are not driven by a general change in the teaching profession.

Does the grant program draw good students from teaching majors in nonprogram colleges to program colleges? This is unlikely to be the case because the nonprogram teacher colleges have relatively lower, albeit comparable, scores and academic rankings than the program colleges. Nevertheless, we construct a test for this hypothesis by applying the DID model (Eq. (2)) to compare the teaching and regular majors in nonprogram colleges. The results for the mean and maximum scores, reported in columns (3) and (4) of Table 4, respectively, show no significant differences in the change in scores between teaching and regular majors in nonprogram colleges. This finding lends further confidence that the assumption underlying the DID model (Eq. (2)) holds.

Table 4
Alternative DID models.

	Program vs. nonprogram colleges		Teaching vs. regular majors	
	Teaching majors only		in nonprogram colleges	
	Mean (1)	Max (2)	Mean (3)	Max (4)
Post*Program colleges	0.555** (0.200)	0.339* (0.140)		
Teaching major (TM)			−0.015 (0.089)	−0.004 (0.052)
Post*TM			−0.042 (0.077)	0.102 (0.068)
ln(*Admission*)	−0.133* (0.055)	0.394** (0.093)	−0.041 (0.074)	0.773** (0.095)
ln(*Admission*)^2	−0.017 (0.016)	0.005 (0.016)	−0.019 (0.014)	−0.038* (0.016)
No. of observations	12,064	12,054	11,064	11,067
R-squared	0.844	0.782	0.902	0.840

Notes: Robust Standard errors in parentheses, clustered at province level; ** $p < 0.01$, * $p < 0.05$.

To account for potential track-specific trends, we further apply the DDD model as specified in Eq. (3). Columns (1)–(3) of Table 5 present the regression results using all majors, majors in the seven disciplines with both teaching and regular majors, and majors in subdisciplines with both teaching and regular majors. As shown in column (3), the increase in the percentiles of teaching majors relative to regular majors in program colleges is 0.545 percentage points higher than that in nonprogram colleges. We further restrict our sample to nonlocal enrollment and reestimate the three regressions in columns (1)–(3). The results, shown in columns (4)-(6), remain robust. The magnitude of the DDD estimates is slightly larger than that of the DID estimates of the program effects discussed in Section 5.1.

Overall, these results suggest that the program successfully attracts academically more capable students to teaching majors. As approximately 10 million students take the CEEs each year, our estimates suggest that the conditional grant on average can attract to teaching majors students who are ranked 54,500 places higher than those who would have been admitted to teaching majors without such programs.

5.3. Does the program draw students from regular majors?

Does the program draw students from regular majors to teaching majors in the same colleges? We address this question by comparing the score changes of regular majors in program colleges versus nonprogram colleges. The results, presented in Table 6, show no significant differences in regular majors between those two types of colleges. In other words, the grant program did not lead to significant changes in the quality of students in regular majors in program colleges. A possible reason is that the enrollment of these six program colleges is relatively small compared to the potential qualified student pool and that the grant program did not change the general equilibrium. This test lends further support to the validity of our identification strategy.

Where the better teacher candidates in program colleges come from remains an open question. A likely source of those candidates are those who would have gone to regular colleges of similar ranks to study in majors that have similar field training as the teaching majors but no teaching commitment. To check whether this is the case, we examine first-tier regular colleges. We define as teaching-like majors with similar field training as teaching majors, including mainly Chinese literature, mathematics, English literature, history, Chemistry, Biology, and so forth. We compare the score changes in teaching-like majors with those in other majors in first-tier regular colleges using the DID model (2). The results on the mean and maximum scores are presented in panels A and B of Table 7, respectively. The results of regressions in columns (1)–(3) are estimated using all majors, majors in similar disciplines, and majors with similar subdisciplines. The results consistently show that scores in the teaching-like majors have a greater decrease than those in other majors in the sample. The results excluding local enrollment (columns (4)–(6)) exhibit the same pattern.

Taken together, these findings suggest that the FEFT program attracts better students who would have gone to regular colleges of similar ranks as program colleges to study majors similar to teaching majors. The grant program attracts people who would have majored in similar fields into the teaching profession.

6. Which province benefits from the grant program?

A major goal of the FEFT program is to mitigate the teacher-staffing problem in economically advantaged areas. However, students from those areas are less willing to return to their home provinces. It is worth questioning whether economically disadvantaged provinces experience greater program effects. To address this question, we divide the sample into two groups based on the GDP per capita of students' home provinces. We apply both the DID and DDD models separately to subsamples of students from provinces with low and high GDP per capita.

Panel A of Table 8 reports the results for the mean scores. The DID results presented in column (1) show that compared to students in

Table 5
DDD effects on standardized scores.

	All enrollment			Nonlocal enrollment		
	All majors	Similar disciplines	Similar majors	All majors	Similar disciplines	Similar majors
	Panel A: Dependent variable: standardized mean score					
	(1)	(2)	(3)	(4)	(5)	(6)
Teaching major (TM)	−0.018	−0.021	0.085	0.016	0.004	0.145
	(0.080)	(0.083)	(0.090)	(0.081)	(0.081)	(0.112)
Post*TM	−0.041	−0.036	−0.133	0.048	0.051	−0.057
	(0.076)	(0.078)	(0.086)	(0.070)	(0.069)	(0.097)
Program college (PC)*TM	0.108	0.085	−0.028	0.076	0.066	−0.077
	(0.090)	(0.092)	(0.105)	(0.108)	(0.100)	(0.112)
Post*TM*PC	0.367*	0.402*	0.545**	0.325*	0.356**	0.494**
	(0.146)	(0.153)	(0.189)	(0.128)	(0.128)	(0.148)
ln(*Admission*)	0.021	0.014	−0.042	0.090	0.105	0.053
	(0.052)	(0.049)	(0.049)	(0.078)	(0.082)	(0.084)
ln(*Admission*)^2	−0.028*	−0.026*	−0.018	−0.068*	−0.071*	−0.062
	(0.012)	(0.013)	(0.011)	(0.030)	(0.032)	(0.031)
No. of observations	30,014	27,235	19,271	24,515	22,241	15,756
R-squared	0.891	0.895	0.900	0.895	0.896	0.903
	Panel B: Dependent variable: standardized maximum score					
	(7)	(8)	(9)	(10)	(11)	(12)
Teaching major (TM)	−0.035	−0.033	0.069	−0.003	−0.014	0.173
	(0.059)	(0.063)	(0.073)	(0.090)	(0.095)	(0.121)
Post*TM	0.098	0.095	0.005	0.065	0.075	−0.075
	(0.067)	(0.067)	(0.077)	(0.084)	(0.084)	(0.095)
Program college (PC)*TM	0.053	0.027	−0.082	0.035	0.027	−0.166
	(0.082)	(0.083)	(0.082)	(0.120)	(0.119)	(0.127)
Post*TM*PC	0.242*	0.288*	0.422**	0.307*	0.336*	0.521**
	(0.108)	(0.114)	(0.137)	(0.124)	(0.128)	(0.148)
ln(*Admission*)	0.550**	0.544**	0.500**	0.732**	0.748**	0.694**
	(0.083)	(0.083)	(0.078)	(0.139)	(0.142)	(0.132)
ln(*Admission*)^2	−0.018	−0.017	−0.009	−0.130**	−0.137**	−0.125**
	(0.013)	(0.013)	(0.012)	(0.042)	(0.044)	(0.039)
No. of observations	30,006	27,225	19,262	24,502	22,226	15,742
R-squared	0.830	0.834	0.838	0.850	0.853	0.864

Notes: All controls and fixed effects as in equation (3). Standard errors are shown in parentheses and are clustered at the province level. Significance levels are indicated as follows: ** $p < 0.01$, and * $p < 0.05$. Constant terms are not reported.

Table 6
Policy effects on regular majors.

	Standardized mean score		Standardized maximum score	
	All (1)	Nonlocal (2)	All (3)	Nonlocal (4)
Post*Program colleges	0.124	−0.110	0.028	−0.143
	(0.182)	(0.277)	(0.127)	(0.214)
ln(*Admission*)	0.084	0.190*	0.681**	0.907**
	(0.052)	(0.086)	(0.084)	(0.152)
ln(*Admission*)^2	−0.056*	−0.086*	−0.046**	−0.159**
	(0.020)	(0.039)	(0.015)	(0.048)
No. of observations	18,010	14,596	18,015	14,597
R-squared	0.824	0.829	0.768	0.783

Notes: Robust standard errors are shown in parentheses and are clustered at the province level. Significance levels are indicated as follows: ** $p < 0.01$, and * $p < 0.05$. Constant terms are not reported.

regular majors, the students in teaching majors exhibit a 0.307 percentage point increase in mean score after the implementation of the program for students from provinces with low GDP per capita. In contrast, column (2) shows that the program effect is statistically insignificant for students from provinces with high GDP per capita. However, the DDD results in columns (3) and (4) show that the positive and significant program effect is only robust for students from provinces with high GDP per capita after controlling for the major-specific trends.

Panel B of Table 8 reports the results for the maximum score. The DID results in columns (1) and (2) show positive program effects on students from provinces with both low and high GDP per capita, while the DDD results in columns (3) and (4) show that provinces with high GDP per capita see greater program effects: relative to those of students in regular majors, the maximum scores of students in teaching majors increased by 0.756 percentage points.

The combined results of the DID and DDD models show that the grant program significantly improves the scores of students attracted to teaching majors relative to those in regular majors, especially for those from relatively wealthy provinces. The program effects on the scores of students from relatively poor provinces are less pronounced in the DDD estimates.

If the program effects arise mainly from the reduction in the individual cost of college education, we would expect to observe greater impacts in locations with more students who are likely to be credit-constrained. Given the high urban-rural income inequality, rural households tend to be more credit-constrained, especially those with more children. Moreover, the preference for sons remains strong in China, especially in rural areas. Parents likely invest more in their sons relative to daughters, other things being equal. Therefore, female students from rural areas who have more siblings are likely to be economically disadvantaged and more sensitive to cost in choosing colleges and majors. Therefore, we utilize the share of high school students who from rural areas and/or are female and the average number of siblings of students in each province as proxies for the share of students who are potentially credit-constrained. We enhance Eq. (3) by including these proxies, the interactions between these proxies and the triple-difference term (*Post*Teaching*Program*).

Table 7
Regular majors in first-tier regular colleges: teaching-like vs. nonteaching-like.

	All enrollment			Nonlocal enrollment		
	All majors	Similar disciplines	Similar majors	All majors	Similar disciplines	Similar majors
	Panel A: Dependant variable: standardized mean score					
	(1)	(2)	(3)	(4)	(5)	(6)
Teaching-like major (TM)	−0.114**	−0.133**	−0.163**	−0.104**	−0.124**	−0.156**
	(0.017)	(0.020)	(0.022)	(0.018)	(0.020)	(0.022)
Post*TM	−0.072**	−0.046**	0.010	−0.077**	−0.052**	0.005
	(0.018)	(0.016)	(0.016)	(0.019)	(0.016)	(0.016)
ln(*Admission*)	0.083**	0.054*	0.036*	0.040	0.020	0.020
	(0.022)	(0.020)	(0.018)	(0.024)	(0.020)	(0.022)
ln(*Admission*)^2	−0.002	−0.006	−0.016**	0.018*	0.010	−0.009
	(0.005)	(0.006)	(0.006)	(0.008)	(0.007)	(0.008)
No. of observations	446,531	306,306	134,973	414,894	284,260	124,927
R-squared	0.880	0.891	0.905	0.877	0.888	0.903
	Panel B: Dependant variable: standardized maximum score					
	(7)	(8)	(9)	(10)	(11)	(12)
Teaching-like major (TM)	−0.101**	−0.115**	−0.134**	−0.093**	−0.108**	−0.129**
	(0.017)	(0.018)	(0.021)	(0.017)	(0.019)	(0.022)
Post*TM	−0.073**	−0.049**	−0.006	−0.079**	−0.056**	−0.011
	(0.019)	(0.016)	(0.019)	(0.021)	(0.017)	(0.020)
ln(*Admission*)	0.627**	0.580**	0.540**	0.657**	0.626**	0.585**
	(0.070)	(0.065)	(0.061)	(0.078)	(0.073)	(0.070)
ln(*Admission*)^2	−0.029**	−0.031**	−0.038**	−0.046**	−0.056**	−0.061**
	(0.009)	(0.009)	(0.009)	(0.016)	(0.016)	(0.016)
No. of observations	446,531	306,306	134,973	414,894	284,260	124,927
R-squared	0.838	0.850	0.868	0.840	0.852	0.870

Notes: All controls and fixed effects as in Eq. (2). Standard errors in parentheses, clustered at province level; ** $p < 0.01$, * $p < 0.05$.

Table 8
Heterogeneous policy effects by student home province's development.

	Grouping by provincial per-capita GDP			
	Low	High	Low	High
	Panel A: Standardized mean score			
	(1)	(2)	(3)	(4)
Teaching major (TM)	0.133	−0.003	0.074	0.090
	(0.080)	(0.092)	(0.114)	(0.141)
Post*TM	0.307*	0.470	0.008	−0.250
	(0.112)	(0.280)	(0.075)	(0.140)
Program college (PC)*TM			0.073	−0.151
			(0.094)	(0.183)
Post*TM*PC			0.280	0.809*
			(0.136)	(0.350)
ln(*Admission*)	−0.053	0.138	−0.051	−0.054
	(0.098)	(0.113)	(0.058)	(0.077)
ln(*Admission*)^2	−0.020	−0.049	−0.003	−0.028
	(0.029)	(0.040)	(0.013)	(0.014)
No. of observations	6437	5552	9890	9381
R-squared	0.858	0.861	0.905	0.892
	Panel B: Standardized maximum score			
	(5)	(6)	(7)	(8)
Teaching major (TM)	0.070	−0.091	0.021	0.104
	(0.088)	(0.110)	(0.131)	(0.080)
Post*TM	0.290**	0.645**	0.115	−0.089
	(0.090)	(0.216)	(0.090)	(0.117)
Program college (PC)*TM			0.055	−0.214
			(0.112)	(0.123)
Post*TM*PC			0.134	0.756**
			(0.092)	(0.249)
ln(*Admission*)	0.433**	0.712**	0.432**	0.611**
	(0.120)	(0.159)	(0.099)	(0.121)
ln(*Admission*)^2	−0.056	−0.064	−0.015	−0.010
	(0.028)	(0.037)	(0.019)	(0.017)
No. of observations	6426	5551	9879	9383
R-squared	0.791	0.815	0.843	0.830

Notes: The results are based on the most refined sample, i.e., majors in sub-disciplines with both teaching and regular majors. All controls and fixed effects are as in equation (3). Scores' standard errors are shown in parentheses and are the clustered at the province level. Significance levels are indicated as follows: ** $p < 0.01$, and * $p < 0.05$.

The results are reported in Table 9. Columns (1) and (2) show that program effects are not significantly stronger for provinces with more rural or rural female students than other provinces. Column (3) shows that the program effect increases by 0.904 percentage points if the average number of siblings a rural student is increased by one. This effect is statistically significant at the 1% level. The program effect is more sensitive to the average number of siblings of rural female students. As shown in column (4), the program effect increases by 1.001 percentage points if the average number of rural female students' siblings increases by one, and this finding is statistically significant at the 1% level.

The above findings support our hypothesis that the program effects arise mainly from relaxing the credit constraints on the economically disadvantaged group. On the one hand, to the extent that the credit-constraint problem is more severe in poor areas, the grant program improves the quality of the prospective teachers in those areas. On the other hand, students from poor provinces are reluctant to commit to returning to their home province. The net program effect may depend on whether the benefit of easing credit constraints outweighs the prospects of migrating to a richer province.

7. Discussion and conclusions

This paper examines a conditional grant program designed to improve the selection of prospective teachers into university-based teacher training programs. Program recipients enjoy a four-year stipend and tuition waivers conditional on them serving as teachers in their home provinces after graduation. We find that this conditional grant program improves the credentials of prospective teachers. The average percentiles of those enrolled in the teaching majors in program colleges increased by 0.37–0.55 percentage points despite the increased enrollment quota in those majors. The program effects are particularly strong for students from provinces with larger shares of economically disadvantaged applicants than elsewhere. The heterogeneity analysis suggests that the likely channel through which the program effect arises is easing of the credit constraints.

We also find evidence suggesting that the grant program draws talent from those who would have attended regular colleges of similar

Table 9
Heterogeneous policy (DDD) effects by student household characteristics.

	Panel A: Standardized mean score			
	(1)	(2)	(3)	(4)
Post*TM*PC (DDD)	−0.050	0.155	−0.587	−0.813
	(0.817)	(0.755)	(0.464)	(0.458)
Rural share*DDD	0.953			
	(1.191)			
Rural female share*DDD		1.363		
		(2.343)		
No. of siblings of rural*DDD			0.904**	
			(0.306)	
No. of siblings of rural female*DDD				1.001**
				(0.279)
No. of observations	19,271	19,271	19,271	19,271
R-squared	0.900	0.900	0.901	0.901
	Panel B: Standardized maximum score			
	(5)	(6)	(7)	(8)
Post*TM*PC (DDD)	0.618	0.665	−0.181	−0.368
	(0.409)	(0.439)	(0.359)	(0.354)
Rural share*DDD	−0.314			
	(0.594)			
Rural female share*DDD		−0.848		
		(1.388)		
No. of siblings of rural*DDD			0.481	
			(0.268)	
No. of siblings of rural female*DDD				0.582*
				(0.245)
No. of observations	19,262	19,262	19,262	19,262
R-squared	0.838	0.838	0.839	0.839

Notes: The results are based on the most refined sample, i.e., majors in subdisciplines with both teaching and regular majors. All controls and fixed effects as in Eq. (3). Scores' standard errors are shown in parentheses and are clustered at the province level. Significance levels are indicated as follows: ** $p < 0.01$, and * $p < 0.05$. Constant terms are not reported.

ranks into teaching majors. This result sheds some light on the potential general equilibrium effects when the program is scaled up to incorporate more lower-rank teachers' colleges. Higher scoring students within each rank would be more likely to study in the teaching majors due to the program. Overall, the quality of teacher candidates would increase relative to the entire pool of applicants.

There is consensus among practitioners and scholars that improving the selection of teachers is crucial to the quality of the teaching force. It is sometimes difficult to trigger profound changes in the existing system of teacher pay such that high-aptitude candidates are attracted to and retained in the teaching profession. Our study examines an alternative policy option that does not require a thorough reform of the existing system and hence is relatively easy to implement.

There are several limitations in our study. First, we use the scores to measure the quality of teacher candidates. Higher scoring candidates may not be better teachers. Future studies would benefit from having better measures on the effectiveness of teaching. Second, given our data, we can only explore the short-run effect of the program. By the end of our data period, the first cohort of grant recipients had not started their mandatory service. Better knowledge of the individual cost incurred by the program may be gathered over time, which may affect the choice of the later cohorts. The long-term effects of such programs merit future research.

CRediT authorship contribution statement

Li Han: Conceptualization, Methodology, Formal analysis, Writing - review & editing. **Jiaxin Xie:** Data curation, Formal analysis, Writing - original draft.

References

Abraham, K. G., & Clark, M. A. (2006). Financial aid and students' college decisions: Evidence from the district of columbia tuition assistance grant program. *Journal of Human Resources, 41*(3), 578–610.
Backes, B., Goldhaber, D., Cade, W., Sullivan, K., & Dodson, M. (2018). Can UTeach? Assessing the relative effectiveness of STEM teachers. *Economics of Education Review, 64*, 184–198.
Chetty, R., Friedman, J., & Rockoff, J. (Friedman, Rockoff, 2014a). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review, 104*(9), 2593–2632.
Chetty, R., Friedman, J., & Rockoff, J. (Friedman, Rockoff, 2014b). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review, 104*(9), 2633–2679.
Cornwell, C., Mustard, D., & Sridhard, D. (2006). The enrollment effects of merit-based financial aid: Evidence from Georgia's HOPE program. *Journal of Labor Economics, 24*(2), 761–786.
Corcoran, R. P., & O'Flaherty, J. (2018). Factors that predict pre-service teachers' teaching performance? *Journal of Education for Teaching, 44*(2).
Dynarski, S. (2000). Hope for whom? Financial aid for the middle class and its impact on college attendance. *National Tax Journal, 53*, 629–661.
Hanushek, E. (2009). School policy: Implications of recent research for human capital investments in South Asia and other developing countries. *Education Economics, 17*(3), 291–313.
Hanushek, E., & Rivkin, S. (2006). Teacher quality. *Handbook of the Economics of Education, 2*, 1051–1078 Eds by Eric Hanushek and Finis Welch
Jackson, K. (2014). Teacher quality at the high school level: The importance of accounting for tracks. *Journal of Labor Economics, 32*(4), 645–684.
Jackson, K., Rockoff, J., & Staiger, D. (2014). Teacher effects and teacher-related policies. *Annual Review of Economics, 6*(1), 801–825.
Jia, Z., Tao, L., & Yu, G. N. (2012). An investigation of tuition-free student teachers' studies. *Teacher Education Research, 24*(2), 69–74 In Chinese
Jacob, B. A. (2016). The power of teacher selection to improve education. Economic Studies at Brookings. *Evidence Speaks Report*, 1(12), March 12, 2016. Retrieved from https://www.brookings.edu/research/the-powerof-teacher-selection-to-improve-education/.
Kane, T. (2003). A quasi-experimental estimate of the impact of financial aid on college-going. NBER Working Paper No. 9703. National Bureau of Economic Research, Inc.
Li, T. (2017). Financial decentralization and geographical stratification of access to higher education in China: The case of shanghai. *China Sociological Review, 49*(3), 212–238.
Linsenmeier, D., Rosen, H., & Rouse, C. (2006). Financial aid packages and college enrollment decisions: An econometric case study. *Review of Economics and Statistics, 88*(1), 126–145.
Liu, C., Zhang, L., Luo, R., Wang, X., Rozelle, S., Sharbono, B., & Glauben, T. (2011). Early commitment on financial aid and college decision making of poor students: Evidence from a randomized evaluation in rural China. *Economics of Education Review, 30*(4), 627–640.
Long, B. T. (2003). How do financial aid policies affect colleges? The institutional impact of the georgia HOPE scholarship. *The Journal of Human Resources, 39*(4), 1045–1066.
Monks, J. (2008). The impact of merit-based financial aid on college enrollment: A field experiment. *Economics of Education Review, 28*(1), 99–106.
Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *Quarterly Journal of Economics, 125*(1), 175–214.
Rothstein, J. (2015). Teacher quality policy when supply matters. *American Economic Review, 105*(1), 100–130.
Wang, Z., & Yang, Y. (2018). Students with free normal education by local government: policy analysis and investigation of present situation. *Educational Research, 460*(5), 76–82 In Chinese
Zhou, W. (2010). The survey and policy analysis on teachers source structure in Shandong high schools. *Teacher Education Research, 22*(3), 61–65 In Chinese