Impact of a Senior High School Tuition Relief Program on Poor Junior High School Students in Rural China

Xinxin Chen,[a] Yaojiang Shi,[b,] * Di Mo,[c] James Chu,[d]
Prashant Loyalka,[d] Scott Rozelle[d]

Abstract

A significant gap remains between rural and urban students in the rate of admission to senior high school. One reason for this gap might be the high levels of tuition and fees for senior high school. By reducing students' expectations of attending high school, high levels of tuition and fees could be reducing student academic performance in junior high schools. In this paper we evaluate the impacts of a senior high tuition relief program on the test scores of poor rural seventh grade students in China. We surveyed three counties in Shaanxi Province and exploited the fact that, while the counties are adjacent to one another and share similar characteristics, only one of the three implemented a tuition relief program. Using several alternative estimation strategies, including difference-in-differences, propensity score matching and difference-in-differences matching, we find that the tuition program has a statistically significant and positive impact on the mathematics scores of seventh grade students. More importantly, this program is shown to have a statistically significant and positive effect on the poorest students in the treatment group compared to their wealthier peers.

Key words: education program evaluation, rural China, tuition relief program
JEL codes: I22, O12, O15

I. Introduction

After more than a decade of effort focused on universalizing access to primary education, policy-makers in developing countries are changing their focus to expanding access to secondary education (UNESCO, 2011). One major challenge for policy-makers in developing countries in expanding access to secondary education (including junior and senior high school) is that students from rural areas are underrepresented in senior high school compared to students from urban areas (Ohba, 2011). Because fewer students from rural areas attend senior high school than students from urban areas, students from rural areas might be less likely to benefit from the high economic returns associated with both senior high school and college (Psacharopoulos and Patrinos, 2004). Thus, the lack of access to senior high school among rural students can, in turn, lead to greater income inequality between urban and rural areas as well as lower economic growth (Sala-i-Martin *et al.*, 2004).

Low educational attainment in poor rural areas is an important and emerging issue in China. Senior high school enrollments are far lower in most poor rural areas than in urban areas in China (Yang, 2006). Nearly 90 percent of students in large cities in China attend senior high school. In contrast, approximately 25 percent of students attend senior high school in poor rural areas (Liu *et al.*, 2009).

In the longer run, this gap in school participation might harm the Chinese economy. When poor rural children end up working in factories or on construction sites to obtain short-run returns rather than studying mathematics, language, English and computers, there are serious concerns that the children are not acquiring the necessary skills to be gainfully employed in China's labor market in future (Rozelle *et al.*, 2012). If China's economy continues to grow over the coming decade, wages (which have been rising fast over the past several years (Cai, 2007)) will almost certainly continue to increase. However, employers will only be willing and able to hire workers who have sets of skills worthy of high hourly rates. Workers that lack such skills will either be forced to find employment in the informal economy (in which returns are low and expectations of future rises of income are negligible) or become unemployed (or drop out of society and even possibly resort to organized crime). Hence, as rich urban students reap high returns, students from poor rural families might be left behind and China could face high rates of chronic unemployment, embedded inequality and even growth-reducing instability.

A major barrier preventing rural students from attending senior high school in developing countries in general is high tuition fees. High tuition fees are obviously a problem because poor, rural families have limited resources to pay for tuition and other direct costs of education (Banerjee *et al.*, 2000). Poor rural families also might be unable to borrow money to pay tuition fees because of limited access to well-functioning credit

markets (Banerjee, 2004). Where well-functioning credit markets do exist, families may still be unable to borrow money, due to market imperfections: for example, human capital cannot be used as collateral for educational loans (Deininger, 2003). Even when families can locate the funds to pay tuition fees, tuition combined with the opportunity costs of going to school can make the total costs of schooling prohibitively expensive (Lincove, 2012).

High tuition fees appear to be a major reason for the low matriculation rates to senior high schools among students from poor rural areas in China. Notably, the tuition fees of rural public senior high schools in China are among the highest in the developing world and are often unaffordable for students from these areas (Liu *et al.*, 2009). Not only are tuition fees high, but students from poor rural areas appear to overestimate the actual costs associated with attending senior high schools (Loyalka *et al.*, 2010). Indeed, the perception of high tuition fees might discourage students in poor rural areas from studying hard in junior high school, which, ultimately, lowers their chance of attending senior high school (Wang *et al.*, 2011).

In this study, our overall goal is to test whether a senior high tuition relief program could counteract the negative impact of high tuition on the academic performance of poor rural students. More precisely, we investigate whether students would respond to a tuition reduction program by (working harder and) performing better. We also attempt to determine whether the poorest students in areas with a tuition relief program benefit more than non-poor students.

The rest of the paper is organized as follows. Section II describes the tuition relief program, and Section III illustrates the data collection and sampling methodology. In Section IV we explain our analytic approach. Finally, in Section V we present the results of the study and discuss their implications.

II. Ningshan County's Tuition Relief Program

A tuition relief program implemented in the County of Ningshan provides us with an opportunity to examine the effect of such programs on the academic performance of junior high students in China's poor rural areas. Ningshan is a nationally-designated poor county, where rural per capita income was 3201 yuan (US$500) in 2009 (Ankang Statistical Bureau, 2010). As in other rural areas, the annual costs of attending senior high school far exceed the income of poor families. According to the official website of the Ankang Prefecture Bureau of Education, tuition for senior high schools in 2010 was 1500 yuan per year.[1] In

---

[1] All the counties studied in this paper (i.e. Ningshan, Shiquan and Hanyin) are located in the same prefecture of Ankang in Shaanxi Province.

addition, there were other fees and miscellaneous expenses. When interviewing the officials of Ningshan, Shiquan and Hanying counties, we were told that approximately 80 percent of students live in the dormitories of their senior high schools. According to a survey of senior high school students at Yulin, Yanan and Shangluo in Shaanxi Province, on average, a senior high student has to pay an additional 3000 yuan per year for accommodation, food and educational materials (Liu *et al.*, 2009).

The Ningshan tuition relief program was announced in July 2009 during the summer vacation. The local government promised to pay annual tuition (1500 yuan) for 3 years of senior high school for those among the top 500 students in the entrance examination to senior high school. As the average annual enrollment in the only senior high school in this county is approximately 550, the program coverage was 91 percent. In effect, the implementation of the program meant that most students enrolled after August 2009 did not need to pay tuition fees. All junior high school students were informed of the tuition relief program.

Although in September 2009 only 15 percent of junior high students knew of the program, when we revisited the schools in March 2010, 100 percent of students that we randomly selected from grade 7 (or the first year of junior high school) were aware of and could generally describe the nature of the program. According to interviews with officials from the Bureau of Education in Ningshan, the government promoted this program among all junior high students in early October 2009. As a result, almost every 7 grader in this county knew about this program right after the fall term began. Hence, whether students were familiar with this program or not, could not change the ultimate effect of this program.

The neighboring counties of Shiquan and Hanyin are located in the same prefecture as Ningshan. In China, students within the same prefecture are usually required to take the same courses, use the same textbooks, take the same entrance examination for senior high school and pay the same amount of tuition. This is true in the case of Ankang Prefecture. There are other similarities between Ningshan and Shiquan/Hanying counties. Like Ningshan, Shiquan and Hanying are poor counties. In 2009, the rural per capita income was 3323 yuan (US$519) in Hanyin and 3338 yuan (US$522) in Shiquan (Ankang Statistical Bureau, 2010). All three counties are extremely mountainous. Per capita fiscal revenues are nearly zero in all three counties, which means that almost all education expenditures are financed by transfers from governmental units in the prefecture, province and national levels. Moreover, more than 98 percent of the rural population in all three counties are Han. In sum, in terms of characteristics such as poverty rate, geography, fiscal capabilities and ethnic make-up, the three counties are almost identical.

The sample students also appear to be similar in terms of what would be expected in a poor, rural setting in China. For example, there are 6 percent more boys than girls, a ratio

similar to that cited in the Ministry of Education's 2010 Annual Yearbook (MOE, 2010): 7 percent more boys than girls. Approximately 98 percent of the seventh graders in our sample are aged between 11 and 15 years.

Although the main sample at the time of the baseline survey (September 2009) included a total of 36 schools and 3121 students, there was some attrition by the time of the endline survey in September 2010 (Figure 1).[2] For various reasons (e.g. dropouts, absences and death), by the time of the endline survey, we were only able to follow up with 2742 students: 672 students in Ningshan County (the treatment group) and 2070 students in Shiquan and Hanyin Counties (the control group). The attrition rate is almost the same (12 percent) in both groups, thus reducing the probability of attrition bias.[3]
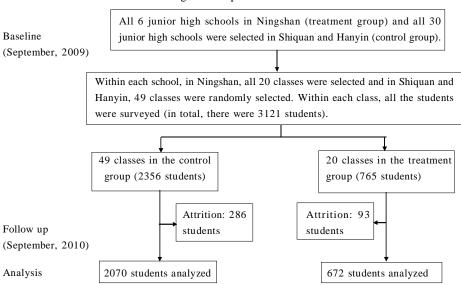
Figure 1. Experiment Profile

[2] It is possible that because of the leaking of information about the program, fewer students in the tuition relief program county schools dropped out prior to the start of grade 7 (i.e. prior to the baseline). Two things indicate that this is not a problem. First, during the baseline survey, we discovered that only 15 percent of the students and their parents knew about the program. Second, consulting records from grade 6 (June 2009) and grade 7 (September 2009), we discovered that the dropout rate between the end of grade 6 (June) and the beginning of grade 7 (September) was low (less than 1 percent). In addition, the dropout rates were almost identical (1.03 percent in the county tuition relief program county schools and 1.09 percent in the control schools).

[3] The attrition rate, while high, is not unusual. In a working paper based on data from a county in Shaanxi Province (Mo *et al.*, 2011), the dropout rate of poor rural students between the first month of the first year of junior high (grade 7) and the first month of the second year of junior high (grade 8) is reported to be 13.3 percent.

Table 1. Sample Average for the Students in the Treatment Group and the Control Group in 2009

		(1) Treatment group	(2) Control group	(3) Difference in mean (1) – (2)
(1)	Raw mathematics test score in 2009 (full score = 100)	54.82	54.29	0.54
		(15.29)	(17.33)	(0.72)
(2)	Age of the student (year)	12.92	13.06	–0.14
		(0.81)	(1.00)	(3.28)***
(3)	Boy student (%)	49.18	53.20	–4.02
		(0.50)	(0.50)	(1.81)*
(4)	Student who attended kindergarten (%)	16.47	15.15	1.31
		(0.37)	(0.36)	(0.82)
(5)	Student who attended preschool (%)	93.26	93.19	0.07
		(0.25)	(0.25)	(0.06)
(6)	Student without any sibling (%)	28.27	25.89	2.38
		(0.45)	(0.44)	(1.21)
(7)	Age of the father (year)	39.48	40.62	–1.14
		(4.78)	(5.08)	(5.02)***
(8)	Age of the mother (year)	36.73	37.64	–0.91
		(4.26)	(4.72)	(–4.34)***
(9)	Father who completed middle school (%)	44.13	38.33	5.79
		(0.50)	(0.49)	(2.65)***
(10)	Mother who completed middle school (%)	36.87	24.52	12.35
		(0.48)	(0.43)	(6.16)***
(11)	Father who mainly worked in agriculture (%)	33.04	30.10	2.94
		(0.47)	(0.46)	(1.43)
(12)	Mother who mainly worked in agriculture (%)	51.34	49.86	1.48
		(0.50)	(0.50)	(0.67)
(13)	Number of family members (person)	4.25	4.47	–0.23
		(1.07)	(1.15)	(4.42)***

Source: Authors' survey.

Notes: Standard deviations are reported in parentheses for columns (1) and (2); absolute values of *t*-statistics are reported in column (3). *** and * denote significance at the 1 and 10-percent level, respectively.

Table 1 shows that the students were balanced in the key dependent variable (raw mathematics test score in 2009: row 1). To be specific, the mean mathematics score in the county of Ningshan (the treatment group) was 54.82 and the mean mathematics score in the counties of Shiquan and Hanyin was 54.29. There was no statistically significant difference between them. In addition, rows 2 to 13 show that there were no statistically significant differences for most control variables, including the preschool history of the student, the sibling dummy for the student and the occupation dummy for the parent.

However, there was one fundamental difference between Ningshan and Shiquan/Hanying. There was no tuition relief program in either Shiquan or Hanyin. Because of this, we believe we perform a quasi-experiment, and we designate students in Ningshan as treatment students and students in Hanyin and Shiquan as control students. That is, unlike

the seventh graders in Ningshan County who knew about the tuition relief program for senior high school, students in the control group who were in their first year of junior high school were well aware of the fact that if they wanted to go to senior high school they (or their parents) would have to pay tuition.

III. Sampling and Data Collection

To evaluate the effectiveness of the tuition relief program, we collected data in Ningshan County and two control counties, Shiquan and Hanyin. All 36 junior high schools in the three study counties were surveyed. In Ningshan County, all seventh grade classes in all six junior high schools were surveyed. In Shiquan and Hanyin Counties, a subset of seventh grade classes in each of the 30 junior high schools was randomly selected for the survey because Shiquan and Hanyin had a larger population (and it would have been too costly to survey all students in all classes). We surveyed all students in each sample class. In total, our survey covered 3121 seventh graders. These students were in 69 classes spread over 36 junior high schools.

Two surveys were conducted. Our baseline survey occurred in early September 2009, at the beginning of the autumn semester, and our evaluation survey occurred 1 year later, in September 2010. During each round of survey, the enumeration team visited each school and conducted a two-block survey.

The first block of the survey was a 30-min standardized mathematics test. This test was given to all sample students in the treatment group and the control group. Because we designed, printed and administered the survey/test ourselves, we know that there was no coaching for the test before our survey. Because the test was conducted at the start of the school year, we also know that neither students nor teachers shifted their attention from other subjects to math in order to prepare for this test. In addition, even if the students had known about the program, rural students seldom take extra tutoring classes during summer vacation. Therefore, the mathematics test scores we collected in early September can reasonably be used as representative of the pre-program outcome. When we administered the standardized tests, both the treatment and control schools were blind to the fact that the surveys were meant to evaluate the effects of the tuition relief program. Only the two lead enumeration team managers were informed of the goal of the survey.[4]

---

[4] We chose mathematics test scores because they are one of the most common outcome variables used to proxy educational performance in rural China (Liu *et al*., 2009; Loyalka *et al*., 2010; Wang *et al*., 2011). The mathematics standardized exam is based on the Trends in Mathematics and Science Study (TIMSS) test.

In our analysis, we report raw mathematics test scores (full score equals 100) without any further manipulation, for ease of interpretation. As a robustness check, we also use the normalized Z-score of the mathematics score. The normalized score is created by subtracting the average test score of all sample students from the raw score for each student and then dividing it by the standard deviation of the test scores of all sample students in the same grade. With this transformation, the normalized test score is interpreted as the units of standard deviations from the mean score of all students in the same grade. When we replicate our empirical analyses using normalized scores, the results are almost the same. Using normalized scores has one advantage of facilitating comparison with other educational programs.

In the second block, enumerators collected data on the demographic and socioeconomic characteristics of students and their families. From this part of the survey we are able to create our control variables. The dataset includes measures of each student's age (measured in years), gender (described by an indicator boy student which is equal to one for boys and zero for girls), sibling information (described by an indicator, one child, which is equal to one for students who had no siblings and zero for students who had siblings), student pre-schooling and kindergarten information (described by the indicator of preschool and kindergarten, which equals one if students attended either preschool or kindergarten), father's and mother's age (measured in years), father's and mother's education level (whether he/she completed at least middle school) and father's and mother's occupation (described by an indicator variable called occupation, which equals one if a student's parent worked in the agriculture sector, and zero if the parent worked in the non-agricultural sector).

As part of the second block, students were also asked to indicate which assets their families own from a list of 30 household asset items. Using these data, we generated an asset index using principal component analysis to measure the wealth of each household. Following the method by Filmer and Pritchett (2001), we use scoring factors from the first principal component to create the asset index. It is, in fact, a weighted average of the observed 30 variables of assets, and variables with higher coefficients have more weight in determining the score. The higher the asset index is, the wealthier the household is. Based on this asset index, we divide the students into five equal-sized groups and create the variables Poorest, Second, Median, Fourth and Richest to represent the students whose household wealth was among the bottom 0–20, top 60–80, top 40–60, top 20–40 and top 20 percent and above.

IV. Analytical Approach

In this section we introduce our analytical approach. In the analysis we use a number of

alternative estimation procedures to examine the impact of the tuition relief program. In particular, we use difference-in-differences (DD), propensity score matching (PSM) and difference-in-differences matching (DDM).

1. Difference-in-differences Model

We are interested in understanding the mean impact of the "treatment on the treated" (TT), which is the average impact of the program among those treated (Smith and Todd, 2005). Thus, we use the following DD model to estimate the average treatment on the treated:

$$\Delta Score_i = \alpha + \delta Program_i + \gamma Score\_09_i + \beta X_i + \varepsilon_i ,\qquad(1)$$

where i is an index for the student, $\Delta Score_i$ is the change in the score of student i between 2009 and 2010, and $Program_i$ is the treatment variable (which makes δ the parameter of interest). In our analysis, $Program_i = 1$ if student i participates in the program ($Program_i = 0$ if student i does not participate in the program). Finally, the term X_i is a vector of covariates that are incorporated to capture the characteristics of a student and his/her parent and household, including age, gender, preschool history, number of siblings of the student, educational attainment, occupation, age of the student's father and mother, and wealth status of the household.

It is important to remember that the identification of the causal effect using DD relies on the "parallel trend" assumption. That is, without the policy change (or intervention of the program in our case), the average change in the outcome variable would have been the same for the treated and the comparison groups.

As might be expected, the effectiveness of DD depends on the validity of this assumption. In this study, the difference in these differences can be interpreted as the causal effect of the tuition relief program under the assumption that in the absence of the program the differences in the test scores of students would not have been systematically different in the treatment and control groups. This identification strategy might be invalid if the pattern of differences in student scores varies systematically across counties.

2. Propensity Score Matching and Difference-in-differences Matching

We apply different methods in the present study to obtain more robust findings. To begin, we use PSM, an approach that does not require the parallel trend assumption. With a sufficient region of support (or common support), it is possible to estimate the propensity scores of all students and to compare the outcomes of students who participated in and students who did not participate in the program that have similar propensity scores.

Propensity score matching is a more general method than standard linear regression because it does not require assumptions about linearity or constant treatment effects, and, thus, bias correction is improved. Moreover, imposing common support[5] when matching the treatment group with the control group in PSM can lead to efficiency improvements, especially when the sample size is small. It should be noted, however, that PSM estimates are only unbiased if the unobservables are correlated with the observables upon which the matching is based.[6]

Even though we control for the individual observable differences in estimating the propensity score, there may still be systematic unobservable differences between the outcomes of students who participated and students who did not participate in the program. The systematic differences could arise, for example, because the student's decision to participate is based on certain unmeasured characteristics. Such differences could violate the identification conditions required for matching (Smith and Todd, 2005).

To eliminate the bias due to time-invariant unobservable differences between students who participated in and students who did not participate in the program, we extend the cross-sectional PSM approach to a longitudinal setting and implement the DDM strategy. With DDM we can exploit the data on students both in 2009 and 2010 to construct the required counterfactual, instead of just using the data in 2009 (as is used in the PSM analysis). The advantage of DDM is that the assumptions that justify DDM estimation are weaker than the assumptions necessary for DD or the conventional PSM estimator. DDM removes time invariant unobservable differences between students who participated in and students who did not participate in the program conditional on the propensity score to participate in the program, a clear advantage over cross-sectional PSM.

When performing DDM, we match by using the log odds ratios and the same nearest-neighbor matching methods with replacement that were used for our PSM approach. In addition, we compute an "adjusted" version, where the control units are weighted by the

---

[5] It ensures that persons with the same characteristics have a positive probability of being both participants and non-participants (Heckman *et al.*, 1999).

[6] We achieve a good balance between the treatment and control groups after matching. The standardized percent bias (Rosenbaum and Rubin, 1985) for each of the covariates across our matched treated and control groups is small, especially for the covariates that have a relatively large difference in treatment and control means (a large bias) in the pre-matched sample. The standardized percent bias (for each covariate) is defined as the percentage difference of the sample means in the treated and non-treated (full or matched) subsamples as a percentage of the square root of the average of the sample variances in the treated and non-treated groups (Leuven and Sianesi, 2003). Full results are available from the authors upon request.

number of times that they are matched to a treated unit. The standard errors are bootstrapped using 1000 replications.

V. Results

1. Effect of the Tuition Relief Program: Difference-in-differences

The results from our difference-in-differences model suggest that the Ningshan tuition relief program had a significant impact on students' academic performance. The baseline test scores of the students in the treatment (54.82) and the control group (54.29) were almost the same in 2009 (Table 2, row 1). However, when comparing the change in the performance of the treatment and control students between 2009 and 2010, control students improved less than treatment students. The rise in the tests scores of the treatment students (18.37) was higher than the rise in the test scores of the control students (15.26: row 3). Moreover, the 3.1-point higher rise in the test scores of the treatment students is statistically different at the 1-percent level of significance (column 4). From these basic descriptive statistics, it appears as if the Ningshan tuition relief program succeeded in stimulating the interest/effort of the average student in Ningshan relative to the average students in Shiquan and Hanying.

We are interested in determining whether students from poor families are likely to be most affected by the program. To test this proposition, we divided the sample of students in the study counties into five equal sized groups (from poorest to richest), based on asset

Table 2. Change in Student Mathematics Score between 2009 and 2010

		Full sample	Treatment group	Control group	Difference (t-statistics in parentheses)	
		(1)	(2)	(3)	(4) = (2) – (3)	
Panel A: Change in raw mathematics score						
(1)	Mean score in 2009	54.42	54.82	54.29	0.53	(0.72)
		(16.85)	(15.29)	(17.33)		
(2)	Mean score in 2010	70.44	73.19	69.55	3.64	(4.96)***
		(16.61)	(15.49)	(16.87)		
(3)	Difference = (2) – (1) (t-statistics in parentheses)	16.02	18.37	15.26	3.11	(2.96)***
		(35.45)***	(21.88)***	(28.71)***		
Panel B: Change in normalized mathematics score						
(4)	Mean score in 2009		0.02	–0.01	0.03	(0.72)
(5)	Mean score in 2010		0.17	–0.05	0.22	(4.96)***
(6)	Difference = (2) – (1) (t-statistics in parentheses)		0.15	–0.04	0.19	(2.99)***
			(2.82)***	(1.45)		

Source: Authors' survey.

Notes: Standard deviations are reported in parentheses in rows (1) and (2); absolute values of t-statistics are reported in parentheses in row (3) and column (4). *** denotes significance at 1 percent.

Table 3. Change in Student Mathematics Score between 2009 and 2010 by Wealth

		Treatment group			Control group			(7)	
		(1)	(2)	(3)	(4)	(5)	(6)		
		Score in 2009	Score in 2010	Difference = (2) − (1)	Score in 2009	Score in 2010	Difference = (5) − (4)	Difference = (3) − (6)	
					Change in raw score by wealth				
(1)	Poorest	54.86	75.00	20.14	51.99	68.31	16.32	3.82	(1.81)*
(2)	Second	55.34	73.53	18.20	52.86	69.88	17.02	1.17	(0.69)
(3)	Median	55.43	73.15	17.72	55.13	69.13	14.00	3.72	(2.35)**
(4)	Fourth	52.68	71.60	18.92	54.30	69.69	15.39	3.52	(2.22)**
(5)	Richest	56.22	73.93	17.70	57.57	70.89	13.32	4.38	(2.39)**
					Change in normalized mathematics score by wealth				
(6)	Poorest	0.03	0.28	0.25	−0.14	−0.12	0.02	0.23	(1.84)*
(7)	Second	0.05	0.18	0.13	−0.09	−0.03	0.06	0.07	(0.71)
(8)	Median	0.06	0.16	0.10	0.04	−0.08	−0.12	0.22	(2.37)**
(9)	Fourth	−0.10	0.07	0.17	−0.01	−0.05	−0.04	0.21	(2.22)**
(10)	Richest	0.11	0.21	0.10	0.19	0.03	−0.16	0.26	(2.39)**

Source: Authors' survey.

Notes: Absolute values of *t*-statistics are in parentheses for column (7). ** and * denote significance at the 5 and 10-percent level, respectively.

index ranking.

While Table 3 demonstrates that students in all wealth categories that enjoyed tuition relief (those in Ningshan County) benefited, the average point rise among some categories of students increased more than others. A little surprisingly, the richest students seemed to benefit most from the program. After the program, the difference in the rise in score between the two groups was 4.38 points (column 7, row 5). This does not mean that the poorest students did not benefit. In fact, ignoring the scores of the students in the richest categories, the poorest students saw their average scores rise the most (the difference in the rise between the treatment group and the control group was 3.82: row 1). The puzzle arises in the nonlinearity from the poorest to the richest. The second poorest students seem to benefit least from the program because the difference in the rise between the treatment and control groups was only 1.17 points (column 7, row 2). After this (as one moves from the second to the fifth group), scores generally rise (rows 3 to 5).

In the next section we seek to control for a number of observable characteristics of students (using the regression model spelled out in model (1)). If we control for factors that differ between rich and poor students as much as possible (e.g. the educational levels of their parents), we might be able to confirm whether students (especially rich students) are benefiting from the tuition relief program.

2. Effect of the Tuition Relief Program:
Difference-in-differences Multivariate Results

If we add control variables to our DD analysis, the results are largely consistent with the

Table 4. Difference-in-differences Regressions: Evaluating the Effects of the Tuition Relief
Program on the Normalized Mathematics Score of the Average Students
and the Poorest Students (Shaanxi Province, China)

	Dependent variable ($\Delta Score_i$)= $Score_{i, 2010}$ − $Score_{i, 2009}$		
		(1)	(2)
(1)	Program dummy (1 = participated in the program)	0.17	0.14
		(5.65)**	(1.50)
(2)	Poorest dummy (1 = poorest)[a]		−0.04
			(0.73)
(3)	Interaction term of poorest and program dummy (Program * Poorest)		0.23
			(2.99)***
(4)	Second poorest dummy (bottom 20%–bottom 40%)[b]	0.03	
		(0.53)	
(5)	Median dummy (bottom 40%–bottom 60%)[b]	−0.01	
		(0.20)	
(6)	Second richest dummy (top 20%–top 40%)[b]	0.01	
		(0.08)	
(7)	Richest dummy (> top 20% quintile)[b]	0.00	
		(0.02)	
(8)	Student controls[c]	Yes	Yes
(9)	Observations	2264	2264
(10)	R^2	0.31	0.31

Source: Authors' survey.

Notes: Absolute values of *t*-statistics are in parentheses. *** and ** denote significance at the 1 and
5-percent levels, respectively. [a]The poorest dummy is the same as for Table 3. It equals 1 if the asset
index of the household is among the bottom one-fifth. [b]The dummies indicating wealth are the same
as in Table 3 and the comparison group is the poorest students (bottom 0–20%). [c]The control
variables are those listed in Table 1. [c]student controls included are student age, gender, whether the
student is the only child of the family, whether student has any kindergarten and preschool experience,
each parent's education level (whether he/she completed at least middle school), age, occupation
(whether he/she works in the agriculture sector) and number of family numbers.

descriptive statistics in terms of the overall impact (Table 4). According to our analysis
(and consistent with the findings in Table 2), there is a positive and statistically significant
impact of the tuition relief program on the test scores of students. In the estimation of
Equation (1), the results demonstrate that the estimated treatment effect of the tuition
relief program on mathematics test scores is equal to 0.17 standard deviations and the
impact is significant at the 5-percent level (row 1, column 1). That is, when we use a DD
approach, we find that the tuition relief program has had a positive and significant effect
on the academic effort (as measured by the scores of the standardized tests) of the
average student.

However, our multivariate analysis shows that the poorest students in the sample
(those with the lowest asset indices) benefit the most.[7] As is evident from Table 4

[7] Because there may be too few clusters to run the regression this way, we executed the bootstrapping tests
as delineated in Cameron *et al.* (2008). In doing so, we find that the bootstrapping results are similar to
the results from the regression approach (the estimated effect of the tuition relief policy is still significant
at the 1-percent level). Full results are available from the authors on request.

(column 2), when we control for all of the covariates and then add an interaction term between the treatment variable (tuition relief program dummy) and a dummy variable representing the poorest students, we find that the program's effect is primarily on the poor. This result may have been "disguised" when only examining the descriptive statistics in table 3. Specifically, while the average treatment effect (which in this case is the average treatment effect for all of the students except the poorest) is still positive (0.14), the standard error is relatively large. In other words, we cannot reject the hypothesis that the tuition relief program's effect on all but the students in the poorest asset category is zero. However, the positive (and large) coefficient on the Poorest * Program interaction variable (0.23) means that the test scores of the poorest junior high students are 0.23 standard deviations higher (and significantly so) than those of other students.

Hence, if the multivariate results are to be taken seriously, the results as a whole (i.e. both descriptive and multivariable) demonstrate that the tuition relief program increased student academic performance in junior high. More importantly, the poorest students, or those who were most likely to be from families that were financially constrained, benefited most from the program.

3. Effect of the Tuition Relief Program: Matching Results

In order to examine the robustness of our results using other approaches, in this subsection we present results from our PSM and DDM analyses. In fact, the results of both PSM and DDM analysis are qualitatively identical to the DD multivariate results (Table 5).

In Table 5, rows 1 to 3 present the estimated average treatment effects on the treated (ATT) of different treatment groups. Columns 1 and 2 show the estimation results from PSM and DDM, respectively. The PSM results show that the program has a positive effect on the mathematics scores of the average students in Ningshan (compared to the average students in the control counties). The effect is 0.18 standard deviations and is significant at the 1-percent level (row 1). Likewise, when using DDM, the average treatment effect is 0.15 standard deviations and significant. The average impact on test scores when using PSM (0.18) and DDM (0.15) is close to the average impact when using DD multivariate regression (0.17).

The results of the PSM and DDM also reveal that the greatest impact of tuition relief program appears to be on the scores of the poorest students (Table 5, rows 2 and 3). When we use PSM, the average test scores of the poorest students in the treatment group improved by 0.28 standard deviations when compared to the poorest students in the control group. Moreover, this result is statistically significant (row 2, column 1). By

Table 5. Evaluating the Effects of the Tuition Relief Program on the Efforts of Students Using Propensity Score Matching and Difference-in-differences Matching (Shaanxi Province, China)

		Propensity score matching[a,b]		Difference-in-difference matching	
		(1) Average treatment effect for the treated	*t*-statistic/ *Z*-value[b]	(2) Average treatment effect for the treated	*t*-statistic/ *Z*-value[b]
(1)	Students in the treatment group versus students in the control group	0.18	(2.95)***	0.15	(2.42)***
(2)	Poorest students in the treatment group versus poorest students in the control group	0.28	(1.80)*	0.28	(1.69)*
(3)	Richest students in the treatment group versus richest students in the control group	0.12	(0.82)	0.08	(0.53)

Source: Authors' survey.

Notes: *** and * denote significance at the 1 and 10-percent level, respectively. [a]The matching method used is the nearest neighbor matching method (random draw version) with replacement. [b]*t*-statistics are reported for propensity score matching. The standard errors were bootstrapped using 1000 replications.

contrast, although the point estimate of the impact of the tuition relief program on the richest students in the treatment group improves by 0.12 standard deviations relative to the richest students in the control group, the gain is statistically insignificant (row 3, column 1).

Echoing these results, the DDM results show that the tuition relief program has a statistically significant impact on the test scores of junior high students. This impact is significant for the poorest students, but the program seems to have no effect on the richest students (column 2, rows 2 and 3).

In sum, all three (DD, PSM and DDM) estimation strategies show that the program improves junior high students' mathematics scores by more than 0.15 standard deviations. Moreover, the poorest students are, indeed, benefiting from the program.

VI. Summary and Discussion

In this paper we exploit a quasi-experiment to examine the effect of a senior high tuition relief program on junior high students in poor rural schools in Shaanxi Province in China. Seeking to understand whether a tuition relief program improved the academic performance of junior high students, we compared seventh grade students in Ningshan County, where a tuition relief program was implemented, to seventh graders in nearby Shiquan and Hanyin Counties. We fielded a survey and administered a standardized mathematics

examination, then analyzed the data using various estimation strategies, including DD, PSM and DDM.

The descriptive and econometric results of the program effect were robust. In general, we find that the Ningshan tuition relief program positively impacted students' academic performance. Indeed, for all of the models, we discovered a statistically significant rise in the change of the mathematics scores between the control and treatment students.

More importantly, we also find (in all of the different approaches using the multivariate analyses) that the tuition relief program had the largest (and only) significant impact on poorest students. In short, and perhaps unsurprisingly, our findings demonstrate that the test scores of poorest students rose more (and significantly so) than that of other non-poorest students. Our data also show that the tuition relief program did not have a statistically positive impact on the richest students, who are seldom financially constrained when deciding whether to go to senior high school. This result renders additional support to the validity of the assumption in the DD analysis.

Taken together, these results might suggest that poor students work harder when they realize that their families can afford high school tuition. However, it is important to realize that other potential mechanisms exist: teachers, whose wages are linked to student matriculation in highly ranked high schools, might invest more time teaching poorer students who would have otherwise dropped out after ninth grade due to financial reasons. In general, the tuition relief program clearly and positively impacts students' academic performance quickly (over just 1 year) and far before ninth grade (in seventh grade).[8]

The results of this study contribute to a broader policy debate about how to effectively invest in rural education. Recently, there has been increasing support from officials in the Ministry of Education for greater investment in rural education. Opinions are divided on how money should be invested. Our results suggest that China's top educational officials should at least provisionally expand tuition relief programs in poor rural areas as an additional way to improve the human capital in rural areas. If future evaluations of tuition relief programs also show results consistent to ours, China should consider waiving senior high tuition for all poor students.

_____

[8] In fact, because our intervention was only over 1 year, any effects on academic performance over the long term are likely to be underestimated. While we cannot make empirically-backed claims regarding the persistence of the benefits, the fact that the intervention had such immediate effects may further underline the impact of the program.

References

Ankang Statistical Bureau, 2010, *Ankang Statistical Yearbook*, Beijing: Chinese Statistical Publishing House (in Chinese).

Banerjee, Abhijit, 2004, "Educational policy and the economics of the family," *Journal of Development Economics*, Vol. 74, No. 1, pp. 3–32.

Banerjee, Abhijit, Suraj Jacob, Michael Kremer, Jenny Lanjouw and Peter Lanjouw, 2000, "Promoting school participation in rural Rajasthan: Results from some prospective trials," mimeo, MIT, Massachusetts.

Cai, Fang, 2007, "Lewis turning point is just ahead," *Zhongguo Shehui Baozhang (Chinese Social Security)*, Vol. 155, No. 5, pp. 24–6.

Cameron, A. Colin, Gelbach Jonah and Miller Douglas, 2008, "Bootstrap-based improvements for inference with clustered errors," *Review of Economics and Statistics*, Vol. 90, No. 3, pp. 414–27.

Deininger, Klaus, 2003, "Does cost of schooling affect enrollment by the poor? Universal primary education in Uganda," *Economics of Education Review*, Vol. 22, No. 3, pp. 291–305.

Filmer, Deon and Pritchett Lant H., 2001, "Estimating wealth effects without expenditure data–or tears: An application to educational enrollments in states of India," *Demography*, Vol. 38, No. 1, pp. 115–32.

Heckman, James., R. LaLonde and J. Smith, 1999, "The economics and econometrics of active labor market programs," in O. Ashenfelter and D. Card,eds, *Handbook of Labor Economics*, Vol.III, pp. 1865–2097, Amsterdam: Elsevier.

Leuven, Edwin and Barbara Sianesi, 2003, "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing," *Statistical Software Components S432001*, Boston College Department of Economics, Boston.

Lincove, Jane A., 2012, "The influence of price on school enrollment under Uganda's policy of free primary education," *Economics of Education Review*, Vol. 31, No. 5, pp. 799–811.

Liu, Chengfang, Linxiu Zhang, Renfu Luo, Scott Rozelle, Brian Sharbono and Yaojiang Shi, 2009, "Development challenges, tuition barriers and high school education in China," *Asia Pacific Journal of Education*, Vol. 29, No. 4, pp. 503–20.

Loyalka, Prashant, Yingquan Song, Jianguo Wei and Scott Rozelle, 2010, "Information, college decision and financial aid: Evidence from a cluster-randomized control trial in China," *Rural Education Action Project Working Paper*, Stanford University, Standford.

MOE (Ministry of Education), 2010, *Chinese Statistics Yearbook of Education*, Beijing: People's Education Publishing House (in Chinese).

Mo, Di, Linxiu Zhang, Hongmei Yi, Renfu Luo, Scott Rozelle and Carl Brinton, 2011,"School dropouts and conditional cash transfers: Evidence from a randomized controlled trial in rural China's junior high schools," *Rural Education Action Project Working Paper*, Stanford University, Standford.

Ohba, Asayo, 2011, "The abolition of secondary school fees in Kenya: Responses by the poor," *International Journal of Educational Development*, Vol. 31, No. 4, pp. 402–08.

Psacharopoulos, George and Patrinos Harry Anthony, 2004, "Returns to investment in education: A further update," *Education Economics*, Vol. 12, No. 2, pp. 111–34.

Rosenbaum, Paul and Donald B. Rubin, 1985, "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score," *American Statistician*, Vol. 39, No. 1, pp. 33–8.

Rozelle, Scott, Linxiu Zhang, Renfu Luo, Hongmei Yi and Yaojiang Shi, 2012, "Avoiding the middle income trap: Rural education, health and building the skills needed in China' 21st century economy," *Rural Education Action Project Working Paper*, Stanford University, Stanford.

Sala-i-Martin, Xavier, Gernot Doppelhofer and Ronald I. Miller, 2004, "Determinants of long-term growth: A bayesian averaging of classical estimates (BACE) approach," *American Economic Review*, Vol. 94, No. 4, pp. 813–35.

Smith, Jeffrey and Petra Todd, 2005, "Does matching overcome Lalonde's critique of nonexperimental estimators?" *Journal of Econometrics*, Vol. 125, No. 1–2, pp. 305–53.

UNESCO Institute for Statistics, 2011, "Global education digest. Comparing education statistics across the world: Focus on secondary education" [online; cited September 2012]. Available from: http://www.uis.unesco.org/Education/Pages/ged-2011.aspx.

Wang, Xiaobing, Chengfang Liu, Linxiu Zhang, Renfu Luo, Thomas Glauben, Yaojiang Shi, Scott Rozelle and Brian Sharbono, 2011, "What is keeping the poor out of college? enrollment rates, educational barriers and college matriculation in China," *China Agricultural Economic Review*, Vol. 3, No. 2, pp.131–49.

Yang, Dongping, 2006, *Equitable Education in China: Dream and Reality*, Beijing: Peking University Publishing House (in Chinese).

(Edited by Jing Qiu)