International Journal of Educational Development 30 (2010) 570–585

Contents lists available at ScienceDirect

International Journal of Educational Development

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

The effect of primary school mergers on academic performance of students in rural China

Chengfang Liu [a,*], Linxiu Zhang [a], Renfu Luo [a], Scott Rozelle [b], Prashant Loyalka [c]

[a] Center for Chinese Agricultural Policy, Institute for Geographical Sciences and Natural Resource Research, Chinese Academy of Sciences, 11A, Anwai Datun Road, Beijing 100101, China
[b] Freeman Spogli Institute, Stanford University, Encina Hall East, E407, Stanford, CA94305, USA
[c] China Institute for Educational Finance Research, Peking University, Beijing 100080, China

ARTICLE INFO

JEL classification:
O12
O15

Keywords:
School merger
Academic performance
Rural China

ABSTRACT

We examine the impact of primary school mergers on academic performance of students using a dataset that we collected using a survey designed specifically to examine changes in the academic performance of students before and after their schools were merged. We use difference-in-differences and propensity score matching approaches and demonstrate that overall the primary school merger has not harmed the academic performance of students, as some have claimed. We do find, however, that the timing of mergers matter; when students are older (e.g., the fourth grade) their grades rise after merging. The grades of younger students, however, fall.

1. Introduction

To achieve the goals set by the 2000 World Education Forum in Dakar to ensure that all children will be in school by 2015, reforms to the education systems in many countries have been undertaken across the developing world (UNESCO, 2006). China is no exception. In 2001 China's leaders enacted a policy called "A Decision to Reform and Develop Primary School Education." Added to the challenge of improving the quality of education for all children in the nation, this policy also was made at a time when demographic changes were leading to sharp declines in the enrollment of primary schools in rural China (National Bureau of Statistics, 2006). One of the main propositions of China's rural education policy was the implementation of a primary school merger program. The key part of the program involved plans to close down smaller schools in more remote villages and merging them with larger "central" schools. The overall goal of the merger program was to use scarce educational resources more economically and efficiently and improve the quality of primary school education for all rural students.

Since the proclamation of the merger policy, education officials have begun to aggressively merge smaller schools into larger central schools across the country. Nationwide, the number of primary schools in rural China has fallen by 24%, from 416,000 in 2001 to 317,000 in 2005 (National Bureau of Statistics, 2006). On average nearly 25,000 primary schools in rural China were closed down each year during this period.

As its implementation has proceeded, the primary school merger program in rural China has become a subject of an intense debate. Proponents of school merger say that the school merger program improves educational quality by enhancing equity through providing poorer children with access to schools that are supported with better educational resources, especially better teachers and facilities (Zhuo, 2006). The program is theoretically supposed to make education more efficient by taking advantage of economies of scale (under the assumption that the government is allocating more investment to the new merged schools—which is part of the merger program package). In contrast, critics of school merger caution that in the case of many primary school mergers, while the number of students in the central schools increases, the number of teachers and the size and quality of the teaching facilities do not rise enough. In such cases, teaching and administration in merged schools suffer; some have even argued that education has not improved (Pang, 2006). Some researchers are also concerned that students whose schools are closed down are bearing the brunt of the negative impacts associated with primary school mergers (Shi, 2004), as experience from some developed countries shows (e.g., Purcell and Shackelford, 2005, in the United States). If this were true, since the families of students in remote schools are almost certainly poorer and have less access to many of the opportunities that are available in China's economy today, the merger program – which was designed to improve equity – might ironically be regressive in its effects.

* Corresponding author. Tel.: +86 10 64888990; fax: +86 10 64856533.
E-mail address: cfliu.ccap@igsnrr.ac.cn (C. Liu).

0738-0593/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijedudev.2010.05.003

Despite the vehemence of the debate and the importance of the issues involving the school merger program, the literature on China's primary school merger program is almost completely lacking in evidence-based empirical work about the effects of school mergers on education quality. Most previous analyses rely only on anecdotes (Pang, 2006; Zhuo, 2006). Moreover, most of these analyses treat students as if they all belonged to a single group. No research – to our knowledge – acknowledges the potential for school mergers to affect both those students whose schools are closed down and those whose schools remain open (and "take over" other schools). Nor have they allowed for heterogeneous impacts across grades. In other words, no study has attempted to raise and answer certain key questions: How has the academic performance of students from schools that experienced mergers changed, especially relative to the performance of students from schools that did not merge? Are the students whose schools were shut down affected positively or negatively? What about students in the schools that were expanded?

The overall goal of this paper is to examine the effect of primary school mergers on the academic performance of students. To meet this goal, we will pursue four specific objectives. First, we compare the distribution of academic performance across three types of students: merger-guest students, merger-host students and non-merger students. We define *non-merger students* as those students who attended a school that was *not* involved in any mergers during the study period. In contrast, those students who attended schools involved in a merger during the study period are called *merger students*. We further differentiate merger students into two types. If the school that students attended for grade one was *closed down* and students were moved to another school that remained open in process of the merger, we call those students *merger-guest students*. In contrast, those students who attended schools that *remained open* during the course of the merger are called *merger-host students*. To meet this first objective, we descriptively compare the mathematics and Chinese language grades of the three types of students and describe how their scores (or grades) vary over time.

Our next three objectives move beyond correlation in order to explore causal impacts of mergers on students' academic performance. First, we examine the impact of school mergers on the test scores of rural children, and compare the direction and magnitude of impact across merger-guest and merger-host students. Second, we seek to identify the sources of the merger impact by examining the effect of specific changes in the ways that schools are run that are associated with school mergers (for example, changes in class sizes; changes in the quality of teachers; differences in class placement; and so on) on the scores of rural children. Finally, we examine how the impact of school mergers varies by the grade of the students.

As with all empirical studies, we face several limitations. First, since our data are limited to two regions of China, we cannot generalize our findings to the whole of China. Second, the school merger program is complex and has led to many other changes – such as the need for some children to stay in boarding schools and eat away from home – that may imply significant, non-academic impacts (including changes to children's health, safety and overall welfare). We do not look at these costs. Third, we realize that there is a rich literature on student's learning achievement and on the sociology of student learning. For example, in the literature (among the many studies) there is work on international education by Harold Stevenson and his colleagues in the United States, Japan, China's mainland, and Taiwan (e.g., Stevenson and Stigler, 1992; Stevenson et al., 1998; Fuligni and Stevenson, 1995). There also are studies by Esther Ho and her coauthors using PISA data (e.g., Ho, 2004, 2006; Yip et al., 2004; Chiu and Ho, 2006). In this paper, while recognizing the value and contribution of these literatures, we adopt a more empirical economics of education approach. Finally, although we realize that in some ways there is a long distance between school mergers and academic performance, we believe that going from educational policy to student outcome may also yield information that will aid in understanding the empirical relationship between a policy level intervention (in our case, China's school merger program) and student academic performance. In fact, in the economics of education literature, research teams often analyze the direct link between policy and educational outcomes (e.g., Angrist and Lavy (2002) examine the direct link between the installation of computers in classrooms on student academic performance in Israel; Glewwe et al. (2009) examine the direct link between textbook provision and student test scores in Kenya).

2. The school merger program in rural China

In the 1990s almost every village in rural China had a primary school (Pang, 2006; Guo, 2007). The primary schools were heterogeneous, reporting different levels of enrollment, facilities and teaching staff. For example, it was reported that primary schools in some villages in one province had less than 10 students whereas other primary schools in the same area of the same province had more than 200 students (Xinhuanet, 2002). In some schools there was only one temporarily hired teacher who was responsible for teaching all subjects and all grades. In other schools, there were not only teachers for each grade, there were even specialized teachers for certain subjects.

Although initially the sharp differences across schools were addressed by a program that tried to equalize resources across all schools, it soon became clear that this policy could not work and would likely lead to a waste of lots of resources. Instead a new idea was launched that began to explore the possibility of concentrating limited resources on a smaller number of schools. It was out of this thinking that education officials enacted a new program called the School Merger Program in 2001.

As soon as the merger policy was enacted, it quickly was extended throughout the country. Each provincial government was charged to set up a primary school merger plan. Different provinces (and the counties below them) took different approaches. For example, according to a set of guidelines developed by the education bureau in a county in Henan province, local officials were supposed to begin to merge schools by the following criteria: (a) schools that have less than 20 students per grade, on average, are to be merged; (b) several neighboring villages that have a combined population of up to 8000 are allowed to retain one primary school; and (c) the number of primary schools should be cut from 238 to 141. Although such targets are written in quite precise figures, in fact, according to one critic (Pang, 2006), nobody really knows where these figures came from. In other words, it is thought that in many cases the merger plan is quite arbitrary. Therefore, despite the high profile of the merger program, there is a lot of local discretion and little centrally directed implementation.

3. Data

The data used in this paper come from a survey we carried out in 2006. The survey was designed specifically to examine the changes in the academic performance of children before and after the schools that students attended for grade one merged. While the survey in part relied on recall data – especially for some of the control variables – we were able to use records and rely on multiple sources of information for our two key variables—the grades of the students (our measure of academic performance) and the merger status of the schools that they attended in grade one.

The sampling strategy was designed to collect data on a random sample of schools and students in the program area. The sample was drawn using a multi-stage, clustering design with random selection procedures employed at each stage. In the first stage 10 counties were selected, six from the whole of 107 counties in Shaanxi province, and four from the whole of 21 counties in Ningxia province. In the second stage two townships were randomly selected in each county as follows. First, all townships in the county were ranked according to per capita income in 2006. One township was then randomly chosen from both the top and bottom halves of the township income distribution. In stage three a list of all primary schools with grades one through six (*wanxiao)* was created in each selected township.[1] Three primary schools were then chosen from this list using the following procedure. First, the central primary school of the township was always included in the sample.[2] Second, if the central primary school in the township had been involved in a merger between 2001 and 2006, a second primary school that was not involved in the merger program was randomly chosen. If instead the central primary school had not been involved in the merger program, the second sample school was randomly selected from those schools that had been involved in a merger. The second sample school was thus either independent of mergers or was the host of students from a school that had been closed over the sample period. Finally, after eliminating the first two sample schools, one additional school was randomly selected from the list of all *wanxiao*.

The sample students were selected during the final stage of the sampling procedure. The sample consists of all students in each of the sample schools that were enrolled in the sixth grade during the 2006–2007 academic year. On average there were 1.7 sixth grade classes per school, ranging from 1 to 4. Since we carried out the survey in September (Shaanxi) and November (Ningxia), the students had just begun their new school year. Therefore, all of the sample students had just completed their fifth grade year during the 2–4 months prior to the study (as the school year in China runs between early September and mid-July).

Enumerators surveyed 62 primary schools in 20 towns in Shaanxi and Ningxia provinces, two of the nation's poorest provinces in northwest China.[3] In total, the sample includes 2446 students and their families. Of the sample, 561 students (23% of them) belonged to the merger-guest group. The sample also included 820 students (or 34%) that belonged to the merger-host group. The largest group, 1065 students (or 43%), was part of the non-merger student group (Fig. 1).

Descriptive statistics generated from our data show that the profile of sample sixth graders is fairly typical of students from rural areas. Forty-seven percent of sample students were girls. In the annual yearbook published by the Ministry of Education (2006), girls in rural China account for the same percentage, namely, 47% percent, of the class.[4] Almost 70% of the students are between 11 and 13 years old. Twenty-seven percent of the students at some point of their primary school education were held back for at least 1 year (see Chen et al., 2010, for a complete discussion of grade retention).

We also elicited information about the students from their homeroom teachers (*banzhuren* in Chinese). In 98% of the cases, the homeroom teacher was also the instructor of the student's Chinese

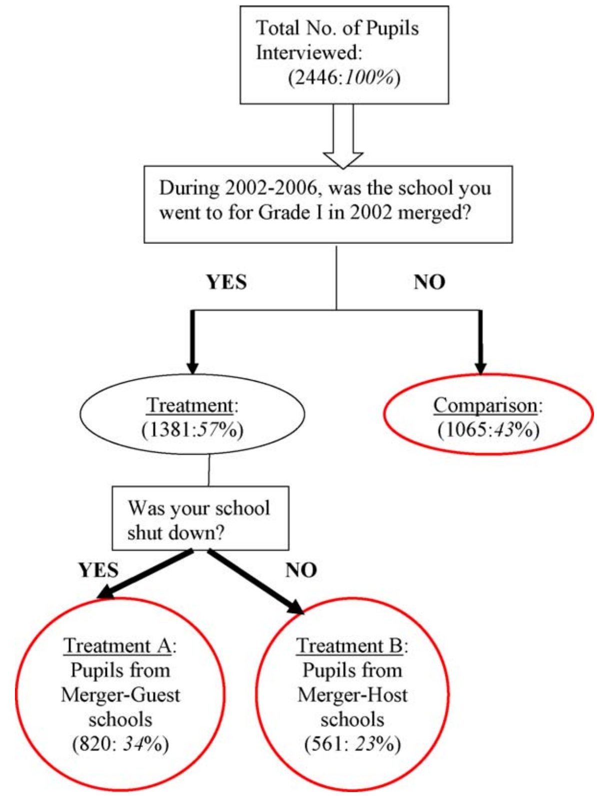

Fig. 1. Research design for evaluating the impact of primary school merger on academic performance.

or mathematics class (or both). In China the homeroom teacher not only teaches the children one or two subjects, he/she also is in charge of administering each student's school program and is the interface between the students and the principal's office and the students and their parents. Many homeroom teachers make a point of visiting the homes of their students. Therefore, in most cases the homeroom teacher was intimately familiar with the academic performance and family life of each student.

To evaluate whether the merger program in rural China is improving the education quality of primary schools in rural China, as it is intended to, we measure education quality by student performance on academic tests, as do many other empirical studies about the effect of policies on educational outcomes (Glewwe and Jacoby, 1994 in Ghana; Glewwe et al., 1995, in Jamaica; Tan et al., 1997, in the Philippines; and Glewwe et al., 2004, in Kenya). Our main measure of academic performance is based on the mathematics and Chinese language test scores of the students from 2001/2002 (their first grade year) to 2005/2006 (their fifth grade year).[5] Fortunately, in China every student in almost every primary school (including all of the schools in our sample) keeps in his/her possession a booklet that contains a comprehensive record of the mathematics and Chinese scores for each semester of his/her schooling, which we call "raw scores". Raw scores are on a 0-to-100 scale. The access to the scores from these written records means that the academic performance variables that we use in our analysis are record-based. In other words, the information on academic performance is not from recall, but from each student's grade booklet. The grades were copied by our enumerators with the assistance of the homeroom teachers.

[1] In some places in Ningxia province, *wanxiao* has five rather than six grades.

[2] In China, there are several *wanxiao* within each township. There is always one central primary school (which is always a *wanxiao*) in each township. The central primary school is usually located in the township seat. The township central primary school oversights the teaching and operation of all primary schools within the township. In almost all cases, township central primary schools are non-merger schools or merger-host schools.

[3] Enumerators in Shaanxi collected data on two extra randomly selected schools, that is why the number of sample schools is 62 rather than 60.

[4] According to our calculation using data published by statistical yearbooks of Shaanxi, Ningxia, Qinghai, Gansu and Xinjiang, in 2007, girls in rural areas of northwest China account for the same percentage, namely, 47%, of the class.

[5] In the case of the students that were held back, we recorded the grades of the first grade year, which unless the child had only been held back for first grade year, was prior to the 2001/2002 academic year.

In this paper, we focus on second semester mathematics and Chinese language scores because in most cases, the scores for these classes are based on a single year-end test that is standardized across the entire township (at least—sometimes it is standardized across the county). Our school survey shows that approximately 90% of the sample schools had a township- or county-wide standardized test in the second semester of each school year between 2001/2002 and 2005/2006. The examinations are standardized in two dimensions. First, the questions are the same for all students within the schools in the same township. Second, the final examinations were graded according to a single set of criteria by a township-wide panel of teachers.

In addition to using test scores from standardized township-wide examinations, we also standardize the raw score variable (Stevenson et al., 1990; Li, 2003; Huang, 2004). To do so, for each individual observation, we subtract the mean grade of the township and divide the result by the standard deviation. We call this newly transformed variable, the standardized score or *Z*-score. A *Z*-score of 0.2, for example, represents someone who scored 0.2 standard deviations above the average in the township.

Because of the nature of the way we constructed our variables, it can be shown that the *Z*-scores and raw scores are highly correlated and can both be used in our analysis as measures of academic performance. Our data show that the correlation coefficient between the *Z*-score and raw score variables is more than 0.80 (and significant at one percent level). Therefore, in the rest of the paper, since interpretation of the raw score measures is easier and more intuitive, we use the raw scores to describe the changes and distributions of the academic performance of students during the study period. In contrast, we use the *Z*-score in the multivariate analysis that evaluates the effect of primary school mergers on the academic performance of students.

Although we primarily look at changes between the second semester scores in the first and fifth grades, we also perform a number of sensitivity exercises to see how robust our findings are. For example, in one set of analyses we use the average grades for the entire year (both semesters) instead of just for the second semester. In such (and other) cases we did not find substantive differences between the results from these alternative approaches and those from our basic analysis (that we report below).

During the survey enumerators also recorded detailed information on the merger histories of each school during the study period (our key independent variable of interest). The principal of each school was asked: Was your school ever involved in a school merger? If yes, when was your school merged with another school? During the merger, which schools were closed down, and which schools remained open? Since these data were checked against several sets of records (and the responses of students and teachers), we believe these are measured with little error.

In addition to questions about academic performance and merging activities, other questions were also asked during the survey that could be used to create variables to control for other observed factors that might be expected to affect academic performance of each student. Three sets of variables were collected.[6] The first includes the following student-level characteristics that enumerators asked students themselves: gender, age, birth order, whether or not he/she was retained for a year at any time between first and fifth grade and whether or not they were student cadres (class monitors). The survey form also included the following questions about students' parents and family: the total number of household members, the educational attainment of each parent, and the household's holdings of land and non-land assets.[7]

Finally, a set of questions was asked to collect information about characteristics of each student's school and his/her teachers. During an interview with the school principals, the survey team collected information on the class size, the pupil-teacher ratio, the nature of the teaching facilities, the distance in kilometers between the student's home and his/her school; the teacher's gender, his/her teaching experience, his/her educational attainment and if the teacher was paid at least in part on the basis of the grades of his/her students (as proxied by a variable indicating whether teachers had ever received any teaching award at some point between 2001/2002 and 2005/2006 academic years, 0 = no, 1 = yes). Since some of these variables (e.g., changes in the distance between the student's home and school) are associated with the merging process, in our initial analysis we do not include them when we include our measure of mergers in order to avoid the problem of over-controlling (Duflo, 2003). In the last part of our analysis (in which we seek to understand some of the channels by which mergers affect academic performance) we drop the merger variable and add a number of these school/teacher characteristics. English translations of the forms for students, parents, teachers, and principals are available upon request.

4. Primary school mergers and academic performance

Reflecting the trends in national data (National Bureau of Statistics, 2006), there was a rapid reduction in the number of primary schools in our sample area since the introduction of the primary school merger program. During the study period, the number of primary schools serving the students in the sample decreased from 119 in 2002 to 62 in 2006, a reduction of almost half. Of the 62 schools that remained open in 2006, 26 schools (or 42% of them) were involved in some type of school merger during the sample period. The rest of the schools (36 schools or 58% of them) were not involved in any merger. Importantly, for our study, the fact that the school merger program was implemented in some schools, but not others, allows us to study the impact of China's school merger policy on the academic performance of students.

In looking at our data, it is understandable why if one were naively to visit rural areas and search out students from merger schools and ask them about their academic performance during the study period, the findings of such an inquiry could raise concerns about the impact of the primary school merger program on academic performance. When comparing the grades of merger schools between the first and fifth grades, the average raw score of students fell, decreasing from 79 in 2002 to 75 in 2006 (Fig. 2). More careful thinking, however, might raise a caution that this fall in grades could be for reasons besides the merger program. In fact, when looking at the change in average raw scores of students from non-merger schools it is clear that they also fell, from 78 in 2002 to 75 in 2006 (Fig. 2). It is possible that the falling scores for merger school students are simply a function of the fact that the material and tests are more difficult in the fifth grade when compared to the first grade.

If the interviewers had sought out students specifically from merger-guest and merger-host groups, the results of interviewing these students may have raised another concern about the possible negative effects of school mergers on merger-guest students. In our sample the students of merger-guest group on average had lower average raw scores during their fifth grade year than those from

[6] We collected information on these three sets of variables for two time periods: the second semester of the 2001/2002 academic year (i.e., before merger), and the second semester of the 2005/2006 school year (i.e., after merger).

[7] For questions about student's parents and family, enumerators first explained these questions to students. Then students brought questions home after school and asked their parents to answer and fill in those questions. When students brought the completed survey forms to school, enumerators checked each survey form carefully. Whenever necessary, enumerators would make calls to student's family to double check.

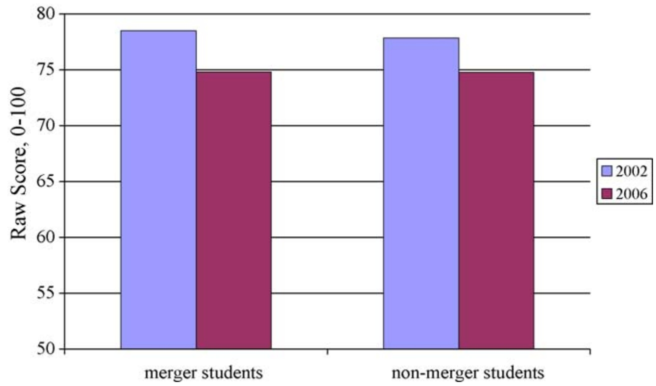

Fig. 2. Average raw scores of math and Chinese over time by merging status.

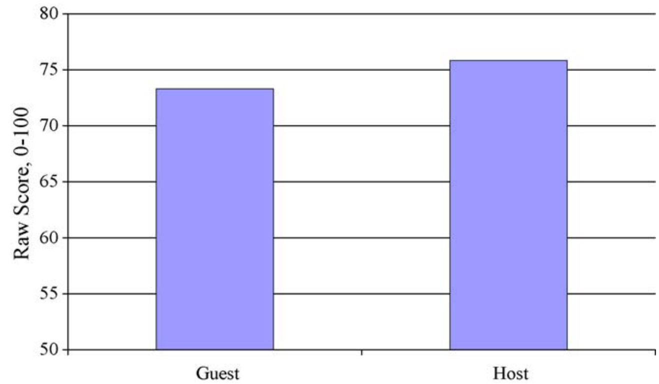

Fig. 3. Average raw scores of Chinese and math in 2006, guest versus host students.

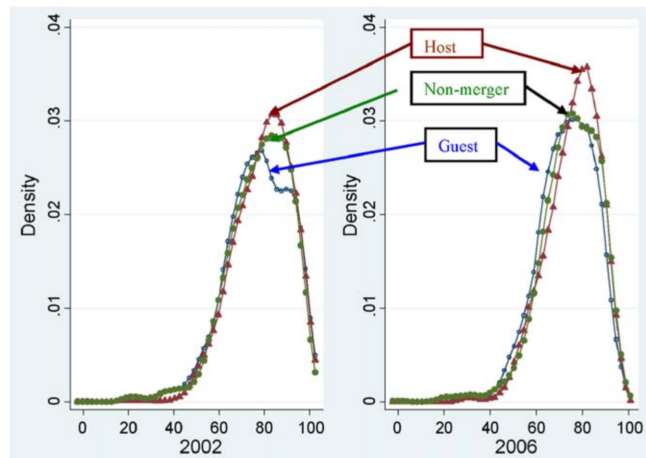

Fig. 4. Kernel density plots of distributions of average raw scores in 2002 and 2006, by merging status.

the merger-host group (Fig. 3). If these differences in scores are due to the school merger program, they would be consistent with some of our interviews in the field that parents of students in guest group were unhappy about the school merger program and sometimes did not believe the educational experience of their children was necessarily improving their academic performance.

The need to exercise caution in interpreting our results and move beyond descriptive statistics to more rigorous multivariate analysis is shown when comparing the distributions of scores of the different types of students from 2002 to 2006 (Fig. 4). Although the scores of fifth grade students from the merger-guest group were lower than those from the comparison group students in 2006, they also were lower in 2002 before the school merger program. In other words, at least on average (and without holding other things constant), the grades of the students were already low in the merger-guest group prior to times that their schools were closed down. In fact, although it is difficult to tell definitively by comparing the distributions in Fig. 4, one might actually infer that the school merger program did not have any negative effect on the students in either the merger-host or merger-guest groups. The distributions seem to maintain their relative positions between 2002 and 2006.[8]

5. Methodology

To examine the effect of the primary school merger program on the academic performance of students, the evaluation exploits two dimensions of variation. The first is temporal and comes from comparing the periods before and after the school merger (i.e., 2002 and 2006). The second is cross-sectional and comes from comparing students from merger and non-merger schools. The strategy therefore consists of comparing students before and after the school merger by merger status. School mergers can be considered as the *treatment* and our sample students need to be divided into two treatment groups and a comparison group. The treatment groups include (a) the merger-host students; and (b) the merger-guest students. The comparison group includes all the non-merger students (as defined in Section 1). With this setup, we employ a difference-in-difference method (DID) to compare the outcomes (i.e., academic performance) before and after a primary school merger for students whose grade one schools were involved in the change (merger students) to students whose grade one schools were not merged during the same period (non-merger students). This comparison produces what we call standard DID estimator.

In addition to the standard DID estimator, we implement three other DID estimators: an "unrestricted" version that includes the academic performance in 2002 or the pre-program outcome as a right hand variable, an "adjusted" version that includes a series of control variables from 2002 and an unrestricted + unadjusted model that combines the features of both the "unrestricted" and "adjusted" model. In summary, the models to be estimated are:

Model (1), restricted and unadjusted: $\Delta ZScore_i = \alpha + \delta_1 MG_i + \delta_2 MH_i + \varepsilon_i$

Model (2), unrestricted and unadjusted: $\Delta ZScore_i = \alpha + \delta_1 MG_i + \delta_2 MH_i + \gamma ZScore\_02_i + \varepsilon_i$

Model (3), restricted and adjusted: $\Delta ZScore_i = \alpha + \delta_1 MG_i + \delta_2 MH_i + \beta X_i + \varepsilon_i$,

Model (4), unrestricted and adjusted: $\Delta ZScore_i = \alpha + \delta_1 MG_i + \delta_2 MH_i + \gamma ZScore\_02_i + \beta X_i + \varepsilon_i$

where, i is an index for the student, $\Delta ZScore_i$ is the change in the second semester Z-score of student i between 2002 and 2006 (that is the final Z-score from the fifth grade minus the final Z-score from the first grade); MG_i and MH_i are the two treatment variables

[8] We did Kolmogorov–Smirnov equality of distributions test in Stata to check whether the distributions of test scores are equal between non-merger, merger-host and merger-guest students prior to merger (i.e., in 2001/2002). Based on the test statistics, we fail to reject the hypothesis that the distributions of test scores of *non-merger* students and *merger* students are equal (p-value = 0.735). We also failed to reject the hypothesis that the distributions of test scores were equal between *non-merger* students and *merger-guest* students (p-value = 0.550). Finally, we failed to reject the hypothesis that the distributions of test scores are equal between *non-merger* students and *merger-host* students (p-value = 0.195). In addition, we also did equality of means test in Stata to check whether the means of test scores are equal between non-merger, merger-host and merger-guest students prior to merger (i.e., in 2001/2002). However, we failed to reject the hypothesis that the means of test scores were equal between among *non-merger* students, merger-host students and *merger-guest* students (Prob > F = 0.7982).

Table 1
Comparison of mean characteristics of treatments and comparisons, before and after the school merger.

	Before (2002)			p-Value of F-test	After (2006)			p-Value of F-test
	Comparison (1)	Guest (2)	Host (3)	H_0: (1) = (2) = (3)	Comparison (4)	Guest (5)	Host (6)	H_0: (4) = (5) = (6)
Student characteristics								
Male, 1 = yes	0.53	0.54	0.53	0.915	0.53	0.54	0.53	0.915
Age, year	7.85	7.79	7.88	0.204	11.85	11.79	11.88	0.204
Has elder sibling? 1 = yes	0.55	0.52	0.52	0.361	0.55	0.52	0.52	0.361
Ever repeated a grade, 1 = yes	0.25	0.28	0.30	0.033	0.25	0.28	0.30	0.033
Student cadre, 1 = yes	0.28	0.28	0.28	0.950	0.30	0.27	0.23	0.009
Parent/household characteristics								
Dad's years of schooling	7.07	7.41	7.50	0.004	7.07	7.41	7.50	0.004
Mom's years of schooling	6.15	6.44	6.54	0.033	6.15	6.44	6.54	0.033
Holding of arable land, μ	6.86	6.21	6.43	0.098	6.98	6.37	6.54	0.123
Household size	4.80	4.67	4.64	0.018	4.80	4.66	4.57	0.000
Purchase value of durable assets, 1000 yuan	6.01	7.25	5.61	0.065	9.85	9.92	8.52	0.266
Distance from home to school, km	1.73	1.74	2.02	0.117	1.84	1.75	3.14	0.000
Teacher/school attributes								
% of male instructors	32.82	21.95	47.86	0.000	43.99	40.55	39.48	0.050
Years of teaching experience	13.82	13.31	19.90	0.000	17.84	18.09	16.88	0.011
Any teaching award ever received by instructors, 1 = yes	0.74	0.75	0.79	0.095	0.88	0.89	0.87	0.365
Classroom building made of brick or better material, 1 = yes	0.61	0.62	0.41	0.000	0.58	0.69	0.71	0.000
Modern teaching facility in the classroom	0.92	0.85	0.60	0.000	1.00	0.97	0.96	0.000
Class size	36.13	34.82	23.57	0.000	36.76	43.15	42.98	0.000
Has student dorm	0.05	0.08	0.00	0.000	0.08	0.12	0.14	0.000
Has student canteen	0.03	0.08	0.02	0.000	0.13	0.12	0.14	0.463
% of teachers with college and above diploma in the school	25.80	30.49	11.23	0.000	65.24	58.05	57.72	0.000
Pupil/teacher ratio	18.82	22.77	20.13	0.000	16.58	18.95	19.16	0.000

(which make δ_1 and δ_2 the parameters of interest). Finally, the term X_i is a vector of covariates that are included to capture the characteristics of students, parents, households and schools (as discussed above in Section 3). Throughout our analysis, X_i also includes a set of township dummy variables.

It is important to remember that the identification of the causal effects using DID relies on the assumption that absent the school mergers the average change in academic performance after and before school merger would have been the same for the treated and the comparison groups. Formally, this is called the "parallel trend" assumption.

As might be expected, the effectiveness of DID in a very real sense depends on the validity of this assumption. Whether or not the assumption is valid, however, depends on the context of the study and on how similar the comparison and treatment groups are. In general, the more similar are the treatment and comparison groups, the more convincing the DID approach. Using our data we find that although non-merger students (the comparison group), merger-guest students and merger-host students (the treatment groups) are not significantly different when comparing most of the children characteristics in 2002, there are differences when comparing certain parent, household, and teacher characteristics (Table 1). This finding suggests the importance of controlling for those characteristics in doing multivariate analysis. We also test these parallel assumptions below and use alternative methods (Propensity Score Matching and DID-Matching) to see if our results are robust.

6. Results of multivariate analysis

The results of the DID analysis using Models (1)–(4) demonstrate that the models perform fairly well and are consistent with our intuition (Table 2). For the version of the model that uses the change in the average Z-scores of the mathematics and Chinese language classes between grade one and grade five as the dependent variable, the goodness of fit measures (R-square ranges up to 0.34) are relatively high for this type of analysis. The coefficients on some of the control variables are also as expected. For example, when we use the unrestricted + adjusted specification of the empirical model (column 4), the Z-scores of students that are older in grade one drop relatively more than those of younger students (row 5). This finding is reasonable since, ceteris paribus, students that are older when entering primary school may have an initial advantage (because they are relatively more mature) that gradually disappears as younger children catch up over the course of primary school, which is consistent with other findings. Such an effect was found by Frederikkson and Öckert (2005): children who start school at an older age do better in school and ultimately end up going on to have more education than their younger peers. Additionally, when a student's mother has a higher level of education, the student's grades improve relatively more over time (row 10). While few papers in the literature have examined the impact of the mother's education on the change of grades, there is a large, related literature that shows the strong, positive correlation between mother's education and the academic performance of her children (e.g., Duncan et al., 1994).

We will focus mainly on the results of the unrestricted + adjusted specification (Model 4). We do so because this regression has a higher goodness of fit (or R-square) statistic. In part, almost certainly, this better fit reflects the importance of capturing beginning grades (and the unobserved ability of a student that is embodied in this measure) and the other covariates. Because of this, in the rest of the analysis, while we report all of the results from Models (1) to (4), we will mostly focus on the results from Model (4).

One of the most important findings in Table 2 is that we reject the hypothesis that primary school merger negatively affects the academic performance of *merger-guest* students. Although the coefficient of the merger-guest treatment indicator is negative in two out of the four specifications, none of them is statistically

Table 2
Difference in differences regression results analyzing the effect of school merger on academic performance of students in China, average *Z*-scores.

	Dependent variable = changes in the 2nd semester average *Z*-scores of Chinese and math between 2002 and 2006 ($\Delta ZScore$)			
	(1)	(2)	(3)	(4)
	Restricted and unadjusted	Unrestricted and unadjusted	Restricted and unadjusted	Unrestricted and adjusted
Treatment variables				
1. Student from a guest school? 1 = yes	−0.016	−0.017	0.002	0.016
	(0.30)	(0.40)	(0.03)	(0.36)
2. Student from a host school? 1 = yes	0.058	0.042	0.074	0.062
	(1.24)	(1.07)	(1.53)	(1.56)
Student characteristics in 2002				
3. Average of *Z*-scores of Chinese and math		−0.570		−0.623
		(33.01)***		(34.36)***
4. Male = 1, female = 0			−0.026	−0.085
			(0.68)	(2.72)***
5. Age, year			0.007	−0.045
			(0.31)	(2.48)**
6. Has elder sibling? 1 = yes			0.017	−0.019
			(0.42)	(0.59)
7. Ever repeated any grade? 1 = yes			−0.062	−0.234
			(1.24)	(5.69)***
8. Student cadre? 1 = yes			−0.190	0.117
			(4.45)***	(3.23)***
Parent/household characteristics in 2002				
9. Dad's years of schooling			−0.013	−0.003
			(1.63)	(0.51)
10. Mom's years of schooling			−0.002	0.015
			(0.26)	(2.40)**
11. Family land holding, μm			−0.001	0.002
			(0.39)	(0.75)
12. Household size, person			−0.010	0.007
			(0.58)	(0.48)
13. Purchase value of durable assets			0.001	0.001
			(0.72)	(0.48)
Teacher characteristics in 2002				
14. % of male instructors			−0.001	−0.000
			(0.83)	(0.05)
15. Ever received teaching award? 1 = yes, 0 = 0			0.251	0.188
			(4.74)***	(4.33)***
Constant	−0.016	−0.010	−0.048	0.332
	(0.52)	(0.40)	(0.17)	(1.41)
Observations	2446	2446	2446	2446
Number of town	20	20	20	20
R-squared	0.00	0.31	0.02	0.34

Note: Absolute value of t statistics in parentheses.
** Significant at 5%.
*** Significant at 1%.

significant (row 1). This means that the grades of the children of the merger-guest group did not fall relatively to the children of the non-merger group. In other words, ceteris paribus, primary school mergers – apart from the effect of the other factors included in the model – did not hurt the academic performance of merger-guest students as some have feared (e.g., Shi, 2004), at least in our sample area.

We also reject the hypothesis that primary school merger negatively affects the academic performance of *merger-host* students. In all four models the coefficient on the merger-host treatment indicator is not negative (row 2). In fact, the coefficients are all positive.

Although our data show that the primary school merger program in rural China generally has no negative impact on the overall academic performance of a typical student (being merger-guest or merger-host student), it is worth noting that taking the average of Chinese language and mathematics might mask the difference between the two subjects. Thus, we need to explore further whether the primary school merger program has any effect on the Chinese language and mathematics academic achievements, respectively. To implement this, we reran Models (1)–(4) using the changes in Chinese language and mathematics *Z*-scores between grade one and grade five, during the study period.

The same basic results hold when using the change in *Z*-scores in Chinese language class; there is no negative effect of primary school merger on academic performance on either merger-guest students or merger-host students (Table 3, Section A). In Table 3 we only report the coefficients of the treatment variables (that is, δ_1 and δ_2). The rest of the results are suppressed for brevity but are available from the authors upon request. In all four models the coefficient on the *merger-guest* treatment indicator is not statistically different from zero, although all are negative (row 1). This means that, ceteris paribus, primary school mergers did not hurt academic performance in Chinese language of merger-guest students, at least in our sample area. In the same version of the model, the coefficients on the *merger-host* treatment indicator are all positive, although not statistically different from zero (row 2). This indicates that primary school merger did not hurt the academic performance in Chinese language of merger-host students, either, at least in our sample area. Additionally, the

Table 3
Difference in differences regression results analyzing the effect of school merger on academic performance of students in China, by-subject *Z*-scores.

	Dependent variable = changes in the 2nd semester *Z*-scores between 2002 and 2006 ($\Delta ZScore$)			
	(1)	(2)	(3)	(4)
	Restricted and unadjusted	Unrestricted and unadjusted	Restricted and unadjusted	Unrestricted and adjusted
Section A: Chinese				
Student from a guest school during merger 2002–2006? 1 = yes	−0.034	−0.038	−0.014	−0.001
	(0.54)	(0.72)	(0.22)	(0.02)
Student from a host school during merger? 1 = yes	0.042	0.040	0.050	0.052
	(0.75)	(0.87)	(0.87)	(1.11)
Observations	2446	2446	2446	2446
Number of town	20	20	20	20
R-squared	0.00	0.31	0.01	0.34
Section B: math				
Student from a guest school during merger 2002–2006? 1 = yes	0.002	0.003	0.018	0.035
	(0.04)	(0.05)	(0.27)	(0.66)
Student from a host school during merger? 1 = yes	0.074	0.040	0.097	0.069
	(1.33)	(0.86)	(1.68)*	(1.46)
Observations	2446	2446	2446	2446
Number of town	20	20	20	20
R-squared	0.00	0.31	0.02	0.34

Note: Absolute value of *t* statistics in parentheses.
* Significant at 10%.

same basic results continue to hold when using changes in the *Z*-scores of *mathematics* as the dependent variable; there is no effect of primary school merger on academic performance in mathematics for neither merger-guest students nor merger-host students. In fact, for the case of merger-host students, the coefficient on the merger-host treatment indicator is even statistically different from zero in one out of the four models (Section B).[9]

So why is it that the primary school merger program has little impact on the academic performance of merger-guest students whereas there is some evidence of a positive impact on the test scores of certain subjects of the merger-host students? Although we cannot definitively say anything on the basis of the results of Tables 2 and 3, one possible reason is that after the merger, merger-host students remain in the same school and thus are not subject to the instability associated with switching schools. Although there might be some crowding associated with the increases in school size and class size, host students might also benefit from other changes in their newly merged schools. Following this argument, it appears as if the benefits of merging might dominate the crowding. Thus on the whole, we observe net positive impacts of the merger on the academic achievements of merger-host students. In contrast, the grade one schools of the merger-guest students were, by definition, closed down during the merger. Although they moved to the merger-host schools (and might suffer costs of dislocation, etc.), merger-guest students might also benefit from interacting with merger-host students and the better studying environment in the host schools. Thus, on the whole, we do not observe any net negative impact of the mergers on the academic performance of merger-guest students.

[9] In the statistically significant case in column 3, the size of the coefficient on the merger-host treatment indicator is 0.097, meaning that, everything else held constant, after their grade one schools were merged between the first and firth grade, the average *Z*-scores of math and Chinese language of the children from the merger-host group actually rose 0.097 points relatively to the children of non-merger group. When converted back to the 0–100 raw grade scale, this is equivalent to 1.5 points. In other words, primary school merger did not hurt academic performance of merger-host students, at least in our sample area.

6.1. Validity of the parallel trend assumption

As discussed in Section 5, in executing the DID model, we have assumed that absent a school merger the average change in the academic performance of those in the merger schools would have been the same as those in the non-merger schools. However, this is an assumption. It is possible, however, that students in the different types of schools would have performed differently (because there was something fundamentally – albeit difficult to measure – different. If the parallel trend assumption was not valid, the DID results would likely be biased.

In this subsection, we follow a simple "check" proposed by Duflo (2003) to test whether or not the parallel trend assumption holds in our analysis. To implement this check, we take a subsample of our data. Specifically, we focus on a subset of students whose grade one schools were merged only after they had begun to attend either grade four or grade five (and compare the shift in grades to the sample of non-merger schools). With this subset of data, we then redo the DID analysis comparing the test scores in grade one and grade three of the students that were still in the same school (even though eventually their schools were merged) with the test scores of students who were in schools which never were merged. If the coefficient on this "placebo treatment variable" in the DID is zero, it is in support for the validity of the parallel trend assumption. The logic is, of course, that if the merger and non-merger schools moved in concert with one another before the merger – apart from the merger effect – the non-merger-related academic performance trends should have moved together.

In fact, results from the placebo test demonstrate that the parallel trend assumption appears to be valid in our sample areas. Specifically, when we compare the change in test scores from grade one to grade three, neither of the coefficients on the merger indicator variables for the merger-guest or merger-host school is significantly different from zero (Table 4). In other words, the scores of our treatment (or guest students/host students) and comparison (or non-merger students) groups seem to be tracking one another fairly closely before the merger occurs. Therefore, it is fair for us to say that the results that we have produced from the

Table 4
Test of parallel trend assumption in difference in differences.

	Dependent variable: change in *Z*-scores between 2002 and 2004		
	Average	Chinese	Math
Student from a guest school during merger 2002–2006? 1 = yes	−0.022	−0.022	−0.036
	(0.41)	(0.33)	(0.55)
Student from a host school during merger? 1 = yes	−0.012	−0.002	−0.019
	(0.23)	(0.03)	(0.30)
Constant	0.345	0.141	0.618
(1.35)	(0.46)	(2.05)**	
Observations	1741	1741	1741
Number of town	20	20	20
R-squared	0.32	0.31	0.34

Note: Absolute value of *t* statistics in parentheses.
** Significant at 5%.

DID analysis are accurate, given that there is no evidence that the parallel trend assumption is violated.

7. Sensitivity analysis with alternative estimators

Despite the preceding analysis, the nature of our question (understanding the effect of primary school merger program on the academic performance of students) is actually so complicated that it could be that even though we control for a large number of observable variables in 2002 in the adjusted + unrestricted versions of the DID estimates, there could be other unobservable factors that may compromise the parallel trend assumption in another direction. Specially, because of the potential existence of unobservable differences between non-merger students and merger students (that could be correlated with mergers and grades), our DID results could be biased. In order to control for some of these unobservable factors, in this section we use propensity score matching (PSM) methods to address our main question and see if our results are robust to the empirical approach. PSM is an approach that does not require the parallel trend assumption.

In addition, we also seek to push the PSM approach further in order to eliminate the bias due to time-invariant unobservable differences between merger and non-merger schools (a shortcoming of the PSM approach). To do so, we extend the cross-sectional PSM approach to a longitudinal setting and implement a difference-in-differences matching (DIDM) strategy.

The results of the cross-sectional propensity score matching (PSM) analysis – regardless of the approach that we use for matching – also reveal that primary school mergers have no significant negative effect on the academic performance of either *merger-guest* or *merger-host* students (Table 5, columns 1 and 2). When examining the effect of school mergers on the academic performance of merger-guest and merger-host students using *Basic Matching* methods, there are no cases in which the coefficient on the merger-guest or merger-host treatment variables are negative and significant (rows A1, B1 and C1). This fundamental finding holds for the academic performance when we examine the effect on Chinese language scores, mathematics scores and the average of Chinese language and mathematics scores. The same is true when using *Multi-dimensional Matching* (rows A2, B2 and C2). In other word, the results from the PSM analysis are quite similar to those from the DID analysis; there is no measured negative effect of primary school merger on the academic performance of either the *merger-guest* or *merger-host* students—at least in our sample areas.

In addition, the findings remain largely consistent when using Difference in Difference Matching (DIDM—Table 5, columns 3 and 4). Regardless of whether we use Basic Matching or Multi-dimensional matching, and regardless whether we examine the average *Z*-scores of Chinese language and mathematics separately or together (using the average of the two), none of the coefficients of any of the *merger-guest* treatment variables are negative and significant. Moreover, our results also show that none of the coefficients of the *merger-host* treatment variable are negative. In

Table 5
Propensity score matching and multi-dimension matching estimators and the effect of school merger on the academic performance of students in rural China, 2002 and 2006.

Matching methods	Propensity score matching		Difference in differences matching	
	Average treatment effect for the *merger-guest* students	Average treatment effect for the *merger-host* students	Average treatment effect for the *merger-guest* students	Average treatment effect for the *merger-host* students
	(1)	(2)	(3)	(4)
Section A: average of Z-scores of Chinese and math				
A1. Basic matching	−0.034 (0.49)	0.078 (1.32)	−0.002 (0.03)	0.04 (0.57)
A2. Multi-dimensional matching	0.04 (0.79)	0.13 (3.08)***	0.094 (1.4)	0.073 (1.24)
Section B: Z-score of Chinese				
B1. Basic matching	−0.023 (0.09)	0.139 (1.9)*	−0.006 (0.10)	0.107 (1.26)
B2. Multi-dimensional matching	−0.005 (0.06)	0.141 (2.63)***	0.08 (0.08)	0.114 (1.63)
Section C: Z-score of math				
C1. Basic matching	−0.033 (0.40)	0.074 (1.08)	−0.038 (0.37)	0.014 (0.17)
C2. Multi-dimensional matching	−0.039 (0.62)	0.016 (0.31)	0.108 (1.36)	0.031 (0.46)

Note: *Z*-value or bootstrapped *t*-value in parentheses.
* Significant at 10%.
*** Significant at 1%.

fact, all the coefficients on the merger-host treatment variables are positive. Hence, whether using DID, PSM or DIDM, there not only is not any evidence that primary school mergers in our sample area have hurt the academic performance of either merger-guest students or merger-host students, in the estimations of many of our models we find that there is a small, positive effect.[10]

8. The sources of the merger–academic performance relationship

Although we found little effect of the school merger program on the academic performance of students, this one single coefficient on the merger indicator is, in fact, measuring simultaneously the many different dimensions of the impact of mergers. It is likely that while this overall net effect is something that policy makers, school administrators and rural communities would like to know, they might also be interested in which of the specific changes in the ways that schools are run that are affected by mergers has either a positive or negative impact on student academic performance. We will call these the "*source variables*".

The literature has examined a number of different source variables. For examples, in China there are a number of case studies – e.g., Yang (2004), Pang (2006), Xu (2006) and Xiong (2007) – that have analyzed the sources of the effect of mergers on academic performance. A close reading of the literature has identified ten possible variables: (a) school quality: including the distance of the school from a student's home (in kilometers); the quality of the classroom building (a dummy variable equals 1 if the building was made of brick and steel or some higher quality material; and zero if not); the availability of modern teaching facilities (proxied by a dummy variable, which is 1 if a school has any of the following teaching facilities: voice recorder, television, desktop computer, and long-distance learning system; and zero otherwise), the availability of boarding facilities (measured by two dummy variables—one indicating whether or not there is student dorm; and the other indicating whether or not there is a student canteen); (b) teacher quality: including the experience of the teaching staff (in average number of years); and the proportion of teachers with a college or above (professional school) diploma; (c) the size of the school, including the size of the class (measured in the number of students); and the pupil/teacher ratio; and (d) the nature of the peer group of each student (which is measured as the number of classmates that came from each student home village, a variable that we call a "*village peer effect*").

In order to examine the effect of these specific dimensions of mergers, we first examine descriptive relationships between some of these source variables and the academic performance of students. Next, we seek to redo the analysis that produced the results in Tables 2 and 3. In this case, however, we replace the merger indicator in Model (4) with these 10 variables of merger dimensions.

Table 6
Relationship between dimensions of merger and academic performance of students.

	No		Yes
1. Category: classroom building made of brick or better material			
Chinese	−0.024		0.018
Math	−0.003		0.003
Average of Chinese and math	−0.014		0.011
2. Category: modern teaching facility			
Chinese	−0.043		0.009
Math	−0.089		0.019
Average of Chinese and math	−0.066		0.014
	Lower	**Middle**	**Upper**
3. Category: years of teaching experience			
Chinese	−0.035	0.077	−0.037
Math	−0.064	0.023	0.046
Average of Chinese and math	−0.050	0.050	0.004
4. Category: class size			
Chinese	0.019	−0.007	−0.012
Math	0.057	0.022	−0.080
Average of Chinese and math	0.038	0.007	−0.046
5. Category: distance from home to school			
Chinese	0.000	−0.040	0.050
Math	0.056	−0.125	−0.009
Average of Chinese and math	0.028	−0.082	0.021
	No		**Yes**
6. Category: student dorm			
Chinese	0.000		−0.007
Math	0.012		−0.245
Average of Chinese and math	0.006		−0.126
7. Category: student canteen			
Chinese	−0.003		0.057
Math	0.006		−0.130
Average of Chinese and math	0.002		−0.037
	Lower	**Middle**	**Upper**
8. Category: % of teachers with college or above diploma			
Chinese	0.005	0.041	0.036
Math	0.003	0.067	0.065
Average of Chinese and math	0.004	0.054	0.050
9. Category: pupil/teacher ratio			
Chinese	−0.017	−0.028	0.046
Math	0.076	0.001	−0.080
Average of Chinese and math	0.029	−0.014	−0.017
10. Category: number of village peers			
Chinese	0.003	0.019	0.021
Math	0.018	0.004	−0.026
Average of Chinese and math	0.010	0.012	0.024

8.1. Descriptive results

While it is important for policy makers to know the net impact of mergers on academic performance, it is also important to try to identify the sources of the effect. To do so, we conduct cross-tabulations between each of the ten variables of merger dimensions and academic performance (Table 6). Our cross-tabulation analysis suggests that certain changes in the way education is organized that are associated with mergers are responsible for the findings in the previous section (Tables 2 and 3). One of the specific elements that are associated with mergers that are correlated with student performance in our descriptive statistics is the quality of classroom building. And the correlation remains positive whether

[10] Although the original study objective was to evaluate the impact of school merger on the academic performance of students, as requested by one of the anonymous referees, we also carried out HLM analysis to detect variations in the change in student academic performance before and after the school merger. Our sampling design leads to a three-level clustered dataset. Specifically, students (Level 1) are nested within randomly sampled schools (Level 2), which are in turn nested within randomly sampled townships (Level 3). We do not consider the class level since of the 62 schools that we visited in 2006, 46 schools (77%) have only one class of 6th graders (Level 1). In other words, we analyze data for 2446 students sampled from 62 schools in 20 townships. Results from HLM analysis show that: (a) there is significant variance in student academic performance among schools nested within townships; (b) student academic performance within townships is not correlated; student academic performance is modestly correlated within schools; (c) student-level covariates are effectively explaining some of the random variation in student academic performance at the student level (as expected). However, there is still unexplained random variation in student academic performance; (d) neither of the fixed effects associated with school- nor township-level covariates helps much to explain the random variation in student academic performance; and most importantly; (e) results from HLM analysis do not provide any evidence that school merger hurts the academic performance of students, which is consistent with what we conclude in the main analysis of our manuscript. In the revised manuscript, we discuss the steps taken to undertake HLM analysis and present the results in Appendix A.

we use the test scores in Chinese language, mathematics or an average of the two subjects (Table 6, Section 1). At least according to these descriptive findings, one of the reasons that school mergers are having a positive effect on academic performance may be due to the investments that are going into school buildings which is one part of the school merger program.

In addition, our descriptive analysis also demonstrates that when students have access to modern teaching facilities, such as voice recorder, television, desktop computer, or long-distance learning system, they tend to perform well. For example, as students move from schools with no access to any of these modern teaching facilities to schools with such access, their Z-scores on Chinese language rise from −0.043 to 0.009. This positive correlation suggests that the merger-related policies of trying to provide modern teaching facilities during the merger process maybe, to some extent, contribute to the success.

Finally, our data also show that teacher quality matters. As students move from classes that are in the lowest tercile of classes (when ranked in terms of the years of teaching experience of their teachers) to the highest tercile, their Z-scores on math rise from −0.064 to 0.046 (Section 3). This positive correlation suggests that the merger-related policies of trying to upgrade the quality of teachers during the merger process may be, at least, in part behind the success.

Unlike the cases of classroom building, teaching facility and teacher experience (which are factors that are associated with mergers that are shown to potentially enhancing academic performance), the descriptive data show that there are factors that are associated with mergers that detracting at the same time. Most prominently, as we showed above, during the process of mergers, class size grows (as we might expect). However, this rise of class size (at least in the descriptive statistics) does not appear to bringing positive economies of scale. Instead, we find that in schools in which class sizes get larger, academic performance deteriorates. If this dynamic is found to continue to hold up in multivariate analysis, then the effects of mergers would have been positive had it not been for the fact that class sizes rises with mergers. In fact, the negative relationship between class size and academic performance should not be surprising. Outside China, the education literature has a large number of research works that documents this negative relationship between class size and academic performance. For example, the negative relationship between class size and academic performance is consistent with what Hanushek (1999) found in Tennessee, the United States.

8.2. Multivariate analysis

In order to see if the descriptive cross-tabulations between the factors that are associated with mergers and academic performance have net effects on scores, we need to undertake a multivariate analysis.[11] The results from the version of Model (4) that replaces the merger indicators with the ten variables that seek to measure specific effects of mergers demonstrate that the model performs as well as the version with merger indicator variables (Tables 2 and 7). Specifically, the goodness of fit measures is up to 0.35. Moreover, the estimates of most of control variables have the similar sign, magnitude and significance level as in the model with merger indicator variables.

Most importantly, the multivariate analysis is consistent in a number of key ways with the cross-tabulation analysis. The most important finding in Table 7 is that we are able to identify a number of the sources of the *net* positive effect of mergers. In particular, our results show that the coefficients on teaching experience of the instructors and the quality of teaching building are positive and significantly different from zero, respectively. These basic results also appear when using the scores of mathematics and Chinese language separately as the dependent variable. In other words, school mergers appear to generate benefits in academic performance at least in part due to the investments that are being made into a better teaching environment and more experienced teachers.

Our results also demonstrate, however, that there are certain changes in the organization of schools that are triggered by mergers that dampen the positive effect of mergers. For example, for the version of Model (4) that uses the average of the mathematics and Chinese language scores as the dependent variable, the coefficient on the class size variable is negative and significant. This result is consistent with the cross-tabulation findings in Table 6. In addition, the number of village peers is also found to have negative impacts on the average test scores of students and their test scores of mathematics. In other words, these negative correlations suggest that the merger-related increase in class size and change in village peers maybe, in part dampen the success. Overall, if it had not been for these sources of the change (rising class room sizes and class placement decisions that fails to keep students from one guest school together in the same classroom in the host school after the merger), the positive effect of mergers would have been even larger.

9. Heterogeneous effects

While we have found no significant negative impacts of primary school mergers on the academic performance of children from merger schools, all of these results have been for the *average student* (that is, for the typical student from merger school). It is possible, however, that although on average there is no negative effect that there could be a negative effect on certain types of students from merger schools. In this subsection we examine whether or not school mergers affect those students that were attending different grades at the time when their schools were merged.

The reason that we focus on this aspect of the heterogeneity effects is because it is possible that school merging activities were undertaken when students were at different points of their primary education: grade two, grade three, grade four or grade five. Although the merging process is broadly similar across the four grades, the treatment effect may vary for several reasons. In particular, when school merging took place at an earlier point in students' primary education, children may have had more time to adapt to the new learning environment before they reached grade five. Alternatively, it is possible that the impact was more severe when it occurred in the case of younger children and may have had a relatively adverse impact. Hence, in this section we provide separate estimates of the treatment effect by including four interaction terms (merger indicators *times* the grade during which the merger occurred).

Unlike the results for the average student reported in Tables 2 and 3, the results from the DID analysis that examines the heterogeneity effects demonstrate that there is a significant negative effect of school merger on the Chinese language test scores of both *merger-host* and *merger-guest* students when merging was undertaken at a time early in primary school years (Table 8, column 1, rows 1 and 2). For the version of the model that uses Chinese language Z-score as the dependent variable, the coefficient on the interaction term between the dummy variable for merging in grade two and the *merger-guest* treatment indicator

[11] In this subsection, we will move beyond our concerns of over-controlling. Because variables of the different dimensions of mergers (e.g., teachers' education and experience and school facilities) might have an impact on school performance, in this section we add them to our model—first, in place of the merger indicator variables; then, simultaneously with them.

Table 7
Results from Model (4) with merger indicators replaced by dimensions of mergers.

	Average	Average	Chinese	Math
Treatment variables				
Student from a guest school? 1 = yes		0.005		
		(0.10)		
Student from a Host school? 1 = yes		0.053		
		(1.22)		
Dimensions of school mergers in 2002				
Distance from homestead to school, km	−0.004	−0.004	0.002	−0.009
	(0.75)	(0.70)	(0.27)	(1.46)
Number of classmates from the same village, person	−0.003	−0.002	−0.001	−0.005
	(1.82)*	(1.49)	(0.72)	(2.31)**
Class size, person	−0.005	−0.005	−0.004	−0.007
	(2.87)***	(2.74)***	(1.66)*	(3.11)***
Pupil/teacher ratio	−0.002	−0.003	0.004	−0.008
	(0.70)	(0.94)	(1.06)	(2.04)**
Years of teaching experience	0.007	0.007	0.007	0.007
	(2.76)***	(2.64)***	(2.40)**	(2.54)**
% of college and plus teachers in the school	−0.001	−0.001	−0.002	0.001
	(0.45)	(0.77)	(1.34)	(0.70)
School had any modern teaching facilities? 1 = yes	0.121	0.121	0.132	0.110
	(2.10)**	(2.02)**	(1.92)*	(1.59)
School had student dorm? 1 = yes	0.124	0.123	0.336	−0.030
	(0.85)	(0.81)	(1.92)*	(0.17)
School had student canteen? 1 = yes	−0.038	−0.035	−0.077	−0.033
	(0.23)	(0.21)	(0.40)	(0.17)
School house was of brick mix structure or better? 1 = yes	0.126	0.126	0.140	0.113
	(2.54)**	(2.46)**	(2.35)**	(1.90)*
Student characteristics in 2002				
Gender, 1 = male	−0.087	−0.087	−0.218	0.032
	(2.79)***	(2.80)***	(5.81)***	(0.86)
Pupil's age, year	−0.048	−0.047	−0.031	−0.073
	(2.67)***	(2.61)***	(1.46)	(3.40)***
Had any elder sibling? 1 = yes	−0.018	−0.017	0.004	−0.045
	(0.54)	(0.52)	(0.09)	(1.14)
Ever repeated any grade? 1 = yes	−0.231	−0.234	−0.230	−0.260
	(5.61)***	(5.68)***	(4.70)***	(5.31)***
A student cadre? 1 = yes	0.125	0.125	0.151	0.151
	(3.46)***	(3.44)***	(3.52)***	(3.51)***
Parent/household characteristics in 2002				
Dad's years of schooling	−0.003	−0.003	−0.001	−0.003
	(0.43)	(0.44)	(0.16)	(0.35)
Mom's years of schooling	0.015	0.015	0.015	0.017
	(2.36)**	(2.40)**	(2.02)**	(2.29)**
Family land holding, μm	0.002	0.002	0.001	0.004
	(0.80)	(0.75)	(0.19)	(1.34)
Household size, person	0.005	0.005	0.005	0.007
	(0.34)	(0.34)	(0.28)	(0.43)
Purchase value of household assets, 1000 yuan	0.001	0.001	0.000	0.001
	(0.47)	(0.45)	(0.30)	(0.41)
Teacher characteristics in 2002				
% of male instructors	−0.001	−0.001	−0.002	−0.001
	(2.20)**	(1.90)*	(2.75)***	(0.98)
Ever received any teaching award? 1 = yes	0.135	0.142	0.081	0.171
	(2.74)***	(2.87)***	(1.37)	(2.91)***
Z-Score in 2002	−0.627	−0.626	−0.684	−0.675
	(34.48)***	(34.39)***	(34.70)***	(34.50)***
Constant	0.483	0.461	0.184	0.819
	(1.89)*	(1.79)*	(0.60)	(2.69)***
Observations	2446	2446	2446	2446
Number of town	20	20	20	20
R-squared	0.35	0.35	0.35	0.35

Note: Absolute value of *t* statistics in parentheses.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

is negative and significant. In the same version of model, the coefficient on the interaction term between the dummy variable for merging in grade two and the *merger-host* treatment indicator is also negative and significant. In addition, coefficient is also negative and significant on the interaction term between the dummy variable for merging in grade three and the merger-host treatment indicator for the version of model that uses the average *Z*-score as the dependent variable.

Heterogeneity effect analysis also shows that primary school merger undertaken later in primary school years might have significant impacts on the academic performance of merger host students (row 6, columns 1–3). The coefficients on the interaction

Table 8
Difference in differences regression results with heterogeneous effects by the grade when merging was undertaken.

	Dependent variable = changes in the 2nd semester *Z*-scores between 2002 and 2006 ($\Delta ZScore$)		
	Chinese	Math	Average of Chinese and math
	(1)	(2)	(3)
1. Merged in grade Two × guest	−0.638	−0.141	−0.402
	(2.51)**	(0.55)	(1.88)*
2. Merged in grade Two × host	−0.886	0.385	−0.213
	(3.53)***	(1.52)	(1.01)
3. Merged in grade Three × guest	0.005	−0.006	−0.006
	(0.08)	(0.08)	(0.09)
4. Merged in grade Three × host	0.112	−0.011	0.050
	(1.74)*	(0.17)	(0.91)
5. Merged in grade Four × guest	0.007	0.058	0.042
	(0.08)	(0.72)	(0.63)
6. Merged in grade Four × host	0.185	0.142	0.160
	(2.49)**	(1.89)*	(2.56)**
7. Merged in grade Five × guest	0.092	0.119	0.096
	(0.85)	(1.09)	(1.05)
8. Merged in grade Five × host	−0.179	0.088	−0.036
	(1.55)	(0.76)	(0.37)
Observations	2446	2446	2446
Number of town	20	20	20
R-squared	0.35	0.35	0.35

Absolute value of *t* statistics in parentheses.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

term between the dummy variable for merging in grade four and the merger-host treatment indicator are positive and significant no matter we use the *Z*-scores of Chinese language, mathematics or their average as the dependent variable. In addition, the coefficient is also positive and significant on the interaction term between the dummy variable for merging in grade three and the merger-host treatment indicator for the version of model that uses *Z*-score of Chinese language as dependent variable. These results mean that, everything else held constant, school merger affects those students who were attending different grades at the time when their schools were merged in a heterogeneous way.

Although it is beyond the scope of this paper to isolate the exact reason why, during our field work a number of interviews revealed insights into what might be driving the finding that there are positive effects for older children and negative effects for younger children. One possible explanation might be that when kids experience school mergers in the early years of their primary education, they, being merger guests or merger hosts, find it difficult to adapt to changes in curriculum, instructor's teaching style and study environment associated with the merger. Thus their academic performance might suffer. However, when school merger happened latter in primary school years, kids might find it easier to adapt to changes associated with the merger—and are able to take advantage of the better school facilities and better teachers. In particular, for the case of merger-host students, the benefits they get from merger through the higher quality teaching environment, teaching facilities and more experienced teachers may be the reason why we see in the statistical analysis that the positive factors dominate the negative impacts associated increase in class size.

10. Summary and conclusions

In this paper we have tried to understand whether or not the academic performance of children suffer when the schools that they went to in grade one of their primary education years were merged. Despite a perception that is commonly found in the literature and the popular press, our results – somewhat surprisingly – show that there is little effect of the process of merging itself on the overall academic performance of the students from merger schools. Comparing the change over time (between the first and fifth grade) of the grades of students from merger-host group and merger-guest group, with those of students from schools that was not involved in any school merger during the study period, we can reject the hypothesis that primary school merger harms the grades of merger-guest and merger-host students when their grade one schools were merged sometime between the first and fifth year of their primary school education. We do find, however, that the timing of mergers matter; when students are older (e.g., the fourth grade) their grades rise after merging. The grades of younger students, however, fall. If this is true, policymakers might want to consider when to conduct primary school mergers in the future. Education officials possibly could consider trying to merge schools step by step, say, putting mergers off until students reach their fourth year in primary school education.

Based on the results, it might be tempting to conclude that since there is no measurable effect of primary school merger on the overall academic performance of average students, policy makers do no need to take any actions. If there was, education officials might want to consider trying to improve the environment in rural schools so that teachers could pay more attention to students in schools in which there were many children after the merger and/or many merger-guest children. This could be done by reducing class size or hiring more experienced teachers or improving classroom buildings. Provision of school buses, school meals and boarding schools might offer some of the services that students originally received from home when studied in the school closest to home before the merger. However, all of these programs are expensive. And, although there might be good reasons to implement such policies anyhow, according to a strict reading of our results, they should not be carried out because primary school merger has a negative effect on academic performance of an average child; at least in our study area there is no evidence that this is true.

It is worthy noting that there may be other effects of mergers that we are not capturing. For example, is there long-lasting health and psychological effect of merging—especially in schools in which students are forced into boarding schools when they are so young? Are there adverse effects on the community? What is the effect of merging on the income and the earning opportunities of the parents when they have so many more responsibilities to take and pick up children when they attend schools that are far from home. Few of these costs are accounted for in our analysis.

Acknowledgements

We acknowledge the financial support of the Ford Foundation (Beijing), and the National Natural Science Foundation of China (70903064).

Appendix A. HLM analysis

To perform the HLM analysis for this three-level dataset, we take the following steps drawing on the approach discussed in Braudenbush and Bryk (2002). First, we fit the initial "unconditional" (also referred to as "means-only" or "variance components") model. The unconditional model includes a fixed intercept, random effects associated with the intercept for schools (Level 2), and random effects associated with the intercept for townships (Level 3) (Column 1). We also fit a model without the random school effects and performed a likelihood ratio test to decide whether the random effects associated with the intercept for schools nested within townships could be omitted. Based on the significant result of the test ($p < 0.05$), we concluded that there was significant variance in student academic performance between schools nested within townships. Therefore, we retained the random intercepts associated with schools. We also retained the random intercepts associated with townships in all subsequent models to preserve the hierarchical structure of the data.

Based on the variance components estimates from this step (Section C, column 1), we estimated the intra-class correlation coefficient (ICC) to describe the similarity (or homogeneity) of observed change in student academic performance within a given level. Our estimates show that the ICC of observations on students within the same township is $0/(0 + 0.046 + 0.831) = 0$ whereas the ICC of observations on students within the same school nested within a township is $(0 + 0.046)/(0 + 0.046 + 0.831) = 0.052$. This tells us that observations on students in the same township are not correlated, while observations on students within the same school are modestly correlated.

In the second step, we added Level 1 (student level) covariates to the ‘‘unconditional'' model and evaluate the reduction in the residual variance. Specifically, we added fixed effects associated with 15 covariates (Section A1, column 2). Results show that the addition of the fixed effects of the student level covariates reduced the estimated between-student variance (i.e., the residual variance) by 38% [=(0.831 − 0.515)/0.831] (Section B, columns 1 and 2). This suggests that the 15 student-level covariates are effectively explaining some of the random variation in the change in student academic performance at the student level of the dataset (as expected). However, the magnitude of the variance components suggests that there is still unexplained random variation in the change in student academic performance of this dataset.

We also used a likelihood ratio test to decide whether we should add the fixed effects associated with all the of 15 student-level covariates. Based on the significant result ($p < 0.001$), we decided to add these fixed effects associated with Level 1 (student level) covariates.

In the third step, we added fixed effects associated with the 10 school-level (Level 2) covariates to see if they helped to explain random variation at different levels of the dataset (Section A2, column 3). We also used a likelihood ratio test to decide whether we should keep the fixed effects associated with the 10 Level 2 covariates. Based on the non-significant test results ($p > 0.1$), we did not retain these Level 2 (school level) covariates in future analysis.

In the last step, we added a fixed effect associated with the only township-level covariate, township population not older than 14 years old (Section A3, column 4). Based on the *z*-test for the fixed effect of township population not older than 14, we did not retain this fixed effect. Therefore, we chose the model with fixed effects associated with Level 1 (student level) covariates (i.e., Model 2) as our final model for the HLM analysis of our dataset (Appendix Table 1).

Appendix Table 1
Results from HLM analysis of variation in change of student academic performance.

	Dependent variable: changes in the 2nd semester *Z*-scores between 2002 and 2006 ($\Delta ZScore$)			
	Estimate (*t*-value)			
	(1) Condition-al model	(2) W/Level 1 covariates	(3) W/Level 2 covariates	(4) W/Level 3 covariates
A. Fixed-effect parameter				
Intercept	0.002	0.509*	−0.202	0.469
	(0.04)	(2.25)	(0.45)	(1.91)
A1. Level 1 (student level) covariates				
1. Student from a guest school during merger 2002–2006? 1 = yes		0.06	0.057	0.06
		(0.94)	(0.82)	(0.94)
2. Student from a host school during merger? 1 = yes		0.049	0.043	0.05
		(0.8)	(0.62)	(0.81)
3. Average of *Z*-scores of Chinese and math in grade one		−0.666***	−0.668***	−0.666***
		(38.09)	(38.15)	(38.09)
4. Male = 1, female = 0		−0.069*	−0.070*	−0.069*
		(2.31)	(2.34)	(2.31)
5. Age, year		−0.057**	−0.058***	−0.057**
		(3.28)	(3.36)	(3.29)
6. Has elder sibling? 1 = yes		−0.027	−0.028	−0.027
		(0.88)	(0.91)	(0.88)
7. Ever repeated any grade? 1 = yes		−0.229***	−0.226***	−0.229***
		(5.83)	(5.74)	(5.83)
8. Are you a student cadre? 1 = yes		0.388***	0.390***	0.388***
		(11.3)	(11.34)	(11.3)
9. Dad's years of schooling		−0.002	−0.002	−0.002
		(0.36)	(0.31)	(0.34)
10. Mom's years of schooling		0.017**	0.017**	0.017**
		(2.85)	(2.86)	(2.85)
11. Family land holding, μm		0.002	0.002	0.002
		(0.74)	(0.77)	(0.73)
12. Household size, person		0.003	0.003	0.003
		(0.21)	(0.19)	(0.2)
13. Purchase value of durable assets, 1000 yuan		0.001	0	0.001
		(0.59)	(0.5)	(0.58)
14. Distance from home to school, km		0.005	0.004	0.005
		(0.85)	(0.67)	(0.86)

Appendix Table 1 (*Continued*)

	Dependent variable: changes in the 2nd semester *Z*-scores between 2002 and 2006 ($\Delta ZScore$)			
	Estimate (*t*-value)			
	(1)	(2)	(3)	(4)
	Uncondition-al model	W/Level 1 covariates	W/Level 2 covariates	W/Level 3 covariates
15. No. of classmates from the same village		0.001 (0.41)	0.001 (0.38)	0.001 (0.38)
A2. Level 2 (school level) covariates				
16. % of male instructors			0.002 (1.65)	
17. Ever received teaching award? 1 = yes			−0.001 (0.79)	
18. Classroom building made of brick or better material, 1 = yes			0.045 (0.52)	
19. Modern teaching facilities in the school			0.313 (0.99)	
20. Years of teaching experience			0.005 (0.86)	
21. Class size			0.005 (1.09)	
22. Has student dorm, 1 = yes			−0.052 (0.16)	
23. Has student canteen, 1 = yes			0.136 (0.47)	
24. % of teachers with college and above diploma in the school			0.001 (0.55)	
25. Pupil/teacher ratio			−0.001 (0.07)	
A3. Level 3 (township level) covariate				
26. Township population not older than 14 years old				0 (0.43)

	Dependent variable: changes in the 2nd semester *Z*-scores between 2002 and 2006 ($\Delta ZScore$)			
	Estimate (SE)			
	(1)	(2)	(3)	(4)
	Uncondition-al model	W/Level 1 covariates	W/Level 2 covariates	W/Level 3 covariates
B. Covariance parameter				
$\sigma^2_{\text{int:township}}$	2.01E−14 (2.19E−13)	7.85E−17 (7.38E−16)	3.56E−13 (2.67E−10)	9.16E−17 (1.17E−15)
$\sigma^2_{\text{int:school}}$	0.046 (0.015)	0.060 (0.017)	0.068 (0.019)	0.061 (0.019)
σ^2 (residual variance)	0.831 (0.024)	0.515 (0.015)	0.514 (0.015)	0.515 (0.015)
C. Model information criteria				
−2 RE/ML log-likelihood	6559.0	5510.3	5568.4	5530.2
AIC	6567.0	5548.3	5626.4	5570.2
BIC	6590.2	5658.6	5794.7	5686.2
#Obs	2446	2446	2446	2446

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

References

Angrist, J., Lavy, V., 2002. New evidence on classroom computers and pupil learning author(s): Joshua Angrist and Victor Lavy Source. The Economic Journal 112 (482), 735–765.

Braudenbush, S., Bryk, A., 2002. Hierarchical Linear Models: Applications and Data Analysis Methods. Sage Publications Inc..

Chen, X., Liu, C., Zhang, L., Shi, Y., Rozelle, S., 2010. Does taking one step back get you two steps forward? Grade retention and school performance in poor areas in rural China. International Journal of Educational Development 30 (6), 544–559.

Chiu, M., Ho, S., 2006. Family effects on student achievement in Hong Kong. Asia Pacific Journal of Education 26 (1), 21–35.

Duflo, E., 2003. Empirical Methods. Handout Prepared for the Course of 14.771 Development Economics: Microeconomic Issues and Policy Models' MIT, http://web.mit.edu/14.771/www/lec-notes.html.

Duncan, G.J., Brooks-Gunn, J., Klebanov, P.K., 1994. Economic deprivation and early childhood development. Child Development 65, 296–318.

Frederikkson, P., Öckert, B., 2005. Is early learning really more productive? The effect of school starting age on school and labor market performance. IZA Discussion Paper No. 1659, Institute for the Study of Labor (IZA).

Fuligni, A., Stevenson, H., 1995. Time use and mathematics achievement among American, Chinese, and Japanese high school students. Child Development 66 (3), 830–842.

Glewwe, P., Jacoby, H., 1994. Student achievement and schooling choice in low-income countries: evidence from Ghana. The Journal of Human Resources 29 (3), 843–864.

Glewwe, P., Grosh, M., Jacoby, H., Lockheed, M., 1995. An electric approach to estimating the determinants of achievement in Jamaican primary education. Economic Review 9 (2), 231–258.

Glewwe, P., Kremer, M., Moulin, S., Zitzewitz, E., 2004. Retrospective vs. prospective analysis of school inputs: the case of flip charts in Kenya. Journal of Development Economics 74, 251–268.

Glewwe, P., Kremer, M., Moulin, S., 2009. Many children left behind? Textbooks and test scores in Kenya. American Economic Journal: Applied Economics 1 (1), 112–135.

Guo, Q., 2007. On the concrete effects of the rural primary secondary school mapping restructure: positive research on the basis of six provinces and regions in the midwest. Education and Economy 2, 45–49.

Hanushek, E.A., 1999. Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. Economic Evaluation and Policy Analysis 21 (2), 143–163.

Ho, S., 2004. Accomplishment and challenges of Hong Kong education system: what we have learned from PISA. Education Journal Winter 2003 31 (2), 1–30.

Ho, S., 2006. High-stakes testing and its impact on students and schools in Hong Kong: what we have learned from the PISA studies. KEDI Journal of Educational Policy 3 (1), 69–87 (Korea, South: Korean Educational Development Institute).

Huang, Y., 2004. Using standardized score to evaluate the academic performance of students. Journal of Liaoning Institute of Technology 6 (4).
Li, F., 2003. Evaluating student's achievement in a scientific way: an application of standardized score. Journal of Liaoning Teachers College (Social Sciences Edition) 1 .
Ministry of Education, 2006. China Education Statistics. , http://www.moe.edu. cn/edoas/website18/1020.htm.
National Bureau of Statistics, 2006. China Statistics Yearbook. China Statistics Press, Beijing.
Pang, L., 2006. Primary and high school consolidation in rural China: issues. Reasons, and Solutions. Education Development Research 2, 1–6.
Purcell, D., Shackelford, R., 2005. An evaluation of the impact of rural school consolidation. What challenges may a new round of rural school consolidations have on the safety, educational performance, and social environment of rural communities? In: A Manuscript Presented to the National Rural Education Association Executive Committee, January 13–15, 2008, http://www.eric.ed.gov/ERIC Docs/data/ericdocs2sql/content_storage_01/0000019b/80/29/90/90.pdf.
Shi, R., 2004. Experience drawn from research on school consolidation in foreign countries. Comparative Education Research 12.
Stevenson, H., Stigler, J., 1992. The Learning Gap: Why Our Schools are Failing and What We Can Learn from Japanese and Chinese Education. Summit Books, New York.
Stevenson, H., Lee, S., Lewis, C., Rohlen, T., 1998. An examination of American student achievement from an international perspective. Brookings Papers on Education Policy 1, 7–52.
Stevenson, H., Lee, S., Chen, C., Lummis, M., Stigler, J., Fan, L., Ge, F., 1990. Mathematics achievement of children in china and the United States. Child Development 61 (4), 1053–1066.
Tan, J.-P., Lane, J., Caustere, P., 1997. Putting inputs to work in elementary school: what can be done in the Philippines? Economic Development and Cultural Change 45 (4), 857–879.
UNESCO, 2006. Education for All Global Monitoring Report 2002. UNESCO, Paris, France.
Xinhuanet, 2002. Focal Talk Online Restructuring of Primary Schools in Rural Villages by Dinghaijun. http://news.xinhuanet.com/misc/2002-12/30/content_672012.htm.
Xiong, X., 2007. Reflection on the present adjustment to the primary and secondary schools distribution in the rural areas. Education and Economy 2, 50–53.
Xu, Y., 2006. Concerns about elementary school mergers. The School Principals 6, 42–44.
Yang, Y., 2004. Rural elementary school mergers: issues and suggestions. Modern Elementary and Secondary Education 3, 3–4.
Yip, D., Chiu, M., Ho, S., 2004. Hong Kong student achievement in OECD-PISA study: gender differences in science content, literacy skills and test item formats. International Journal of Science and Mathematics Education 2 (1), 91–106 (National Science Council, Taiwan).
Zhuo, M., 2006. Primary school consolidation in rural China: advantages, disadvantages and solutions. Education Forum, The Education Department of Fujian Province. http://218.85.72.34/newsInfo.aspx?pkId=10103.