LICOS Centre for Institutions and Economic Performance
Centre of Excellence

LICOS Discussion Paper Series

Discussion Paper 282/2011

Transfer Paths and Academic Performance: The Primary School Merger Program in China

Di Mo, Hongmei Yi, Linxiu Zhang, Yaojiang Shi,
Scott Rozelle and Alexis Medina

Katholieke Universiteit Leuven

LICOS Centre for Institutions and Economic Performance
Waaistraat 6 – mailbox 3511
3000 Leuven

BELGIUM

TEL:+32-(0)16 32 65 98
FAX:+32-(0)16 32 65 99
http://www.econ.kuleuven.be/licos

Transfer Paths and Academic Performance: The Primary School Merger Program in China

Di Mo
LICOS Centre for Institutions and Economic Performance, K.U. Leuven

Hongmei Yi and Linxiu Zhang
CCAP, IGSNRR, Chinese Academy of Sciences

Yaojiang Shi, Xibei University
College of Economic Management

Scott Rozelle and Alexis Medina
Stanford University

December 2010

Abstract

In the late 1990s and early 2000s, China's Ministry of Education embarked on an ambitious program of elementary school mergers by shutting down small village schools and opening up larger centralized schools in towns and county seats. The goal of the program was to improve the teacher and building resources in an attempt to raise the human capital of students in poor rural areas, although it was recognized that students would lose the opportunity to learn in the settings of their own familiar villages. Because of the increased distances to the new centralized schools, the merger program also entailed building boarding facilities and encouraging or mandating that students live at school during the week away from their family. Given the magnitude of the program and the obvious mix of benefits and costs that such a program entails there has been surprisingly little effort to evaluate the impact of creating a new system that transfers students from school to school during their elementary school period of education and, in some cases, making student live in boarding facilities at school. In this paper, our overall goal is to examine the impact of the Rural Primary School Merger Program on academic performance of students using a dataset from a survey that we designed to reflect transfer paths and boarding statuses of students. We use OLS and Propensity Score Matching approaches and demonstrate that there is a large "resource effect" (that is, an effect that appears to be associated with the better facilities and higher quality of teachers in the town and county schools) that appears to be associated with the transfers of students from less centralized schools (such as, village schools) to more centralized schools. Boarding, however, is shown to have negative impacts on academic performance. However, students who transfer to county school benefit from the transfer no matter where they start and whether they board or not.

Acknowledgements:

The authors gratefully acknowledge the collaboration of Zigen, an NGO that is working in the region on a series of educational projects. The authors wish to provide a special thanks to Jade Chien, Funden Wang and his wife (and their foundation, IET) and LICOS (KULeuven) for the financial and moral support for the evaluation part of the project. Carl Brinton, Renfu Luo, Ai Yue and Lei Liu also deserve thanks for their assistance with data collection and useful collaboration throughout the project.

Transfer Paths and Academic Performance:
The Primary School Merger Program in China

1. Introduction

In 2001 as part of its effort to improve the overall level of primary education and address the educational disparities between urban and rural areas, China's State Council announced the Rural Primary School Merger Program (Ministry of Education, 2001). The Program essentially involves shutting down remote village schools and "teaching points" (*jiaoxuedian*)—one-room type school houses located in villages that offer schooling to students in grades 1 through 3 or 4—and merging them into centralized town or county schools. Implementation of the Merger Program accelerated in the early and mid-2000s. For example, 31,700 primary schools were closed down and merged in 2004 (MOE, 2005). The number of primary schools in rural China fell by nearly 24 percent between 2001 and 2005 (National Bureau of Statistics, 2006).

The Merger Program has created a lot of debate in the literature about whether the policy is benefiting children in poor rural areas. In theory, students are supposed to benefit from improved educational quality by having access to larger, more centrally located educational facilities which can be built in such a way as to take advantage of scale economies. Better teachers can be hired. Facilities can be built to higher quality standards and equipped better. In larger schools, teachers are able to focus on students in a single grade and, in many cases, on a single course. In contrast, teaching points, which are remotely located schools sometimes accommodating fewer than 10 students, typically have only one teacher per school, who is responsible for teaching several different grades and acting as the managing staff for the whole school. The curriculum is often restricted to math and Chinese language—with little supplementary teaching of science, art, music or other types of courses. Central schools are supposed to offer a richer curriculum, including science, English, art, music and other subjects. In the rest of the paper, when schools have better facilities and higher quality teaching staff, we will call this the *resource effect*.

While there may be many potential benefits to the Merger Program, there also are potential costs. Pang (2006) describes a number of the detrimental effects. Because the new centralized schools are often located far from the homes of students, children must be boarded at school. In some cases this means that six and seven year olds have to leave the comfort and familiarity of their homes and care of their parents to live in dormitories far away from their friends and family. Safety is certainly a concern. The new living environment may take a toll on the psychological and physical health of students and thus affect learning (Luo et al., 2009).

Another set of potential costs of the Merger Program stems from student transfers out of schools that are being shut down and into the new centralized schools. Similar costs are incurred even when students are not caught up in mergers but when parents transfer their children from one school (say a school in the village or town) to another school (for example, a school in the county seat). We term the ways in which students have been shuffled around the school system (or the different channels by which students move from grade 1 to grade 6) *transfer paths*. Possible transfer paths depend on the local government's implementation of the Merger Program and on decisions made by the parents themselves. For instance, Shanxi Province piloted its Merger Program in 2001. The program was expanded in 2004. By the end of 2004, 829 primary schools had been shut down. These efforts were part of their plan to remove all teaching points and cut village schools from 2224 in 2007 to 800 in 2010 (MOE, Lüliang County). Students who attended the teaching points or village schools that closed were transferred to a town, county or other village school. Liu et al. (2010) show that each provincial and county government has taken a different approach to setting up the Merger plan. These have included: merging all sub-village teaching points into one school for each (one or two) village(s) if the village has a population above a certain threshold (Henan Province); shutting down teaching points that have only one teacher and merging them into surrounding village schools (Yunnan Province); establishing large, centralized town and county boarding schools to receive students from nearby villages and teaching points (Qinghai Province); and many others. In our interviews in China's poor northwest region, we often find that all of these different transfer paths can exist in a single county. As a result, specific transfer paths differ by student even within the same county. Since each transfer path has its own unique set of benefits (resource effects vary across schools) and costs (abrupt changes in environment, embodied in different transfer paths, can affect students differently depending on age, etc.), it is possible that different transfer paths will have different impacts on the educational performance of students.

Empirically, the success of the Merger Program has been mixed. There is no doubt that centralized schools have better teachers, facilities and curriculum offerings (Zhuo, 2006). However, there are documented costs as well. Shi (2004) has shown that when boarding schools are poorly managed, children perform worse in school. Some studies have found that the poor nutrition and health in boarding schools (relative to the home environment) are correlated with poor educational performance (Luo et al., 2009; Luo et al., 2010). Shi et al. (2009) has evidence that students who transfer from their own village's teaching points into boarding facilities in a centrally located township schools have more behavioral and psychological problems. To date, only one research team (to our knowledge) has attempted to empirically disaggregate the costs and benefits to determine the net effect of the Merger Program on students. Using data from a large sample in Shaanxi province, Liu et al. (2010) find that the overall effect of transferring students from a village school or teaching point closed under the Merger Program to a larger, more central school is neutral; that is, the benefits from the improved resource effect are similar in magnitude to the costs. The question of whether and how different transfer paths lead to better academic performance has not yet been quantitatively analyzed.

The key questions we attempt to address in our study include: What transfer paths are students taking as a result of the Merger Program and other educational policy shifts? How are student test scores affected by the nature of the transfer path? Does the Merger Program lead to improved academic performance? Are there any negative impacts of the

Merger Program? How are these related to the different transfer paths or whether students are boarded?

In order to answer the above questions, we first outline the transfer paths that students have taken and the distribution of these transfer paths in our sample. We then compare standardized math test scores of students who took different paths. We also identify characteristics of the educational experiences of students that were (may have been) affected by the Merger Program and examine their impact on students that took different transfer paths. From this analysis, we make general assessments of the academic costs and benefits that the Merger Policy imposes on students.

In the next section, we describe the data collection process and features of the dataset. In Section 3 we conduct a descriptive analysis of transfer paths and student academic performance. The econometric model is specified in Section 4 and the results are discussed in Section 5. In the last section we conclude.

2. Data

The dataset we use is generated from a survey carried out by the Chinese Academy of Sciences and Xibei University of Xian in September, 2009 in a poor county in Shanxi Province. The county, located on the Yellow River in northwest Shanxi province, is an appropriate place to study the impact of transfer paths on the educational performance of students in poor areas for several reasons. The active Merger Program and other policy changes allowing for various transfer options make our study county a place where transfer paths have changed significantly over the past several years. The county is one of the poorest county in Shanxi Province and one of the poorest in the nation, making it representative of poor areas in China. In 2008 the average yearly income of farmers in the county was 1024 Yuan (150 USD at nominal exchange rates). In 2008 less than 10 percent of China's population earned under 0.57 USD per day. The county also has many features common to other regions in Northwestern China, a region of the country defined based on geographical and cultural characteristics (Han, Regional Geography of China, 2000). It is located in the mountainous region of the Loess Plateau. Although there are natural resources (e.g., coal—which by policy belongs to the State), it has scarce agricultural resources due to climate, soils and access to markets.

Our sample consists of the entire seventh grade (first year of junior high school) population in the county in 2009. The entire primary school experience of these students took place during the implementation of the Merger Program, which started in 2001 and accelerated afterwards. All ten junior high schools in the county were visited. A total of 1507 students participated in the survey, with the participation rate exceeding 99 percent. The students that were surveyed have characteristics that are typical of rural seventh graders in China. There are around 6 percent more boys than girls, a similar ratio to that cited in the Ministry of Education's 2006 Annual Yearbook. Approximately 95 percent of the students are aged between 11 and 14 years of age. Around 23 percent of the students had been held back one or more grades during primary school (see Chen et al., 2009, for a complete discussion on retention).

The survey instrument included three main blocks. The first block focused on the schooling histories of students. We asked students a series of questions about their time in their elementary schools as a way to re-create each student's transfer path from grade 1 to grade 6. Specific questions included primary schools ever attended, which grades were spent in each school, school location, reasons for each transfer and boarding status. We produce from these questions several variables: ***student transfer path***; ***boarding status***; and a number of controls for pre-primary educational experience.

The second section was a 30 minute standardized math test. This test is used as a measure of each student's educational performance. Using tests on basic skills, such as math, to serve as a measurement of academic performance is a common practice in the literature (Reynolds, 1991; Glewwe et al., 1995; Tan et al., 1997; Gruman et al., 2008 etc.). Because we administered the survey/test ourselves, we know that there was no coaching for the test before our survey. Since the test is administered at the start of the school year, we also know that neither students nor teachers shifted their efforts from other subjects to math. The test was scored on a scale from 0 to 100. The results we obtained closely approximate a normal distribution with a mean score of 56 points and a standard deviation of 17 points. We keep the scores without any further manipulation for the ease of interpretation.

The third and final section of the survey contained a number of questions on each student's personal and family characteristics. These questions gathered data on each student's age, gender, household registration (*hukou*) and ethnicity. Information on the socio-economic background of students was also obtained through questions about the number of members in each student's family and each family member's hukou status, age, employment status and schooling history. The answers to detailed questions about household assets were used to generate a variable measuring the value of the household durable assets to represent household socioeconomic status or wealth. All of the control variables in our econometric model are produced from the above information.

3. Transfer paths and academic performance

In part because of the closing and/or merging of a large number of schools, nearly half (49 percent) of our sample transferred from one school to another at some point during their primary school years. Our data contain many unique starting and ending points for the transfer experiences of students which we use to identify a variety of student transfer paths. In this section, we describe these transfer paths, identify the most common paths and link them (descriptively) with academic performance.

Student transfer paths

In examining the starting school (first school attended between grades 1 and 6) and ending school (last school attended between grades 1 and 6) of each student's primary school experience, a number of student transfer path patterns emerge (Table 1). Our data show that more students transfer *to* town and county schools than transfer *from* them. Likewise, more students transfer *from* teaching points and village schools than transfer *to* them. Indeed, no students in our sample transferred *to* a teaching point. This pattern suggests that the activities in our sample counties are consistent with the goals of the

Merger Program; that is, students are being encouraged to transfer from teaching points or village schools to more centralized town and county schools.

Our analysis focuses on the student transfer paths of students who started primary school in teaching points or village schools and transferred to more centralized schools, as these students are the target population of the Merger Program and account for about 71 percent of all transfer experiences. These specific student transfer paths also form one of the bases for our analysis.

Of students who started school in teaching points (25 percent of all students), the length of stay in the teaching point varies but does not exceed four years (Table 2). This is because teaching points, despite being an important component of the traditional rural primary education system, usually do not provide education beyond the fourth grade. Our data reflect this fact: no students in our sample complete their primary education at a teaching point and no students remain enrolled at a teaching point beyond the fourth grade (although students can theoretically spend more than four years at a teaching point if they fail to matriculate to the next grade after one academic year).

The ending schools vary for students with identical starting points. The majority of students who started in teaching points eventually transferred to town schools (Table 3—around 56 percent). Another 30 percent transferred to county schools. Most of the students who started in village schools also transferred to a more centralized school, either town or county schools. Only a small share of students transferred to village schools. Only 13 percent of students who started in teaching points and 14 percent of students who started in village schools transferred to a (another) village school. This movement away from village schools is likely (at least in a significant part) because of the closing of village schools under the Merger Program.

Academic performance

Our data show that mean math scores are correlated with different student transfer paths, with the direction of correlation seemingly determined by the resource effect (Table 3). All scores over 60 are associated with either starting or ending education in county schools. Moreover, when the starting points are held constant, test scores decrease with the level of centralization of the ending points. In other words, students who have attended county schools (as their ending schools) have the highest scores; students who have attended town schools achieve the second highest scores; and students who have attended village schools have the lowest scores.

Using kernel density plots (Figures 1 and 2), we can provide distributional evidence on the impact of transfer paths on math scores beyond the mean comparisons. Figure 1 includes the plots using information from the group of students that started in teaching points; Figure 2 includes plots using information from the group of students that started in village schools. The figures show that the mean difference is caused not by a small group of extremely high-achieving students but by overall improvements in scores (across the distribution). Figure 2 shows that the mean scores increase and the distribution better approximates a normal distribution as the students ending school change from village schools to town schools to county schools. Overall, then, our data indicate that students perform better academically when they transfer from less-centralized schools to more-centralized schools, and that their performance increases most when they transfer to county schools. This trend is true of all students, regardless of where they started schooling.[1]

Boarding status also can be shown to be correlated with math scores (Table 4). According to our data, non-boarding students have higher mean math scores than do boarding students. The difference between boarding and non-boarding student reaches 10 points (or about 0.6 standard deviations) and is significant at the 1% level. This trend holds true for mean math scores within each student transfer path. In fact, boarding students, on average, never score higher than non-boarding students.

Other student characteristics

Other characteristics—beyond their transfer paths and boarding school status—also may affect academic performance. According to the literature (Shariff, 1998; Gibson, 2001; Borooah, 2005; Linnemayr, Alderman, & Ka, 2008; Chen & Li, 2009, Liu et al 2010 etc.) individual student characteristics, such as gender, age, hukou identity, kindergarten and preschool attendance and the number of elder siblings may affect educational performance. Parental characteristics (age, education and occupation) and household characteristics (e.g., household size and wealth) also have been shown to affect academic performance.

---

[1] One exception to the ending school trend: students who started in a county school earned average test scores of 61 or 62 regardless of where they ended.

Descriptive statistics show that some of these variables seem to be associated with student test scores in our sample (Table 5). For example, students with higher scores seem to be younger, have kindergarten experience, have no elder siblings, have better educated parents with off-farm jobs and come from non-rural and richer households. These findings underline the importance of conducting multivariate analysis and including parental and household characteristics in the analysis as control variables since they may also be correlated with student transfer paths.

4. Multivariate model

The data and descriptive analysis presented in the previous section show substantial differences in math scores across student transfer paths. However, based on a simple comparison of means it is impossible to satisfactorily attribute the differences in scores to the different student transfer paths. In this section we present an econometric analysis to address this issue. We first present different estimators and specifications and we then discuss how we intend to perform robustness and sensitivity checks. The results are presented in Section 5.

4.1 Basic estimator—Ordinary Least Squares (OLS)

In estimating the impact of student transfer paths and boarding status on math test scores, we first use OLS—controlling (at least in part) for selection bias (and endogeneity due to unobserved heterogeneity) by including a large set of observable covariates in the regression of key independent variables on math scores:

$$y_i = \alpha + \beta' P_i + \gamma B_i + \delta' X_i + \varepsilon_i \tag{1}$$

where, the dependent variable y_i indicates the math score of student i; P_i is a vector that includes six student transfer paths of interest: a.) from teaching points to village schools; b.) from teaching points to town schools; c.) from teaching points to county schools; d.) from village schools to other village schools; e.) from village schools to town schools and f.) from village schools to county schools. The symbol, B_i , is the our boarding status indicator variable, which takes a value of 1 if the student has ever boarded during the years that he/she was in elementary school and 0 if the student has never boarded. Finally, the term X_i is a vector of covariates (or other control variables) that is included to capture the effect on the dependent variable of the characteristics of students, parents and households. To increase efficiency, we compute White's heteroskedasticity-robust standard errors in all regressions.

4.2 Alternative estimator—Propensity score matching (PSM)

Rather than directly correcting for a large number of relevant covariates, adjustments can be made based on a propensity score—defined as the conditional probability of receiving "treatment" (Rosenbaum and Rubin, 1983, Imbens, 2004; Dehejia and Wahba, 2002, Liu et al., 2010). In our setup, the treatments are defined to be the different student transfer paths and boarding statuses. Specifically, we compare (the characteristics of) students who attended teaching points with those who did not; those who transferred from village schools to town schools with those who stayed in village schools; those who transferred from village schools to county schools with those who stayed in village schools; and those who boarded (or boarding status=1) with those did not. We are ultimately interested in estimating the average treatment effects on the treated (ATT) of attending teaching points, transferring from village schools to town schools, transferring from village schools to county schools, and boarding status on academic performance. The propensity score (i.e. the conditional probability of "receiving" these treatments) is calculated by estimating a logit model with student, parental and household characteristics as the independent variables (Appendix 1).

We estimate the ATTs with a propensity score matching (PSM) method where matching involves pairing treatment and comparison units with similar propensity scores (Abadie and Imbens, 2002). In other words, ATTs are calculated as a weighted average of the outcome difference between treated and matched controls. PSM is a more general method than standard linear regression since it does not require assumptions about linearity or constant treatment effects, and thus improves bias correction. Moreover, imposing common support in PSM can lead to efficiency improvements, especially when the sample size is small. It should be noted, however, that PSM estimates are only unbiased if the unobservables are correlated with the obsevables upon which the matching is based.

In our paper we use several different matching algorithms. Specifically, we first use Nearest Neighbor Matching where matching is done with replacements in order to ensure that each treatment unit is matched to the comparison unit nearest to it in propensity score (which is one way to maximize the reduction of selection bias—Imbens, 2004):

$$ATT = \frac{1}{N^T}\sum_{i\in T} y_i^T - \frac{1}{N^T}\sum_{j\in C} y_j^C \tag{2}$$

where, y_i^T indicates the math score of student *i* in the treatment group (T) and y_j^C is the "nearest neighbor" *j* in the control group (C) that is matched to *i*; N^T and N^C denotes the number of treated units and that of control units respectively.

To serve as robustness checks, we also use Kernel Matching and Stratification Matching, because they incorporate trade-offs between quality and quantity of matches differently than the Nearest Neighbor Matching (Becker & Ichino, 2002). Kernel Matching estimates the ATT using:

$$ATT = \frac{1}{N^T}\sum_{i\in T}\left\{ y_i^T - \frac{\sum_{i\in C} y_j^C G\left(p_j - p_i/h_n\right)}{\sum_{k\in C} G\left(p_k - p_i/h_n\right)}\right\} \tag{3}$$

where every treated unit *i* is matched with a weighted average of all control units *j* with weights that are inversely proportional to the distance between their scores p_j - p_i; *G()* is a kernel function with h_n the bandwidth parameter.

Stratification Matching estimates the ATT using:

$$ATT = \sum_{q=1}^{Q}\left(\frac{\sum_{i\in I(q)} y_i^T}{N_q^T} - \frac{\sum_{i\in I(q)} y_i^C}{N_q^C}\right)\frac{\sum_{i\in I(q)} D_i}{\sum_{\forall i} D_i} \tag{4}$$

where, observations are divided by blocks Q defined over intervals of the propensity score; in each block q treated unites and control unites have balanced covariates; ATT in each block q is then weighted to generate the overall ATT with the block weighting function $\frac{\sum_{i \in I(q)} D_i}{\sum_{\forall i} D_i}$.These methods are all implemented with common support, a logit model for calculating the propensity score, and bootstrapped standard errors. The joint consideration of the three methods offers a way to assess the robustness of the estimates.

4.3 Sensitivity analysis

The applied regression (OLS) and matching methods can yield unbiased estimates of ATT subject to the crucial assumption of conditional independence (CIA): conditional upon observable covariates, the receipt of treatment is independent of the potential outcomes with and without treatment (Imbens, 2004). This assumption is not directly testable with non-experimental data (Imbens, 2004), but Ichino, Mealli and Nannicini (2006) proposed a method for testing the sensitivity of matching estimates against the assumption. The method simulates an unobserved binary confounder that is suspected to affect both academic performance and transfer paths/boarding status. We use the method with confounders calibrated to mimic observable binary covariates as in Ichino et al. (2006). We will discuss the simulated confounders and results in the next section.

5. Results and discussion

The estimation results of the basic estimator using equation (1) are presented in Table 6. Column (1) to (3) of Table 6 differ in the independent variables that are included in estimation: column (1) only includes the student transfer path variables (with no covariates); in column (2) we add the boarding status variable; and in column (3) we include the boarding status variable and all of the covariates. The model performs better as we move from column (1) to column (3) as the R-square grows and covariates are shown to effectively capture more of the variation in math scores. Therefore, in the rest of our discussion we mostly focus on the results in column (3).

The results in Table 6 can be seen to be largely consistent with the descriptive analysis. There are three main results (based on column (3)). First, holding other factors, students who started primary school in teaching points or village schools in general have lower math scores if they transferred to village or town schools and did not transfer to county schools. The negative effect of transferring from a teaching point to a village or town school is 6.8 and 7.3 points respectively (row 1 & 2). [2] The negative effect of transferring from a village school to another village school is 9.7 points (row 4). Keeping the starting point constant, students who transferred to county school have significant and larger positive transfer effects. Transferring from a teaching point to a county school has a positive effect of 4.3 points (row 3) and transferring from a village school to a county school has a positive effect of 8.0 points (row 6). These effects seem to add up to a

[2] The coefficients on the dummy variables which measure some of the common transfer paths of students in the sample are compared with the base (excluded) group. In this table, the base group is all students that started elementary school in either a town or county school and the students who did not transfer.

difference of 11.1 points for teaching point starters who ended in village school compared to those that ended up in a county school (row 1 & 3). The difference is 17.7 points for those that started school in a village school and ended in village school compared to those that started in a village school and ended in a county school (row 4 & 6).

Second, our results also show that boarding status matters. In particular, holding all other factors constant (including the student transfer path), when a student stays in a boarding facility there is a significant negative effect (at 1% level) on his/her math scores. The results show that the boarding student's score is reduced by 3.7 points (row 7).

Third, many of the covariates are shown to affect academic performance as expected. For example, the older students perform worse than younger students (significant at 1% level, row 9); rural hukou has a negative effect (significant at 5% level, row 10); attending kindergarten helps increase math score (at 1% level, row11); having elder sibling reduces math score (at 1% level, row 13); students that have mother working in agriculture score lower (at 1% level, row 19).

Propensity Score Matching

The results of the PSM analysis are shown to be qualitatively identical and quantitatively similar with the OLS results and that the results are similar across the sets of results generated by the three alternative PSM estimation strategies (Table 7). Rows 1 to 3 present the ATTs estimated using Nearest Neighbor Matching, Kernel Matching and Stratification Matching, respectively. Column 1 shows that teaching points has a negative effect on the math scores of students and the effect is 3.6 points which is significant at 1%

level in Kernel and Stratification Matching (Row 2 & 3). Column 2 shows that for students who started primary education in a village school and then transferred to town school improves his/her scores by 7.6-9.0 points when compared to the students that stayed in their own village schools (Row 3 and 2). Column 3 also shows that village school starters who transferred to county schools can make progress as large as 19.0 to 20.5 points (Row 1-3). Column 4 shows that boarding status has a negative effect of 5.8-6.4 points (Row 1-3), which is slightly larger than the OLS estimates. In general, estimates of Kernel Matching and Stratification Matching have lower standard errors, which is likely due to a larger number of control units that these methods take into account.

Assessing the Assumption of Conditional Independence (CIA)

Despite the preceding analysis, the transfer paths are so diverse that it could be that even though we control for a large number of observable variables, there could be other unobservables that may have simultaneously affected the transfer paths and academic results of students (violating the assumption of conditional independence of treatment). Following Ichino et al. (2006), we assess the validity of the conditional independence assumption by simulating an unobserved confounder that is used as additional matching factor.

We calibrate the confounder to mimic mother's education level and students' plan to go to high school (Appendix 1) to simulate students' capability and taste for schooling. We show in Table 8 that the estimators with binary confounder differ less than 5% from the previous PSM results in Table 7.[3] This is an indication of the robustness of the ATT estimates and validity of the CIA assumption as far as we can test.

7. Discussion and Conclusions

In this paper we have tried to understand how the Merger Program may have affected the academic performance of students by analyzing a set of transfer paths that students have taken during primary education. Despite the controversies about the benefits and costs of the Merger Program, our results show that at least in our study county, there is positive resource effect that is gain when students transfer from less centralized schools (such as teaching points or village schools) to more centralized schools (such as town schools and county schools). This positive effect, however, may be partially offset by boarding. When students stay in boarding schools, there is a large measured negative effect. Hence, if a student transfers from a village school (or teaching point) to a town (or county) school, but has to stay in the school's boarding facilities, the positive resource effect may, at least in part, be reduced by the negative boarding school effect. However, by comparing the transfer effect with the boarding effect, we find that even if students board after transfer, they still benefit academically from transferring to county school no matter whether they started primary education in teaching point or village school.

[3] By including the additional confounder, the effect of transferring from village to town school is shown to be significant (Table 8). Without the confounder it is not significant (Table 7, Column 2, Row 1). However, the estimates in Table 8 are very similar with the ones using the other two PSM methods (Table 7, Column 2, Row 2 & 3).

Policywise, our paper has several implications. First, the results confirm that the additional resources that the government was hoping would come to bear by the Merger Program appears to be true. In other words, the results are a vindication of the decision to shut down local elementary schools and create centralized schools at the town and county level. The results, however, give extra impetus to the finding of Shi et al. (2009) and his findings that the nation should put extra emphasis on managing boarding schools. If a way could be made to attenuate the negative boarding school effect, students might be able to take more advantage of the additional resources—teaching and facilities—that are being made available by the Merger Program.

References

Abadie, A., Imbens, G., 2002. Simple and Bias-Corrected Matching Estimators for Average Treatment Effects, NBER technical working paper No 283, Cambridge, Massachusetts, U.S.A.

Becker, S., Ichino, A., 2002. Estimation of Average Treatment Effects Based on Propensity Scores. *The Stata Journal* 2(4), 358-377.

Borooah, V. K., 2005. The height-for-age of Indian children. *Economics and Human Biology*, 3, 45–65.

Chen, X., Huang, Q., Rozelle, S., Zhang, L., Shi, Y., 2009. Effect of Migration on Children's Educational Performance in Rural China. *Comparative Economic Studies* 51, 323-343.

Chen, Y., Li, H. B., 2009. Mother's education and child health: Is there a nurturing effect? *Journal of Health Economics*, 28, 413-426.

Dehejia, R.H., Wahba, S., 2002. Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics* 84(1), 151-161.

Gibson, J., 2001. Literacy and intra-household externalities. *World Development* 29,155–166.

Glewwe, P., Grosh, M., Jacoby, H., and Lockheed, M., 1995. An Eclectic Approach to Estimating the Determinants of Achievement in Jamaican Primary Education. *Economic Review* 9 (2), 231-258.

Gruman, D. H., Harachi, T. W., Abbott, R. D., Catalano, R. F., & Fleming, C. B., 2008. Longitudinal effects of student mobility on three dimensions of elementary school engagement. *Child Development*, 79, 1833-1852.

Ichino, A., Mealli F., Nannicini, T., 2006. From Temporary Help Jobs to Permanent Employment: What Can we Learn from Matching Estimators and their Sensitivity? CEPR Discussion Paper No 5736, London, UK.

Imbens, G., 2004. Nonparametric Estimation of Average Treatment Effects under Exogeneity: a Review. *The Review of Economics and Statistics* 96(1), 4-29.

Linnemayr, S., Alderman, H., Ka, A., 2008. Determinants of malnutrition in Senegal: Individual, household, community variables, and their interaction. *Economics and Human Biology*, 6 , 252–263.

Liu, C., Zhang, L., Luo, R., Rozelle, S., and Loyalka, P., 2010. The Effect of Primary School Mergers on Academic Performance of Students in Rural China. *International Journal of Educational Development* 30 (6), 570-585.

Luo, R., Zhang, L., Liu, C., Rozelle, S., Sharbono, B., Adams, J., 2010. Behind Before They Start: The Challenge of Early Childhood Education in Rural China. REAP working paper No. 209, CCAP, Beijing, China.

Luo, R., Shi Y., Zhang L., Liu C., Rozelle S., Sharbono B., 2009. Malnutrition in China's Rural Boarding Schools: the Case of Primary Schools in Shaanxi Province. *Asia Pacific Journal of Education* 29 (4), 481-501.

Ministry of Education, Lüliang County, 2005. *The "Eleventh 5 Year Plan" for Educational Development.* http://www.lljyj.gov.cn/lljy.htm

Ministry of Education, 2005. *2004 Communiqué on National Educational Development.* http://news.xinhuanet.com/edu/2005-03/02/content_2634435.htm

National Bureau of Statistics, 2006. *China Statistics Yearbook*. Beijing: China statistics press.

Pang, L., 2006. Primary and High School Consolidation in Rural China: Issues, Reasons, and Solutions. *Education Development Research*, 2, 1-6.

Reynolds, A. J., Bezruczko, N., 1993. School adjustment of children at risk through fourth grade. *Merrill-Palmer Quarterly*, 39, 457-48.

Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects, *Biometrika* 70, pp. 41-55.

Shi, R., 2004. Experience Drawn from Research on School Consolidation in Foreign Countries. *Comparative Education Research,* 12.

Shariff, A., 1998.Women's status and child health. In: Krishnaraj, M., Sudarshan, R.M., Shariff, A. (Eds.), Gender, Population and Development. Oxford University Press, New Delhi, 185–219.

Shi, Y., Zhang, L., Liu C., Li, H., Rozelle, S., Sharbono, B., 2009. Community Service, Social Responsibility and Educational Performance in Rural China's Middle Schools: Evidence from a Case Study of Northwest China. REAP Working Paper No 208, CCAP, Beijing, China.

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.

Zhuo, M., 2006. Primary School Consolidation in Rural China: Advantages, Disadvantages and Solutions. *Education Forum*, the Education Department of Fujian Province.

Table 1. Distribution of sample students starting in and graduating from different types of primary schools (Shilou County, Shanxi Province, China 2009)

	Starting school		Ending school	
Type of school	No.	%	No.	%
County school	649	43.1	873	57.9
Town school	237	15.7	465	30.9
Village school	248	16.5	169	11.2
Teaching point	373	24.8	0	0
Total	1507	100	1507	100

Table 2. Number of years that sample students spent in teaching points and mean score (Shilou County, Shanxi Province, China 2009)

Years of schooling spent in teaching point	Obs.	
	No.	%
0	1118	74.2
1	38	2.5
2	60	4.0
3	134	8.9
4	157	10.4
Total	1507	100

Table 3. The distribution and mean math scores of sample students by transfer path (Shilou County, Shanxi Province, China 2009)

Transfer paths		Obs.		
		No.	%	score
1	Teaching point-village school	50	3.3	45.3
2	Teaching point-town school	210	13.9	46.7
3	Teaching point-county school	113	7.5	60
4	Village-village school	14	0.9	46.4
5	Village-town school	41	2.7	55.3
6	Village-county school	100	6.6	64.1
7	Town-village school	5	0.3	50
8	Town-town school	44	2.9	55.1
9	Town-county school	39	2.6	68.2
10	County-village school	22	1.5	61.6
11	County-town school	6	0.4	61.7
12	County-county school	97	6.4	60.6
13	No transfer	766	50.8	57.5
Total		1507	100	56.5

Table 4. The distribution and mean math scores of sample students by boarding status (Shilou County, Shanxi Province, China 2009)

Transfer paths		Boarding			Non-boarding		
		No.	%	Score	No.	%	Score
1.	Teaching point-village school	29	58	45.3	21	42	45.2
2.	Teaching point-town school	105	50	46.9	105	50	46.6
3.	Teaching point-county school	19	16.8	58.4	94	83.2	60.1
4.	Village-village school	3	21.4	40	11	78.6	48.2
5.	Village-town school	12	29.3	48.8	28	70.7	58
6.	Village-county school	18	18	58.1	82	82	65.4
7.	Town-village school	2	40	40	3	60	56.7
8.	Town-town school	5	9.1	53	40	90.9	55.4
9.	Town-county school	12	30.8	62.9	27	69.2	70.6
10.	County-village school	1	4.5	50	21	95.5	62.1
11.	County-town school	0	0		6	100	61.7
12.	County-county school	16	16.5	56.6	81	83.5	61.4
13.	No transfer	62	8.1	46.8	704	91.9	58.5
Total		284	18.8	49.5	1223	81.2	58.1

Table 5. Decomposed student characteristics and mean math scores (Shilou County, Shanxi Province, China 2009)

Variables	Value range	Mean math scores	Std. Dev
Student characteristics			
Gender	female	57.3	17.5
	male	55.8	17.2
Age, year	[9,12]	59.0	17.7
	(12,13]	55.5	17.1
	(13,16]	54.3	16.9
Rural hukou identity	no	63.7	16.3
	yes	54.9	17.2
Attended kindergarten	no	52.9	16.4
	yes	57.3	17.5
Attended preschool	no	56.2	17.8
	yes	57.1	16.4
Have elder sibling	no	59.5	17.3
	yes	54.7	17.2
Parental characteristics			
Age of father	[30,38]	56.4	17.3
	(38,41]	58.2	17.5
	(41,62]	54.6	17.2
Age of mother	[28,36]	57.1	17.0
	(36,39]	57.2	18.2
	(39,55]	54.9	16.8
Father holding middle school diploma	no	54.3	17.3
	yes	58.9	17.1
Mother holding middle school diploma	no	54.5	17.3
	yes	59.6	17.1
Father working in agriculture	no	58.4	17.5
	yes	52.6	16.4
Mother working in agriculture	no	60.2	17.4
	yes	52.0	16.3
Household characteristics			
Household size	[1,4]	59.6	17.5
	(4,5]	54.2	17.0
	(5,9]	54.9	16.9
Household durable assets value (1000 yuan)	[0,6.5]	55.1	17.8
	(6.5,12]	56.3	17.2
	(12,218]	58.0	17.0

Table 6. Multivariate regression results analyzing transfer paths and their impact on students' academic achievement (Shilou County, Shanxi Province, China 2009)

Dependent variable: standardized math score (0-100 pts)			
	(1)	(2)	(3)
Path variables and boarding status			
1. Transfer from teaching point to village school, 1= yes	-12.9***	-10.5***	-6.8***
	[-6.5]	[-5.0]	[-3.1]
2. Transfer from teaching point to town school, 1= yes	-11.5***	-9.5***	-7.3***
	[-10.3]	[-7.6]	[-5.6]
3. Transfer from teaching point to county school, 1= yes	1.6	1.9	4.3**
	[0.9]	[1.1]	[2.4]
4. Transfer from village school to village school, 1= yes	-11.8**	-11.2**	-9.7*
	[-2.4]	[-2.3]	[-2.0]
5. Transfer from village school to town school, 1=yes	-3.0	-2.0	0.9
	[-1.1]	[-0.7]	[0.3]
6. Transfer from village school to county school, 1= yes	5.9***	6.2***	8.0***
	[3.6]	[3.9]	[4.7]
7. Boarding status, 1=boarded		-5.0***	-3.7***
		[-4.2]	[-3.2]
Student characteristics			
8. Male=1,female=0			-1.0
			[-1.2]
9. Age, year			-1.3***
			[-2.7]
10. Hukou identity, 1=rural			-3.2**
			[-2.4]
11. Kindergarten, 1=attended			3.7***
			[3.4]
12. Preschool, 1=attended			0.4
			[0.5]
13. Having elder sibling, 1=yes			-3.5***
			[-3.5]
Parental characteristics			
14. Age of father, year			0.0
			[0.0]
15. Age of mother, year			0.1
			[0.2]
16. Father holding middle school diploma, 1=yes			0.7
			[0.7]
17. Mother holding middle school diploma, 1=yes			0.6
			[0.6]
18. Father working in agriculture, 1=yes			-0.1
			[-0.1]
19. Mother working in agriculture, 1=yes			-3.2***
			[-2.9]
Household characteristics			
20. Household size			-0.1
			[-0.2]
21. Household durable assets value (1000 yuan)			0.0
			[0.4]
Constant	58.2***	58.7***	76.2***
	[105.2]	[105.6]	[10.0]
Observations	1507	1507	1507
R-squared	0.08	0.09	0.15

Note: 1) t statistics in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.
2) The coefficients on the dummy variables which measure some of the common transfer paths of students in the sample are compared with the base (excluded) group. In this table, the base group is all students that started elementary school in either a town or county school and the students who did not transfer.

Table 7. PSM results analyzing transfer paths and their impact on students' academic achievement (Shilou County, Shanxi Province, China 2009)

	Average treatment effect of the treated (ATT)			
	Teaching point students (1)	Village school students who transfer to town schools (2)	Village school students who transfer to county schools (3)	Boarding students (4)
1. Nearest Neighbor Matching	-2.4	5.2	20.5***	-5.8***
	[1.9]	[11.8]	[5.6]	[1.8]
2. Kernel Matching	-3.6***	9.0***	19.0***	-6.4***
	[1.2]	[3.1]	[2.3]	[1.1]
3. Stratification Matching	-3.6***	7.6***	19.2***	-6.4***
	[1.2]	[3.1]	[2.6]	[1.1]

Note: bootstrapped standard errors in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8. Simulation-based sensitivity analysis for propensity score matching estimators[1]

	Estimated ATT using Nearest Neighbor Matching	Outcome effect[2]	Selection effect[3]
Treatment (1): Attending teaching points			
A. Confounder calibrated to mimic mother's education level	-2.4	1.5	0.3
B. Confounder calibrated to mimic student's plan for high school	-2.8	2.8	0.9
Treatment (2): Transfer from village school to town school			
A. Confounder calibrated to mimic mother's education level	8.9*	0.8	0.9
B. Confounder calibrated to mimic student's plan for high school	9.0**	4.1	1.1
Treatment (3): Transfer from village school to county school			
A. Confounder calibrated to mimic mother's education level	18.9***	0.9	1.8
B. Confounder calibrated to mimic student's plan for high school	18.4***	4.5	3.5
Treatment (4): Boarding			
A. Confounder calibrated to mimic mother's education level	-6.3***	1.4	0.4
B. Confounder calibrated to mimic student's plan for high school	-5.9***	2.7	0.7

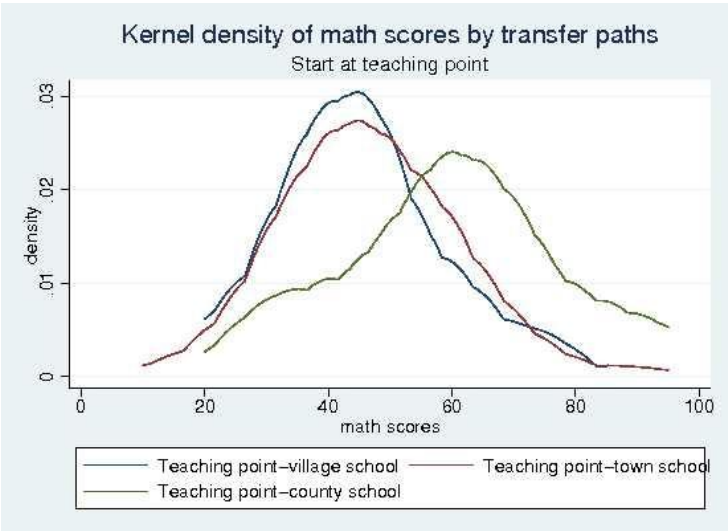
Note: * significant at 10%; ** significant at 5%; *** significant at 1%.

[1] The method is described by Ichino, Mealli and Nannicini (2006) and builds on Rosenbaum and Rubin (1983) and Rosenbaum (1987). The method simulates this binary confounder in the data that is used as an additional matching factor. A comparison of the estimates obtained with and without matching on the simulated confounder informs to what extent the estimator is robust to this specific source of failure of the conditional independence assumption.

[2] The outcome effect measures the estimated effect of the simulated binary confounder on the outcome variable—math score.

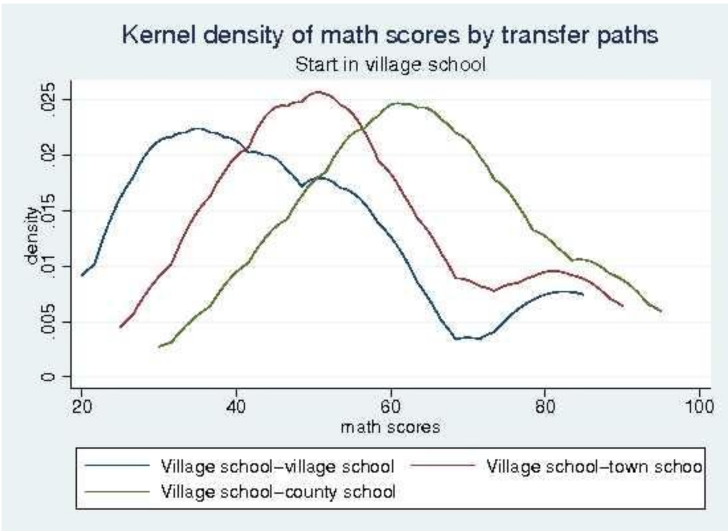
[3] The selection effect measures the estimated effect of the simulated binary confounder on the selection into treatment.

Figure 1



Kernel density of math scores by transfer paths
Start at teaching point
density
0
.01
.02
.03
0
20
40
60
80
100
math scores
Teaching point–village school
Teaching point–town school
Teaching point–county school

Figure 2


Kernel density of math scores by transfer paths
Start in village school
density
0
.005
.01
.015
.02
.025
20
40
60
80
100
math scores
Village school–village school
Village school–town schoo
Village school–county school

Appendix 1 Summary of mean characteristics of students (Shilou County, Shanxi Province, China 2009)

Variables	Mean	Std. Dev.	Min	Max
Transfer paths and boarding dummies				
Teaching point-village school	0.0	0.2	0	1
Teaching point-town school	0.1	0.3	0	1
Teaching point-county school	0.1	0.3	0	1
Village-village school	0.0	0.1	0	1
Village-town school	0.0	0.2	0	1
Village-county school	0.1	0.2	0	1
Stay in village school	0.1	0.2	0	1
Boarding status, 1=yes	0.2	0.4	0	1
Student characteristics				
Male, 1=yes	0.5	0.5	0	1
Age, year	12.9	0.9	9	16
Hukou identity, 1=rural	0.8	0.4	0	1
Kindergarten, 1=yes	0.8	0.4	0	1
Preschool, 1=yes	0.3	0.5	0	1
Having elder sibling, 1=yes	0.6	0.5	0	1
Parental characteristics				
Age of father, year	39.6	4.3	30	62
Age of mother, year	37.9	4.1	28	55
Father holding middle school diploma, 1=yes	0.5	0.5	0	1
Mother holding middle school diploma, 1=yes	0.4	0.5	0	1
Father working in agriculture, 1=yes	0.3	0.5	0	1
Mother working in agriculture, 1=yes	0.5	0.5	0	1
Household characteristics				
Household size	4.9	1.0	1	9
Household wealth	1.0	0.2	0	1
Coundoufer for simulation for sensitivity analysis				
Mother holding middle school diploma, 1=yes	0.4	0.5	0	1
Student's plan for high school, 1=go to high school after graduation	0.7	0.4	0	1