Comparative Economic Studies, 2009, 51, (323–343)

www.palgrave-journals.com/ces/

Symposium Article

Effect of Migration on Children's Educational Performance in Rural China

XINXIN CHEN[1], QIUQIONG HUANG[2], SCOTT ROZELLE[3], YAOJIANG SHI[4] & LINXIU ZHANG[5]

[1]College of Economics, Zhejiang Gongshang University, 149 Jiaogong Road, Hangzhou 310035, China.
[2]Department of Applied Economics, University of Minnesota, 1994 Buford Avenue, St Paul, MN 55108, USA.
[3]Freeman Spogli Institute, Stanford University, Encina Hall East, E301, Stanford, CA 94305, USA.
[4]School of Economics and Management, Northwest University, 1 XueFu Road, Xi'an 710127, China.
[5]Center for Chinese Agricultural Policy, Chinese Academy of Sciences, No. Jia 11, Datun Road, Anwai, Beijing 100101, China.

Migration is one of the main ways of alleviating poverty in developing countries, including China. However, there are concerns about the potential negative effects of migration on the educational achievement of the children that are left behind in villages when one or both of their parents out-migrate to cities. This paper examines changes in school performance before and after the parents of students out-migrate. Surprisingly, we find that there is no significant negative effect of migration on school performance. In fact, we find that educational performance improves in migrant households in which the father out-migrates.
Comparative Economic Studies (2009) **51,** 323–343. doi:10.1057/ces.2008.44;
published online 21 May 2009

Keywords: migration, school performance, rural China

JEL Classifications: I22, O12, O15

INTRODUCTION

Migration is widely known by researchers and policy makers as one of the main ways of alleviating poverty in developing countries (Todaro, 1989). Many positive effects have been identified. Having a migrant may increase a household's income per capita significantly (eg, Du *et al*., 2005). Migrant

remittances have been shown to help reduce income inequality in countries such as Mexico (eg, Benjamin *et al.*, 2005; McKenzie and Rapoport, 2007). Increases in out-migration can lead to increased investment in assets related to agricultural production and other investments in source communities (de Brauw and Giles, 2007). Giles (2006) shows that access to migration offers a risk-coping mechanism that allows households to reduce the variability of income caused by shocks affecting agricultural production.

Migration itself, however, is not costless. In recent years, a group of researchers have raised concerns about the potential negative effects of migration on the educational outcome of the school-aged children that are left behind in rural areas when one or both of their parents out-migrate to cities for work. McKenzie and Rapoport (2006) found that children in migrant households are less likely to be attending school and complete fewer total years of schooling than children in non-migrant households. One of the main reasons may be that migration results in significantly less parental supervision of school attendance and the loss of any positive influence through learning at home (Hanson and Woodruff, 2004). There may be other effects. For example, students may have less time to spend on studying because they have to do more housework when parents out-migrate. Similar findings of adverse effects of migration on children's schooling are also observed in Philippines (Battistella and Conaco, 1998).

On the other hand, researchers have also found a positive relationship between migration and education of migrant children. Remittance transfers, by relaxing a household's liquidity constraint, allow investment in education and thereby can increase educational attainment of children in migrant households. For example, Cox Edwards and Ureta (2003) found that receiving remittances reduces the likelihood of quitting school among individuals aged 6–24 years in El Salvador. Similar arguments are also found in Glewwe and Jacoby (2004). In addition, migration, by increasing household income, can contribute positively to child development (Blau, 1999; Duncan *et al.*, 1994). Overall, migrant children could benefit from the positive income effect of migration.

The impact of migration on the educational performance of children that are left behind is also an important and emerging issue in China. The migrant labour force has been growing rapidly since the 1990s in China, surpassing 100 million individuals (de Brauw *et al.*, 2002). Migrants also are moving further away from home and leaving for a longer period of time (Rozelle *et al.*, 1999). Since most of China's migration is by individuals instead of entire households, in most cases the school-aged children are left behind in the village when their parents move to the city for work (Wu, 2004). Researchers have claimed that school performance of migrant children is

adversely affected (Li, 2004; Tan and Wang, 2004; Wang and Wu, 2003; Zhou and Wu, 2004).

If migration indeed has a negative impact on education of the next generation, the government may want to respond. In fact, in the long run, Heckman (2005) actually argues that the government probably should start to modify education policies to favour children of migrants.

The overall goal of this paper is to examine the effect of migration activities of parents on the educational performance of their children and provide policy makers with information about whether or not they need to do anything to change the school systems and childcare systems in China's rural and urban areas. To meet the overall goal, we will pursue two specific objectives. First, we compare the distribution of children's scores across different types of households (migrant and non-migrant households) and over time (before and after one or both parents out-migrate). Second, we examine whether migration negatively affects the school performance of migrant children.

DATA

The data used in this paper come from a survey conducted by the authors in 2006. The sample was drawn from 36 primary schools in 12 townships in Shaanxi province, one of the nation's poorest provinces. The sample was drawn using a four-stage clustering design with random sampling procedures employed at each stage. In each stage, we randomly chose six counties that were selected from 93 counties in Shaanxi province; two townships from each county; and three primary schools were randomly chosen from a list of all primary schools with six years of schooling (or all *wanxiao*) in each township.

The sample students were selected during the final stage of sampling. The sample included all students that were in the entering year of the sixth grade classes in each of the sample schools. On average, there were 1.4 sixth grade classes per school, ranging from one to three. Since the survey was conducted in September and the school year in China runs between early September and mid-July, all of the sample students had just completed the fifth grade about 2 months ago. In total, the sample included 1,649 children and their families. Approximately 45% of the sample students were girls. The ages of the students ranged between 10 and 16 years.[1] In addition, since we sampled

[1] Although the ages of children ranged from 6 to 16, only few children in the sample were older than 14.

from the population of students that made it to the sixth grade, we did not track down those that either dropped out or accompanied their parents to the city before the sixth grade. Fortunately, according to the information from the principal questionnaire-based interviews, this is not a serious problem. The drop-out rate in our sample is low. More than 98% of the children that started first grade were still in school in the sixth grade.

In addition to interviewing students directly, we also elicited information about the students from their homeroom teachers (or *banzhuren*). In more than 90% of the cases, the homeroom teacher had been with the students for at least 2 years. In China, homeroom teachers are in charge of administering students' school programme in addition to teaching one or two subjects. For many reasons, the homeroom teacher was intimately familiar with the school performance and family life of each student.

The measure of one of the key variables, educational achievement, is based on the math and Chinese language scores of the students from the calendar year 2001/2002 (the year in which the students were in the first grade) to 2005/2006 (the year students were in the fifth grade). Fortunately, students in all the sample schools keep in their possession a booklet that records their math and Chinese scores for each semester during their schooling. The scores were copied by our enumerators with the assistance of the homeroom teachers. Therefore, the educational achievement variables are measured with great accuracy since they are record-based.

In our analyses, we use the second term (or the spring semester) math and Chinese language scores to measure educational achievement because they are based on standardised tests (and not any other work during the year). The tests are standardised in two dimensions. First, the questions are the same for all schools within the same township. Second, the final exams were graded according to the same set of criteria by a township-wide panel of teachers.[2]

In order to measure another key variable, the migration status, we collected detailed information on the migration histories of each student's family. The first set of information came from the survey questionnaire that was filled by students and their family with the supervision of enumerators. In the questionnaire, we have a section that asked for the migration status of each parent during the first grade and during the fifth grade. If the parents

[2] The results of our analyses, however, do not depend on the choice of using the second term scores from the first and fifth grades. As a robustness check, we also used average scores for the whole year instead of just for the second semester. In another check, we compared scores that averaged scores from first and second grade to scores that averaged scores from fourth and fifth grade. Our results remain largely the same in these cases.

were both out of the village, we called one of the parents and asked them these questions over the telephone. As a way of cross checking, the homeroom teacher was asked to verify the information on migration status.

On the basis of migration status, there are two types of households in this study: *migrant households* (or those households in which at least one parent out-migrated during the period between 2002 and 2006) and non-migrant households. Recognising that the effect of migration on student performance may be affected by *which family member* out-migrates (ie, father, mother or both), we further subdivided the migrant households into six types of households: *Any Parent Migrated* households (ie, households in which both parents lived at home in 2002 and at least one parent – either the father; mother or both parents – out-migrated by 2006); *Father Migrated Only* (*or mother-stayed-at-home*) households where only the father out-migrated by 2006 but was at home in 2002; *Father Migrated* (*Unconditional*) households where the father was at home in 2002 but out-migrated by 2006 (including households in which the mother was either at home or not at home in 2006); *Mother Migrated Only* (*or father-stayed-at-home*) households where only the mother out-migrated by 2006 but was at home in 2002; *Mother Migrated* (*Unconditional*) households where the mother was at home in 2002 but out-migrated by 2006 (including households in which the father was either at home or not at home in 2006); and *Both Parents Migrated* households where both parents were at home in 2002, but out-migrated by 2006. It should be noted that the six types of households are not mutually exclusive. For brevity, when we talk about all of these households as a group, we call them *New Migrant* households to distinguish them from households that were already in the migrant labour force in 2002 (which are not included in our study). In addition, we define *Never Migrant* households as those in which both parents stayed at home in both 2002 and 2006.

In addition to educational achievement and migration status, we also collected information on variables that can help us explore whether the effects of migration on the school performance of children are heterogeneous across households that are different in several aspects. First, as family wealth may improve the learning outcomes of students (Brown and Park, 2002), we asked the parents whether their house was worth more than 5,000 yuan or not as a proxy for family wealth.[3] Admittedly, this is a crude measure of

[3] There are several issues to discuss when considering our measure of income. The value of the house only includes the part of the house used for domicile purposes and the value of assets that were used for farming and non-farm businesses was not included. Yuan is the Chinese currency. One dollar was about 7.6 yuan during the time of our survey. Finally, we admit that we only have a rough proxy for income. Because of this it is possible that we will not be able to identify the impact of income on grades (since the coefficient of the variable could be biased down to zero). The cross

wealth. However, given that our sample size is larger than 1,600, financial and time constraints dictated that we used this measure instead of implementing a long questionnaire to collect detailed information on income and assets from all different sources as well as information on consumption. Second, since previous studies have also documented the effect of the number of children in a family on the school performance of each child (eg, Hanushek, 1992; Steelman and Mercy, 1980), we collected information on whether a student had any siblings or not and the number of siblings.

Finally, we also collected information on other variables that may affect school performance. We collected information on each student's gender, age, whether they were student cadres and whether they had mentors to help them study. The survey questionnaire also included questions on the characteristics of parents and households: each parent's age, educational attainment, the household's land holdings and the total number of other household members.

MIGRATION AND EDUCATIONAL PERFORMANCE

Similar to the migration status in many other poor rural areas in China (Rozelle *et al*., 1999), many households were already in the migrant labour force in 2002, the first year of our sample. In 236 households (about 15% of the 1,594 sample households), either one parent or both parents migrated (Table 1, column 1, rows 1–3). In most migrant households (149), the father was the parent that out-migrated (while the mother stayed at home).

In addition, similar to the migration trend in the rest of China (as reported in de Brauw *et al*., 2002), the number of new migrant households rose rapidly during the period between 2002 and 2006. Among the 1,358 non-migrant households in 2002 (column 1, row 4), one or both parents in 220 households entered the migrant labour force between 2002 and 2006 (row 4, columns 2–4). After subtracting the 81 households that migrated in 2002 but returned to the village in 2006 from the 220 New Migrant households (column 5, rows 1–3), the total migrant households rose to 375 households in 2006 (row 5, columns 2–4), a 9% rise from the 2002 migration level.

sectional variation for income (using our measure), however, does show that there is at least a negative correlation (richer households have children with higher grades). In fact, there are reasons to expect a positive effect of income on grades. The literature (eg, Kandel and Kao, 2001) has shown that the positive relationship between the father's migration and the school performance partly stems from the financial resources provided migration, which lowers the likelihood of children's labor force participation and increases resources for consumption of education-related goods.

Table 1: Patterns of migration in sample households in 2002 and 2006, Shaanxi Province, China

	Migration status in 2002	Migration status in 2006			
	(1) Number of households in 2002[a]	(2) *Father Migrated Only* (*mother stayed home*)	(3) *Mother Migrated Only* (*father stayed home*)	(4) *Both Parents Migrated*	(5) Return migrants *(rows 1–3)*[b] Never migrant *(row 4)*
(1) Only father migrated	149	94[c]			55
(2) Only mother migrated	18		9[c]		9
(3) Both parents migrated	69	7	5	40[c]	17
(4) Neither parent migrated	1,358	131[d]	35[d]	54[d]	1,138
(5) Total number of households	1,594	232	49	94	1,219

Data source: Authors' survey

[a] Column (1)=Column (2)+Column (3)+Column (4)+Column (5).

[b] The households in column 5, rows 1, 2 and 3 are return migrants (or those households in which households had a migrant in 2002 and by 2006 had returned home. These households are dropped from the multivariate analysis.

[c] The diagonal elements in the first three rows of the 2006 matrix (row 1, column 2; row 2, column 3; row 3, column 4) are *Always Migrant* households. These households are dropped from the multivariate analysis.

[d] Total new migrants (or those households in which the parents did not migrate in 2002 and migrated by 2006) is found in row 4 by summing columns 2, 3 and 4).

More than 70% of the households did not participate in migration activities at all during this period (column 5, row 4). The existence of these non-migrant households as well as new migrant households offers a unique opportunity to examine the relationship between the migration activity of parents and the school performance of children. In the rest of the paper, we will focus on comparing the school performance of the children of the 1,138 Never Migrant households and those of the 220 New Migrant households.

If one were naively to have sought out parents of New Migrant households and asked them for the record of their children's scores over the years, one would likely have found that the scores have fallen since the first grade. It is understandable how the findings of such an inquiry could raise concerns about the potential negative impact of migration on school performance. However, the falling scores may not be a problem that can be solely blamed on migration. As our data show, not only have the scores of the students from the New Migrant households fallen (by about 1 point – from 71.6 to 70.8), but also those from other households including never migrant households have fallen (by about 3 points from 73.7 to 70.6). According to our data, this is true for both math and Chinese scores. When asking teachers about this trend, we were told the pattern of falling scores is easily explained by two factors: The materials covered in the fifth grade are much more advanced and difficult than those in the first grade; and, in general, the fifth grade teachers grade harder than the first grade teachers. Since these two factors affect students from both New Migrant and Never Migrant households, the general trend of falling scores is not surprising and clearly cannot be solely attributed to parents' migration activities.

If the interviewers had sought out Both Parents Migrated households or Mother Migrated households, the results of interviews might raise an additional source of concern about the effect of migration on the grades of the children of new migrants. In our sample, students from Both Parents Migrated households had lower average test scores during their fifth grade year (69.9 points) than those from Never Migrant households (70.9 points). Similar results were found in a number of Chinese studies (eg, Li, 2004; Tan and Wang, 2004; Wang and Wu, 2003; Zhou and Wu, 2004). Although the difference is not statistically significant, the fifth grade scores of the children of Mother Migrated (Unconditional) households (70.6 points) were also lower than those of Never Migrants. While we will explore this result further in the forthcoming analyses, it may be that it is these types of findings, which appear in our cross section of fifth grade households, that have made the effect of migration on school performance a high-profile issue. Interestingly, if the interview team had chosen Father Migrated (Unconditional) households (72.0 points), they would have found that on average scores of children from

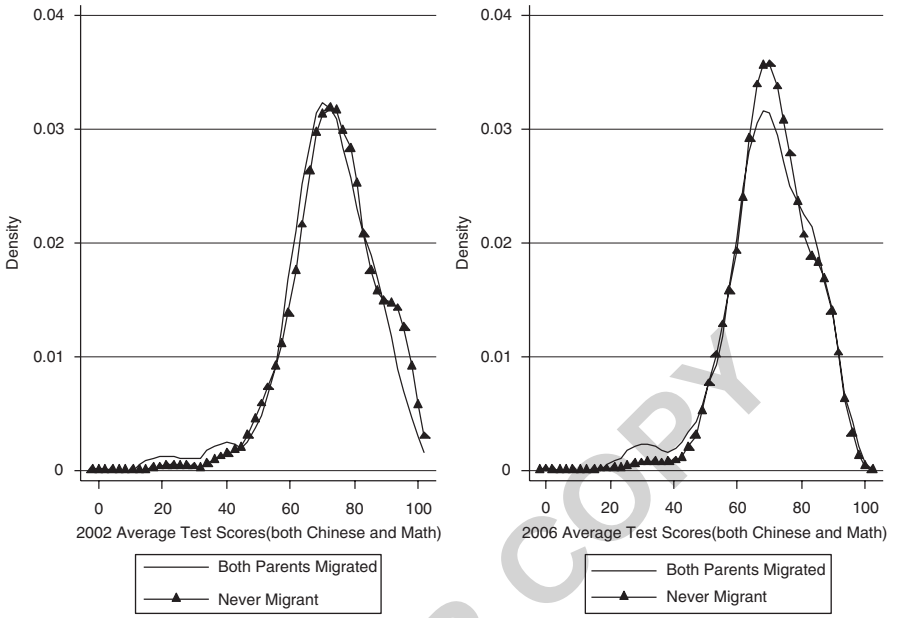

Figure 1: Distribution of average test scores in Never Migrant households and in Both Parents Migrated households in 2002 and 2006 using Kernel Density, Shaanxi Province, China. *Data source*: Authors' survey

these households were slightly higher than those from Never Migrant households. The differences in the relationship between students' scores and parents' migration activities among different types of households indicate that the effect of migration on school performance is a complicated issue and those relying exclusively on cross-sectional data should exercise caution in any interpretation.

The need to exercise caution is reinforced when we compare the first grade scores in 2002 in addition to comparing the fifth grade scores in 2006. Although students from Both Parents Migrated households scored lower than those from the Never Migrant households in 2006, Figure 1 shows they already scored lower in 2002 when they were in the first grade. In other words, on average the scores of the students from Both Parents Migrate households were already lower *before their parents migrated*. This finding from our panel data indicates we should not jump to the conclusion that migration hurt children's school performance. Moreover, the distribution of the scores of the non-migrant children appears to actually move slightly closer to that of the migrant children over years. The distribution of the scores of non-migration students shifts to the left in contrast to its distribution in

2002 (a higher peak in the middle and thinner tail at the upper end). Although the distribution of the scores of migrant children has a thicker tail at the lower end in contrast to its distribution in 2002, its distribution does not shift as much as that of non-migrant children. Overall, it appears that the gap between the scores of migrant children and non-migrant children actually was narrowing slightly between 2002 and 2006. From this seemingly narrowing gap, one might infer that migration helps improve school performance. However, *t*-tests show that the means of the two distributions are not statistically significantly different either in 2002 or in 2006.

Further analyses of our data reveal that school performance may be explained by many factors other than migration activities. These factors may change over time and differ between migrant and non-migrant households. More importantly, these variations may be masking the relationship between migration and school performance. For example, as mentioned earlier, the difficulty of class materials and teachers' grading attitude change over time, which affect students' scores. In addition, school performance may differ by a household's wealth – or more specifically in our study, the value of the housing assets of the household.[4] According to our data, students from wealthier households score systematically higher than those from poorer households (on average about 2 points higher). This result is consistent with previous findings that the grades of children from better-off families often are higher since these children have access to better nutrition and better studying facilities, including access to extra reading materials and exercise books (Princiotta *et al*., 2006).[5]

If either a household's wealth or some of other factor differs systematically with a household's migration status, two-way correlations between a family's migration status and the grades of its children could be misleading. For example, de Brauw and Giles (2007) find that migrant households, while poorer, improve their family's income status after migration. Higher income could have a positive effect on the grades of migrant children over time that might offset any other adverse effects. Therefore, further analyses are needed to explore the impact of migration on educational achievement while holding as many other factors constant as possible.

[4] The term, wealth, when used in the rest of the paper will refer to the value of housing assets only.

[5] Scores may also differ among households with different household demographic compositions. According to our data, students from households in which there are no siblings (70.3 points) scored slightly lower than those from households with siblings in 2006 (71.5 points). Such a finding is consistent with Brown and Park's study (2002), which found that children with older siblings have significantly higher test scores than their peers.

Methodology

The objective of this study is to examine the effects of migration on the student's educational performance. In order to evaluate the effects of migration, conceptually we are making migration the treatment. In other words, our sample students are divided into a treatment group (those that were in households in which the parent(s) out-migrated) and a control group (those in Never Migrated Households). With this set-up, we are interested in understanding the mean impact of 'treatment on the treated,' which is the average impact of grade retention among those treated (Smith and Todd, 2005):

$$\begin{aligned} TT &= E((Y_1 - Y_0)|X, D = 1) \\ &= E(Y_1|X, D = 1) - E(Y_0|X, D = 1) \end{aligned} \tag{1}$$

where we denote Y_1 as the outcome (the scores of students – in our case) after the parent of the student out-migrated and Y_0 as the outcome if a student's parent did not out-migrate. In equation 1, our treatment is denoted by $D = 1$, which stands for the students whose parent(s) out-migrated and for whom Y_1 is observed and $D = 0$ stands for the student whose did not out-migrate and for whom Y_0 is observed. As in reality we do not observe either the counterfactual mean, $E(Y_0|X, D = 1)$, or the mean outcome for the students had they not been retained in a grade after they were retained, we need to employ a difference-in-difference estimation approach (DD). Using the DD approach allows us to compare the outcomes before and after a student repeated a grade with students not affected by the treatment (those who were not retained).

In addition to the standard DD estimator (Smith and Todd, 2005), we implement three other DD estimators: an 'unrestricted' version that includes $Y_{t'}$ as a right-hand variable, an 'adjusted' version that includes other covariates in addition to the treatment variable (in our case they are a series of control variables from 2002 or the pre-programme period), and an unrestricted/adjusted model that combines the features of both the 'unrestricted' and 'adjusted' model. The unrestricted and adjusted DD estimators relax the implicit restrictions in the standard DD estimator that the coefficient associated with $Y_{t'}$ (pre-programme outcome) and covariates in t' (pre-programme period) equals one. The combination of unrestricted and adjusted DD estimators relaxes both of these assumptions. In summary, the models to be estimated are as follows:

Model (1), Restricted and Unadjusted:

$$\Delta Score_i = \alpha + \delta MIGRATE_i + \varepsilon_i$$

Model (2) Restricted and Adjusted:

$$\Delta Score_i = \alpha + \delta MIGRATE_i + \beta X_i + \varepsilon_i$$

Model (3) Unrestricted and Unadjusted:

$$\Delta Score_i = \alpha + \delta MIGRATE_i + \gamma Score\_beforemigrate_i + \varepsilon_i$$

Model (4) Unrestricted and Adjusted:

$$\Delta Score_i = \alpha + \delta MIGRATE_i + \gamma Score\_beforemigrate_i + \beta X_i + \varepsilon_i$$

In addition to the set of DD estimators, we also use a matching approach to check and see whether our results are robust to our choice of estimators. Matching estimators match a student in the treatment group with an observably similar student from the control group and interpret the difference in their school performance as the effect of the parents' migration activities. The key assumption underlying the matching estimator is the Conditional Independence Assumption (CIA), which states that non-participation outcome, $Score_0$, is independent of parents' migration status after being conditional on a set of observable characteristics (Rosenbaum and Rubin, 1983).

In order to implement our matching estimator, we follow a series of well-established steps (Caliendo and Kopeinig, 2008). First, since matching is only justified over the common support region, we check whether there is a large overlap in the support of the covariates, Z, between the New Migrant and Never Migrated households. In our study, we use propensity scores as a tool to enforce a common support. Fortunately, the common support is fairly wide in our sample.[6] In the second step, we choose the method of matching. In our analysis, we use the nearest neighbour matching method with replacement. Following Smith and Todd (2005), we match students based on the log odds ratio and standard errors are bootstrapped using 1,000 replications. The last step is to assess the matching quality. Since we do not condition on all covariates but on the propensity score alone in PSM, it has to be checked whether the matching procedure is able to balance the distribution of the relevant covariates in both the control and treatment group. To do so, we use balancing tests described in Dehejia and Wahba (1999, 2002). The balancing tests were satisfied for all covariates. In order to guard against the potential

[6] Graphs of distributions of propensity scores that show the common support are available upon request.

source of bias (shown by Abadie and Imbens, 2002), we also implement the *Bias-Corrected Matching* estimator developed by Abadie and Imbens (2006).[7]

Finally, since all matching methods only match observations based upon observable covariates, they do not account for all unobservable covariates. To control for part of the unobservable factors, in particular, those factors that are time-invariant, we extend the cross-sectional matching estimator to a longitudinal setting and implement a difference-in-differences matching (DDM) estimator. With DDM we can exploit the data on the Migrant households in 2002 to construct the required counterfactual, instead of just using the data in 2006 (as is used in matching). When implementing DDM, we use both PSM matching and bias-corrected matching. In our PSM estimators, we match using the log odds ratios. In both PSM and bias-corrected matching, the nearest neighbour matching methods with replacement is used. In addition, we also compute the 'adjusted' version where observations in the control group are weighted by the number of times they are matched to an observation in the treatment group. For more information about the exact specification and the theory of our approaches, see a more complete working paper at www.reap-china.com.

RESULTS OF MULTIVARIATE ANALYSIS

Although we do not report the full version of the regressions from models 1 to 4 for brevity purposes, the DD analyses perform fairly well. The results from the DD regressions (and throughout the rest of the analyses) show that the estimates of the treatment effect (ie, the coefficient on the Any Parent Migrated dummy variable) are largely the same across all four specifications (models 1–4). However, the Unrestricted and Adjusted specification (model 4) generates a much higher goodness-of-fit statistic (or R^2) than other specifications (almost certainly because of the importance of capturing beginning scores, which embodies the unobserved ability of a student, and other covariates). Therefore, in the rest of the analyses, when reporting the results, we will mostly focus on the results from the Unrestricted and Adjusted model. The DD analyses also produce estimates with expected signs and significance. For example, the scores of older students drop relatively more than those of younger students (row 4, column 4). This finding is reasonable since students that enter primary school at an older age may have an initial advantage because they are relatively more mature (Fredriksson and

[7] This is achieved by using the STATA command 'nnmatch.'

Öckert, 2005), but the initial advantage gradually disappears as younger children catch up over the course of primary school.

The most important finding from the basic regression models is that we reject the hypothesis that migration affects school performance negatively (Table 2).[8] In all four models, the coefficient on the Any Parent Migrated household dummy variable is *not* negative. In fact, the coefficients are all positive and significantly different from zero. The magnitudes of the coefficients range from 1.16 to 3.18, meaning that, everything else held constant, after any parent in a household out-migrated between the first and fifth grade, the scores of the children of the migrants actually rose relative to the children of Never Migrant households. In other words, unlike claims made by some researchers, migration did not hurt school performance. At least in the migrant households in our sample area, migration has improved school performance.

The results hold when we examine other types of migrant households: no negative effect of migration on school performance is found. In the rest of Table 2, for each of the four specifications, we look at the effect of migration on school performance in all six types of migrant households.[9] In 20 out of the 24 cases the coefficient is positive. The coefficients are only negative for Mother Migrated Only households (row 4) and Mother Migrated (Unconditional) households (row 5) when the Restricted and Adjusted (column 3) or Unrestricted and Adjusted specification (column 4) is used. In each of these four cases, however, the *t*-statistic is smaller than 0.50, indicating there is no statistically significant effect of migration on school performance. Interestingly, as in the case of Any Parent Migrated households (row 1), when the father out-migrates (rows 2–3), the scores of migrant children improve.

So why is it that migration does not appear to have a negative effect on the scores of migrant children, and in some cases even appears to have a positive effect? Although we cannot answer this question from our analyses, one possible reason is that the income effect is relatively large compared to the adverse effect of less parental supervision. If migration leads to higher income, as found in Du *et al.* (2005), the migrant households that experience rising incomes may be able to provide better nutrition, improved access to educational supplies and burden their children with less housework. This

[8] In Table 2 we only report the coefficients on the treatment variable. The rest of the results are suppressed for brevity but are available from the authors upon request. We report the results for 24 different regressions.

[9] For completeness in Table 3, we include the results of the effect of Any Parent Migrated on school performance, but, in fact, this is a duplication of the results from Table 2, row 1.

may have a positive effect on school performance. The positive income effect is probably behind our finding that the largest positive effects are found in the Father Migrated Only households (Table 2, row 2). This result may arise since not only would children in such households benefit from higher incomes from migration, they would also suffer relatively less from falling parental care since the mother is still at home. Such an interpretation is also consistent with other findings. For example, Kandel and Kao (2001) found a positive relationship between fathers' migration and students' grades.

Matching results

The results of cross-sectional matching analysis, regardless of the method of matching, also reveal that migration has no significant negative effect on the

Table 2: Evaluating the effects of migration activities of parents on school performance of students in all six types of migrant households using difference-in-difference, Shaanxi Province, China[a]

Treatment variable (MIG_i)[b]	Outcome variable ($\Delta Score_i$)=$Score_{i,\,2006}-Score_{i,\,2002}$			
	(1)	(2)	(3)	(4)
	Restricted and unadjusted	Unrestricted and unadjusted	Restricted and adjusted[c]	Unrestricted and adjusted[c]
(1) *Any_Parent_migrated*	3.183	2.327	2.169	1.164
	(3.72)***	(3.03)***	(2.58)**	(1.65)*
No. of observations	1,575	1,575	1,549	1,549
R^2	0.01	0.27	0.10	0.43
(2) *Father_Migrated_Only* (*mother stayed home*)	4.634	3.812	3.630	2.356
	(4.27)***	(4.09)***	(3.45)***	(2.73)***
No. of observations	1,577	1,577	1,549	1,549
R^2	0.01	0.28	0.10	0.43
(3) *Father_Migrated* (*Unconditional*)	3.812	2.879	2.984	1.508
	(4.10)***	(3.52)***	(3.24)***	(1.98)**
No. of observations	1,595	1,595	1,551	1,551
R^2	0.01	0.27	0.10	0.43
(4) *Mother_Migrated_Only* (*father stayed home*)	0.839	0.156	−0.861	−0.121
	(0.45)	(0.08)	(0.45)	(0.07)
No. of observations	1,576	1,576	1,549	1,549
R^2	0.00	0.27	0.09	0.43
(5) *Mother_Migrated,* (*Unconditional*)	0.903	0.444	−0.147	−0.541
	(0.73)	(0.37)	(0.12)	(0.48)
No. of observations	1,587	1,587	1,551	1,551
R^2	0.00	0.27	0.09	0.43

Table 2: (continued)

Treatment variable (MIG_i)[b]	Outcome variable ($\Delta Score_i = Score_{i,\,2006} - Score_{i,\,2002}$)			
	(1)	(2)	(3)	(4)
	Restricted and unadjusted	Unrestricted and unadjusted	Restricted and adjusted[c]	Unrestricted and adjusted[c]
(6) *Both_parents_migrated,*	1.367	0.615	1.040	−0.536
	(0.79)	(0.38)	(0.58)	(0.35)
No. of observations	1,575	1,575	1,549	1,549
R^2	0.00	0.27	0.09	0.43

Data source: Authors' survey

[a] Four versions of specifications are used in the difference-in-difference estimation. An unrestricted model includes $Score_{i,2002}$ as a right-hand side variable. This removes the restriction in the standard DD model that the coefficient on $Score_{i,2002}$ equals one. An adjusted model includes other covariates in addition to the treatment variable. Model (1) is the standard DD model and is restricted and unadjusted. Mathematically, model (1) is expressed as $\Delta Score_i = \alpha + \delta MIG_i + \varepsilon$, where i is the index for students $\Delta Score_i$ is the before – after change in the school performance of student I, that is, scores from the fifth grade minus scores from the first grade: MIG_i is the treatment variable and δ is the parameter of interest, which measures the treatment effect. Model (2) is unrestricted and unadjusted, which is expressed as $\Delta Score_i = \alpha + \delta MIG_i + \gamma Score_{i,2002} + \varepsilon_i$. Model (3) is unrestricted adjusted. Model (3) is expressed as $\Delta Score_i = \alpha + \delta MIG_i + \boldsymbol{X_i\beta} + \varepsilon_i$, where $\mathbf{X}_i$ is a vector of covariates that include the characteristics of students, parents and households, and a set of township dummy variables. Model (4) is unrestricted and adjusted and is expressed as $\Delta Score_i = \alpha + \delta MIG_i + \gamma Score_{i,2002} + \boldsymbol{X_i\beta} + \varepsilon_i$.

[b] The treatment variable MIG_i takes the following six forms:. *Any Parent_Migrated*, which is a dummy variable that is equal to 1 if both parents lived at home in 2002 and at least one parent (either the father; mother or both parents) out-migrated by 2006. *Father_Migrated_Only* (*mother stayed at home*) is a dummy variable that is equal to 1 if only the father out-migrated by 2006 but was at home in 2002. *Father_Migrated* (*Unconditional*) is a dummy variable that is equal to 1 if the father was at home in 2002 but out-migrated by 2006 (including households in which the mother was either at home or not at home in 2006). *Mother_Migrated_Only (Father stayed at home*) is a dummy variable that is equal to 1 if only the mother out-migrated by 2006 but was at home in 2002; *Mother_Migrated* (*Unconditional*) is a dummy variable that is equal to 1 if the mother was at home in 2002 but out-migrated by 2006 (including households in which the father was either at home or not at home in 2006). *Both_Parents_Migrated* is a dummy variable=1 if both parents were at home in 2002, but out-migrated by 2006.

[c] The coefficients on the township dummy variables are not reported here for the sake of brevity.

[d] Robust t-statistics in parentheses.

*Significant at 10%; **significant at 5%; ***significant at 1%.

school performance of students. When Propensity Score Matching is used to examine the effect of migration on school performance for all six types of New Migrant households, there are no cases in which the coefficient on the treatment variable is negative and statistically significant (Table 3, column 1, rows 1a, 2a, 3a, 4a, 5a and 6a). The same is true when Bias-Corrected Matching is used (column 1, rows 1b, 2b, 3b, 4b, 5b and 6b). In fact, results from matching are quite similar to those from the DD analyses. When we use Bias-Corrected Matching, which perhaps generates better estimates and standard errors, we find that the coefficients on the treatment variables in the Father Migrated Only household model and Father Migrated (unconditional)

Table 3: Evaluating the effects of migration activities of parents on school performance of students in all six types of migrant households using matching and difference-in-difference matching, Shaanxi Province, China[a]

Treatment variable[b]	Matching[a]		Difference-in-difference matching	
	Average treatment effect for the treated	*t*-stat/*z*-value[c]	Average treatment effect for the treated	*t* stat/*z*-value[c]
Any_parent_migrated				
(1a) Propensity score matching	1.16	(1.02)	0.31	(0.28)
(1b) Bias corrected matching	1.57	(1.60)	2.12	(1.86)*
Father_Migrated_Only (*mother stayed home*)				
(2a) Propensity score matching	2.04	(1.36)	1.12	(0.77)
(2b) Bias corrected matching	3.59	(2.96)***	3.12	(1.93)**
Father_migrated, (*Unconditional*)				
(3a) Propensity score matching	1.57	(1.20)	2.35	(1.93)**
(3b) Bias corrected matching	2.19	(2.04)***	2.52	(1.99)***
Mother_Migrated_Only (*father stayed home*)				
(4a) Propensity score matching	−0.63	(−0.22)	−1.1	(−0.39)
(4b) Bias corrected matching	−0.94	(−0.43)	1.93	(0.58)
Mother_migrated (*Unconditional*)				
(5a) Propensity score matching	−0.45	(−0.26)	−1.51	(−0.88)
(5b) Bias corrected matching	−0.46	(−0.32)	0.82	(0.48)
Both_parents_migrated				
(6a) Propensity score matching	−0.22	(−0.09)	−0.56	(−0.23)
(6b) Bias corrected matching	−0.28	(−0.13)	0.97	(0.43)

Data source: Authors' survey

[a] Propensity scores are estimated using the same set of covariates as in Table 2.

[b] The treatment variables are described in note b in Table 2.

[c] *t*-Statistics are reported for propensity score matching and *z*-values are reported for bias-corrected matching in parentheses.

*Significant at 10% level; **significant at 5% level; ***significant at 1% level.

household model are positive and statistically significant and the magnitudes are similar to those from the DD analyses. In addition, and importantly, the findings remain largely the same when the DDM estimator is used (results not reported for brevity sake – please see our working paper at www.reap-china.com).

CONCLUSIONS

In this paper, we have tried to understand whether or not the school performance of children suffers when their father, mother or both parents migrate from the village into the city. Despite the perception that is commonly found in the literature and the press, our results, somewhat surprisingly, show that there is no effect of migration on the school performance of the children from migrant households. Comparing the change in the grades before and after parents out-migrate between children from migrant households and those from non-migrant households, we can reject the hypothesis that migration harms the grades of migrant children. In fact, in the analysis of some migrant households, especially in those in which the father out-migrates, migration is shown to have a statistically significant and positive effect on the performance of migrant children. We also find that the effects of migration on children's school performance are not systematically different for households that are more or less wealthy households. Neither are the effects different across households that have one or more than one child.

Based on these results, it might be tempting to conclude that policy makers do not need to take any action since there is no measurable effect of migration on school performance. If there were, education officials might want to reduce class sizes or hire more qualified teachers to improve the mentoring programme in schools in which there were many children of migrants. Boarding schools might offer some of the services that parents originally carried out before they entered the migrant labour force. Ultimately, measures can be promoted to offer the children of migrants who lived in China's cities better access to urban schools so parents would not have to leave their children behind. However, all of these programmes are costly. Although there might be good reason to implement such policies anyhow, according to our results, they should not be carried out on the ground of the negative effect of migration on school performance.

Although we have tried a number of alternative approaches to identify the effect of migration, and although the findings are largely robust, if the

assumptions underlying our methodologies were not valid, our estimates could be bias. Even though we control for many observed and time-invariant unobserved factors, there still may be factors that are known to the parents of migrants and potential migrants but are not be observable to the econometrician. For example, it may be that all parents who were in the village with their children in 2002 worry about whether or not their migration decision would negatively affect the school performance of their children. If it is the case that those parents who – though having an opportunity to migrate – believed that the grades of their children would suffer decided not to migrate, while those that believed their children's grades would not suffer decided to migrate, then our results would be subject to selection bias.

If there was, in fact, such a selection bias and we did not account for it (as we were unable to – due to the absence of any effective instrumental variable), would our results be useless? We believe not. We believe even if there was a selection bias our results are showing that when rural parents out-migrate, the grades of their children do not suffer. It is true that part of the reason for the zero effect may be exactly this selection effect – parents do not go when they believe the grades of the children would suffer. But, from society's point of view, there is less cost in terms of school performance of its children due to migration.[10]

Acknowledgements
The authors would like to thank Chengfang Liu, Renfu Luo who have spent uncountable days coordinating the survey and cleaning data. A special thanks to all the enumerators, school principals and students. We are also grateful for the useful comments from Belton Fleisher and two anonymous referees. We acknowledge grants to support field research from The Ford Foundation (Beijing), Chinese Academy of Sciences (KSCX2-YW-N-039) and support for follow-up research from the National Natural Science Foundation of China (70803047). The Natural Science Foundation of Zhejiang Province (Y607420) and the Social Sciences Foundation of Zhejiang Province (07CGLJ005YBQ).

[10] There is also another potential source of endogeneity that we are not able to account for in the analysis. It is possible that unaccounted for shocks, either in the local economy or in individual households, affect both parents' migration activities and students' grades. If these shocks systematically affect all the households, then it is possible that our coefficients also are biased due to the fact that we did not account for this type of unobservable heterogeneity. In this case, it is difficult to determine the direction of the bias. These shocks could be either negative (eg, the family suffers a crop failure or family sickness) or positive (eg, the family receives an inheritance or enjoys a bumper crop) and can lead to negative or positive bias in our estimates.

REFERENCES

Abadie, A and Imbens, GW. 2002: *Simple and bias-corrected matching estimators*. Technical report, Department of Economics, UC Berkeley.

Abadie, A and Imbens, GW. 2006: Large sample properties of matching estimators for average treatment effects. *Econometrica* 74: 235–267.

Battistella, G and Conaco, MCG. 1998: The impact of labour migration on the children left behind: A study of elementary school children in the Philippines. *Journal of Social Issues in Southeast Asia* 13: 220–241.

Benjamin, D, Brandt, L and Giles, J. 2005: The evolution of income inequality in rural China. *Economic Development and Cultural Change* 53: 769–824.

Blau, DM. 1999: The effect of income on child development. *Review of Economics and Statistics* 81: 261–276.

Brown, PH and Park, A. 2002: Education and poverty in rural China. *Economics of Education Review* 21: 523–541.

Caliendo, M and Kopeinig, S. 2008: Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys* 22: 31–72.

Cox Edwards, A and Ureta, M. 2003: International migration, remittances, and schooling: Evidence from El Salvador. *Journal of Development Economics* 72: 429–461.

de Brauw, A and Giles, J. 2007: *Migrant labor markets and the welfare of rural households in the developing world: Evidence from China*. Working paper, Michigan State University.

de Brauw, A, Huang, J, Rozelle, S, Zhang, L and Zhang, Y. 2002: The evolution of China's rural labor markets during the reforms. *Journal of Comparative Economics* 30: 329–353.

Dehejia, RH and Wahba, S. 1999: Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association* 94: 1053–1062.

Dehejia, RH and Wahba, S. 2002: Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics* 84: 151–161.

Du, Y, Park, A and Wang, S. 2005: Migration and rural poverty in China. *Journal of Comparative Economics* 33: 688–709.

Duncan, GJ, Brooks-Gunn, J and Klebanov, PK. 1994: Economic deprivation and early childhood development. *Child Development* 65: 296–318.

Fredriksson, P and Öckert, B. 2005: *Is early learning really more productive? The effect of school starting age on school and labor market performance*. IZA Discussion papers, 1659, Institute for the Study of Labor.

Giles, J. 2006: Is life more risky in the open? Household risk-coping and the opening of China's labor markets. *Journal of Development Economics* 81: 25–60.

Glewwe, P and Jacoby, HG. 2004: Economic growth and the demand for education: Is there a wealth effect? *Journal of Development Economics* 74: 33–51.

Hanson, GH and Woodruff, C. 2004: *Emigration and educational attainment in Mexico*. Working paper, University of California, San Diego.

Hanushek, EA. 1992: The trade-off between child quantity and quality. *The Journal of Political Economy* 100: 84–117.

Heckman, JJ. 2005: China's human capital investment. *China Economic Review* 16: 50–70.

Kandel, W and Kao, G. 2001: The impact of temporary labor migration on Mexican children's educational aspirations and performance. *International Migration Review* 35: 1205–1231.

Li, X. 2004: The investigation on the rural migrant children (in Chinese). *Journal of Women Study in China* 10: 35–37.

McKenzie, D and Rapoport, H. 2006: *Can migration reduce educational attainment? Evidence from Mexico*. The World Bank Policy Research Working paper series no. 3952.

McKenzie, D and Rapoport, H. 2007: Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico. *Journal of Development Economics* 84: 1–24.

Princiotta, D, Flanagan, KD and Hausken, EG. 2006: *Fifth grade: Findings from the Fifth grade follow-up of the early childhood longitudinal study, Kindergarten class of 1998–99*. National Center for Education Statistics No. 2006–2038.

Rosenbaum, PR and Rubin, DB. 1983: The central role of the propensity score in observational studies for causal effects. *Biometrika* 70: 41–55.

Rozelle, S, Guo, L, Shen, M, Hughart, A and Giles, J. 1999: Leaving China's farms: Survey results of new paths and remaining hurdles to rural migration. *The China Quarterly* 158: 367–393.

Smith, J and Todd, P. 2005: Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics* 125: 305–353.

Steelman, LC and Mercy, JA. 1980: Unconfounding the confluence model: A test of sibship size and birth-order effects on intelligence. *American Sociological Review* 45: 571–582.

Tan, S and Wang, X. 2004: The study on the migrant children (in Chinese). *Journal of Education in Hubei* 20: 11–12.

Todaro, MP. 1989: *Economic Development in the Third World*. Longman: New York.

Wang, Y and Wu, X. 2003: A case study of the migrant children (in Chinese). *Investigation of the Youth* 4: 7–10.

Wu, N. 2004: The problems on the rural migrant children (in Chinese). Unpublished Manuscript, China National Institute for Education Research.

Zhou, Q and Wu, H. 2004: A mother of 13 years old. *XinhuaNet*, May 26.