Migration, Schooling Choice, and Student Outcomes in China

Xiaobing Wang
Yu Bai
Linxiu Zhang
Scott Rozelle

The fast pace of economic growth in China is in no small part attributed to the massive movement of migrant workers from rural to urban areas. It is estimated that in 2014 more than 168 million migrants were living and working in China's cities (NBSC 2015). In China, as elsewhere, migration imparts significant benefits to individuals through the higher returns to work; it can also have strong and transformative impacts on both the origin and destination communities (Taylor, Rozelle, and de Brauw 2003; Du, Park, and Wang 2005; Gibson and McKenzie 2012).

The same benefits do not always accrue to other family members, particularly the children of migrants. Researchers internationally have documented the downside of migration on the children of migrants (Spera 2005; Lahaie et al. 2009). In China, the potential impact of the decision of parents to migrate may be especially large, given the country's unique institutional and policy environment. Because of the *hukou* household registration system, migrant children are not unconditionally entitled to enroll in urban public schools (Lai et al. 2014). Although policies have fluctuated over time, two-thirds of migrant children in many of China's large cities are not eligible to attend public schools (Chen and Feng 2013).[1]

When rural parents debate whether to move to the city, they may consider the impact of that decision on family members. Migrant parents who are unable to enroll their children in local urban (or, often distant suburban) public schools have to make one of two choices. In one case their children can attend rural public schools. These children then live with their grandparents (or other relatives/caregivers) and attend the local public school. Students in this situation are called left-behind children (LBC). According to an All-China Women's Federation report (ACWF 2013) based on the 2010 Population Census, there are more than 61 million LBC (including children living with their father or mother in their hometown) aged 17 years or younger in China. These left-behind children comprise 22 percent of all children in China and 38 percent of rural children.

Alternatively, migrant parents who choose to bring their children to the city with them but are unable to enroll them in urban public schools can enroll them in private, for-profit schools (which we refer to here as migrant schools) that have emerged for the purpose of educating the children of migrants. In this case, migrant children are educated in schools with other migrant children (but no other urban children). Much has been written about migrant children and the conditions of their schools (Goodburn 2009; Chen and Feng 2013; Lai et al. 2014). Tuition can be expensive and teacher turnover high. There is little, if any, regulation or oversight by urban education officials (Yiu 2016). Further, the shift of a child from his or her rural community often disrupts the child's progress in school (Hanushek, Kain, and Rivkin 2004).

There are potentially offsetting benefits and costs of the decision to send a child to either a migrant urban school or a rural public school. Above all, parents must consider the quality of education in each type of school. Rural students attend regulated rural public schools where recent investments in teachers and facilities have resulted in low teacher turnover (Lai, Sadoulet, and Janvry 2011). Public education policy stipulates that all teachers must have a minimum level of education and certification (Chen and Feng 2013) and teach a standard, approved curriculum (Zhao et al. 2014). Facilities in rural schools also are much improved in recent years owing to large, centrally funded investments. When compared to the informal nature of migrant schools, rural public schools may deliver a more consistent education.

Beyond the differences in the quality of rural public schools and migrant schools, there are other factors that migrant parents must consider. If migrant parents take their children with them, the children may benefit from parents' higher income in cities and from daily parental care. However, due to the family's often relatively poor economic status, the home environment in which migrants live and the neighborhoods in which their children grow up are not always safe or nurturing. In contrast, if migrants choose to have their children remain in the village and attend rural public schools, the children will lack the day-to-day care of parents. However, since most LBC live with their grandparents, they will have regular caretakers and will be living in familiar and safe surroundings.

Of course, parents can also decide not to migrate and instead stay at home and continue to work on the farm. In this case children living with their parents (CLP) attend rural public schools like LBC. CLP, like their migrant counterparts in urban areas, are able to return from school each day to their parents. However, although each rural household in China has access to cultivated land (both through contracts with the village and through rental agreements), the average operational size of a farm is only 0.6 hectares (NBSC 2014; Wang et al. 2016). Hence, given the nature of China's small-scale farming sector, when husband and wife both live at

home (or one spouse lives at home and the other migrates), family income can be significantly lower than when both parents migrate to cities. With lower incomes, the scarcer resources available to the family to support their child's education, health, and nutrition can lead to poorer academic performance (Chen, Huffman, and Rozelle 2009; Zhang et al. 2014).

In short, there are tradeoffs in schooling and employment/residency decisions. Migrant children live in families with higher incomes than those rural families with no off-farm sources of income and also live with their parents. However, almost all of them must attend migrant schools and do not always live in good, supportive neighborhoods. If parents decide to send their child to a rural public school, the tradeoffs differ when deciding between working as a migrant and living away from home and living at home and farming. In the case of LBC, migrant parents will have higher earnings, but will also provide less care. In the case of CLP, parents will have lower earnings, but will be able to provide more care and children will be living at home and attending school in a familiar environment.

In this study we compare the schooling outcomes of migrant students, LBC, and CLP by examining the impacts of migration and schooling choice on academic performance and learning anxiety. During a one-month period in May 2014, we collected data (including giving students a standardized math exam) on a random selection of students in 87 Shanghai and Suzhou migrant schools. During the same time period we also visited the home province (Anhui) of about 57 percent of our sample of migrant children (and the home towns in five counties in Anhui of around 24 percent of our sample) and collected data on (and gave the same standardized math test to) a sample of students in 30 rural public schools in these five counties. This is the first study to use this type of matched data.

We have two objectives. First, we compare the distributions of students' standardized math scores and the prevalence of learning anxiety among migrant students, LBC, and CLP. We believe, based on our unique dataset, that this will provide an accurate estimate of the achievement and learning anxiety gaps between migrant students and CLP/LBC. Second, we use multivariate analysis (employing both ordinary least squares and propensity score matching) to examine whether parental migration affects the academic performance of migrant students, LBC, and CLP and seek to identify the sources of the achievement gap. We identify three potential sources of any such gap. First, school outcomes between migrant students and LBC/CLP may differ due to differences in the quality of schools in urban and rural communities. Beyond school quality, part of the gap may be due either to a selection effect (that is, parents of children with certain characteristics or with particular occupations are more likely to take their children to the city with them) or to a migrant living environment effect (that is, children's studies might be disrupted when they are removed from their rural home environment and relocated in an urban migrant community).

In analyzing the sources of the achievement gaps, we first isolate the selection effect and measure the magnitude of the rest of the gap (which would be made up of a school quality effect and a living environment effect). We use a dataset that includes migrant students and CLP. With this subset of data, we match students on a number of observable characteristics and limit our matching analysis to children who have similar living arrangements. Finally, we compare the schooling outcomes of migrant children and CLP.

To further decompose the achievement gap, we examine the impact of the duration of attendance at migrant schools on schooling outcomes. This analysis allows us to determine whether there is a school quality effect. We compare migrant students who migrated to the city at the beginning of primary school to those students who migrated to the city at the beginning of grade 5. If there is a difference in their outcomes, we will assume that it is due to the quality of migrant schools.

Our main contribution is to compare the schooling outcomes of migrant students who were originally from a certain part of the country with the schooling outcomes of students who attend rural public schools in the same part of the country. The matching of data on students in migrant communities who live with their parents and on children living with their parents in rural source communities also helps us gauge the academic performance of the migrant students, accounting for the role of both observable student/family characteristics and the living arrangements of families. The analysis of the different components of the performance gap also identifies the factors that may be responsible for the poor outcomes of migrant students in China.

Data and methodology

We began by conducting a survey to choose a sample of schools in two suburban areas around central Shanghai (in Suzhou, Jiangsu province and Shanghai's outlying districts and counties). Unlike public urban and rural schools, no official list of private migrant schools is available in Shanghai or Suzhou. To collect a comprehensive list of private migrant schools, we contacted all educational and research institutes and non-profit organizations in the two cities that might have contact information for such schools. We then called each school to confirm that it was still operating. Our list contained 87 schools, and all private migrant schools in that sampling frame were part of our overall sample.

We visited the private migrant schools and randomly chose one fifth grade class in each school. We chose fifth grade students because we believe children mostly aged 10–12 are able to answer questions about their home environment and their parents' migrant status and educational background. There were 3,755 migrant students in fifth grade classes in the 87 migrant schools. All students were the children of migrants.

In our initial survey, we asked students to identify the community (by town and county name) that they (or their parents) came from. Of the 3,755 fifth graders in the migrant school sample, 914, or 24 percent, came from three prefectures in Anhui province: Fuyang, Lu'an, and Haozhou. We refer to these prefectures as our three core study areas.

The second part of the sampling protocol was developed to sample towns, schools, and students from the three core study areas. We randomly chose two counties in Fuyang prefecture, two counties in Lu'an prefecture, and one county in Haozhou prefecture. The survey team obtained a list of all schools in each of the sampled counties and narrowed this list to the elementary schools that offered six full grades. From this list, we randomly selected six schools in each of the five counties, for a total of 30 rural public schools.

Next, we created a sample of students. In each of the schools, we selected all students in one randomly chosen fifth grade class. A total of 1,514 students were included in our sample of rural public schools in five counties in three prefectures in Anhui province. On average, there were 50 students per class.

In order to make comparable assessments of levels of learning and learning anxiety, we carried out the data collection in both the urban and rural study areas during the same two-week period in May 2014. This ensured that the students in the two school samples were at the same point in their schooling in terms of years, months, and weeks of schooling. In each of the sample migrant and rural schools, enumerators collected information on student characteristics (sex, age, etc.) and family characteristics (parental education levels, whether the student was an only child, and household assets).

After the individual interviews, the enumerators administered standardized math tests designed to be appropriate for fifth grade students. As stated above, identical tests were given during the same two-week period in both the private migrant schools and the rural public schools. Local educators in both the migrant and the rural communities assisted with the selection of questions from items developed for the Trends in International Mathematics and Science Study.[2] The test, administered on paper, was timed (25 minutes), proctored by the enumerators at each school, and graded by our research team. For analysis, we normalize scores using the grade distribution of the control group (students in rural public schools).

This portion of the survey also employed an internationally recognized psychological scale, the Mental Health Test (MHT), to measure students' learning anxiety. The MHT was administrated and proctored by our survey team in the classroom. Learning anxiety is most closely related to academic performance and has been used extensively across China as a measure of the mental health of primary school students in both urban and rural areas (Deng, Lei, and Cao 2002; Wang et al. 2015; Shi et al. 2016). A higher

TABLE 1 Summary of the datasets

		Number of students				
			Migrant schools			
Data-set	Sample composition	Total	Shanghai and Suzhou	Shanghai	Suzhou	Rural schools
1	Whole sample	4,770	3,380	2,428	952	1,390
2	Migrant students from Anhui and students who live in Anhui	2,881	1,491	1,125	366	1,390
3	Migrant students from three core study areas in Anhui and students who live in Anhui	2,263	873	693	180	1,390
4	Rural students who live in three core study areas in Anhui	1,390				1,390

SOURCE: Authors' dataset.

score (which ranges from 0 to 15) indicates higher earning anxiety and a possible need for professional help. A score higher than 7 indicates that the student is considered to have severe learning anxiety. Based on this standard, we categorized the learning anxiety score into a dummy variable (1 = higher learning anxiety; 0 = otherwise) by setting the cutoff at 8. Besides this dummy variable, we also use the learning anxiety score (a continuous variable) in this study.

The third portion of the survey provided data on the key independent variable, parental migration. This information was derived from the survey questionnaire filled out by students under the supervision of enumerators. As a way of cross-checking, homeroom teachers were asked to verify the information on parental migration status provided by each student. From this information, we determined the type of households that children lived in when attending either private migrant schools or rural public schools. We use only the sample observations in which either "both parents have migrated" or "neither parent has migrated," since most children in Shanghai and Suzhou were living with both parents.

Constructing four datasets

We constructed four datasets. The first three differed according to the origin (or source community) of the migrant students (Table 1). Dataset 1 contains all migrant students and all rural school students living in Anhui; dataset 2 contains all migrant students from Anhui and students who live in Anhui; dataset 3 contains all migrant students from the three core study areas and students who live in those areas; and dataset 4 contains only students from the rural sample—that is, all LBC and all CLP. Dataset 4 excludes one type of household that was fairly common: father-only migration households.

These are rural households in which students lived with their mother while only their father migrated and did not live at home. Father-only migration households accounted for 19 percent of the Anhui sample. Although we excluded these households to make the comparison between rural CLP and migrant students (who almost always lived with both parents) more precise, the basic findings do not change when we included these households in our sample (results available on request).

Empirical methodology

We use ordinary least squares (OLS) and propensity score matching (PSM) to examine whether parental migration affects students' math scores and learning anxiety. We divide our sample students into three treatment groups and three comparison groups according to parents' migration status: a) migrant students who live with both parents in Shanghai and Suzhou/LBC whose parents both migrated; b) migrant students living with both parents in Shanghai and Suzhou/CLP living with both parents in rural areas; and c) LBC whose parents both migrated/CLP living with both parents in rural areas. The treatment/comparison sets for a and b are estimated based on datasets 1, 2, and 3 separately. The treatment/comparison set for c is estimated based on dataset 4. In total, seven treatment/comparison sets are analyzed separately.

To test the impact of parental migration, we regress the outcome variables (standardized math scores and learning anxiety scores) on a dummy variable of the treatment status. The model we estimate is:

$$Score_{is} = \alpha + \beta \cdot Treat_{is} + \delta \cdot X_{is} + \varepsilon_{is}$$

where the dependent variable, $Score_{is}$, indicates the measures of the math score, learning anxiety, or the severity of learning anxiety of student i in school s. $Treat_{is}$ is the treatment variable (1 = treatment group; 0 = control group), and β is the parameter of interest. X_{is} is a vector of covariates that are included to capture the characteristics of students and their households, including sex, age, whether the student is an only child (1 = yes; 0 = no), education attainment of father and mother, and household assets.

In addition to OLS, we used matching approaches to check whether our results are robust to the choice of estimators. Rosenbaum and Rubin (1983) proposed propensity score matching to reduce the bias in the estimation of treatment effects with observational datasets. PSM allows us to match a student in the treatment group with a similar student from the comparison group and interpret the difference in their outcomes as the effect of parental migration status when observable characteristics are continuous, or when the set of explanatory factors that determine parental migration contains multiple variables (Rosenbaum and Rubin 1985). We estimate the propensity scores of sample individuals and compare the outcomes of

students in the treatment groups to students in the comparison groups who have similar propensity scores.[3]

Analytical strategy

Our datasets allow us to set up eight treatments to identify the migration effects and/or school effects on the academic performance and learning anxiety of migrant and rural students. The first set of three treatments compares the schooling performance and learning anxiety of migrant students with those of LBC. Treatment 1 includes all migrant children and LBC; Treatment 2 includes all migrant children from Anhui and LBC; and Treatment 3 includes all migrant children from the three core study areas in Anhui and LBC.

We take two approaches to isolate the selection effect from the school quality effect and the living environment effect on the standardized math scores, the prevalence of learning anxiety, and the severity of learning anxiety. First, we use a second set of three treatments. In this part of the analysis (that is, in Treatments 4, 5, and 6) we hold the nature of the living arrangements constant by comparing migrant students with CLP. The treatment variable is migrant students; the base reference group is CLP students. Analogous to Treatments 1, 2, and 3, Treatment 4 includes all migrant children and CLP; Treatment 5 includes all migrant children from Anhui and CLP; and Treatment 6 includes all migrant children from the three core study areas in Anhui and CLP. In each of these treatments, we then use individual observable characteristics to match migrant students and CLP as the final part of our effort to hold the selection effect constant.[4]

Second, we use only data from migrant schools, calling this school quality analysis Treatment 7. With these data, we compare the math scores of students who attended migrant schools for five years (that is, those who spent all of their primary school years in migrant schools) with the scores of students who attended rural public schools until the year prior to our survey (that is, those who were in migrant schools for only one year). If there is a difference, we assume that it is mostly due to the poor quality of migrant schools, although it might also be, in part, a living environment effect.

Using Treatment 8, we seek to understand one additional impact of migration on the test scores and learning anxiety of students. Using the sample of rural students in Anhui, we control for school effects by comparing students who live in different types of families but who attend the same school. The treatment group is LBC; the base reference group is CLP. With this dataset, we can isolate a pure care effect caused by the absence of parental care. In other words, this specification allows us to test the effect of having both parents migrate on academic performance and learning anxiety.

TABLE 2 Migration status of students by source communities

		Father is a migrant	
		Yes	**No**
Migrant students in Shanghai and Suzhou ($N = 3{,}380$)			
[1] Mother is a migrant	Yes	2,934[a] (86.8)	161 (4.8)
[2]	No	195 (5.8)	90 (2.7)
Migrant students in Shanghai and Suzhou who come from Anhui ($N = 1{,}491$)			
[3] Mother is a migrant	Yes	1,291[b] (86.6)	64 (4.3)
[4]	No	87 (5.8)	49 (3.3)
Migrant students in Shanghai and Suzhou who come from three core study areas in Anhui ($N = 873$)			
[5] Mother is a migrant	Yes	764[c] (87.5)	34 (3.9)
[6]	No	49 (5.6)	26 (3.0)
All students in rural Anhui ($N = 1{,}390$)			
[7] Mother is a migrant	Yes	727[d] (52.3)	43 (3.1)
[8]	No	263 (18.9)	357[e] (25.7)

NOTES: Numbers of students are presented and, in parentheses, the percentage of students by the migration status of parents in each of the groups.
[a]Migrant children who live with both parents for all migrant students in Shanghai and Suzhou.
[b]Migrant children who live with both parents for migrant students in Shanghai and Suzhou who come from Anhui.
[c]Migrant children who live with both parents for migrant students in Shanghai and Suzhou who come from the three core study areas in Anhui.
[d]LBC: children who attend rural public schools and live with relatives or other caretakers while both parents migrated to the city.
[e]CLP: children living with both parents in rural areas and attending rural public schools.
SOURCE: Authors' dataset.

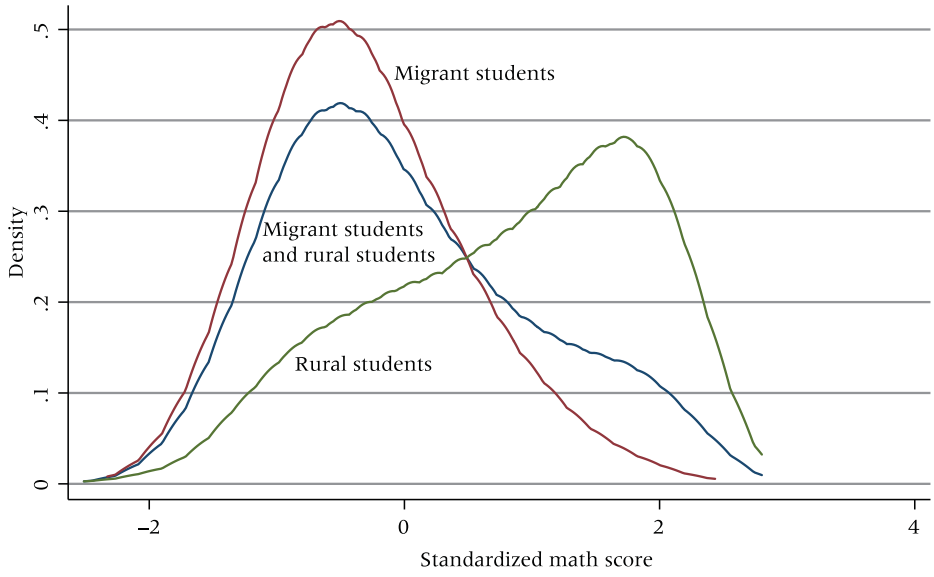
Results

Prevalence of migrant households

We observe distinct patterns of working and living arrangements in rural China (Table 2), similar to the results of other research (e.g., Bai et al. 2017; Shi et al. 2016). In our migrant sample, most children in Shanghai and Suzhou were living with both parents, regardless of whether we look at all migrants (from dataset 1, rows 1 and 2), all migrants from Anhui (from dataset 2, rows 3 and 4), or all migrants from the three core study areas in Anhui (from dataset 3, rows 5 and 6). In the three samples/subsamples, 87–88 percent of the migrant children lived with both parents. The remaining migrant children lived either with one parent or with a relative.

Our data revealed a more diverse set of living arrangements in rural Anhui (Table 2, rows 7 and 8). Among the 1,390 households in the Anhui sample, 52 percent of rural public school children were living in households in which both parents had migrated and could be categorized as LBC. At the same time, 26 percent of the sample students lived with both parents and could be categorized as CLP. The rest of the students lived with one parent while the other migrated (we do not include these students in our analysis because they comprise a small share of the total sample).

FIGURE 1 Kernel density estimates of students' standardized math scores



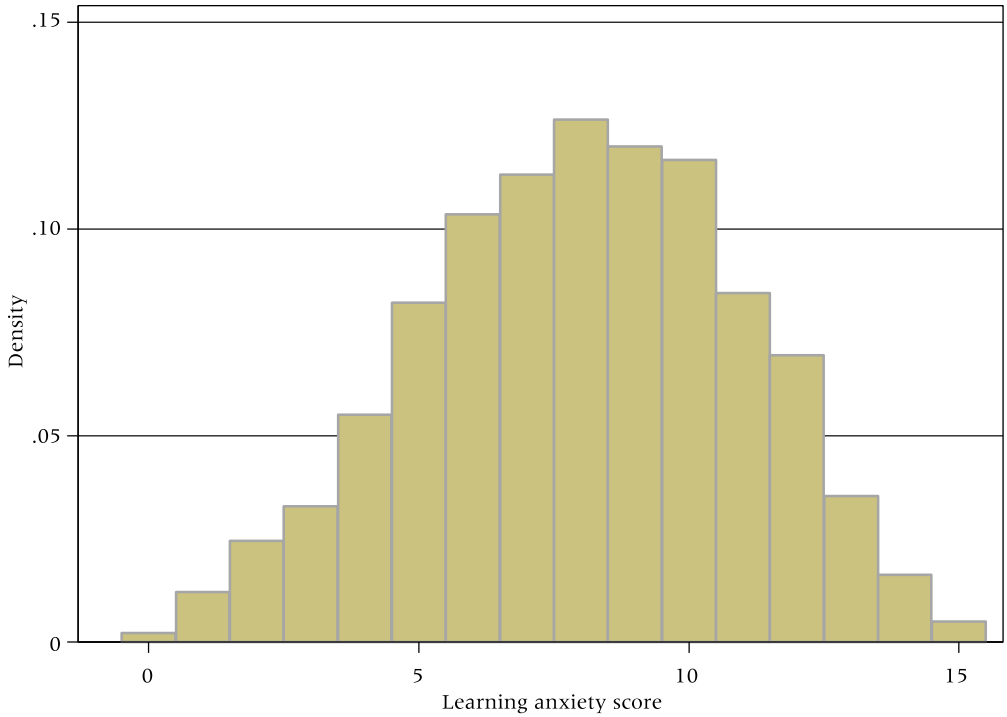
Schooling outcomes

Distribution of math scores. In the design of the standardized math tests, it is critical that the distribution of the scores are normal and not subject to either ceiling effects (too many students with perfect scores) or floor effects (too many students with zero or very low scores). Figure 1 shows that the distribution of the entire sample is close to normal and the distributions of the migrant and rural student samples are both near normal. Interestingly, the raw scores of students from Anhui are more than 1 standard deviation higher than the raw scores of migrant students.

Although not shown in the figure, the mean scores of LBC (0.81) and CLP (0.90) are similar, with no statistically significant difference. While this result may be surprising to some observers of China's education system, given the frequent discussion of the poor performance of LBC in rural schools, the results are consistent with a recent paper by Zhou et al. (2015), which shows that the performance of LBC and CLP is statistically the same.

Prevalence of learning anxiety. Figure 2 displays the kernel density estimation of students' scores from the learning anxiety test that we administered to the sample students. We find that 61 percent of students in migrant schools and 50 percent of rural public school students showed signs of moderate to severe learning anxiety. In addition, the mean learning anxiety score of migrant students (8.2 points) is higher than that of LBC and CLP (7.5 points). Taken together, these findings suggest that anxiety is more severe and more prevalent among migrant students than among rural

FIGURE 2 Density estimation of learning anxiety for all students



NOTE: Higher score (from 0 to 15) indicates higher learning anxiety.

students. In the descriptive statistics, the difference in learning anxiety between migrant and rural students is statistically significant (Table A1, row 16, columns 3 and 6).[5] However, no significant differences in learning anxiety are found between LBC and CLP in rural Anhui (row 16, column 9).

Multivariate analysis

Math test scores. According to the results of the multivariate analysis, the standardized math test scores of migrant students are far below those of LBC (Table 3). In the OLS results, the math scores of all migrant students are 1.21 standard deviations below those of LBC (row 1, column 1), holding all else constant. The measured gap using the matching analysis and the whole sample is slightly wider at 1.26 standard deviations (row 1, column 2). When using the other subsamples of migrant students (Anhui migrant students and migrant students from the three core study areas), the math scores of migrant students are still far below those of LBC in rural Anhui (from 1.25 to 1.35 standard deviations lower—rows 2 and 3, columns 1 and 2).

TABLE 3 OLS and PSM estimates of the school effects on the outcomes of students in migrant schools compared with outcomes of left-behind children attending rural public schools

	Standardized math score		Learning anxiety score		Dummy of high learning anxiety (1 = yes; 0 otherwise)	
	OLS (1)	PSM (2)	OLS (3)	PSM (4)	OLS (5)	PSM (6)
Dataset 1						
[1] Treatment 1	–1.21***	–1.26***	0.70***	0.86***	0.11***	0.13***
	(0.04)	(0.08)	(0.13)	(0.21)	(0.02)	(0.04)
Obs.	3661	3554	3661	3554	3661	3554
R-sq	0.25		0.02		0.02	
Dataset 2						
[2] Treatment 2	–1.25***	–1.35***	0.72***	0.99***	0.10***	0.11***
	(0.04)	(0.09)	(0.15)	(0.25)	(0.03)	(0.04)
Obs.	2018	1776	2018	1776	2018	1776
R-sq	0.31		0.03		0.03	
Dataset 3						
[3] Treatment 3	–1.25***	–1.31***	0.76***	0.63**	0.11***	0.09**
	(0.05)	(0.090)	(0.16)	(0.25)	(0.03)	(0.05)
Obs.	1491	1147	1491	1147	1491	1147
R-sq	0.32		0.03		0.03	

NOTE: See text for descriptions of datasets and Treatments.
*significant at 10%; **significant at 5%; ***significant at 1%.
SOURCE: Authors' dataset.

To put into perspective the large gap in academic performance between migrant and rural students, Ma et al. (2016) found that the average student's test score rises about 0.5 standard deviations in a year of study. If this is true, the gaps we find in math test scores between migrant and rural children suggest that by grade 5 migrant students are about 2 school years behind rural public school students.

To uncover the source of this gap, in Table 4 we examine the school effect by comparing the math test scores of migrant students and CLP. We hold constant the observed part of the selection effect, since in our matching results we are comparing students who have the same observable individual and family characteristics and who live with their parents. The scores of migrant students are even lower than those of rural CLP. When using OLS, the gap increases slightly in comparison to the migrant/LBC gap to 1.28 standard deviations (row 1, column 1). When using PSM, the gap is also larger than the gap between migrant and LBC students (1.38 standard deviations—row 1, column 2). Similar results are found when comparing migrants from Anhui (or the three core study areas from Anhui) with CLP (rows 2 and 3, columns 1 and 2).

TABLE 4 OLS and PSM estimates of the school effects on the outcomes of students in migrant schools compared with outcomes of children living with parents and attending rural public schools

	Standardized math score		Learning anxiety score		Dummy of high learning anxiety (1 = yes; 0 otherwise)	
	OLS (1)	PSM (2)	OLS (3)	PSM (4)	OLS (5)	PSM (6)
Dataset 1						
[1] Treatment 4	–1.28***	–1.38***	0.69***	0.50	0.13***	0.13**
	(0.05)	(0.11)	(0.17)	(0.30)	(0.03)	(0.05)
Obs.	3291	3267	3291	3267	3291	3267
R-sq	0.22		0.02		0.02	
Dataset 2						
[2] Treatment 5	–1.32***	–1.36***	0.70***	0.69**	0.12***	0.12**
	(0.05)	(0.10)	(0.19)	(0.28)	(0.03)	(0.05)
Obs.	1648	1561	1648	1561	1648	1561
R-sq	0.32		0.03		0.03	
Dataset 3						
[3] Treatment 6	–1.32***	–1.37***	0.74***	0.53	0.13***	0.12**
	(0.06)	(0.12)	(0.20)	(0.33)	(0.03)	(0.06)
Obs.	1121	992	1121	992	1121	992
R-sq	0.37		0.03		0.04	

NOTE: See text for descriptions of datasets and Treatments.
*significant at 10%; **significant at 5%; ***significant at 1%.
SOURCE: Authors' dataset.

Because the gap in math test performance is wide even when only examining students who live with their parents, the gap may arise from one of two (or both) effects: the quality of schools in migrant communities and the migrant living environment. Such a finding is consistent with studies by Lai et al. (2014) and Wang et al. (2017), who find that the quality of migrant schooling appears to be so low that it is seriously affecting the learning of migrant students.

While it is possible that there is a negative quality effect, it is not possible to say from the analysis in Table 4 whether all of the gap is due to the poor quality of schools. Part of the measured gap, holding living arrangements constant, may also be due to the fact that migrant students live in an environment that is not conducive to learning.

To generate a purer measure of the school quality effect, we present the results of an analysis that uses a sample containing only migrant students. We then compare test scores of migrant students who had been attending school in the city for different lengths of time. If the poor quality of migrant schools hurts student performance, we would expect that, among migrant students in the same grade, those who recently moved to cities might not be as negatively affected as those who have been in the city longer. We

TABLE 5 OLS and PSM estimates comparing the schooling outcomes of migrant students who have lived in the city for at least five years with outcomes of students who migrated at the start of grade 5

	Standardized math score		Learning anxiety score		Dummy of high learning anxiety (1 = yes; 0 otherwise)	
	OLS	PSM	OLS	PSM	OLS	PSM
	(1)	(2)	(3)	(4)	(5)	(6)
[1] Treatment 7	–0.21***	–0.22**	–0.53***	–0.35	–0.06*	–0.04
	(0.06)	(0.09)	(0.24)	(0.32)	(0.04)	(0.05)
Obs.	1925	1915	1925	1915	1925	1915
R-sq	0.05		0.02		0.02	

NOTE: See text for description of Treatment 7.
*significant at 10%; **significant at 5%; ***significant at 1%.
SOURCE: Authors' dataset.

TABLE 6 OLS and PSM estimates comparing the schooling outcomes of left-behind children with outcomes of children living with their parents

	Standardized math score		Learning anxiety score		Dummy of high learning anxiety (1 = yes; 0 otherwise)	
	OLS	PSM	OLS	PSM	OLS	PSM
	(1)	(2)	(3)	(4)	(5)	(6)
[1] Treatment 8	–0.08	–0.03	–0.05	0.09	0.02	0.06
	(0.07)	(0.10)	(0.19)	(0.27)	(0.03)	(0.05)
Obs.	1084	977	1084	977	1084	977
R-sq	0.06		0.02		0.02	

NOTE: See text for description of Treatment 8.
*significant at 10%; **significant at 5%; ***significant at 1%.
SOURCE: Authors' dataset.

compare migrant students who lived in the city for at least five years (that is, since the beginning of primary school) to migrant students who moved to the city at the beginning of grade 5.

Our findings are reported in Table 5. Students who have attended migrant schools for longer periods of time show deteriorating performance. The estimated achievement gap between students who migrated to the city at least five years ago and those who migrated at the beginning of grade 5 is 0.22 standard deviations and is statistically significant (row 1, column 2). Hence, while the estimated gap is smaller than the coefficients reported in Table 4, this analysis suggests that at least part of the observed gap between migrant students and rural students is due to a school quality effect.

In addition to the three effects that we have focused on so far, we examine the impact of the presence or absence of adult care. Somewhat surprising (given the current scholarship—e.g., Antman 2012; Zhou, Murphy, and Tao 2014), there is little difference between LBC and CLP who attend rural public schools in Anhui (Table 6, row 1, columns 1 and 2). Whether using OLS or PSM, we find no significant differences in math test scores.

Learning anxiety. According to our multivariate analysis, the learning anxiety scores of migrant students are higher than those of LBC (Table 3, row 1, column 3). After holding the control variables constant, the learning anxiety scores of migrant students are 0.70 standard deviations higher than those of LBC when using the full sample. The measured gaps using the matching analysis and the overall sample are slightly wider, 0.86 standard deviations (row 1, column 4). When using the other subsamples of migrant students, the learning anxiety of migrant students is still far higher than that of LBC (0.63 and 0.99 standard deviations—rows 2 and 3, columns 3 and 4). The results are quite similar if we consider the dummy of high learning anxiety as our outcome variable. The coefficient on the treatment dummy is significantly positive, ranging from 0.09 to 0.13 percentage points, suggesting that migrant students are more likely to have high levels of learning anxiety (columns 5 and 6).

To assess the school effect on learning anxiety, Table 4 compares the learning anxiety of migrant students and CLP. When both groups of students live with both parents, the learning anxiety scores of migrant students are even higher than those of CLP. When using OLS, the gap is 0.69 standard deviations (row 1, column 3); when using PSM, the gap is 0.50 standard deviations, but this difference is not significant (row 1, column 4). Similar results are found when comparing migrants from Anhui (or the three core study areas in Anhui) with CLP (rows 2 and 3, columns 3 and 4). The fact that the gap in learning anxiety scores is wide even when we control for the living arrangements of children suggests that some characteristics of migrant communities are contributing to the higher levels of anxiety among these children. It is reasonable to conclude that the reason for this gap in levels of learning anxiety is the same as the reason for the gap in math test scores: the quality of schools in migrant areas is lower than that in rural areas.

Conclusion

We have sought to determine the effects of parents' rural-to-urban migration on math scores and learning anxiety among grade 5 students in urban and rural China. We compared three groups of students: migrants, LBC, and CLP. Our results show that with either LBC or CLP as the control group, migrating with parents and attending private migrant schools in urban areas has a large and significant negative effect on students' math scores. Migrant students also exhibit higher levels of learning anxiety. These findings differ from those in Xu and Xie (2015), the most relevant study in the literature. The authors found that migrant students are preforming better in math than their peers remaining in the countryside. After considering the nature of the samples in the two studies, we believe that the more plausible reason for the differences in results is that most migrant children in the Xu and Xie study are living with parents who are well integrated in the urban

communities and, as a result, their children are attending higher-quality public urban schools. In contrast, our sample is fully drawn from schools that cater to migrants who have been unable to enroll their children in public urban schools and instead attend low-quality private schools.

Self-selection may also be at work, as migrant parents whose children are better students are more successful in enrolling them in urban public schools. Lai et al. (2014) found that migrant students in public schools perform much better than migrant students in migrant schools and are doing as well as Beijing resident students in public schools. With CLP as the reference group, parental migration has no statistically significant effects on either the math or the learning anxiety scores of LBC, a result consistent with Xu and Xie (2015). We also show that there are differences in levels of learning anxiety between migrant students and students attending rural schools.

Our results suggest that it is not parental migration per se that negatively influences cognitive and non-cognitive outcomes of students; rather attendance at private migrant schools in the city and living in an urban environment appear to be the main sources of negative impacts on migrant children. Although we cannot distinguish fully between the magnitudes of the effect of school quality and the effect of the migrant living environment, our analysis strongly suggests that at least part of the gap between migrant students and rural students is due to the poor quality of private schools in migrant communities.

Migrant schools have been unsuccessful in providing quality education to migrant children. One step that could be taken to improve the academic performance of migrant students is to expand access to urban public schools. Doing so would provide public education to all children living in China's cities, regardless of *hukou* status. Although this is an expensive proposition, providing migrant students with a high-quality education should be a priority. Expanding access to education in this manner offers an opportunity to raise the human capital of millions of students, who one day will become key players in China's labor market. Even if China makes a policy decision to provide public education for all migrant students, however, change will not happen overnight. In the meantime policymakers should also make increased investment in migrant schools to improve teacher quality and the overall quality of education.

Our results may also point to problems related to the living environment in migrant communities. Even if children were allowed to attend urban public schools, there might be negative effects associated with being removed from the familiarity of one's home community and placed in a new living environment. Another factor that may account for the negative effect on migrant students is the poor living and working conditions of migrant parents. For example, migrants lack social protection for their jobs and families (Wong at al. 2007; Huang and Yi 2015). If

migrant parents believe they are being marginalized, this could negatively affect learning and the subjective well-being of migrant students (Wong, Chang, and He 2009). If public opinion and discrimination are part of the reason for this (Yiu 2016), public relations and community awareness campaigns may need to be undertaken to eliminate or soften such public perceptions.

Notes

Corresponding author is Yu Bai: Yubai@snnu.edu.cn. The authors acknowledge funding support from the National Natural Science Foundation of China (71373255; 71333012) and the Chinese Academy of Sciences (KSZD-EW-Z-021-1). The authors are also supported by the 111 Project (Grant Number B16031).

1 The majority of rural migrants in cities lack a local household registration (*hukou*), thus are excluded from full access to pension, health care, public education, and other social benefits at the place they live. Migrant children who do not possess *hukou* are rarely permitted to attend public schools.

2 Details of the Trends in International Mathematics and Science Study can be found at http://timss.bc.edu/home/pdf/TP_About.pdf.

3 To implement the matching estimator, we follow 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 between the treatment and comparison groups of students. Intuitively, wide common support means there is a fairly large overlap in the propensity scores. Because the common support is fairly wide in our sample (see Appendix Figure A1), we can estimate the average treatment effect for the treated of a large portion of the sample. (Appendix is available at the supporting information tab at wileyonlinelibrary.comjournalpdr.)

Second, we use nearest neighbor matching with replacement. The standard errors are bootstrapped using 1,000 replications in order to assess the matching quality. Since we do not condition on all covariates but on the propensity score alone in PSM, we check whether the matching procedure is able to balance the distribution of the relevant covariates in both the comparison and treatment groups. We use balance tests described in Dehejia and Wahba (1999, 2002). The balancing tests were satisfied for all covariates. To guard against a potential source of bias (Abadie and Imbens 2002) we also implemented the bias-corrected matching (BCM) estimator developed by Abadie and Imbens (2006). To minimize geographic mismatch, we enforce exact matching by county. Each treatment observation is matched to three control observations with replacement, which is few enough to permit exact matching by county for nearly all observations but enough to reduce the asymptotic efficiency loss significantly (Abadie and Imbens 2006). Matching is based on a set of six time-invariant covariates: sex, age, whether the student is an only child, father has completed junior high school or above, mother has completed junior high or above, and household asset value. The descriptive statistics of the independent variables by students in the migrant schools and in the rural public schools are presented in Table A2. The weighting matrix uses the Mahalanobis metric, which is the inverse of the sample covariance matrix of the matching variables.

4 The occupation of the migrant parents may affect the decision to take children to the city. Including the occupation might help us facilitate our understanding through a selection effect. Unfortunately, we do not have information on the occupations of the parents in our survey.

5 Appendix is available at the supporting information tab at wileyonlinelibrary.comjournalpdr.

References

Abadie, A. and G.W. Imbens. 2002. "Simple and bias-corrected matching estimators for average treatment effects," NBER Technical Working Paper No. 283.

———. 2006. "Large sample properties of matching estimators for average treatment effects," *Econometrica* 74(1): 235–267.

All-China Women's Federation (ACWF). 2013. "Research report into the situation of rural left behind children and rural to urban migrant children in China."

Antman, F.M. 2012. "Gender, educational attainment, and the impact of parental migration on children left behind," *Journal of Population Economics* 25: 1187–1214.

Bai, Y. et al. 2017. "Effect of parental migration on the academic performance of left behind children in North Western China," *The Journal of Development Studies*, pp. 1–17.

Caliendo, M. and S. Kopeinig. 2008. "Some practical guidance for the implementation of propensity score matching," *Journal of Economic Surveys* 22(1): 31–72.

Chen, Y. and S. Feng. 2013. "Access to public schools and the education of migrant children in China," *China Economic Review* 26: 75–88.

Chen, Z., W.E. Huffman, and S. Rozelle. 2009. "Farm technology and technical efficiency: Evidence from four regions in China," *China Economic Review* 20(2): 153–161.

Dehejia, R.H. and S. Wahba. 1999. "Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs," *Journal of the American Statistical Association* 94(448): 1053–1062.

———. 2002. "Propensity score-matching methods for nonexperimental causal studies," *Review of Economics and Statistics*, 84(1): 151–161.

Deng, W., L. Lei, and B. Cao. 2002. "A survey on mental health of urban middle-school students in Jiangxi Province," *Health Psychology Journal* 10(4): 292–293.

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

Gibson, J. and D. McKenzie. 2012. "The economic consequences of 'brain drain' of the best and brightest: Microeconomic evidence from five countries," *Economic Journal* 122(560): 339–375.

Goodburn, C. 2009. "Learning from migrant education: A case study of the schooling of rural migrant children in Beijing," *International Journal of Educational Development* 29(5): 495–504.

Hanushek, E.A., J.F. Kain, and S.G. Rivkin. 2004. "Disruption versus Tiebout improvement: The costs and benefits of switching schools," *Journal of Public Economics* 88(9): 1721–1746.

Huang, Y. and C. Yi. 2015. "Invisible migrant enclaves in Chinese cities: Underground living in Beijing, China," *Urban Studies* 52(15): 2948–2973.

Lahaie, C., J.A. Hayes, T.M. Piper, and J. Heymann. 2009. "Work and family divided across borders: The impact of parental migration on Mexican children in transnational families," *Community, Work & Family* 12: 299–312.

Lai, F., E. Sadoulet, and A.D. Janvry. 2011. "The contributions of school quality and teacher qualifications to student performance: Evidence from a natural experiment in Beijing middle schools," *Journal of Human Resources* 46(1): 123–153.

Lai, F. et al. 2014. "The education of China's migrant children: The missing link in China's education system," *International Journal of Educational Development* 37: 68–77.

Ma, X. et al. 2016. "Improving learning by improving vision: evidence from two randomized controlled trials of providing vision care in China," REAP Working Paper.

NBSC (National Bureau of Statistics of China). 2014. *China Statistical Yearbook*. Beijing: China Statistic Press (in Chinese).

———. 2015. *China Statistical Yearbook*. Beijing: China Statistic Press (in Chinese).

Rosenbaum, P. R. and D.B. Rubin. 1983. "The central role of the propensity score in observational studies for causal effects," *Biometrika* 70(1): 41–55.

———. 1985. "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score," *The American Statistician* 39(1): 33–38.

Shi, Y., Y. Bai, Y. Shen, K. Kenny, and S. Rozelle. 2016. "Effects of parental migration on mental health of left-behind children: Evidence from northwestern China," *China and World Economy* 24(3): 105–122.

Spera, C. 2005. "A review of the relationship among parenting practices, parenting styles, and adolescent school achievement," *Educational Psychology Review* 17(2): 125–146.

Taylor, J.E., S. Rozelle, and A. de Brauw. 2003. "Migration and incomes in source communities: a new economics of migration perspective from China," *Economic Development and Cultural Change* 52(1): 75–101.

Wang, H., C. Yang, F. He, Y. Shi, Q. Qu, S. Rozelle, and J. Chu. 2015. "Mental health and drop out behavior: A cross-sectional study of junior high students in Northwest rural China," *International Journal of Education Development* 41: 1–12.

Wang, X., R. Luo, L. Zhang, and S. Rozelle. 2017. "The education gap of China's migrant children and rural counterparts," *Journal of Development Studies*, 1–17.

Wang, X., F. Yamauchi, K. Otsuka, and J. Huang. 2016. "Wage growth, landholding and mechanization in Chinese agriculture," *World Development* 86: 30–45.

Wong, F.K.D., Y.L. Chang, and X.S. He. 2009. "Correlates of psychological wellbeing of children of migrant workers in Shanghai, China," *Social Psychiatry and Psychiatric Epidemiology* 44(10): 815–824.

Wong, K., D. Fu, C.Y. Li, and H.X. Song. 2007. "Rural migrant workers in urban China: Living a marginalised life," *International Journal of Social Welfare* 16(1): 32–40.

Xu, H. and Y. Xie, 2015. "The causal effects of rural-to-urban migration on children's well-being in China," *European Sociological Review* 31(4): 502–519.

Yiu, L. 2016. "The dilemma of care: A theory and praxis of citizenship-based care for China's rural migrant youth," *Harvard Educational Review* 86(2): 261–288.

Zhang, H., J.R. Behrman, C. Fan, X. Wei, and J. Zhang. 2014. "Does parental absence reduce cognitive achievements? Evidence from rural China," *Journal of Development Economics* 111: 181–195.

Zhao, X. et al. 2014. "Left-behind children in rural China experience higher levels of anxiety and poorer living conditions," *Acta Paediatrica* 103(6): 665–670.

Zhou, C. et al. 2015. "China's left-behind children: Impact of parental migration on health, nutrition, and educational outcomes," *Health Affairs* 34(11): 1964–1971.

Zhou, M., R. Murphy, and R. Tao. 2014. "Effects of parents' migration on the education of children left behind in rural China," *Population and Development Review* 40: 273–292.