LEI WANG, WENBIN MIN, SIQI ZHANG, YAOJIANG SHI, AND SCOTT ROZELLE

Math and Chinese-Language Learning

Where Are China's Vulnerable Subpopulations?

ABSTRACT

This paper seeks to understand the learning outcomes that prevail across key subpopulations in China today. Data from a nationally representative survey show that rural youth are two years behind urban children in math and Chinese. Non-Han minorities, children in poorer counties, and children with less-educated parents are the most vulnerable.

KEYWORDS: China, learning outcomes, inequality, math, Chinese language

INTRODUCTION

After 30 years of rapid economic growth, China has entered a period of economic slowdown.[1] In the last three decades, China's economic growth rate increased by an average of more than 10% per year, but in the past five

LEI WANG is an Associate Professor in the International Business School at Shaanxi Normal University, Xi'an, Shaanxi, China. She would like to acknowledge the support of the Gates Foundation's Healthy Birth, Growth and Development program and the 111Project (grant number B16031). The input of Shasha Jumbe has been invaluable. Lei Wang also thanks Professor Xie Yu and his staff at the office of the China Family Panel Survey for their support. WENBIN MIN is a Ph.D. Candidate in the Center for Experimental Economics in Education, Shaanxi Normal University, Xi'an. SIQI ZHANG is a Ph.D. Candidate in the Center for Experimental Economics in Education, Shaanxi Normal University, Xi'an. YAOJIANG SHI is a Professor at the Center for Experimental Economics in Education, Shaanxi Normal University, Xi'an. SCOTT ROZELLE is a Professor in the Rural Education Action Project, Stanford University, Stanford, CA, USA. Emails: <wangleiml@163.com>, <minwenbincee@163.com>, <zhangsiqiceee@163.com>, <shiyaojiang7@gmail.com>, <rozelle@stanford.edu>.

1. Barry Eichengreen, Donghyun Park, and Kwanho Shin, "When Fast-Growing Economies Slow Down: International Evidence and Implications for China," *Asian Economic Papers* 11:1 (2012): 42–87.

Asian Survey, Vol. 58, Number 5, pp. 797–821. ISSN 0004-4687, electronic ISSN 1533-838X. DOI: https://doi.org/10.1525/AS.2018.58.5.797.

years that rate has fallen by more than four percentage points, according to the *China Statistical Yearbook 1979–2015*.[2] This trend has aroused the concern that, should growth continue to slow, the nation might fall into the middle-income trap.[3] This term refers to the experience of a country whose average income level falls too low to sustain transformative economic development, resulting in slow economic growth or stagnation.[4]

One of the most basic steps a country can take to avoid economic stagnation is to encourage human capital development.[5] The experiences of countries that have successfully avoided the middle-income trap demonstrate that giving large shares of the labor force the skills and abilities necessary for innovation and productivity can help prevent economic slowdown.[6] For example, South Korea's transition from a middle- to a high-income country in the 1980s and 1990s is thought to be due partly to its high rate of investment in education and health.[7] Investing in education—especially upper secondary education[8]—seems to be particularly key to increasing human capital. Investments in education can raise labor force productivity, improve general welfare, and foster economic growth.[9]

Unfortunately, according to Khor et al., only 24% of the Chinese labor force had graduated from upper secondary school, and less than 10% had

2. National Bureau of Statistics, People's Republic of China, *China Statistical Yearbook*, 1999–2016, <http://www.stats.gov.cn/tjsj/ndsj/>, accessed August 19, 2018; National Bureau of Statistics, People's Republic of China, *China Statistical Yearbook* (Beijing: China Statistics Press, 1979–1998).

3. Barry Eichengreen, Donghyun Park, and Kwanho Shin, "Growth Slowdowns Redux: New Evidence on the Middle Income Trap," Paper no. w18673, National Bureau of Economic Research, Cambridge, MA, 2013; Cai Fang and M. Y. Wang, "China's Income Gap and Its Risk of Falling into Middle-Income Trap," *Journal of the Renmin University of China* 3:3 (2014): 2–7.

4. I. S. Gill and H. Kharas, "An East Asian Renaissance: Ideas for Economic Growth," World Bank, Washington, DC, 2017: 17–18.

5. Zhang Linzie, Hongmei Yi, Renfu Luo, Chengfang Liu, and Scott Rozelle, "The Human Capital Roots of the Middle Income Trap: The Case of China," *Agricultural Economics* 44:1 (2013): 151–62.

6. Yilmaz, Gökhan, "Turkish Middle Income Trap and Less Skilled Human Capital," *Iktisat Isletme ve Finans* 30:346 (2015): 9–36.

7. Eichengreen et al., "Growth Slowdowns Redux."

8. Niny Khor, Lihua Pang, Chengfang Liu, Fang Chang, Di Mo, Prashant Loyalka, and Scott Rozelle, "China's Looming Human Capital Crisis: Upper Secondary Educational Attainment Rates and the Middle-Income Trap," *China Quarterly* 228 (2016): 905–26.

9. Yazid Dissou, Selma Didic, and Tatsiana Yakautsava, "Government Spending on Education, Human Capital Accumulation, and Growth," *Economic Modeling* 58 (2016): 9–21; Nikos Benos and Stelios Karagiannis, "Do Education Quality and Spillovers Matter? Evidence on Human Capital and Productivity in Greece," *Economic Modeling* 54 (2016): 563–73.

graduated from college, as of 2010.[10] This upper secondary education attainment rate is less than a third of the average rate of OECD countries and substantially lower than even the OECD countries with the lowest rates, Mexico (36%) and Turkey (31%). China's rates are also relatively low for younger cohorts in the labor force, for example 25-to-34-year-olds: the rate for workers of this age range in China is 36%, whereas in Brazil it is 53%, and in Russia, 91%. Since education attainment is strongly associated with learning ability,[11] it is useful for China to measure the levels of and the rates of growth of children's learning outcomes.

Greater investments in education improve education attainment, and the question China now faces is how to target these investments to best improve its human capital. Unfortunately, and perhaps surprisingly, there is little research that can help us answer this critical question. To our knowledge there are no nationally representative, published studies that use standardized instruments to examine differences in human capital among subpopulations or to identify vulnerable (or thriving) groups.

There are, however, papers that report the learning outcomes of students in certain subpopulations. Many of these papers present conflicting findings. Lai et al. found that students in urban public schools outperformed those in migrant schools.[12] Migrant children, in turn, outperformed children attending rural public schools. But Wang et al. found the opposite: students in a rural public school sample from a single province in Central China outperformed migrant children attending school in suburban communities of Shanghai and Suzhou.[13]

Also contradictory are the findings of papers that compare the learning outcomes of left-behind children (LBCs, or children who are left with a relative in their rural home while their parents migrate to work and live in the city) and children living with their parents (CLPs) in rural areas. Zhang et al. found that LBCs performed worse than CLPs when assessed using either

10. Khor et al., "China's Looming Human Capital Crisis."

11. Richard J. Herrnstein and Charles Murray, *The Bell Curve* (New York: Free Press, 1994); Isaac M. Marks, *Fears and Phobias* (Cambridge, MA: Academic Press, 2013).

12. Fang Lai, Chengfang Liu, Renfu Luo, Linxiu Zhang, Xiaochen Ma, Y. Bai, B. Sharbono, and S. Rozelle, "Private Migrant Schools or Rural/Urban Public Schools: Where Should China Educate Its Migrant Children?" Working Paper, Rural Education Action Program, Stanford University, 2011.

13. Xiaobing Wang, Renfu Luo, Linxiu Zhang, and Scott Rozelle, "The Education Gap of China's Migrant Children and Rural Counterparts," *Journal of Development Studies* 53:1 (2017): 1–17.

standardized math or Chinese-language tests.[14] But Bai et al. and Zhou et al. found that LBCs outperform CLPs.[15] And Hu and Li found no difference between the learning outcomes of LBCs and CLPs.[16]

At the same time, there is a dearth of papers that examine key subpopulations at the national level. Although papers looking at the educational attainment of rural students compared to urban students would lead one to believe that wide gaps exist between them in learning outcomes, to our knowledge there are no published papers that measure this gap using nationally representative data.[17] The same is true for other subpopulations, such as Han and non-Han minorities (although there are papers that examine the Han–minority learning gap for small, geographically focused subpopulations).[18]

To date, there have been only three studies that used a nationally representative sample to examine learning outcomes. All three used the first round of the China Family Panel Survey (CFPS), from 2010.[19] None made an effort to determine which subpopulations are likely to have better or worse learning outcomes. And no work has yet used the most recent wave of CFPS data, collected in 2014, or taken advantage of the panel nature of the CFPS data to examine actual differences in the rates of learning among

14. Hongliang Zhang, Jere R. Behrman, C. Simon Fan, Xiangdong Wei, and Junsen Zhang, "Does Parental Absence Reduce Cognitive Achievements? Evidence from Rural China," *Journal of Development Economics* 111 (2014): 181–95.

15. Yu Bai, Linxiu Zhang, Chengfang Liu, Yaojiang Shi, Di Mo, and S. Rozelle, "Effect of Parental Migration on the Academic Performance of Left Behind Children in North Western China," *Journal of Development Studies* 53:6 (2017): 1–17; Chengchao Zhou, Sean Sylvia, Linxiu Zhang, Renfu Luo, Hongmei Yi, Chengfang Liu, Yaojiang Shi, Prashant Loyalka, James Chu, Alexis Medina, and Scott Rozelle, "China's Left-Behind Children: Impact of Parental Migration on Health, Nutrition, and Educational Outcomes," *Health Affairs* 34:11 (2015): 1964–71.

16. F. Hu and S. Li, "Effect of Parental Migration on Academic Performance among Left Behind Children in Rural China," *Management World* 2 (2009): 67–74.

17. Hongbing Li, Prashant Loyalka, Scott Rozelle, Binzhen Wu, and Jieyu Xie, "Unequal Access to College in China: How Far Have Poor, Rural Students Been Left Behind?" *China Quarterly* 221 (2015): 185–207.

18. Yunfan Yang, Huan Wang, Linxiu Zhang, Sean Sylvia, Renfu Luo, Yaojiang Shi, Wei Wang, and Scott Rozelle, "The Han-Minority Achievement Gap, Language, and Returns to Schools in Rural China," *Economic Development and Cultural Change* 63:2 (2014): 319–59.

19. Dandan Zhang, Xin Li, and Jinjun Xue, "Education Inequality between Rural and Urban Areas of the People's Republic of China, Migrants' Children Education, and Some Implications," *Asian Development Review* 32:1 (2015): 196–224; Guoying Huang, Yu Xie, and Hongwei Xu, "Cognitive Ability: Social Correlates and Consequences in Contemporary China," *Chinese Sociological Review* 47:4 (2015): 287–313; Hongwei Xu, and Yu Xie, "The Causal Effects of Rural-to-Urban Migration on Children's Well-Being in China," *European Sociological Review* 31:4 (2015): 502–19.

subpopulations (which can be used to create a measure of the seriousness of the differences among subpopulations). These nationally representative data are crucial for determining where in Chinese society learning outcomes are deficient today and developing policy solutions to help guide the investment decisions of policymakers.

The overall goal of this paper is to understand the pattern of math and Chinese-language learning outcomes that prevail across subpopulations in China today, with the purpose of understanding where levels of human capital are lower or higher. To do this, we use math and Chinese-language test scores as measures of children's learning outcomes. Specifically, we compare the outcomes of urban and rural children and try to identify the sources of the measured urban–rural gaps. We also seek to understand the differences in outcomes between major subpopulations in the rural (urban) economy, including differences between Han and non-Han-minority children; children in poor and non-poor counties; children in families with parents that have high and low levels of education; boys and girls; LBCs and CLPs; and children living in different parts of China (north vs. south; and eastern vs. central vs. western). Using our findings, we will discuss various investment strategies that China could use to improve its human capital in the future. In the next section we describe the data and our sampling approach. The subsequent section presents our results. The final section summarizes and draws conclusions.

DATA, MEASURES, AND APPROACH

Data

This study's data come from the first and third waves of the CFPS, conducted in 2010 and 2014.[20] The CFPS is a nationally representative, longitudinal

20. The CFPS data we used cover 25 provinces in China, representing 95% of the total population in Mainland China. To evaluate the representativeness of the sample for the population, Yu and Hu (see end of this note) compared the age-sex structure using the 2010 CFPS data and the 2010 census data. They found that the two population pyramids are almost identical. Xu and Yu ("Causal Effects") also compared the 2010 CFPS with the 2010 census. They found that the distributions of age, sex, rural–urban stratification, educational attainment, and marital status in the CFPS closely resemble those in the census. Finally, we compared the gender structure of children aged 10 to 15 using the 2014 CFPS data and the China Statistical Yearbook (2015 data). The structure of sex in the 2014 CFPS is almost the same as the data in the 2015 China Statistical Yearbook. Yu Xie and Jingwei Hu, "An Introduction to the China Family Panel Studies (CFPS)," *Chinese Sociological Review* 47:1 (2014): 3–29.

social survey that collects data at three levels: individual, family, and community. The CFPS includes several sections related to education outcomes and family dynamics. The survey is administered biennially by the Institute of Social Science Survey of Peking University in Beijing.

In both 2010 and 2014, the CFPS surveyed respondents in 25 of the 31 administrative divisions of mainland China (excluding Xinjiang, Tibet, Qinghai, Inner Mongolia, Ningxia, and Hainan). The sampling frame represents 95% of China's total population. The first survey, conducted in 2010, included 14,789 families, covering 33,600 adults and 8,990 children. The third survey wave, conducted in 2014, included 13,946 families, covering 37,147 adults and 8,617 children.

Sampling Protocol

The CFPS used a probability-proportional-to-size sampling strategy with multi-stage stratification. In this sampling method, administrative units and socioeconomic status were used as the main stratification variables. Within each administrative unit, local GDP per capita served as the main socioeconomic-status variable. If the GDP per capita of the administrative unit was not available, the proportion of the non-agricultural population or the non-agricultural population density was used instead.

To generate nationally and provincially representative samples, the CFPS used a three-stage sampling process. In the first stage, 144 county-level units were randomly selected. In the second stage, 640 village-level units (villages in rural areas and neighborhoods/communities in urban areas) were selected from among the county-level units. In the third stage, 14,000 households from the village-level units were selected according to the study's systematic sampling protocol. All members of each household were interviewed, except those who were not at home at the time of survey administration.

Sampling Weights

To determine the appropriate weights for the sample, we referred to the sampling design of the CFPS. In this protocol, there are two types of provinces in the sample: "large provinces," which are representative at the regional level, and "small provinces." To make up an overall sampling frame representative of the national population, half of the sample in the CFPS was generated by oversampling from five large provinces (Shanghai, Liaoning, Henan, Gansu, and Guangdong). The other half of the sample was from an independent sampling frame composed of 20 small provinces. The sampling weights of the 20 small provinces are the same (1 for each), and the weights of the five large provinces are 10.28, 4.45, 2.04, 7.30, and 2.02, for Shanghai, Liaoning, Henan, Gansu, and Guangdong, respectively. We take these oversampling weights as the sampling weights for each observation in the data.

Sample of Children

To focus on the learning outcomes of children in this paper, our study focused on a sample of children aged 10 to 15 (inclusive) in the 2010 and 2014 CFPS. All children in this age range were required to complete the child questionnaire independently. Our final study sample included 5,676 children, of which 3,314 were surveyed in 2010 and 2,362 were surveyed in 2014; 687 were surveyed in both 2010 and 2014. Of the 2014 total, 460 children were chosen from the neighborhood/community sampling units in urban areas, and 1,902 children from villages in rural areas.

Using these data, we constructed three different samples, which are used in the rest of the paper. One sample includes all children (aged 10 to 15) in the 2014 sample, including both the urban and rural sampling units, for a total of 2,362 children. We call this the *urban and rural sample*. The two main uses for the data from this sample are to measure the rate of learning over time (when used in conjunction with the 2010 urban and rural sample—see below for more details) and to assess the differences in levels of learning between rural and urban children in 2014.

The second sample (the *rural-only sample*) includes two subsamples of children. It includes all of the children (aged 10 to 15) in the 2014 rural sample (1,902 children from village units of the sample); and it includes the 185 children from the urban-only sample (the third sample—see next paragraph) that have rural *hukou* (household registration/right of residence). In the rest of this paper, we call these children *migrant children*. The third sample (the *urban-only sample*) includes all of the children (aged 10 to 15) that have an urban *hukou*. The 2014 urban-only sample includes 460 children. We include no migrant children in this sample.

To assess changes in learning outcomes (which we use to assess the relative sizes of learning differences among the subsamples in 2014), we supplemented our 2014 data with data from the 2010 CFPS wave. Any 2010 observations we used came from those individuals in the 2014 wave that were also part of the 2010 sample. Necessarily, this means that our 2010/2014 sample only includes those sample children that were 10 and 11 years old in 2010 and 14 and 15 years old in 2014 and were present during both waves of the survey. There are 687 children in this supplemental sample. Using these data, we were able to measure learning over time (separately for math and Chinese). When we put the measured differences in math and Chinese-language scores between 2010 and 2014 into a per-year-of-learning (average annual rate of learning) basis, we are able to better assess which subpopulation gaps are relatively wide (and which we sometimes call "what matters") and which are more narrow (what does not matter or what matters less).

Measures

To evaluate the learning outcomes of children in the sample, we used the results of the math and Chinese-language tests of the sample in both waves of the CFPS.[21] Both tests were created by the CFPS staff and were designed to evaluate the learning of 10-to-15-year-old children.[22] All items in both tests were drawn from standard primary and secondary school curricula and sorted in ascending order by degree of difficulty. In total, there are 24 questions in the math test and 34 questions in the Chinese-language test.

In the Chinese-language test, Chinese characters were presented to the children in the form of flash cards (one character per card) that children read aloud. For the math test, children were required to answer questions in a paper examination. Both tests were administered as follows. The question with which children began depended on their level of education. For example, primary school students (e.g., the 10-and-11-year-old sample) began at the

21. According to Catell, the types of measures used in the 2010 and 2014 CFPS waves can be used to assess the verbal and mathematical achievements of respondents. To produce such measures, all test items were drawn from the standard curricula in primary and secondary schools. Measures such as these represent "crystallized intelligence," which is knowledge acquired through learning, experience, and education. Raymond B. Cattell, *Intelligence: Its Structure, Growth, and Action* (Amsterdam: North-Holland, 1987).

22. The measures of achievement have been widely used in the literature. There are already at least four other published papers that use these two test measures. Huang et al. ("Cognitive Ability") used these two measures as scales of cognitive ability. Zhang et al. ("Education Inequality") analyzed the education inequality between rural and urban areas taking these two test scores as their main measures. Xu and Yu ("Causal Effects") used these two test scores as measures to analyze the causal effects of rural-to-urban migration on children's educational performance and well-being in China. Jordan et al. used these two test scores as language and math achievement outcomes to measure the learning ability of migrant children: Lucy P. Jordan, Qiang Ren, and Jane Falkingham, "Youth Education and Learning in Twenty-First Century China: Disentangling the Impacts of Migration, Residence, and Hukou," *Chinese Sociological Review* 47:1 (2014): 57–83.

beginning of the test; junior high school students (e.g., the 14-and-15-year-old sample) began with the ninth question of the Chinese test and the 13th question on the math test. If the child failed to correctly identify three characters in a row or failed to answer three consecutive math questions, the test ended. The question number of the most difficult question the child answered correctly was assigned as his/her final score.[23]

To measure the rate of improvement in learning outcomes over time for the whole population, we identified all individuals that were in the sample in both 2010 and 2014 and considered the math and Chinese-language scales during both time periods. We used these observations to compare the average change in the scores of the sample, which allowed us to measure the average amount of learning over a four-year period. We thereby determined the average amount of learning over time, expressed in years, and used it as a basis for expressing the width of a learning gap. The formula we used to convert subgroup difference in learning outcomes into "years behind" is:

$$\begin{aligned}&\text{learning points gained per year}\\&\quad= (\text{mean of test scores in } 2014 - \text{mean of test scores in } 2010)/4\text{years}\end{aligned}$$

$$\begin{aligned}&\text{learning gap(years)}\\&\quad= \text{difference in points between two subpopulations / learning points gained per year}\end{aligned}$$

Subpopulations and Control Variables

We analyzed the relationship between slightly different sets of subpopulations in the urban and rural sample, the rural-only sample, and the urban-only sample. We used the urban and rural sample representing all children across China for one key comparison: the differences between urban and rural children. When we looked at the difference among rural children in China, using the rural-only sample, we compared the differences among six subpopulation pairs: Han and non-Han-minority children; children in poor and non-poor counties; children in families with parents (separately for

23. If a child could not correctly answer the first question in their batch of questions (for example, item 9 on the Chinese language test for a junior high school student), then he/she would next be asked to answer the first question in the batch for students in the lower educational level (for example, item 1, where primary school students started). If the student could not answer that item correctly, then the test would end, and the student would receive a final score of zero.

fathers and mothers) that have high and low levels of education; migrant children and the children who live in rural villages; boys and girls; and LBCs and CLPs. We also measured the difference in rural children living in different parts of China (north vs. south; and eastern vs. central vs. western). We used the same pairs to analyze differences among key subpopulations in China's cities using the urban-only sample (with the exception of two subpopulations: we did not analyze the urban-only sample for differences between migrant and rural children or between LBCs and CLPs).

To create these groups (see below for the descriptive analysis) and to generate a set of variables for the multivariate analysis, we used the following definitions. The variable *rural children* takes the value of 1 when the child has a rural *hukou* (regardless of whether the child was in the urban-only or the rural-only sample), and 0 when the child has an urban *hukou*. *Ethnicity* is 1 when the child is a non-Han minority, and 0 when the child is Han. *Children in poor counties* is 1 when the child is from a county in the lowest quartile of the CFPS sample counties when the counties are ranked by GDP per capita, and 0 otherwise. *Father's education* (*mother's education*) is 1 when the father (mother) did not graduate from junior high school; and 0 if the father (mother) at least graduated from junior high school. *Gender* is 1 for boys and 0 for girls.

In the rural-only sample, we use two additional variables. *Children living in rural villages* is 1 when the child lives in a rural village sampling unit, 0 when the child is a migrant. *CLPs* is 1 if the child is living with one or both parents, 0 if the child is an LBC. To be clear, using this definition, in this paper LBCs are children who live in households in which both parents have out-migrated to work in the city.

We also define four control variables: three geographic variables, and one to control for a child's age. *South* (denoting children living in Southern China) is 1 if the child is from one of 14 provinces/municipalities (Anhui, Chongqing, Fujian, Guangdong, Guangxi, Guizhou, Hubei, Hunan, Jiangsu, Jiangxi, Shanghai, Sichuan, Yunnan, or Zhejiang), 0 if the child is from any of the other 11 provinces/municipalities (Beijing, Gansu, Hebei, Heilongjiang, Henan, Jilin, Liaoning, Shandong, Shaanxi, Shanxi, or Tianjin). *Central* is 1 if the child is from one of eight provinces (Anhui, Heilongjiang, Henan, Hubei, Hunan, Jiangxi, Jilin, or Shanxi). *West* is 1 if the child is from one of seven provinces (Chongqing, Gansu, Guangxi, Guizhou, Shaanxi, Sichuan, or Yunnan). *Age* is the age of the child. In our sample, it was set to the age of the child in 2014, at the time of survey. The range is 10 to 15, inclusive.

Analytical Approach

To examine the relationship between subpopulations, we compared the scores of each subpopulation (e.g., Han vs. non-Han-minority children, or boys vs. girls) for the math and Chinese-language tests separately. Tables 1 and 2 describe gaps among the different subpopulations in math and Chinese-language test scores. To help the reader determine when the two groups differ in a meaningful way, we distinguish in each table between gaps that "matter" (more than one year of learning, as defined by average annual rate of learning) and gaps that do not (less than one year of learning). We use all three samples in the descriptive analysis (the urban and rural sample for Figures 1 and 2; the rural-only sample for Tables 1 and 4; and the urban-only sample for Table 2).

For both the 2014 urban and rural sample and the 2014 rural-only sample we also examined the conditional statistical relationship (or *coefficient* measuring the relationship) between each of the subpopulation and control variables and the outcome measures. The estimates of the coefficients were produced using ordinary least squares regression analysis. In the regression using the urban and rural sample to estimate the coefficients of the relationships between math and Chinese-language test scores and the subpopulation variables (Model 1, Table 3), we included rural children as the independent variable of interest (as we want to use this analysis to understand the difference between rural and urban children, holding control variables constant). We use nine control variables: ethnicity, children in poor counties, father's education, mother's education, gender, three variables to control for geographic location (northern, central, and western), and age.[24]

In the regression using the rural-only sample to estimate the coefficients of the relationships between test scores and the subpopulation variables (Model 2), we include six subpopulation variables: ethnicity, children from poor counties, father's education, mother's education, gender, and CLPs; and we include four control variables: north, central, western, and age. The basic model we used in both Model 1 and Model 2 is

$$\text{math (Chinese-language) score}_i = \beta_0 + \beta_1(\text{subpopulations}) + X_i\delta + \varepsilon_i$$

24. We do not need to run the regressions without control variables, because the results would be same as the results that are currently in Table 1 of this manuscript.

TABLE 1. Differences (and Gaps) in Test Scores among Subpopulation Pairs of Rural Children in the Rural-Only Sample, 2014

Subpopulation pairs	*Math test scores (points)*	*Gap (years)*	*Chinese-language test scores (points)*	*Gap (years)*
What Matters[a]				
Non-Han minority	8.76	1.1***	18.03	1.7***
Han	10.07		20.76	
Children in poor counties	9.22	1.3***	19.75	1.2***
Children in non-poor counties	10.75		21.62	
Father's education (less than jr. high)	9.43	1.0***	20.08	0.8***
Father's education (jr. high and above)	10.64		21.36	
Mother's education (less than jr. high)	9.64	1.0***	20.12	1.4***
Mother's education level (jr. high and above)	10.84		22.37	
Children living in rural villages	9.75	2.1***	20.41	1.7***
Migrant children (rural hukou)	12.16		23.15	
What Does Not Matter[b]				
Boys	9.92	0.03	20.19	0.6
Girls[c]	9.95		21.10	
Children living with parents (CLPs)	9.73	0.2	20.27	0.1
Left-behind children (LBCs)	10.01		20.45	
South	10.08	0.3	20.24	0.3
North	9.88		20.74	
West[d]	9.59	0.9	20.32	0.7
East	10.72		21.49	
Central[d]	9.86	0.7	20.30	0.7
East	10.72		21.49	

***$p < 0.01$; **$p < 0.05$.

SOURCE: By authors

NOTE: Data from the China Family Panel Survey. We assign sampling weights for each observation. The weight for provinces Shanghai, Liaoning, Henan, Gansu, and Guangdong are 10.28, 4.45, 2.04, 7.30, and 2.02, respectively; for the rest of the provinces, the weights are all 1.

[a]That is, the gaps between the test scores of children in the different subpopulations are greater than one year of learning.

[b]That is, the gaps between the test scores of children in the different subpopulations are less than one year of learning.

[c]We did not indicate that there is a statistical difference between boys and girls because the gaps are less than one year of learning.

[d]Although statistically significant, the gaps in the test scores of children in western, central and eastern regions are less than one year of learning, so they "do not matter."

TABLE 2. Differences (and Gaps) in Test Scores among Subpopulation Pairs of Urban Children in the Urban-Only Sample, 2014

Subpopulation pairs	*Math test scores (points)*	*Gap (years)*	*Chinese-language test scores (points)*	*Gap (years)*
What Matters[a]				
Mother's education (less than jr. high)	11.89	0.6***	23.16	1.1***
Mother's education (jr. high and above)	12.57		24.99	
Boys	12.04	0.6	23.83	1.0**
Girls	12.78		25.51	
What Does Not Matter[b]				
Non-Han minority	12.90	0.5	25.13	0.2
Han	12.28		24.75	
Children in poor counties	11.80	0.7	23.91	0.6
Children in non-poor counties	12.58		24.90	
Father's education (less than jr. high)	11.72	0.7	23.65	0.7
Father's education (jr. high and above)	12.56		24.86	
South	12.79	0.6	24.47	0.3
North	12.09		24.94	
West	11.53	1.1	24.88	0.03
East	12.76		24.82	
Central	12.45	0.3	24.17	0.4
East	12.76		24.82	

***$p < 0.01$; **$p < 0.05$.

SOURCE: By authors

NOTE: Data from the China Family Panel Survey. We assign sampling weights for each observation. Specifically, the weight for provinces Shanghai, Liaoning, Henan, Gansu, and Guangdong are 10.28, 4.45, 2.04, 7.30 and 2.02, respectively; the rest of the provinces, the weights are all 1.

[a]That is, the gaps between the test scores of children in the different subpopulations are greater than one year of learning.

[b]That is, the gaps between the test scores of children in the different subpopulations are less than one year of learning.

where the dependent variable, math (Chinese-language) score$_i$, indicates the test scores of child i. The subpopulations variable represents different subgroups of interest. X_i is a vector of covariates that are included to capture the characteristics of children and their households. We account for the clustering at the district level.

TABLE 3. Correlations between Tests Scores of Urban and Rural Children in the Urban and Rural Sample in 2014

	Math test score (1)	*Chinese-language test score* (2)
Panel A: Descriptive Statistics[a]		
Rural children	–2.50***	–4.10***
	(0.35)	(0.46)
Panel B: Correlations		
Subpopulation variables		
Rural children	–1.24***	–2.54***
	(0.30)	(0.43)
Controls included[b]	yes	yes
Constant	–1.79**	3.85**
	(0.87)	(1.34)
N	2182	2182
R^2	0.29	0.28

***$p < 0.01$; **$p < 0.05$.

SOURCE: By authors

NOTE: Data from the China Family Panel Survey. Standard deviation in parentheses.

[a]The descriptive statistics in this table (row 1, cols. 1 and 2) are the same the figures shown in Figure 2.

[b]Control variables include ethnicity, children in poor counties, father's education, mother's education, gender, children in southern counties, children in central counties, children in western counties, and age. We also control for district fixed effects. In all regressions, we also assign sampling weights for each observation. The weights for provinces Shanghai, Liaoning, Henan, Gansu, and Guangdong are 10.28, 4.45, 2.04, 7.30, and 2.02, respectively; for the rest of the provinces, the weights are all 1.

RESULTS

Measuring Changes in Math and Chinese-Language Scales over Time

Before we examined the gaps in learning outcomes among subpopulations in our sample, our first step was to measure the rate of improvement in learning outcomes over time in the population as a whole. Examining how the test scores of the entire sample changed over time, we saw a marked shift in the distribution. The average change in the math scores of children that were 10/11 years old in 2010 and 14/15 years old in 2014 was an increase of 4.65 points over 4 years (Figure 1, Panel A). The average change in Chinese-language scores between 2010 and 2014 was an increase of 6.46 points (Figure 1, Panel B). Using these results, we created a metric of the average change in learning per year. Assuming

TABLE 4. Correlations (Generated by OLS Regression Analysis) between Tests Scores of Rural Children and Selected Children Characteristics in the Rural-Only Sample, 2014

	Math test score (1)	*Chinese-language test score* (2)
Subpopulation variable		
Ethnicity (1 = non-Han minority)	–0.96***	–2.24***
	(0.37)	(0.78)
Children in poor counties[a]	–1.05***	–1.19**
	(0.30)	(0.47)
Father's education	–0.96***	–0.92**
(1 = less than jr. high)	(0.29)	(0.43)
Mother's education	–0.45	–1.69***
(1 = less than jr. high)	(0.35)	(0.51)
Gender	–0.00	–0.96**
(1 = boy)	(0.27)	(0.42)
CLPs [b]	−0.30	0.24
(1 = yes, 0 = LBCs)[b]	(0.44)	(0.73)
Control variables		
South	–0.33	–0.91**
	(0.22)	(0.40)
Central	–0.50*	0.77
	(0.29)	(0.51)
West	–0.16	0.35
	(0.31)	(0.49)
Age	1.00***	1.79***
(years)	(0.08)	(0.13)
Constant	–0.65	1.16
	(1.10)	(1.74)
N	1588	1588
R^2	0.20	0.22

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

SOURCE: By authors

NOTE: Data from the China Family Panel Survey. Standard deviation in parentheses. In all regressions, we control for district fixed effects. We also assign a sampling weight for each observation. The weights for provinces Shanghai, Liaoning, Henan, Gansu, and Guangdong are 10.28, 4.45, 2.04, 7.30, and 2.02, respectively; the rest of provinces are each weighted 1.

[a]Children from counties in the lowest quartile of the CFPS sample counties when the counties are ranked by GDP per capita.

[b]CLPs are children are living with their parents in rural villages; LBCs are children left behind in rural villages while their parents have out-migrated.

the change in the scores is linear over time, we can say that the average gain in learning per year for the sample is 1.16 points for math and 1.62 points for Chinese language. In the rest of the study we call this the *average annual rate of learning*.

Distribution of Learning Outcomes in the Urban and Rural Sample

The distribution of learning outcomes across the nation is uneven when comparing the tests scores (in both math and Chinese language) of urban children aged 10 to 15 with those of rural children the same age (Figure 2). On average, children in urban areas performed significantly better than those in rural areas on both tests. The overall gaps between the urban and rural sample groups in both test scores were wide: 2.5 points ($p < 0.01$) for the math test and 4.1 points ($p < 0.01$) for the Chinese-language test. According to our average annual rate of learning metric (calculated and discussed above), the average urban child in our sample is more than two years ahead of the average rural child in both math and Chinese language.

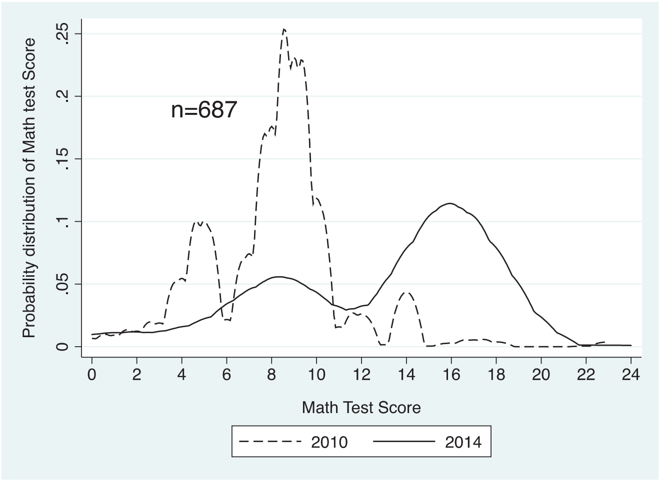
Using the urban and rural sample (Model 1), the results of the OLS analysis (Table 3) are largely consistent with the descriptive results (Figure 2). Holding all the other subpopulation and control variables constant, the math test scores of rural children are lower than those of urban children, with a coefficient of –1.24. In other words, conditional on all other factors, rural children score 1.24 points lower than urban children in math learning. The gap in Chinese-language test scores is also big: 4.1 points. Being rural, regardless of other characteristics, is enough to make a child fall more than a year behind an urban child in both math and Chinese.

Because of these large, fundamental differences between the rural and urban samples, in the subsequent analysis we look at differences between the subpopulations within the rural-only and urban-only subsamples separately. This splitting of the sample is necessary because the rural-only sample is so much larger than the urban-only sample. If we examined the gaps between subpopulations for all of China (using the urban and rural sample), the gap between the rural subpopulations might be expected to dominate.

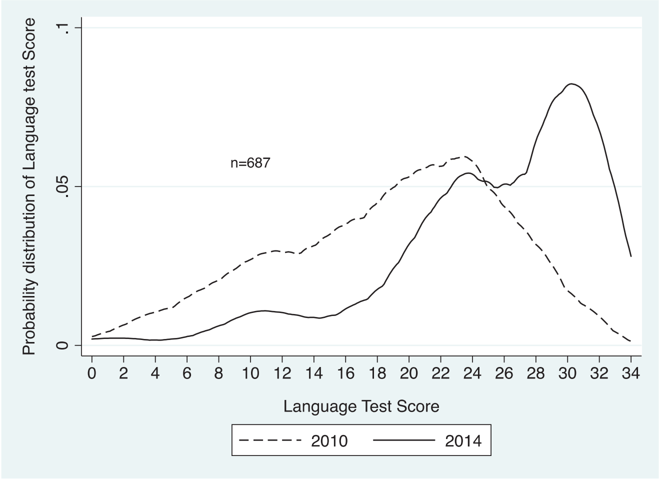
Distribution of Learning Outcomes among Subgroups of Rural Children

Given that rural children have relatively low outcomes, in this section we compare learning outcomes using the rural-only sample to identify poorly

FIGURE 1. Children's Learning in Math and Chinese Language over Time, Ages 10/11 in 2010 and 14/15 in 2014



A. Math Test Scores

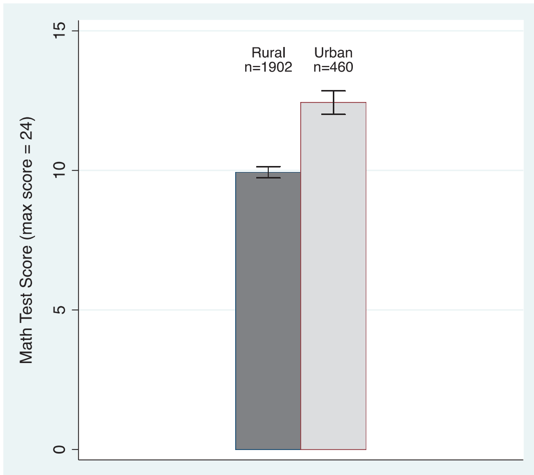


B. Chinese-Language Test Scores

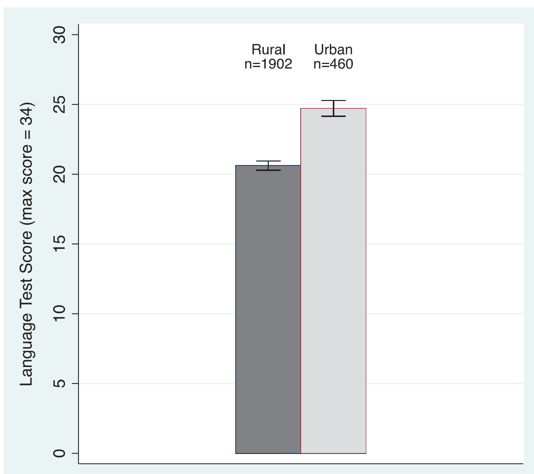
SOURCE: By authors

NOTE: Data from the China Family Panel Survey. The sample for these figures includes the same 687 urban and rural children in 2010 and 2014. We calculate the means of test scores and graph the figures by assigning sampling weights to each observation. The weights for provinces Shanghai, Liaoning, Henan, Gansu, and Guangdong are 10.28, 4.45, 2.04, 7.30, and 2.02; for the rest of the provinces, the weights are all 1. In Panel A, mean(2010) = 8.16 and mean(2014) = 12.81. This means that the children gained 4.65 points over four years, or 1.16 points per year. In Panel B, mean(2010) = 18.81, and mean(2014) = 25.27: the children gained 6.46 points over four years, or 1.62 points per year.

FIGURE 2. Differences in Math and Chinese-Language Test Scores between Rural and Urban Children, Ages 10–15, in 2014, with 95% Confidence Intervals



A. Math Test Scores



B. Chinese-Language Test Scores

SOURCE: By authors

NOTE: Data from the China Family Panel Survey. The sample for these figures includes both urban and rural children in 2014. The size of the sample (2,362) is the same as that of the total sample in 2014. We calculate the means of test scores by assigning sampling weights for each observation. The weights for provinces Shanghai, Liaoning, Henan, Gansu, and Guangdong are 10.28, 4.45, 2.04, 7.30, and 2.02, respectively; the weights for the rest of the provinces are all 1. In Panel A, mean(rural) = 9.93; mean(urban) = 12.43; difference = 2.5 ($p < 0.01$). This means that rural children are 2.2 years behind urban children in math learning. In Panel B, mean(rural) = 20.62; mean(urban) = 24.71; difference = 4.10 ($p < 0.01$). This means that rural children are 2.5 years behind urban children in Chinese-language learning.

performing subpopulations of rural children. We look first at subpopulations that "matter," that is, where the gap between the test scores of children in the different subpopulations is greater than one year of learning. After this we will look at the cases where the gaps are smaller ("do not matter").

We see sharp inequality in learning outcomes between Han and non-Han-minority children in rural areas (Table 1, Row 1). The difference in math test scores between Han and non-Han-minority children is 1.31 points ($p < 0.01$; Table 1, row 1, col. 1). The difference in Chinese-language test scores is 1.87 points ($p < 0.01$; Table 1, row 1, col. 3). The gaps are also large in terms of years of learning. For math it is 1.1 years; for Chinese language, 1.7 years (Table 1, row 1, cols. 2 and 4).

Rural children in poor areas and rural children with parents that have less education consistently perform worse. Specifically, rural children in poor counties in the rural-only sample score lower in both math and Chinese-language learning than rural children in non-poor counties (Table 1, row 2). The "poverty gap" for both math and Chinese language is approximately 1.3 years of learning. We also find similar gaps when we compare the test scores of children based on the level of education of their father or mother (Table 1, rows 3 and 4). In the most general terms, rural children whose parents have less education also have lower test scores. The above gaps are all greater than one year of learning.

We also see a wide gap between migrant children and children who attend rural public schools (Table 1, row 5). Since our sample came from CFPS data, the migrants that are included tend to be those who have lived for many years in the city (sometimes referred to as formal, registered migrants) and have become relatively integrated into city life. According to our data, migrant children perform significantly better on math and Chinese-language tests than children who attend rural schools in rural areas. The gap for math is 2.1 years; the gap for Chinese language is 1.7 years.

Interestingly, when we look at the differences in test scores by gender, female children are not a vulnerable subpopulation when it comes to learning. The test scores of all rural girls are actually slightly higher than those of boys—although the width of the gap is not large when measured in years of learning (only 0.03–0.6 years). Since the differences are less than one year, in the rest of the paper, we will say that they do not matter (that is, both boys and girls are nearly the same). There are also gaps between LBCs and CLPs (Table 1, row 7). Surprisingly, it is the LBCs that have higher math and

Chinese-language scores than CLPs. But both gaps are narrow: only 0.1–0.2 years (one or two months).

While there are clear differences in test scores for a number of the subpopulations in rural China, geography per se does not seem to play a role. When we divide the sample between rural children living in Northern China and those living in Southern China, the gaps for both math and Chinese-language scores are narrow (0.3 years—Table 1, row 8). However, rural children in Northern China have higher scores in math and lower scores in Chinese. Likewise, when we divide the sample among rural children living in Eastern, Central, and Western China, the gaps are also narrow: all are less than 0.9 years (Table 1, row 9).

The signs and levels of significance of the subpopulation variables in the regression results for the rural-only sample (Table 4) are nearly identical to the results in the descriptive statistics (Table 1). The main differences are the magnitudes of the coefficients (which measure conditional gaps, instead of the unconditional gaps—which are measured in Table 1). In other words, almost certainly because the coefficients are measuring conditional relationships, the coefficients for math scores (Table 4, col. 1) on all subpopulation variables (ethnicity, children from poor county, father's education, and mother's education) all suggest that there is less than one year's difference between the intra-subpopulation pairs (rows 1 to 4). The smallest coefficient is associated with the gender variable (col. 1, row 5), which is also consistent with the descriptive results (which found that gender did not really matter with respect to math or Chinese-language test scores).

Distribution of Learning Outcomes in Urban Settings

Looking at the results for the urban-only sample, the most salient observation is that many of the directions of the gaps between subpopulations are similar to what we found in the rural-only sample (although in most cases the within-urban gaps are narrower than the within-rural gaps). The only exception in terms of the direction of the gap is found in the ethnicity (Han/non-Han) comparison. In the rural-only sample, Han children outscored non-Han children in math by 1.31 points and in Chinese language by 1.87 points, while in the urban-only sample non-Han children have higher math and Chinese-language test scores than Han children (although the gaps are not statistically significant—Table 2, row 3).

The directions of the gaps between the children in the other subpopulation pairs are the same in the rural-only and urban-only samples. The poverty gaps (children in poor counties score lower than children in non-poor counties—Table 2, row 4); gaps based on the education of parents (children of parents with less education have lower scores than children of parents with more education—Table 2, rows 1 and 5); gender gaps (boys score lower than girls—Table 2, row 2); and geography (Table 2, rows 6–8) are in the same direction. For the magnitudes of the differences between subpopulations in urban areas, there are only two larger than one year (for children whose mothers have more versus less education, and girls versus boys). For the other subpopulations in urban areas, the differences in both math and Chinese-language tests scores are less than one year.[25]

DISCUSSION AND CONCLUSION

Our study seeks to provide insight into the distribution of learning outcomes that prevail across a number of key subpopulations in China. We also seek to build a repository of learning outcomes for children ages 10 to 15. Specifically, we describe the overall levels of learning outcomes (in math and Chinese language) across the country in rural and urban China. We hope to understand more concretely the differences in math and Chinese-language outcomes among major subpopulations in the rural economy.

We find that the distribution of academic performance in the total sample is uneven across a number of subpopulations in the rural and urban sample, the rural-only sample, and the urban-only sample. Perhaps the most fundamental result is that children in urban areas perform much better in both math and Chinese-language tests than those in rural areas. In fact, urban children are more than two years ahead of rural children in both subjects. This is a huge and problematic gap for a competitive, exam-based education system like China's.

In the rural-only sample, Han children, less-poor children, and the children of educated parents performed better on both math and Chinese-language tests. Migrants (that is, formal, registered migrants) also

25. The results of the empirical analysis (at least in the case of our major subpopulation comparisons—e.g., LBCs to CLPs and Han to non-Han) are robust to our use of fixed effects (versus OLS) and the use of fixed effects at different levels of aggregation (e.g., counties, provinces, or districts).

outperformed children living in rural villages. Gender and LBC/CLP status did not matter as much. Interestingly, the results suggest that geography does not really matter, either.

Because of the study's broad scope, large number of observations, and nationally representative data, our results can be used to assess the validity of a number of previous studies. In comparing the gendered learning outcomes of rural children, our results validate those reported by Lai et al.[26] In a sample of 7,235 middle school students in Beijing, Lai et al. found that girls outperformed boys throughout primary and middle school and in each quartile of the performance distribution. Girls also had a more positive school experience than boys, and boys had a higher dropout rate by the end of middle school. While our results differ from those of Wang et al.'s study of migrant children,[27] this might be explained by the fact that the migrant children in the CFPS sample were mainly children of formal, registered migrants, while those in the Wang et al. paper were mainly children of more temporary, unregistered migrants.

In the current report, using more observations than all previous studies combined, we find the somewhat surprising result that the learning outcomes of LBCs are better than CLPs. But in fact, other studies have shown this. While we are not able to identify in this paper the mechanism for *why* LBCs perform better than CLPs, the results should generate more interest in the findings of Bai et al., which suggest that there is an income/care trade-off when parents out-migrate and leave their children behind.[28] Often (our results suggest more often than not) the income from out-migration more than offsets any negative care effect. Our nationally representative findings call into question the representativeness of the studies that find LBCs are performing less well than CLPs. We believe that there are regions in which this is true (and these regions may have been the sampling targets of the papers that find LBCs perform better than CLPs), but overall our results suggest that if the national government wants to improve learning outcomes, focusing on CLPs could be more productive.

We also shed light on other work that has examined the differences in learning outcomes between other subpopulations. For example, our results

26. Lai et al., "Private Migrant Schools."
27. Wang et al., "Education Gap."
28. Bai et al., "Effect of Parental Migration."

for the learning outcomes of Han and non-Han ethnic-minority rural children validate those reported by Yang et al.[29] In a smaller sample (300 schools), Yang et al. found that non-Han-minority students performed worse than Han students in math and Chinese. Our study demonstrates that the results of these smaller studies can be generalized over a larger part of rural China.

In terms of the learning outcomes of rural children in richer and poorer counties, our results are consistent with the findings of Shi et al., who report that socioeconomic conditions are related to student academic achievement, documenting the poor performance in education of students in their study areas in poor rural China.[30] In this way, our research fills a significant gap (at least in terms of showing that Shi et al.'s finding holds in a nationally representative sample) in mapping the relative levels of learning outcomes in rural areas across China. Poor outcomes in learning are related to poverty.

In the urban-only sample, there is only one result different from what we found in the rural-only sample. In China's cities, non-Han-minority children on average have higher test scores on the study's math and Chinese-language tests than Han children. This, of course, is quite different from the rural-only sample, where Han outperformed non-Han-minority children by 1.1 years of learning in math and 1.7 years in Chinese. A possible reason for the high performance of non-Han-minority children in cities is that many of the non-Han-minority children in urban China are Man (or Manchurian) or Hui (a predominantly Muslim minority from north and midwestern China). Possibly because they have relatively high incomes and social status, their children outperform the average Han urban children in both the math and Chinese-language tests, according to our nationally representative sample.

The overall results mean that if policymakers are really interested in trying to improve the learning outcome of vulnerable groups in China, they need to focus on rural populations, and especially subpopulations that are poor and/or live in minority regions. Special attention needs to be paid to the children of less educated parents. There need be less concern about targeting girls and LBCs.

One of the most salient results for policymakers in China who have targeted poverty by geography (through providing aid to all residents in any county that is labeled "poor") is that targeting by geographic regions (e.g., west versus east)

29. Yang et al., "Han-Minority Achievement Gap."

30. Yaojiang Shi, Linxiu Zhang, Yue Ma, Hongmei Yi, Chengfang Liu, Natalie Johnson, James Chu, Prashant Loyalka, and Scott Rozelle, "Dropping Out of Rural China's Secondary Schools: A Mixed-Methods Analysis," *China Quarterly* 224 (2015): 1046–69.

appears not to be useful. Instead, our results suggest that there are households within both poor and non-poor areas that require targeting in terms of education. Such a strategy shift would require changing many of China's geography-based policies to target households based on more meaningful high-risk characteristics. These kinds of policy changes could have educational benefits for some of China's most vulnerable populations.

While our paper is motivated by looking for ways to improve the overall human capital of China by identifying (and working with) low-performing subgroups (non-Han minorities, those at the lower end of the income distribution, etc.), it may be better to target inequality than to target human capital. In other words, it could be that if raising human capital were the nation's only goal, education officials might not want to be put too much effort into addressing the learning difficulties of the depressed groups. Of course, as we have shown, if we did promote learning among the depressed groups in China, human capital would rise (and inequality would fall). There may be more effective ways to promote human capital (if that is our only goal), for example by focusing on less-depressed groups, which might be able to learn more. Given the high levels of education enjoyed by some of the non-depressed groups in China, however, focusing on the more-depressed groups may be the best way both to reduce inequality and to improve overall human capital.

APPENDIX

Share of Key Subpopulations in Total Population in 2014

	Full sample (urban + rural)		*Rural-only sample*		*Urban-only sample*	
Variables	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Rural children	0.78	0.41	1.00	0.00	0.00	0.00
Ethnicity (1 = non-Han-minority children)	0.09	0.28	0.09	0.28	0.08	0.27
Children in poor counties	0.45	0.50	0.54	0.50	0.19	0.39
Father's education (1 = less than jr. high)	0.48	0.50	0.58	0.49	0.11	0.31
Mother's education (1 = less than jr. high)	0.60	0.49	0.73	0.45	0.15	0.36
Children living in rural villages	0.72	0.45	0.92	0.26	—	—
CLPs	0.69	0.46	0.90	0.30	—	—
LBCs	0.07	0.26	0.10	0.30	—	—
Boy	0.52	0.50	0.53	0.50	0.47	0.50
South	0.29	0.46	0.24	0.43	0.48	0.50
East	0.33	0.47	0.26	0.44	0.59	0.49
Central	0.18	0.39	0.18	0.39	0.19	0.39
West	0.49	0.50	0.56	0.50	0.22	0.42
Age (years)	12.51	1.74	12.53	1.73	12.45	1.80
Observations	2,362		1,902		460	

SOURCE: By authors

NOTE: Data from the China Family Panel Survey. Standard deviation in parentheses. In all regressions, we control for district fixed effects. We also assign a sampling weight for each observation. The weights for provinces Shanghai, Liaoning, Henan, Gansu, and Guangdong are 10.28, 4.45, 2.04, 7.30, and 2.02, respectively; the rest of provinces are each weighted 1. *Rural children* are children who have rural hukou. *Children from poor counties* are children from a county that is in the lowest quartile of the CFPS sample counties when the counties are ranked by GDP per capita. CLPs are children who are living with their parents in rural villages. LBCs are children who are left behind in rural villages while their parents have out-migrated.