China Economic Review xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

China Economic Review

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

Can reading programs improve reading skills and academic performance in rural China?

Qiufeng Gao[a], Huan Wang[a,b,*], Di Mo[b,c], Yaojiang Shi[a], Kaleigh Kenny[b], Scott Rozelle[b]

[a] Center for Experimental Economics in Education, Shaanxi Normal University, 620 West Chang'an Avenue, Xi'an, China
[b] Rural Education Action Program, Freeman Spogli Institute for International Studies, Stanford University, 616 Serra St., Stanford, CA, United States
[c] LICOS Center for Institutions and Economic Performance, University of Leuven, Louvain, Belgium

ABSTRACT

In this paper, we attempt to evaluate the effectiveness of reading programs at improving the reading skills and academic achievement of primary school students in rural China. Using survey data on 4108 students, we find that students exhibited low levels of reading achievement, independent reading quantity, and reading confidence in the absence of any treatment. However, our results also suggest that properly designed treatments may improve the reading and academic outcomes of students. Specifically, we found that increased access to independent reading materials coupled with effective teacher training led to significant improvements in student reading skills, math test scores, and Chinese test scores. We believe that these improvements are due to changes in reading instruction and the attitudes of teachers toward reading. These findings indicate that encouraging higher reading quantity and providing high-quality reading instruction are important components for programs that seek to improve student outcomes in developing country settings.

1. Introduction

The development of adequate reading skills at an early age can have an enormous influence on the academic achievement of students throughout the course of their educational careers (National Reading Panel, 2000; Slavin, Lake, Chambers, Cheung, & Davis, 2009). However, the reverse is also true. Students may fall behind in school and encounter poor employment and social outcomes later in life if they are unable to develop their reading skills early in their years of schooling (Good, Simmons, & Kame'enui, 2001; Slavin et al., 2009; Whitehurst & Lonigan, 2001). For this reason, it is commonly believed that reading challenges should be addressed early to decrease the likelihood that developmental delays will impact students over the long-term.

Previous research has found that reading skill development is related to both the quantity of reading (volume of books) and the quality of reading instruction provided to students (Cunningham & Stanovich, 2001; Kassow, 2006; Topping & Sanders, 2000). Specifically, research shows that increased reading volume combined with effective reading instruction can lead to reading skill improvements (Topping, Samuels, & Paul, 2007). Elley (1996) also found that providing reading instruction and access to more books can improve student reading skills in developing countries. Without an adequate focus on improving the ability of students to read with higher levels of comprehension, increased time reading may only be able to moderately improve the reading skills of students.

In acknowledgement of the importance of reading, many developed countries have established programs to encourage reading and promote the reading skill development of young students (Adler & Fisher, 2001; Lonigan & Shanahan, 2009; Slavin et al., 2009; Slavin, Cheung, Groff, & Lake, 2008). These programs vary in content and design. Some programs focus on improving the quality of reading instruction provided by teachers and parents; others attempt to increase the quantity of reading materials available to children (Lonigan & Shanahan, 2009).

* Corresponding author at: Rural Education Action Program, Freeman Spogli Institute for International Studies, Stanford University, 616 Serra Street, Stanford, CA 94305, United States.
E-mail address: huanw@stanford.edu (H. Wang).

https://doi.org/10.1016/j.chieco.2018.07.001
Received 24 August 2017; Received in revised form 14 June 2018; Accepted 4 July 2018

Research has provided evidence on the effectiveness of reading programs that have been designed and implemented in developing countries (Abeberese, Kumler, & Linden, 2014; Friedlander & Goldenberg, 2016). Much like the programs implemented in developed countries, most of these programs focus on either increasing reading frequency, improving reading instruction, or both. For example, Friedlander and Goldenberg (2016) determined that when reading programs both provided books and delivered effective reading instruction, there was a positive impact on the reading achievement of students in Rwanda. Abeberese et al. (2014) found similar results in the Philippines when students were provided with appropriate reading materials combined with instruction from well-trained teachers. These findings are important for developing countries because evidence suggests that poor children are particularly prone to developmental delays arising from reading challenges (Roskos, Strickland, Haase, & Malik, 2009). If reading programs can reverse the vicious cycle brought on by early reading difficulties, they may, in turn, offer a mechanism through which countries can support future educational development.

Although these reading programs—and their measured impacts on reading ability—are noteworthy, there is an absence of empirical evidence on the ability of such programs to improve other educational outcomes (e.g., test scores or grades in other academic subjects). The rationale behind this potential relationship is that improved reading skills could improve academic performance in other subject areas, such as math or language, by developing skills such as reading comprehension and critical thinking (Good et al., 2001; Rutherford-Becker & Vanderwood, 2009; Slavin et al., 2009). To date, however, the results from the evaluations from the few existing randomized evaluations of reading programs have found different impacts of reading programs on the academic outcomes of students, as some studies have found positive impacts while others have found no impact. For example, Lucas, McEwan, Ngware, and Oketch (2014) found that training teachers on how to teach reading, as well as providing teaching materials, significantly improved student language test scores (in this case, writing and oral literacy test scores) in Ugandan primary schools. Abeberese et al. (2014), however, found that a program providing age-appropriate reading materials and teacher training to students in the Philippines did not significantly improve the math or social studies test scores of students. Borkum, He, and Linden (2012) also found that supplying better reading resources, providing a librarian, and conducting reading activities in Indian schools did not have any significant impacts on the math and science test scores of students. Due to the mixed results of these evaluations, additional empirical research is needed to determine whether reading programs (and what aspects of reading programs) are effective at improving both student reading skills and learning in academic subjects.

Increased focus on reading may be particularly valuable in China, where severe educational inequality between urban and rural areas threatens the nation's continued economic growth and social cohesion (Wang, Liu, Zhang, Shi, & Rozelle, 2013; Zhang, Yi, Luo, Liu, & Rozelle, 2013). It is possible that placing an emphasis on reading in rural schools could help narrow this educational gap, as it seems to have done in certain developed contexts (Kim, 2006; Kim & Quinn, 2013). However, advancing a new reading agenda may not be easy, since, in many traditional settings, independent reading is often thought to be detrimental to student performance core curriculum classes (either because reading distracts students or it takes time away from studying for other subjects).

Fortunately, at least at higher levels, China's government appears to have recognized the developmental potential of improved literacy when it announced a new, national focus on reading in its 2014 annual work report (People.cn, 2015). As part of this effort, the government launched a *nationwide reading initiative* to encourage people of all ages and professions to read in 2015. Following this recent initiative, several provincial and city governments, as well as non-government organizations, have established programs to encourage reading across the country, particularly in rural areas (GMW.cn, 2015).

However, little is known about the reading achievement, reading behaviors, and attitudes of Chinese students—or that of their parents and teachers—toward reading, particularly in rural areas of the country. There are virtually no published papers in the English literature on the issue. In the few papers in the Chinese literature that do exist, educators typically present case studies that primarily describe practices observed in a particular classroom or small set of classrooms (Sun & Xu, 2003; Wang, Sun, & Wang, 2013). Without further information on the reading practices of Chinese students and evidence on the success of programs at encouraging reading/improving educational outcomes, it will be difficult for China's government to develop more effective policies to support reading skill development.

The overall goal of this paper is to better understand the reading practices and to evaluate the effectiveness of reading programs at improving the reading skills and academic achievement of primary school students in rural China. To meet this goal, we have three specific objectives. First, we document the current levels of reading achievement, the reading behaviors, and the attitudes toward reading of primary school students in rural China. Second, we evaluate three different reading program treatments to see if certain treatments (i.e., different components of a reading program) are associated with higher levels of reading skills or higher levels of math and/or Chinese language achievement. Finally, we seek to identify the mechanisms through which an effective reading treatment may lead to improved educational outcomes.

The rest of the paper is organized as follows. In the next section, we describe our sampling procedure, reading program treatments, data collection, and analytical approach. Our results are discussed in section three. We conclude in section four.

2. Experimental design and data collection

2.1. Sampling

In this paper, we evaluate three alternative reading programs that were implemented in rural Guizhou province beginning in 2012. Guizhou is located in southwest China and has a population of about 35 million people, meaning that the province is home to 2.6% of the total population of China (Statistics Bureau of Guizhou Province, 2015). Although the per capita GDP of Guizhou (12,371 RMB per person) was less than the national average in 2014 (20,167 RMB per person), it was only slightly lower than that of other

western provinces in China (13,919 RMB per person – National Bureau of Statistics of China, 2015). Additionally, the GDP growth rate was 10.8% in 2014, which was higher than the national GDP growth rate of 7.4% in the same year (National Bureau of Statistics of China, 2015; Statistics Bureau of Guizhou Province, 2015). In addition, the unemployment rate of Guizhou (3.3%) was slightly lower than the national average in 2014 (4.1% —National Bureau of Statistics of China, 2015).

Although our study only sampled in one prefecture in Guizhou, we find that this area is comparable to other areas of the country in terms of several key socioeconomic indicators. The average net per capita income in 2014 in our rural study area was 9788 RMB, which was higher than the average of 8295 RMB for all rural areas in western China and only slightly lower than the average of 10,489 RMB for all rural areas in China (National Bureau of Statistics of China, 2015; Statistics Bureau of Guizhou Province, 2015). Also, the net enrolment ratio of school-age children in primary schools was 99.9% in our sample area in 2014, which was similar the national average of 99.8% for all of China (National Bureau of Statistics of China, 2015; Statistics Bureau of Guizhou Province, 2015). Taken together, these statistics suggest that our sample site is generally representative of rural areas in China.

In total, we included 128 classes from 30 rural primary schools and the study is sufficiently powered with 128 classes. Using rural primary school data from previous studies, we assumed an intra-class correlation coefficient of 0.20 and an R-squared of 0.50. As is standard in much of the social science literature, we set alpha = 0.05 and Power = 0.80. We then calculated that we required at least 32 individuals per class and 110 classes to detect a standardized effect size of 0.20 under a balanced experimental design. We chose two neighboring counties with similar economic, cultural, and geographical characteristics as the treatment county. We then randomly selected 15 treatment schools in the treatment county, and paired them with 15 control schools of equal size and similar teacher and school characteristics from the lists of schools in neighboring counties to increase the statistical power of our analyses (Imai, King, & Nall, 2009). Our data show that the treatment and control schools are statistically similar in regard to all observable characteristics. In order to ensure that we have enough sample left after trimming to the “common support” in propensity scores, we further increased our sample size by about 20% (to 128 classes) to ensure sufficient power.

Our carefully set-up sample selection protocol ensured that the treatment group and control group were comparable. The sampling protocol consisted of two steps (Fig. 1). To choose sample schools, we first visited the county where the reading program treatments were implemented. From the county education bureau, we obtained a list of schools that had at least one of the reading programs that we were evaluating (see below). In total, there were 73 schools in 15 townships in our treatment school sampling frame. Next, assuming neighboring counties have similar economic, cultural, and geographical characteristics, we obtained lists of schools from two neighboring counties where no school had implemented any reading program treatments. The list of the schools in these two counties became the sampling frame for the control schools in our study. As can be seen in Appendix Table A, our treatment and control schools are statistically similar on all observable characteristics.

After identifying our treatment and control sampling frames, we then selected our sample schools. First, we randomly selected one treatment school from each of the 15 sample townships. In each treatment school, we conducted surveys with students from grade 3 to grade 6. In general, there was typically only one or two classes in each grade. If a school had more than two classes in a grade, we then randomly selected two classes from that grade for inclusion in our sample. In total, 2533 students in 79 classes from 15 schools were included in our treatment group. After we randomly selected 15 treatment schools in the treatment county, we paired them with

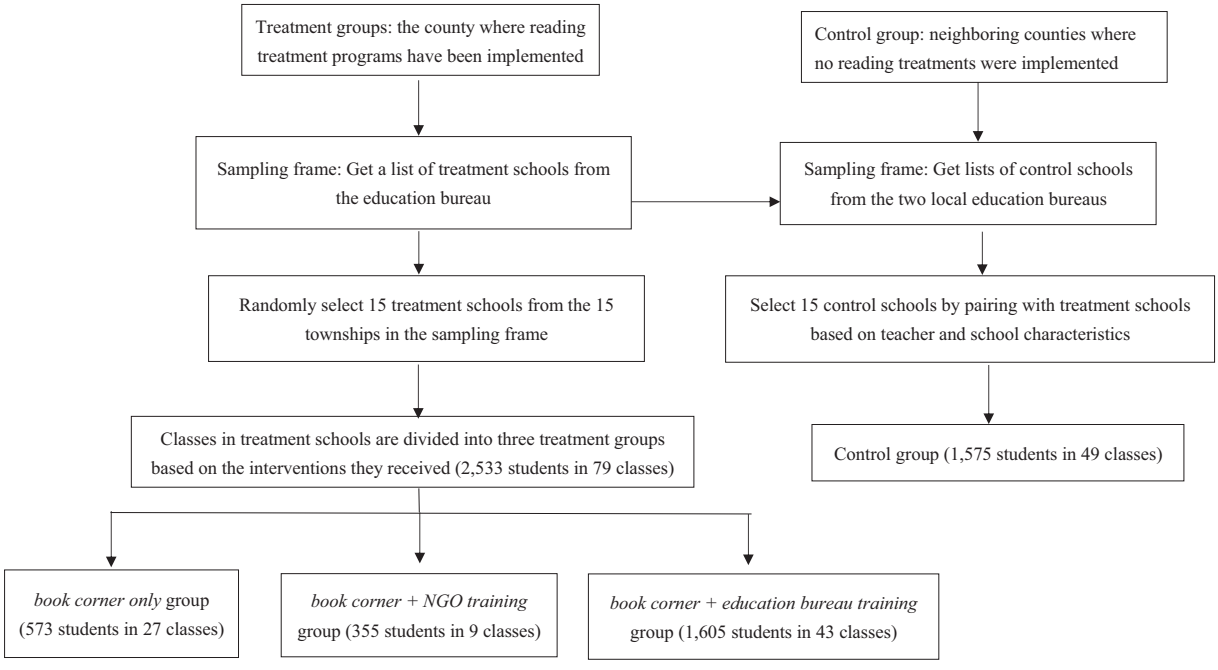


Fig. 1. Sample selection procedure.

15 control schools from the two neighboring control counties with similar teacher and school characteristics listed in Appendix Table A. In each control schools, we conducted surveying with students in grade 3 through grade 6 in the same manner as in our treatment schools. In total, 1575 students in 49 classes were included in our control group. Appendix Table B shows that students from the treatment and control schools are balanced on the set of observable characteristics collected from our survey.

2.2. The reading program treatments

In our study, we evaluated three alternative reading intervention arms (Fig. 1). The first intervention arm provided book corners (and only book corners) with extracurricular books to all classrooms in participating elementary schools, and we refer to this as our *book corner only* treatment. Each book corner had two shelves and was stocked with 70 extracurricular books. The books varied in content based on the age and reading levels of students. The program provided books that were chosen to be interesting to students and that covered content beyond what was taught in school, such as literature, picture, and natural science books. In total, 573 students from 27 classes in 10 schools received this type of intervention.

The second intervention arm combined the same type of classroom book corner with teacher training provided by the local education bureau (and, we refer to this treatment as *book corner + education bureau training*). This training provided by the education bureau focused primarily on how to instruct Chinese language classes in the manner that mandated by the national curriculum. Specifically, this type of training focused on topics such as reading with a rhythmic pace, analyzing key passages, vocabulary development, and summarizing conclusions. In other words, this type of training focused on teaching specific reading skills in isolation, rather than focusing on the overall practice of reading. In total, 43 teachers from 43 classes in 14 schools participated in the education bureau trainings, and 1605 students that were included in the *book corner + education bureau training* treatment group.

The third intervention arm also provided classroom book corners and teacher training, except these trainings were provided by an NGO instead of by the local educational bureau (and, therefore, we refer to this treatment as *book corner + NGO training*). Unlike the *book corner + education bureau training* treatment, the training provided by the NGO focused on instructing students on how to read independently. That is, reading books beyond those used in core curriculum classes. The training informed teachers on how independent reading promotes reading skill development and provided methods to encourage positive reading behaviors among students. Specifically, the NGO provided information on what kinds of books children should read at different educational stages and what sorts of reading activities teachers could organize in their classes to promote independent reading. A total of 9 teachers from 9 schools participated in the NGO trainings and 355 students from 9 classes were included in the *book corner + NGO training* treatment group. As can be seen in Table 2, the students in the three different types of treatment groups and the control group students are balanced based on observable characteristics.

2.3. Data collection

In April 2015, we conducted a three-block survey of students and teachers in all sampled treatment and control schools. In the first block of the survey, we asked a series of questions about student reading habits/attitudes, access to books, and in-school reading instruction/encouragement. Specifically, we asked whether the student spends > 30 min or > 60 min reading independently per day; whether the student believes reading has a positive effect on his/her math or Chinese performance; whether his/her school has a library; whether the student borrows books from the school library every week; whether the student is able to borrow interesting books from the school library; whether there are independent reading books in his/her household; whether the his/her parents buy books for him/her; whether the student receives reading instruction in school; and whether teachers encourage him/her to borrow books every week. We also asked a series of questions designed to evaluate the reading confidence and enjoyment of students, which were based off questions asked in the Progress in International Reading Literacy Study (PIRLS) assessment.

In the first block of the survey, we also collected data on student and family background characteristics. These questions gathered data on each student's gender, age, parental migration status (whether his/her father/mother primarily lives at home), parental education (whether his/her father/mother graduated from primary school), and the quantity of books in the child's home (whether there are > 10 books in the child's home; whether there are > 10 children's books in the child's home). The information gathered from these questions are used as control variables in our analysis.

The second block of the survey was a 30-min standardized test that evaluated reading skills and a 30-min test in math or Chinese language. The reading tests were constructed by trained psychometricians by using test items from the Progress in International Reading Literacy Study (PIRLS) test, an international test of reading comprehension that is widely used throughout the world (Caygill & Chamberlain, 2004; Cheung, Tse, Lam, & Ka Yee Loh, 2009; Mullis, Martin, & Gonzalez, 2004; Tunmer, Chapman, Greaney, Prochnow, & Arrow, 2013). The test questions were carefully translated according to the PIRLS translation guidelines and the content validity was reviewed by a panel of experts and local teachers with knowledge on China's education system. The translated reading tests then went through several rounds of pilot testing in third to sixth grade classes in rural Chinese schools. The psychometric properties of the test were then validated using data from the extensive pilot testing to ensure good distributional properties (no bottom or top-coding, for example). In the analyses, we normalized reading achievement scores using the mean and distribution in the control group. Estimated effects are therefore expressed in standard deviations.

The tests in math and Chinese language were carefully designed with assistance from educators in the local education bureau to ensure compliance with the national curriculum. We pre-tested the exam multiple times to ensure its relevance and that time limits were appropriate. When we administered the exam in the sample schools, it was timed carefully and closely proctored by trained enumerators. In our study, all the students took the reading test. Additionally, half of our sample classrooms were randomly chosen to

take a standardized math test, while the other half took a standardized Chinese language test. In other words, half of the sampled students took reading and Chinese language tests, while the other half took reading and math tests. For analysis, we first calculated the percentage of students whose reading test scores were above the half of possible total score, which we consider to be a passing score for our analysis. Then, we normalized all test scores according to the distribution of scores in each grade.

The third and final block of the survey collected information on the characteristics of teachers and schools in our sample. In this part of the survey, we asked Chinese teachers about their attitudes toward reading (whether Chinese teachers believe reading has a positive effect on the math performance of students, whether Chinese teachers believe reading has a negative effect on the math performance of students). In addition, the survey also collected information on teacher and school characteristics that are included as control variables in our analysis. We asked Chinese teachers about their gender and education (specifically, whether they graduated from college). We also gathered data on the size of their school from the school principal (specifically, area in square meters).

2.4. Empirical strategy

To identify which reading treatments are more effective at improving student outcomes, we first use an OLS model. This model is capable of partially controlling for selection bias and potential endogeneity by including a large set of observable covariates. We use Eq. (1) to estimate the impact of the three alternative reading program treatments separately. The model used is as follows:

$$y_{ijc} = \alpha + \beta' P_c + \gamma' X_{ijc} + \mathscr{E}_{ijc} \tag{1}$$

where the dependent variable y_{ijc} indicates the standardized test score of student i in school j and class c; P_c is a vector of three treatment dummy variables that represent three alternative reading program treatments: *book corner only* classes; *book corner + education bureau training* classes; and *book corner + NGO training* classes. In other words, the treatment variable for *book corner only* is a dummy variable that takes on the value of 1 if the class only received a book corner and no form of teacher training, and takes on the value of 0 otherwise. The term X_{ijc} is a vector of covariates that is included to capture the effect of different student, parent, teacher, and school characteristics on the dependent variable, which are shown in Table 1. In all regressions, we accounted for the clustered nature of our sample by constructing Huber-White standard errors corrected for class-level clustering.

In our robustness check specification, we also include a school dummy variable to control for school-level fixed effects that may be correlated with the treatment variable and affect the dependent variable. Because there are certain schools where two of our reading program treatments were implemented (that is, the *book corner + NGO training* and *book corner + education bureau training* treatments), by controlling for school fixed effects we are able to identify the program impacts of these two reading treatments separate from any confounding unobservable school characteristics. The school-level fixed effects, P_c is equal to 1 if a class receives the *book corner + NGO training* treatment, and is equal to 0 if a class receives the *book corner + education bureau training* treatment.

Another way to correct for many relevant covariates is Propensity Score Matching (PSM). The propensity score (i.e., the conditional probability of receiving treatment) is calculated by estimating a logit model with student, parent, teacher, and school characteristics as the independent variables (Table 1). In our study, the PSM method we employ is Nearest Neighbor Matching with replacement to ensure that each treatment unit is matched to the comparison unit with the propensity score closest to its own. This method is implemented with common support, a logit model for calculating the propensity score, and bootstrapped standard errors. Also, this method offers a way to assess the robustness of the estimates from the OLS regressions.

3. Results

3.1. Descriptive statistics and balance test

Summary statistics for our sample are presented in Table 1. From this table, we can see that our sample is almost evenly split

Table 1
Average characteristics of sample students (Guizhou Province, China 2015).

Variables	Mean	Std. Dev.
Student and family characteristics (n = 4108)		
1. Gender (1 = female)	0.49	0.50
2. Age (1 = elder than mean)	0.50	0.50
3. Mother primarily lives at home (1 = yes)	0.51	0.51
4. Father primarily lives at home (1 = yes)	0.42	0.49
5. Mother graduated from primary school (1 = yes)	0.27	0.45
6. Father graduated from primary school (1 = yes)	0.45	0.50
7. Number of books in household (1 = > 10 books)	0.40	0.49
8. Number of children's books in household (1 = > 10 books)	0.28	0.45
Teacher characteristic (n = 128)		
9. Gender of Chinese teacher (1 = female)	0.51	0.50
10. Chinese teacher graduated from college (1 = yes)	0.39	0.49
School characteristics (n = 30)		
11. School size (1 = larger than mean)	0.43	0.50

Table 2
Covariate balance test among the three treatment groups and the control group.[a]

Variables	Control Group	"Book corner only" Group	"Book corner + Edu bureau training" Group	"Book corner + NGO training" Group	Difference: (2)–(1)	Difference: (3)–(1)	Difference: (4)–(1)	*P*-value of F-test[b]
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Coef (Se)	Coef (Se)	Coef (Se)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1. Gender (1 = female)	0.49 (0.50)	0.47 (0.50)	0.49 (0.50)	0.48 (0.50)	−0.01 (0.01)	−0.00 (0.01)	−0.01 (0.04)	0.648
2. Age (1 = older than mean)	0.47 (0.50)	0.40 (0.49)	0.55 (0.50)	0.60 (0.49)	−0.02 (0.04)	0.04 (0.05)	0.13 (0.18)	0.452
3. Mother primarily lives at home (1 = yes)	0.51 (0.50)	0.46 (0.50)	0.53 (0.50)	0.53 (0.50)	−0.02 (0.01)	0.01 (0.02)	0.02 (0.07)	0.206
4. Father primarily lives at home (1 = yes)	0.40 (0.49)	0.44 (0.50)	0.44 (0.50)	0.43 (0.50)	0.01 (0.01)	0.02 (0.02)	0.03 (0.04)	0.515
5. Mother graduated from primary school (1 = yes)	0.26 (0.44)	0.25 (0.44)	0.28 (0.45)	0.31 (0.46)	−0.00 (0.01)	0.01 (0.01)	0.05 (0.06)	0.804
6. Father graduated from primary school (1 = yes)	0.47 (0.50)	0.38 (0.49)	0.46 (0.50)	0.51 (0.50)	−0.03 (0.01)	−0.01 (0.01)	0.04 (0.05)	0.137
7. Number of books in household (1 = > 10 books)	0.43 (0.50)	0.34 (0.47)	0.40 (0.49)	0.41 (0.49)	−0.03 (0.01)	−0.01 (0.01)	−0.02 (0.05)	0.122
8. Number of children's books in household (1 = > 10 books)	0.30 (0.46)	0.23 (0.42)	0.28 (0.45)	0.29 (0.46)	−0.03 (0.01)	−0.01 (0.01)	−0.01 (0.03)	0.198
9. Gender of Chinese teacher (1 = female)	0.63 (0.48)	0.30 (0.46)	0.47 (0.50)	0.51 (0.50)	−0.11 (0.04)	−0.08 (0.06)	−0.12 (0.20)	0.106
10. Chinese teacher graduated from college (1 = yes)	0.43 (0.50)	0.41 (0.49)	0.31 (0.47)	0.48 (0.50)	−0.01 (0.05)	−0.06 (0.06)	0.05 (0.20)	0.693
11. School size (1 = larger than mean)	0.60 (0.49)	0.39 (0.49)	0.31 (0.46)	0.30 (0.46)	−0.07 (0.05)	−0.14 (0.06)	−0.30 (0.20)	0.064

[a] The three treatment groups refer to the *book corner only*, *book corner + education bureau training*, and *book corner + NGO training* groups.

[b] To check the balance between the control and three treatment groups, we run a regression for each of the control variables. For example: Gender = a1*book corner only + a2* Edu bureau training +a3*NGO training. Then, we conduct a joint F-test: test (a1 = 0) (a2 = 0) (a3 = 0). From this we get the the p-value.

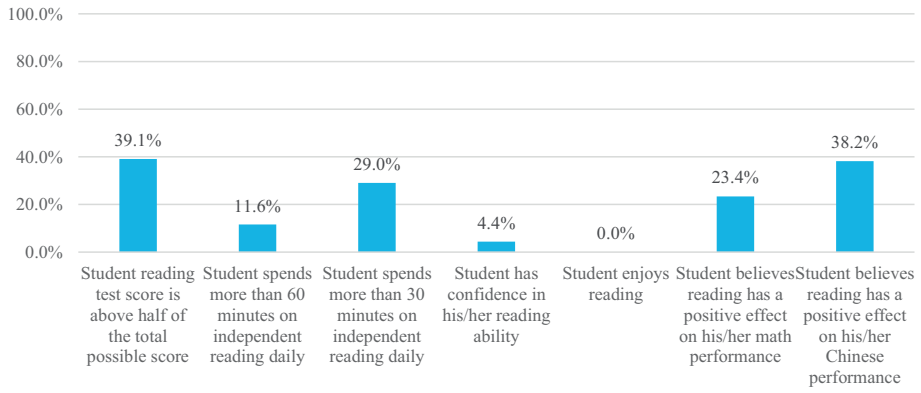


Fig. 2. Student reading achievement, reading behavior, and attitudes toward reading in rural primary schools in Guizhou Province, 2015.
Note: These are data from our control group sample.

between male and female students (51% and 49% of the sample, respectively—Row 1). We also found that 42% of the fathers and 51% of the mothers of sample children primarily live at home (Rows 3 & 4). A higher proportion of fathers graduated from primary school (45%) than mothers (27%—Rows 5 & 6). Also, in our sample of teachers, 51% are female (Row 9) and 39% have a college degree (Row 10).

We also examined whether our treatment and control groups are balanced on these observable characteristics. As can be seen from Table 2, the treatment and control groups show no statistically significant differences in student and family characteristics (Column 8, Row 1–8). Student also do not differ across treatment and control groups in terms of teacher and school characteristics (Column 8, Row 9–10), except for school size ($P = .064$, Column 8, Row 11). From these results, we believe our matching strategy was able to produce a good counterfactual (control group) against which we can assess the effectiveness of our treatments.

3.2. Reading achievement, behavior, and attitudes of students in rural China

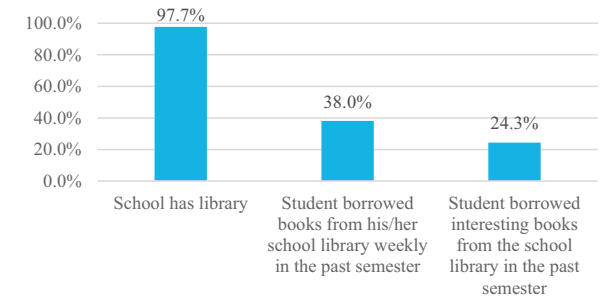
To better understand the state of the reading achievement, reading behaviors, and attitudes toward reading of students in rural China, we examined the case of the 1575 students in our control schools that did not participate in any reading treatments. We do this to evaluate the state of reading in rural China in the absence of any intervention. The results are presented in Figs. 2 and 3.

According to our data, rural Chinese students in our sample exhibit low levels of reading achievement. Only 39.1% of sample control students (615 out of 1575 students) scored above 50% on the reading exam (Fig. 2). In other words, > 60% of the students in the control schools displayed extremely low levels of reading. These findings are similar to those found by Authors (2017), who evaluated the reading achievement of a representative sample of rural Chinese students (including, but not limited to, those in the Guizhou data) against international data supplied by the Progress in International Reading Literacy Study (PIRLS). According to the findings of Authors (2017), the reading achievement of mainland Chinese students is lower than that of students from any country or region that participated in PIRLS.

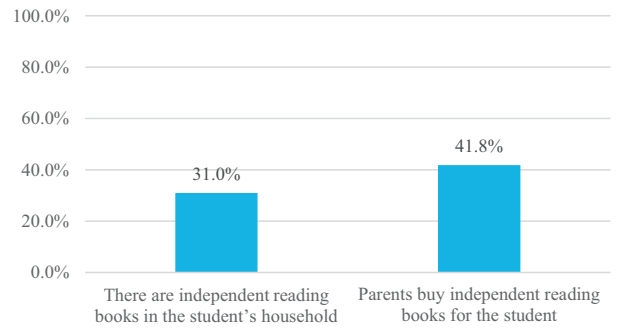
The results also show that rural students spend little time reading, lack confidence in reading, and hold negative attitudes toward reading (Fig. 2). Among our control sample, only 11.6% of students read for > 60 min per day and only 29.0% read for > 30 min per day. Given the little time that sample students spend reading, it is not surprising that none of the students in our control sample (not one) reported that they enjoyed reading. While we do not know if it is due to small amount of time spent reading or to negative attitudes toward reading, findings from the survey suggest that few of the sampled students are confident in their reading skills. Specifically, only 4.4% of respondents indicated that they are confident in their reading ability.

One potential reason students may not be interested in reading is the competitive nature of China's school system and the focus that is placed on performance in math, Chinese language, and other subjects that are tested on standardized examinations. According to our data, only 23.4% of students believe that independent reading has a positive effect on their math performance. Similarly, only 38.2% of students believe that improved reading skills would benefit their Chinese language performance (Fig. 2). This lack of understanding that reading skills are tied to academic performance in other subjects stands in contrast to international research. Specifically, it has been found that improved reading skills offer a wide-range of benefits to the achievement of students in a variety of academic subjects (Rutherford-Becker & Vanderwood, 2009; Thurber, Shinn, & Smolkowski, 2002).

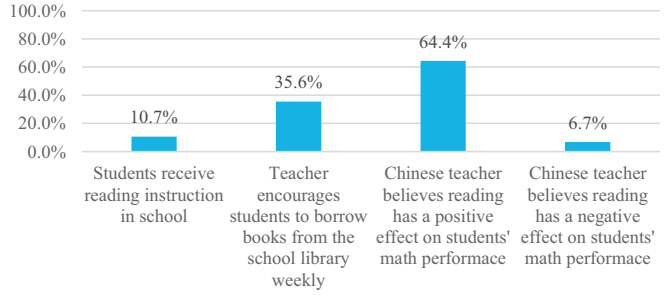
There are several factors that may contribute to the low levels of reading achievement, time spent reading, and reading confidence among students in our control sample. First, even in cases where schools have libraries, students lack adequate access to independent reading books. Considering that the government in China has supported investments in school libraries since 2003 (Ministry of Education of the People's Republic of China, 2003), it is not surprising that the vast majority of schools in our sample have libraries (97.7%— Fig. 3, Panel A). However, from interviews conducted during our survey, it became clear that many school libraries were often locked, making it difficult or impossible for students to access the books inside. This lack of access to school libraries is supported by our finding that only 38.0% of students reported that they could borrow any book from their school library on a weekly basis within the last semester. Even if some school libraries in our sample were accessible, 75.7% of sample students expressed that



A.



B.



C.

Fig. 3. Attitudes/behaviors toward reading and access to reading resources among students and teachers in rural primary schools in Guizhou Province, 2015.
Note: These are data from our control group sample.
Panel A. Library resources and student borrowing behaviors.
Panel B. Reading resources at home.
Panel C. Attitudes and behaviors toward reading of teachers.

they are unable to find interesting books in their school library. Without adequate access to books that are appropriate for the reading levels and match the interests of students, it is understandable that students invest little time in reading, hold negative attitudes toward reading, and have underdeveloped reading skills.

Even if books are unavailable in school, if students can access reading materials at home they may be more likely to invest time in reading, and subsequently improve their reading skills and confidence. Unfortunately, this is not the case. Our data demonstrate students in our control schools also have inadequate access to books at home. For example, 69.0% of students in our control sample did not have any (i.e., zero) independent reading books in their homes. Additionally, over half (58.2%) of the sample students report that their parents have never bought any books for them (Fig. 3, Panel B). These findings are consistent with those presented in Lin (2007) and Liu (2015), which found that rural households lacked independent reading materials. According to these studies, this is because many rural parents and caregivers also hold negative attitudes toward reading with the understanding that reading will take away time from studying.

Even if students had access to books, it is not clear that they would be able to use them effectively. In part, this is because many

Table 3
OLS regression of treatment effects (for *book corner only*, *book corner + education bureau training*, and *book corner + NGO training* treatments) on the reading skills and academic achievement of sample students.

Dependent variable	Standardized reading score (SD)		Standardized math score (SD)		Standardized Chinese score (SD)	
	(1)[a]	(2)	(3)[a]	(4)	(5)[a]	(6)
1. *Book corner only* (1 = yes)	−0.09		−0.03		−0.20	
	(0.10)		(0.10)		(0.12)	
2. *Book corner + Edu bureau training* (1 = yes)	0.09		−0.02		−0.04	
	(0.07)		(0.07)		(0.09)	
3. *Book corner + NGO training* (1 = yes)	0.61***	0.43***	0.35***	0.34***	0.29**	0.57***
	(0.10)	(0.07)	(0.13)	(0.10)	(0.13)	(0.10)
4. Student and family characteristics controlled	Yes	Yes	Yes	Yes	Yes	Yes
5. Teacher characteristic controlled	Yes	Yes	Yes	Yes	Yes	Yes
6. School characteristics controlled	Yes	Yes	Yes	Yes	Yes	Yes
7. Grade dummy	Yes	Yes	Yes	Yes	Yes	Yes
8. School dummy	No	Yes	No	Yes	No	Yes
9. Constant	−0.61***	−0.57***	−0.39***	−0.36**	−0.56***	−0.87***
	(0.11)	(0.08)	(0.10)	(0.15)	(0.16)	(0.14)
10. Observations	4108	1960	2604	1286	1504	674
11. R-squared	0.109	0.170	0.097	0.150	0.122	0.198

Note: 1. Robust standard errors clustered at the class level are presented in parentheses, *** $p < .01$, ** $p < .05$, * $p < .1$.
2. We analyze each type of reading treatment (*book corner only*, *book corner + education bureau training*, and *book corner + NGO training*) on the reading skills and academic achievement of students (scores on standardized reading exam, math exam, and Chinese language exam). We also have controlled for the same covariates displayed in Table 1.
[a] All the sample students took the reading exam (4108 students). Then, students in our sample took either a standardized math (2604 students) or Chinese language exam (1504 students).

teachers are not supportive of student independent reading (Fig. 3, Panel C). Only 10.7% of students reported to have received reading instruction during the last semester. Also, 64.4% of sample students reported that their teachers had not encouraged them to borrow independent reading books on a weekly basis during the last semester (Fig. 3, Panel C). In our data set, 35.6% of sample teachers indicated that they do not believe independent reading would have a positive effect on student learning. There were even a number of teachers (6.7% of the sample) that believed independent reading would have a negative effect on the academic performance of students (Fig. 3, Panel C). These findings are consistent with previous studies that found rural Chinese students do not receive sufficient reading instruction in school and that many rural teachers do not recognize the importance of reading (Wang, 2012; Zhang, 2004).

3.3. Impacts of three alternative reading program treatments

In this subsection, we seek to evaluate whether any of the three reading program treatments are effective at improving student reading skills, and whether certain treatments are more effective than others. We find that students in classes that received the *book corner only* treatment did not improve their reading skills relative to students in the control group (Table 3, Column 1, Row 1). Likewise, the reading skills of students in classes that received the *book corner + education bureau training* treatment appeared to improve relative that those of students in the control group; however, this difference is insignificant (Table 3, Column 1, Row 2).

However, we find that the reading skills of students improved when their classes received the *book corner + NGO training* treatment. As can be seen from Table 3, classes that received the *book corner + NGO training* intervention experienced an average improvement of 0.61 SD in standardized reading scores compared to students in the control group (significant at the 1% level—Column 1, Row 3). Although we cannot say for certain why we found statistically significant results for only the *book corner + NGO training* treatment and not the others, it is likely that some element (or elements) unique to the NGO teacher training drive these differential results. Unlike the government's training program, the NGO program trained teachers about the importance of independent reading and methods to promote independent reading, such as providing graded reading materials and organizing classroom reading activities. In contrast, the other interventions offered no training or training that only focused on teaching specific reading skills that are regularly tested on standardized Chinese language exams.

Table 3 also displays that these results are robust to the choice of analytical approach. Specifically, when we control for school fixed effects and compare outcomes between just the *book corner + education bureau training* and *book corner + NGO training* treatment groups, we find that the standardized reading test scores of students in *book corner + NGO training* group are nearly 0.43 SD higher than those of the *book corner + education bureau training* group (significant at the 1% level—Column 2, Row 3). The result is close to the estimated difference of 0.52 SD (significant at the 1% level) when school fixed effects are not used (Column 1, Rows 2 & 3). In other words, the consistency of our results provides evidence that the main OLS regression results are robust to alternative specifications.

The effects of the different reading treatments on the academic achievement of students are also reported in Table 3. The results find that the *book corner only* and the *book corner + education bureau training* interventions have no significant impacts on the

Table 4
PSM Results of treatment effects (for *book corner only*, *book corner + education bureau training*, and *book corner + NGO training* treatments) on the reading skills and academic achievement of sample students.

	ATT		
	Standardized reading score (SD)	Standardized math score (SD)	Standardized Chinese score (SD)
1. *Book corner only* group	−0.15	−0.03	−0.16
	(0.09)	(0.11)	(0.15)
2. *Book corner + Edu bureau training* group	0.05	0.05	−0.03
	(0.07)	(0.08)	(0.10)
3. *Book corner + NGO training* group	0.62***	0.35***	0.37*
	(0.11)	(0.12)	(0.19)

Note: We compare the three treatment groups with the control group. The sample size is the same as in our OLS regression and all students took the reading exam (4108 students) and either a standardized math (2604 students) or Chinese language exam (1504 students). In the control group, there are 1575 students, of whom 988 took the standardized math exam and 587 took standardized Chinese language exam. In the *book corner only* group, there are 573 students, of whom 330 took the standardized math exam and 243 took the standardized Chinese language exam. In the *book corner + education bureau training* group, there are 1605 students, of whom 1005 took the standardized math exam and 600 took the standardized Chinese language exam. In the *book corner + NGO training* group, there are 355 students, of whom 281 took the standardized math exam and 74 students took the standardized Chinese language exam.

standardized mathematics test scores of students (Column 3, Rows 1 & 2). However, it is important to note here that, while there is no positive effect, there also is no negative impact. In other words, reading in our sample did not hurt the math scores of students.

Although we find no significant results for the other two interventions, our results clearly demonstrate that the standardized math test scores of students in classes that received the *book corner + NGO training* intervention were 0.35 SD higher than those of control group students (significant at 1% level—Column 3, Row 3). Our results remain consistent when we control for school fixed effects and compare only students in the *book corner + education bureau training* and the *book corner + NGO training* groups. When no school fixed effects were used, the difference between these two treatment groups is 0.37 SD (significant at the 1% level—Column 3, Rows 2 & 3), which is close to the estimate of 0.34 SD in the model with school fixed effects (significant at the 1% level—Column 4, Row 3).

Our results also find a significant, positive relationship between receiving the *book corner + NGO training* intervention and Chinese language achievement (Table 3). Specifically, the standardized Chinese language test scores of students in classes receiving the *book corner + NGO training* intervention were 0.29 SD higher than those of control students (significant at the 5% level—Column 5, Row 3). When controlling for school-level fixed effects, students in the *book corner + NGO training* group improved their standardized Chinese test scores by 0.57 SD compared to students in the *book corner + education bureau training* group (significant at the 1% level—Column 6, Row 3).

We also find that the results from the OLS and fixed effects analyses are robust to our choice of analytical tool given the results found using PSM analysis (Table 4). Using the PSM analysis, we find that implementing a classroom book corner in conjunction with teacher training provided by the NGO had positive and significant impacts on student outcomes. Specifically, students in classrooms that received the *book corner + NGO training* treatment improved their standardized reading scores by 0.62 SD, on average (significant at the 1% level—Column 1, Row 3). Our results also find that, on average, standardized math test scores increased by 0.35 SD among students in this treatment group compared to the control group (significant at the 1% level –Column 2, Row 3). In addition, there is an estimated coefficient of 0.37 SD on the standardized Chinese scores of children (significant at the 10% level – Column 3, Row 3). As in the OLS analysis above, we also find no impact on classrooms receiving the *book corner only* or the *book corner + education bureau training* treatment.

The implications of these findings are important. When the appropriate treatment is implemented in the proper manner, our results clearly show that not only do the reading skills of students improve, but so do their math and Chinese language scores. These results stand in contrast to the opinions of many rural educators and may be one of the reasons reading is not taught, libraries are not used, and, subsequently, academic outcomes are so poor. The results support the findings that programs providing reading instruction have positive and significant effects on the reading skills and academic achievement of students (Brynes, 2000; Friedlander & Goldenberg, 2016; Manning & Manning, 1984; Topping et al., 2007). For example, Lucas et al. (2014) who found that a reading program providing teacher training and instructional materials improved student academic achievement in written and oral literacy in Uganda.

Although the pathways through which improvements in reading skills lead to improvements in Chinese language performance are relatively intuitive, it is less clear how reading leads to improvements in math performance. Because better reading skills allow students to develop their vocabulary and reading comprehension skills, it follows that these programs also improve the Chinese language outcomes of students. However, there is much less evidence on the relationship between reading skill level and math achievement; the existing evidence shows that reading skills may support math skill development (Jordan, Hanich, & Kaplan, 2003; Rutherford-Becker & Vanderwood, 2009; Thurber et al., 2002). These studies suggest that there is a link between better reading comprehension and critical thinking skills. It may also be as simple as better reading skills allow students to read and understand “word problems” better.

Table 5
OLS Regression of treatment effects of the *book corner + education bureau training*, and *book corner + NGO training* treatments on student reading behaviors and the attitudes and behaviors of teachers toward reading.

Dependent variable	Student spends > 30 min on independent reading daily (1 = yes)	Student borrowed books from his/her school weekly in the past semester (1 = yes)	Teacher encourages students to borrow books from school weekly (1 = yes)	Student received reading instruction in school (1 = yes)	Chinese teacher believes reading has a positive effect on students' math performance (1 = yes)
	(1)	(2)	(3)	(4)	(5)
1. "Book corner + Edu bureau training" Group (1 = yes)	0.21***	0.20***	0.21***	0.10	0.12
	(0.05)	(0.06)	(0.05)	(0.06)	(0.11)
2. "Book corner + NGO training" group (1 = yes)	0.17***	0.18**	0.20**	0.36***	0.36***
	(0.06)	(0.08)	(0.08)	(0.12)	(0.10)
3. Student and family characteristics controlled	yes	yes	yes	yes	yes
4. Teacher characteristic controlled	yes	yes	yes	yes	yes
5. School characteristics controlled	yes	yes	yes	yes	yes
6. Grade dummy	yes	yes	yes	yes	yes
7. Constant	0.38***	0.66***	0.27***	0.03	0.54***
	(0.07)	(0.09)	(0.08)	(0.06)	(0.14)
8. Observations	3535	3535	3535	3535	3535
9. R-squared	0.112	0.129	0.476	0.178	0.096
10. P-value for coefficient [1] = [2]	0.450	0.848	0.971	0.037	0.007

Note: 1. Robust standard errors clustered on class level in parentheses, *** $p < .01$, ** $p < .05$, * $p < .1$.
2. We compare the *book corner + education bureau training*, and *book corner + NGO training* treatments with the control group (total sample = 3535 students). We also have controlled for the same covariates presented in Table 1.

3.4. The mechanisms of the effect of the reading treatment

In this section, we seek to understand the mechanisms through which the types of teacher training resulted in differences in program impacts on reading skills and academic performance. Specifically, we want to explore why the *book corner + NGO training* program benefited educational outcomes while the *book corner + education bureau training* program did not. To do so, we examined student reading behavior and the reading instruction practices of teachers to determine why the two types of training led to different results for student outcomes.

The results show that, although there are improvements in student reading behavior in both the *book corner + education bureau training* and *book corner + NGO training groups*, the positive impacts on the reading skills and academic performance of students in classes that received the *book corner + NGO training* intervention likely arose from changes in how teachers provided reading instruction and the attitudes of teachers toward reading. As seen in Table 5, relative to the control group, students in both treatment groups were more likely to read at least 30 min per day (Column 1, Rows 1 & 2), were more likely to borrow books from school (Column 2, Rows 1 & 2). Additionally, teachers were more likely to encourage students to borrow books (Column 3, Rows 1 & 2). However, these improvements are not significantly different between the *book corner + education bureau training* and the *book corner + NGO training groups* (Columns 1–3, Row 10).

From the information we gathered, we believe the biggest difference between the impacts of the education bureau and NGO trainings is that NGO training prompted more teachers to provide in-class reading instruction as well as induced more positive attitudes toward reading among sample teachers. When compared to students in the control group, students in classes that received NGO training were 0.36 SD more likely to receive reading instruction, and this coefficient is significant at the 1% level (Table 5, Column 4, Row 2). However, we find that students in the *book corner + education bureau training* intervention were only 0.1 SD more likely than the control group to receive instruction, and this value is insignificant (Column 4, Row 1). This coefficient difference between the *book corner + education bureau training* and the *book corner + NGO training* interventions is significant at 5% level (Column 4, Row 10). In addition, we found that the proportion of Chinese teachers in classes receiving the *book corner + NGO training* intervention who believe reading has a positive impact on the math performance of students was 0.36 SD higher than that of teachers in our control group sample (significant at the 1% level—Column 5, Row 2). In contrast, this value was only 0.12 SD higher than that of control group teachers and insignificant for teachers receiving the *book corner + education bureau training* intervention. Therefore, we find that the ability to change the attitudes of teachers toward reading are likely different between the *book corner + education bureau training* and *book corner + NGO training* interventions (significant at 5% level—Column 5, Row 10).

These findings suggest that reading instruction and teacher attitudes play an important role in improving student reading skills and academic performance. According to our interviews with the students and teachers (both before the evaluation survey and after), after the NGO training, most of the treated teachers organized reading classes once or twice a week. In these reading classes, teachers taught students how to read extracurricular books, organized students into groups to discuss book content with their classmates, and encouraged students to present book reports in class. The literature has shown that reading instruction can directly affect the reading skills of students by improving their reading comprehension levels (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; National Research Council;, 1998). Hence, it appears that with improved comprehension levels, students may have been able to read more efficiently and built on their reading skills more effectively.

The belief of teachers about the importance of reading has also been shown to have a large impact on student reading behavior (Richardson, 1996; van Uden, Ritzen, & Pieters, 2014). If teachers support independent reading and incorporate this attitude into their teaching, students have been shown to become more engaged readers. This is also true according to our data. In addition, other studies have shown that positive teacher attitudes can change the way students read. Instead of skimming books and mainly looking at the pictures, they learn to read carefully and learn more efficiently due to their increased independent reading volume (Elley, 1996; Foorman et al., 1998; Wigfield & Guthrie, 2000). For these reasons, we believe it is the ability of the *book corner + NGO training* intervention to both improve in-class instructional practices (for reading skills) as well as the change the attitudes of teachers about the importance of reading that led to better student outcomes.

4. Conclusion

In this paper, we attempt to understand whether reading program treatments that encourage independent reading can improve the reading skills and academic achievement of primary school students in rural China. Our results suggest that, although rural children in our sample had poor reading outcomes at the time of the baseline survey, reading treatments may offer the potential to improve the reading skills and academic achievement of students in rural China. Specifically, we found that increased student access to independent reading materials coupled with effective teacher training led to significant increases in student reading skills, standardized math test scores and standardized Chinese test scores.

The results of our study highlight the importance of program design to achieve the goals of reading treatments. Although we found that student outcomes in classrooms that received a book corner and NGO teacher training improved significantly, no such effects were found in classrooms that implemented only the book corner or implemented the book corner and received teacher training from the local education bureau. These results suggest that training teachers about the importance of independent reading and methods to promote independent reading among students is more effective at improving student reading skills than instruction focusing on skills regularly tested on Chinese language exams. The results also lead us to believe that trainings on independent reading led to changes in reading instruction and teacher attitudes toward reading. It is likely that reading instruction directly improved student comprehension levels, but changes in the attitudes toward reading of teachers improved student outcomes

indirectly. For instance, the positive attitudes of teachers toward reading may have helped students become more engaged and careful readers. As a consequence, the amount of time students spends reading on their own may have increased and allowed students to develop their reading skills more effectively. Taken together, these results support the idea that it is not only reading quantity (reading volume) that is important for improving reading skills, but the quality of reading instruction is also intrinsic to building effective reading programs in developing countries (Elley, 1988; Friedlander & Goldenberg, 2016).

In terms of future policy, our paper has several implications. First, our research has revealed that reading achievement levels are low and access to independent reading resources are deficient in rural Chinese schools. For this reason, decision makers in China's education system must carefully consider how to improve reading skills in rural China. Second, improving reading skills may require more than just adequate access to books. Providing effective teacher training on independent reading appears to be an important input to improving the reading skill levels of rural students. For this reason, we suggest that the Chinese government develop teacher trainings on both the importance of reading and methods of reading instruction that can improve reading skills and boost student interest in reading. We believe that when the reading skills of students improve, improvements in academic achievement in other subject areas will follow.

Acknowledgement

We would like to acknowledge the financial support of the 111 Project (grant number B16031) and the Fundamental Research Funds for Central Universities (grant number 2016CBZ011)

Appendix A. Appendix

Table A
Characteristics that are used for matching between treatment schools and control schools.

Variables	Control group	Treatment group	Difference: (2)–(1)	p-Value of F-test[a]
	Mean (SD)	Mean (SD)	Coef (Se)	
	(1)	(2)	(3)	(4)
1. Pupil-teacher ratio (1 = larger than mean)	0.59 (0.49)	0.49 (0.50)	−0.10 (0.25)	0.676
2. School area (1 = larger than mean)	0.60 (0.49)	0.33 (0.47)	−0.27 (0.24)	0.267
3. Distance from township government (1 = farther than 1 km)	0.41 (0.49)	0.68 (0.47)	0.27 (0.25)	0.280
4. Number of teachers with college degree (1 = > 18 teachers)	0.15 (0.36)	0.16 (0.37)	0.02 (0.20)	0.939
5. Number of Chinese teachers (1 = > 12 teachers)	0.49 (0.50)	0.48 (0.50)	−0.01 (0.25)	0.965
6. Principle graduated from college (1 = yes)	0.29 (0.45)	0.32 (0.47)	0.03 (0.22)	0.877
7. The subject that principle teaches (1 = Chinese language)	0.58 (0.49)	0.52 (0.50)	−0.06 (0.25)	0.809

Note: 1. In the treatment schools, there are 2533 students. In the control schools, there are 1575 students.

[a] To check the balance between the control schools and treatment schools, we run a regression for a set of teacher and school characteristics. For example: Pupil-teacher ratio = a1*treatment school. Then, we test that the coefficients on all control variables were jointly zero. From this we get the the *p*-value.

Appendix B. Appendix

Table B
Covariate balance test among the treatment group and the control group in rural primary schools in Guizhou Province.

Variables	Control Group	Treatment Group	Difference: (2)–(1)	*p*-Value of F-test[a]
	Mean (SD)	Mean (SD)	Coef (Se)	
	(1)	(2)	(3)	(4)
1. Gender (1 = female)	0.49 (0.50)	0.48 (0.50)	−0.01 (0.02)	0.457
2. Mother primarily lives at home (1 = yes)	0.51 (0.50)	0.51 (0.50)	0.00 (0.05)	0.930
3. Father primarily lives at home (1 = yes)	0.40 (0.49)	0.44 (0.50)	0.03 (0.04)	0.377

(*continued on next page*)

Table B (*continued*)

Variables	Control Group	Treatment Group	Difference: (2)–(1)	*p*-Value of F-test[a]
	Mean (SD)	Mean (SD)	Coef (Se)	
	(1)	(2)	(3)	(4)
4. Mother graduated from primary school (1 = yes)	0.26 (0.44)	0.28 (0.45)	0.01 (0.04)	0.722
5. Father graduated from primary school (1 = yes)	0.47 (0.50)	0.45 (0.50)	−0.02 (0.04)	0.578
6. Number of books in household (1 = > 10 books)	0.43 (0.50)	0.39 (0.49)	−0.04 (0.04)	0.277
7. Number of children's books in household (1 = > 10 books)	0.30 (0.46)	0.27 (0.44)	−0.03 (0.02)	0.164
8. Gender of Chinese teacher (1 = female)	0.63 (0.48)	0.44 (0.50)	−0.19 (0.17)	0.282
9. Chinese teacher graduated from college (1 = yes)	0.43 (0.50)	0.36 (0.48)	−0.07 (0.09)	0.427
10. School size (1 = larger than mean)	0.60 (0.49)	0.33 (0.47)	−0.27 (0.24)	0.267

Note: 1. In the treatment schools, there are 2533 students, In the control schools, there are 1575 students.

[a] To check the balance between the control and treatment group, we run a regression for each of the control variables. For example: Gender = a1*treatment group. Then, we test that the coefficients on all control variables were jointly zero. From this we get the p-value.

References

Abeberese, A. B., Kumler, T. J., & Linden, L. L. (2014). Improving reading skills by encouraging children to read in school: A randomized evaluation of the Sa Aklat Sisikat reading program in the Philippines. *The Journal of Human Resources, 49*(3), 611–633.

Adler, M. A., & Fisher, C. W. (2001). Center for the improvement of early reading achievement: Early reading programs in high-poverty schools: A case study of beating the odds. *The Reading Teacher, 54*(6), 616–619.

Authors (2017).

Borkum, E., He, F., & Linden, L. L. (2012). *School libraries and language skills in Indian primary schools: a randomized evaluation of the Akshara library program (NBER working paper no. 18183)*. Cambridge, MA: National Bureau of Economic Research. Retrieved from: http://www.nber.org/papers/w18183.pdf.

Brynes, J. P. (2000). Using instructional time effectively. In L. Baker, M. J. Dreher, & J. T. Guthrie (Eds.). *Engaging young readers* (pp. 188–208). New York: Guilford Press.

Caygill, R., & Chamberlain, M. (2004). Progress in international reading literacy study (PIRLS): New Zealand's year 5 student achievement 2001. Comparative education research unit, research division, new Zealand Ministry of Education. Retrieved from https://www.educationcounts.govt.nz/publications/series/2539/PIRLS-2001/5287.

Cheung, W. M., Tse, S. K., Lam, J. W., & Ka Yee Loh, E. (2009). Progress in International Reading Literacy Study 2006 (PIRLS): Pedagogical correlates of fourth-grade students in Hong Kong. *Journal of Research in Reading, 32*(3), 293–308.

Cunningham, A. E., & Stanovich, K. E. (2001). What reading does for the mind. *Journal of Direct Instruction, 1*(2), 137–149.

Elley, W. B. (1996). Lifting literacy levels in developing countries: Some implications from an IEA study. In V. Greaney (Ed.). *Promoting reading in developing countries* (pp. 39–54). Newark, DE: Interactional Reading Association.

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Education & Psychology, 90*(1), 37.

Friedlander, E., & Goldenberg, C. (2016). *Literacy boost in Rwanda: Impact evaluation of a 2-year randomized control trial.* Stanford, CA: Stanford University. Retrieved from https://rwanda.savethechildren.net/sites/rwanda.savethechildren.net/files/library/LB%20Rwanda%20-%202%20Year%20Impact%20Evaluation_1.pdf.

GMW.cn (2015). Reading initiative in China. Retrieved June 4, 2017, from http://epaper.gmw.cn/gmrb/html/2015-01/06/nw.D110000gmrb_20150106_3-11.htm.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*(3), 257–288.

Imai, K., King, G., & Nall, C. (2009). The essential role of pair matching in cluster-randomized experiments, with application to the Mexican universal health insurance evaluation. *Statistical Science, 24*(1), 29–53.

Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with comorbid mathematics and reading difficulties. *Child Development, 74*, 834–850.

Kassow, D. Z. (2006). Parent-child shared book reading: Quality versus quantity of reading interactions between parents and young children. *Talaris Research Institute, 1*(1), 1–9.

Kim, J. S. (2006). Effects of a voluntary summer reading intervention on reading achievement: Results from a randomized field trial. *Educational Evaluation and Policy Analysis, 28*(4), 335–355.

Kim, J. S., & Quinn, D. M. (2013). The effects of summer reading on low-income children's literacy achievement from Kindergarten to grade 8: A meta-analysis of classroom and home interventions. *Review of Educational Research, 83*(3), 386–431.

Lin, T. (2007). Analysis on the present situation of rural primary school reading. *Journal of Ningbo University (Educational Science Edition), 29*(2), 146–147 (in Chinese).

Liu, Y. (2015). How to improve rural primary school students' extracurricular reading. *Chinese Journal of Off-Campus Education, 7*(2) (in Chinese).

Lonigan, C. J., & Shanahan, T. (2009). Developing early literacy: Report of the national early literacy panel. Washington, D.C.: National Institute for Literacy. Retrieved from https://eric.ed.gov/?id=ED508381.

Lucas, A. M., McEwan, P. J., Ngware, M., & Oketch, M. (2014). Improving primary school quality across countries: Experimental evidence from Kenya and Uganda. *Journal of Policy Analysis and Management, 33*(4), 950–976.

Manning, G. L., & Manning, M. (1984). What models of recreational reading make a difference? *Literacy Research and Instruction, 23*(4), 375–380.

Ministry of Education of the People's Republic of China. Rules for libraries in primary and secondary schools in China. (2003). [accessed June 4, 2017]. Retrieved from: http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_35/201006/88596.html (in Chinese) .

Mullis, I. V., Martin, M. O., & Gonzalez, E. J. (2004). *International achievement in the processes of reading comprehension: Results from PIRLS 2001 in 35 countries.* Chestnut Hill, MA: International Association for the Evaluation of Educational Achievement, Boston College. Retrieved from: https://timssandpirls.bc.edu/pirls2001i/pdf/Processes.pdf.

National Bureau of Statistics of China (2015). China statistical yearbook 2015. Beijing: China Statistics Press.

National Reading Panel (US), National Institute of Child Health & Human Development (US) (2000). Report of the national reading panel: Teaching children to read:

An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups. *National Institute of Child Health and Human Development, National Institutes of Health*. Retrieved from: https://www.nichd.nih.gov/publications/pubs/nrp/documents/report.pdf.

National Research Council (1998). *Preventing reading difficulties in young children.* Washington, D.C.: National Academies Press.

People.cn (2015). Report on the work of the government. Retrieved from: http://www.people.com.cn/n/2015/0305/c347407-26643598.html, Accessed date: 5 June 2017.

Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula (Ed.). *Handbook of research on teacher education*(second edition). New York: Macmillan Publishing.

Roskos, K., Strickland, D., Haase, J., & Malik, S. (2009). *First principles for early grades reading programs in developing countries. International reading association.* Washington, D.C.: USAID. Retrieved from: http://www.equip123.net/docs/e1-EarlyGradesToolkit.pdf.

Rutherford-Becker, K. J., & Vanderwood, M. L. (2009). Evaluation of the relationship between literacy and mathematics skills as assessed by curriculum-based measures. *California School Psychologist, 14*, 23–34.

Slavin, R. E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis. *Reading Research Quarterly, 43*(3), 290–322.

Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research, 79*(4), 1391–1466.

Statistics Bureau of Guizhou Province (2015). *Guizhou statistical yearbook 2015.* Beijing: China Statistics Press.

Sun, Y., & Xu, Y. (2003). *Reading methods in recent ten years. Curriculum, Teaching Material and Method. Vol. 6*, 31–37 (in Chinese).

Thurber, R. S., Shinn, M. R., & Smolkowski, K. (2002). What is measured in mathematics tests? Construct validity of curriculum-based mathematics measures. *School Psychology Review, 31*(4), 498.

Topping, K. J., Samuels, J., & Paul, T. (2007). Does practice make perfect? Independent reading quantity, quality and student achievement. *Learning and Instruction, 17*(3), 253–264.

Topping, K. J., & Sanders, W. L. (2000). Teacher effectiveness and computer assessment of reading relating value added and learning information system data. *School Effectiveness and School Improvement, 11*(3), 305–337.

Tunmer, W. E., Chapman, J. W., Greaney, K. T., Prochnow, J. E., & Arrow, A. W. (2013). Why the New Zealand National Literacy Strategy has failed and what can be done about it: Evidence from the Progress in International Reading Literacy Study (PIRLS) 2011 and Reading Recovery monitoring reports. *Australian Journal of Learning Difficulties, 18*(2), 139–180.

van Uden, J. M., Ritzen, H., & Pieters, J. M. (2014). Engaging students: The role of teacher beliefs and interpersonal teacher behavior in fostering student engagement in vocational education. *Teaching and Teacher Education, 37*, 21–32.

Wang, L. (2012). Study on the practices of student independent reading in primary schools. *China Academic Journal Electronic Publishing House, 8*, 30–33 (in Chinese).

Wang, S., Sun, Y., & Wang, B. (2013). An indexing system for evaluating children reading promotion programs in public libraries. *Journal of Library Science in China, 6*, 41–52 (in Chinese).

Wang, X., Liu, C., Zhang, L., Shi, Y., & Rozelle, S. (2013). College is a rich, Han, urban, male club: Research notes from a census survey of four tier one colleges in China. *The China Quarterly, 214*, 456–470.

Whitehurst, G. J., & Lonigan, C. J. (2001). Emergent literacy: Development from pre-readers to readers. In S. B. Neuman, & D. K. Dickinson (Eds.). *Handbook of early literacy research* (pp. 11–29). New York, NY: Guilford Press.

Wigfield, A., & Guthrie, J. T. (2000). Engagement and motivation in reading. In S. B. Neuman, & D. K. Dickinson (Eds.). *Handbook of early literacy research* (pp. 403–422). New York, NY: Guilford Press.

Zhang, L., Yi, H., Luo, R., Liu, C., & Rozelle, S. (2013). The human capital roots of the middle income trap: The case of China. *Agricultural Economics, 44*(s1), 151–162.

Zhang, X. H. (2004). Current state of extracurricular reading and reading guidance among rural primary school students. *Scientific Consultation, 11*(45), 87–88 (in Chinese).