

## Many Children Left Behind? Textbooks and Test Scores in Kenya<sup>†</sup>

By PAUL GLEWWE, MICHAEL KREMER, AND SYLVIE MOULIN\*

*A randomized evaluation in rural Kenya finds, contrary to the previous literature, that providing textbooks did not raise average test scores. Textbooks did increase the scores of the best students (those with high pretest scores) but had little effect on other students. Textbooks are written in English, most students' third language, and many students could not use them effectively. More generally, the curriculum in Kenya, and in many other developing countries, tends to be oriented toward academically strong students, leaving many students behind in societies that combine a centralized educational system; the heterogeneity in student preparation associated with rapid educational expansion; and disproportionate elite power. (JEL O15, I21, I28, J13)*

Many economists argue that increasing educational expenditure will have a limited impact on learning in distorted educational systems (e.g., Eric A. Hanushek 1995 and Lant Pritchett and Deon Filmer 1999). Yet even skeptics of the impact of education spending in such systems believe that providing textbooks to schools where they are scarce can substantially increase test scores (see reviews by Stephen P. Heyneman, Joseph P. Farrell, and Manuel A. Sepulveda-Stuardo 1978; Bruce Fuller 1986; Marlaine E. Lockheed and Hanushek 1988; and Fuller and Prema Clarke 1994). Indeed, one political economy model of distortions in education suggests that spending on nonteacher inputs will raise student performance much more

\* Glewwe: University of Minnesota, Department of Applied Economics, 1994 Buford Ave., St. Paul, MN 55108 and World Bank (e-mail: [pglewwe@umn.edu](mailto:pglewwe@umn.edu)); Kremer: Harvard University, Department of Economics, Littauer Center M20, Cambridge, MA 02138, Brookings Institution, Center for Global Development, and NBER, (e-mail: [mkremer@fas.harvard.edu](mailto:mkremer@fas.harvard.edu)); Moulin: Rabat American School, BP 120 c/o American Embassy, Agdal, Rabat, Morocco (e-mail: [smoulin-slimi@ras.ma](mailto:smoulin-slimi@ras.ma)). This project was a collaboration by many people. The data were collected and analyzed by Charles Asoka, Najy Benhassine, Marcos Chamon, Edward Drozd, Daniel Haar, Nauman Ilias, Dan Levy, Irene Majale, Sean May, Ted Miguel, Robert Namunyu, Caroline Nekesa, Stacy Nemeroff, Jonah Rockoff, Jaypee Sevilla, Michael Wambetsa, Polycarp Waswa, Stanley Watt, and Maureen Wechuli. Invaluable assistance was provided by the staff of International Christelijk Steunfonds: Chip Bury, Jos Huizinga, Paul Lipayah, Japheth Mahagwa, Rebecca Mbaisi, and Susan Walji. We are grateful to Joshua Angrist, Angus Deaton, Paul Gertler, Claudia Goldin, Douglas Gollin, Eric Hanushek, Lawrence Katz, Marlaine Lockheed, Richard Murnane, Steve Pischke, and Duncan Thomas for advice and comments. Finally, we are grateful to Joash Patrice Munala, the district education officer of Busia; to the late George Buluma, the inspector of primary schools in Busia; and to the headmasters, teachers, and students of the participating schools. The cost of this evaluation was covered by the National Science Foundation and the World Bank research committee. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank, its executive directors, or the countries they represent.

<sup>†</sup> To comment on this article in the online discussion forum visit the articles page at:  
<http://www.aeaweb.org/articles.php?doi=10.1257/app.1.1.112>

than increased spending on teachers (Pritchett and Filmer 1999). Policy makers appear to agree with this view. For example, when the World Bank increased loans to Kenya after the Moi regime ended, one of its first loans was for a major textbook supply program.

We report the results of a randomized trial program that provided textbooks to rural Kenyan primary schools. Unlike previous studies, we find that textbooks had little effect on the typical student. The results do not appear to be statistical artifacts. The treatment and comparison schools were similar in geographic location, enrollment, and pre-program test scores. Neither selection nor attrition bias appears to drive the results.

Why did textbooks not raise test scores? One clue is that textbooks appear to have raised the scores of students with higher pretest scores. An interaction term between pretest scores and assignment to the textbook program has a highly significant positive impact on post-test scores. Consistent with the hypothesis that textbooks helped only strong students, students who made it to the selective final year of primary school (grade 8) were more likely to enter secondary school if they were in schools that received textbooks. Moreover, a randomized evaluation of a program that gave grants to similar schools, half of which were spent on textbooks, also found that most of the benefits accrued to the strongest students. There is little evidence that textbooks reduced grade repetition or dropping out, consistent with the finding of no impact on weaker students.

The finding that the textbooks provided were of little use to many students is plausible. English is the medium of instruction in Kenyan schools but is the third language of many pupils, including those examined here. Moreover, pupil and teacher absence rates are high, so many pupils fall behind the official curriculum. Once they do, it may be difficult to catch up, since our data show that many students cannot effectively read and comprehend the English textbooks.

This raises a larger issue. It is difficult for Kenya's centralized, uniform education system to serve the entire population, given the vast heterogeneity in the educational and economic backgrounds of students generated by a rapidly expanding education system. The legacy of colonial education and the political economy of post-independence Kenya may have led to an educational system that favors the most advantaged students.

Many other developing countries appear to have a similar mismatch in curricula. Particularly after the early primary years, many developing countries instruct students in a language that is not their mother tongue, often that of the former colonial power or a local majority group. Mother tongue instruction is the norm for the first three years of primary education in Africa but not thereafter (Ayo Bamgbose 2004). In India, most primary schools teach in the most common regional language, but 87 percent of secondary schools teach in another language (National Council of Educational Research and Training 2002).

Poor performance on national and international standardized tests and high repetition and dropout rates suggest that many developing country pupils fall behind the curriculum. For example, in Sri Lanka, academic tests given to grade 5 students revealed that only 37 percent had mastered grade 4 reading skills (in Sinhala or Tamil), and only 38 percent had mastered grade 4 math skills (World Bank 2004). In

developing countries, each year 6.2 percent of primary students repeat a grade, compared to 0.8 percent in developed countries; and 25.5 percent drop out of primary school, compared to 2.3 percent in developed countries (UNESCO 2006). The mismatch in curricula may contribute to high dropout rates and low learning in many developing nations. Our results suggest that distortions in education systems due to political economy factors may go beyond overspending on teachers relative to non-teacher inputs, compromising even those policies with apparently clear benefits such as textbook provision.

This paper is organized as follows. Section I describes primary education in Kenya and explains the design of the textbook program. Section II presents evidence that the program had little impact on average scores. Section III argues that the program benefited only students with the highest initial achievement. Section IV interprets these results in terms of the political economy of education and concludes.

### **I. Background: Primary Education in Kenya and the School Assistance Program**

This section describes Kenya's primary education system and how the textbook program was implemented. It then discusses the selection of schools, the tests, initial conditions in textbook and comparison schools, and the program's impact on textbook access and on pedagogy.

#### *A. Primary Education in Kenya*

In the years covered in this study, Kenya's Ministry of Education set the curriculum, administered national and district exams, and hired all teachers. Local parent-run school committees had to pay for almost all other school costs. Major capital expenses such as construction were financed by large fundraising drives called *harambees*. Recurrent costs (minor repairs, chalk, books for teachers, etc.) were covered by school fees. In practice, parents and headmasters often bargained over how much of the fees parents had to pay.

Almost all Kenyan children start primary school. Grades 1, 2, and 3 are taught in English, Kiswahili, and the local language (Kiluhya in two-thirds of our sample, Ateso in the rest). After grade 3, all instruction is in English. Average class size in our sample ranges from 49 students in grade 1 to 18 students in grade 8. At the end of grade 8, students take the national Kenya Certificate of Primary Education (KCPE) exam, which determines the secondary schools they can attend. To maintain high average scores on the KCPE, some primary schools promote only strong students to grade 8; other students must either repeat grade 7 or drop out. Students also drop out in earlier grades. In our sample, 35 percent of the students in grade 3, when the program began, had dropped out after three years. Schools usually had textbooks for teachers but few for students. A 1990 Ministry of Education survey found a pupil to textbook ratio of 17 to 1 in primary schools.

When the project started, the schools had very few textbooks. About 80 percent of the students in the sample were in classrooms with less than one English textbook for every 20 students. The analogous figures for math and science texts were 78 percent and 89 percent, respectively. In response, some parents purchased textbooks.

In our sample 80 percent to 90 percent of the textbooks students had were bought by their parents. Even after including textbooks purchased by parents for students in grades 3, 4, and 5, only one out of six students had English and math textbooks. In grades 6 and 7, one in four had these textbooks. Very few had textbooks in other subjects. Students in grade 8 had more textbooks. About 40 percent of grade 8 students had math and English textbooks. Access to textbooks was somewhat better, since two or three students usually shared a bench, and if one had a textbook they would share it.

### *B. Textbook and Grant Provision*

In 1995, the Ministry of Education district office selected 100 of the 333 primary schools in Kenya's Busia and Teso districts to participate in the School Assistance Program (SAP) funded by International Christelijk Steunfonds (ICS), a Dutch non-profit organization. Busia and Teso schools are typical of others in Kenya. In 1995, average KCPE scores in these districts were roughly at the median for Kenya as a whole. In early 1996, 25 out of 100 schools were chosen randomly to receive the official government textbooks. (Kenya's school year begins in January and ends in November, with long spring and summer breaks.) Students in grades 3–7 received English textbooks, and students in grades 3, 5, and 7 received math textbooks. Grade 8 students received science textbooks since many already had math and English texts. In early 1997, math textbooks were given to students in grades 4 and 6, and agriculture texts to students in grade 8. The associated teacher's guide was also given in each grade and subject where textbooks were given.

The Kenyan textbooks used cost \$2–\$3, while per capita GDP in Kenya in 1997 was \$330. In rural areas, incomes are substantially lower than average, so the annual cost of buying a set of textbooks (one for each subject) could approach 10 percent of a rural household's per capita income.

Textbooks were given at less than a one-to-one ratio since Heyneman, Dean T. Jamison, and Xenia Montenegro (1984) found little difference in test scores between Philippine schools randomly selected to receive one textbook for every two pupils and other schools randomly selected to receive one textbook for each pupil. A 60 percent textbook per pupil ratio was used for English and science, and a 50 percent ratio for math. Pupils in grades 3–5 could not take textbooks home. Pupils in grades 6–8 were put into pairs to share textbooks by taking them home on alternate days.

In 1997, another 25 of the 100 schools were selected to receive grants equal to \$2.65 per student or, on average, \$727 per school. After transport costs, 43 percent of this money was spent on new textbooks, 46 percent on construction, and the remainder on equipment and supplies. The short-run effect of grants is likely to primarily reflect textbook purchases because classroom construction in Kenya often takes years, and the grants were too small to pay for an entire classroom.

### *C. School Selection*

The 100 SAP schools were chosen because they were deemed to be needy and (with one exception) were not part of an earlier textbook project (discussed in Section

II). In 1995, the median school average test score among SAP schools on the district exam for grades 6 and 7 was at the 40th percentile of the distribution of school average test scores in Busia. On the grade 8 exam, the median SAP school was at the 33rd percentile of the distribution for all schools in Busia. Thus, student performance in the 100 SAP schools was somewhat below the average for all 333 schools in Busia and Teso.

The 100 SAP schools were randomly divided into four groups as follows. Schools were listed alphabetically within geographic divisions. These lists were combined, into a single list in alphabetical order by division names. From this list, every fourth school, starting with the first, was assigned to group 1. Similarly, every fourth school starting with the second, third, and fourth was assigned to groups 2, 3, and 4, respectively. Group 1 schools received textbooks in early 1996. In early 1997, group 2 schools received grants for the purchase of educational materials (including textbooks). Group 3 and 4 schools received similar grants in early 1998 and 2000, respectively. Henceforth, 1995, 1996, 1997, 1998, 1999, and 2000 are referred to as years 0, 1, 2, 3, 4, and 5 of the program, respectively.

#### D. Description of Tests

The Ministry of Education administers district-wide exams to upper-grade primary school students to measure their knowledge of the official curriculum. As explained above, grade 8 students also take the KCPE exam.<sup>1</sup> In years 1–4 of the program, ICS administered additional tests in the 100 SAP schools. These were modeled on the district and KCPE exams by Ministry of Education staff. At the start of year 1, ICS administered baseline tests in English, math, and science for grades 3–8 in the 25 group 1 schools (which received textbooks at the start of year 1) and the 25 group 4 schools (which were not assisted until year 5). At the end of years 1 and 2, ICS administered exams in all 100 schools to grades not participating in the district exams (grades 3 and 4 in year 1 and grades 3–7 in year 2). In years 3 and 4, ICS gave grades 3–8 exams even though most grades had district exams.

There are three sets of comparison schools for the 25 schools that received textbooks in year 1. At the end of year 1, they can be compared to all schools that had not yet received assistance—groups 2, 3, and 4. We call these the 75-school comparison group. In year 2, they can be compared to the 50 schools in groups 3 and 4 that had not yet been assisted, called the 50-school comparison group. In years 3 and 4, they can be compared to the 25 group 4 schools that were assisted in 2000, called the 25-school comparison group. Any results using the year 1 (January 1996) pretest scores must compare the textbook schools to the 25-school comparison group, since only those schools have pretest scores.

<sup>1</sup> District exams are given in October for grades 4 through 7 and in July for grade 8. The KCPE is given in November. Unless otherwise stated, we use October district exam results for grades 4 through 7 and KCPE results for grade 8. In year 2, no October district exams were given due to a national decree unrelated to the textbook program. In year 3, they were given only in Busia, not in Teso.

### E. Initial Conditions

The 25 textbook schools and the 25-school comparison group had very similar pre-program test scores (Table 1). For each grade and subject combination, scores were normalized by subtracting the mean in the comparison group and then dividing by the comparison group standard deviation. All regression estimates use school random effects since students in the same school may share common effects such as headmaster quality.

Averaging over grades, pre-program differences in English and math scores between group 1 and group 4 schools are never more than 0.05 standard deviations, and are never statistically significant. This is also true for science when averaged over all grades. But for grade 8 science (the only grade given science textbooks) the difference is larger, 0.17, and statistically significant (10 percent level). Regressions combining all grade-subject combinations show small, insignificant differences, 0.06 standard deviations for grades given textbooks and 0.02 for all grades.

### F. Impact on Textbook Availability and Pedagogy

In years 1 and 2, the program greatly increased the supply of textbooks in the textbook schools. Yet, the program impact fell over time as books depreciated and comparison schools obtained books from other sources. In year 1, the ratio of school-owned books per pupil for grade-subject combinations that received textbooks was 0.65 in textbook schools but only 0.04 in the comparison schools (Table 2). In contrast, in the grade-subject combinations not given textbooks, the ratios were identical for both groups. In grade-subject combinations that received textbooks, textbook schools had fewer privately owned textbooks than comparison schools, 0.10 versus 0.18. Perhaps the program crowded out 0.08 private books per student in year 1. (Private textbook ownership is correlated slightly with initial test scores, ranging from 0.02 to 0.08.) Yet, this crowding out is small relative to the difference in school-owned books. Combining school and private books, the ratio was 0.75 for textbook schools and 0.22 for comparison schools, a gap of 0.53. In contrast, in the grade-subject combinations not given textbooks the ratios were very close, 0.11 and 0.13.

Similar results hold in later years, but the gap between textbook and comparison schools narrows over time in grade-subject combinations that received textbooks. At the end of year 2, a new program that distributed textbooks to 21 of the 100 schools further narrowed the gap in textbook availability between treatment and comparison schools. Moreover, in year 2, there were some curriculum changes and new editions of some textbooks. The changes were modest, but some teachers may not have wanted pupils to use a version of the text older than the one they were teaching from, so differences across schools in effective numbers of textbooks may have declined. The receipt of new textbooks by some schools at the end of year 2 and the curriculum and textbooks changes in year 2 lead us to focus our analysis on years 1 and 2.

Students in textbook schools report having much more access to school-owned textbooks in grade-subject combinations that received textbooks. Sixty-two percent of textbook school students in year 2 report having access to a school-owned book in class for the grade-subject combinations given textbooks, compared to only 8 percent

TABLE 1—DIFFERENCES IN NORMALIZED PRETEST SCORES BETWEEN TEXTBOOK SCHOOLS AND 25-SCHOOL COMPARISON GROUP

Subject	English		Math		Science		All subjects combined	
	Grades with texts (3–7)	All grades (3–8)	Grades with texts (3, 5, 7)	All grades (3–8)	Grades with texts (8)	All grades (3–8)	Grades with texts	All grades
Difference between textbook schools and comparison schools	0.046 (0.105)	0.033 (0.101)	0.056 (0.090)	0.054 (0.085)	0.173 (0.105)	−0.017 (0.088)	0.061 (0.091)	0.023 (0.087)
Observations	8,516	9,332	5,069	9,302	816	9,276	14,401	27,910

Notes: Each column represents a regression of pretest scores from January of year 1 on a constant and a dummy variable for being in a textbook school, with school random effects. The sample consists of all students from the 25 textbook schools and the 25-school comparison group who took the pretest in January of year 1. Columns 1–6 combine different grades and include dummy variables for each grade. Columns 7 and 8 combine subjects and grades and have dummy variables for each grade/subject combination. Columns 1, 3, 5, and 7 exclude grade/subject combinations that did not receive textbooks. Standard errors in parentheses.

TABLE 2—AVAILABILITY OF TEXTBOOKS PER PUPIL

Program year	Subject/grade given textbooks?	School-owned books		Privately-owned books		Total	
		Textbook schools	Comparison schools <sup>a</sup>	Textbook schools	Comparison schools	Textbook schools	Comparison schools
1	Y	0.65	0.04	0.10	0.18	0.75	0.22
	N	0.03	0.03	0.08	0.10	0.11	0.13
2	Y	0.55	0.04	0.09	0.17	0.64	0.21
	N	0.04	0.03	0.08	0.08	0.12	0.12
3	Y	0.52	0.11	0.09	0.14	0.61	0.25
	N	0.11	0.09	0.09	0.09	0.20	0.19
4	Y	0.43	0.10	0.05	0.11	0.48	0.21
	N	0.10	0.08	0.05	0.06	0.17	0.14

Notes: Textbook availability is calculated using school questionnaire data collected at the start of each school year and data on privately owned textbooks from a pupil questionnaire given to pupils in grades 6–8 (data on privately owned textbooks for pupils in grades 3–5 are from the school questionnaire). Results for years 1, 2, and 3 aggregate over grades 3–8 and over three subjects: English, math, and science. For year 4, results are only for grades 6 to 8, and only in math and English.

<sup>a</sup> Comparison schools: 75-school group in year 1, 50-school group in year 2, and 25-school group in years 3 and 4.

in the 50-school comparison group (Table 3, columns 1 and 2). By year 3, the difference is smaller, 72 percent for students in textbook schools versus 28 percent for comparison school students. Over half of textbook school students report that they can take home school texts in subjects for which textbooks were given (Table 3, columns 3 and 4) compared to less than 10 percent of comparison school students.

Finally, trained observers visited classrooms to see whether the program affected pedagogy. There were few noticeable effects. The main difference is that pupils used textbooks in class more often in textbook schools than in comparison schools, although differences were modest and dissipated by year 3. Specifically, in year 2 textbooks were used in 62 percent of the classroom observation sessions in the textbook schools but only in 46 percent of the observation sessions in the comparison schools, a difference significant at the 5 percent level. By year 3 this difference was smaller (46 percent in textbook schools versus 37 percent in comparison schools) and statistically insignificant. The difference in textbook use in year 2 reflects differences in both teachers' presence in the classroom and use of textbooks conditional

TABLE 3—STUDENT REPORTING ON AVAILABILITY OF SCHOOL-OWNED TEXTBOOKS IN GRADES 6–8

Year	Type of subject/grade combination	School issued you a textbook to use in class?		School allowed you to take the textbook home?	
		Textbook schools (1)	Comparison schools (2)	Textbook schools (3)	Comparison schools (4)
2	Textbooks provided	62.4 %	7.7 %	52.8 %	2.5 %
	Textbooks not provided	8.6 %	7.1 %	5.4 %	1.9 %
3	Textbooks provided	72.0 %	28.3 %	63.5 %	9.4 %
	Textbooks not provided	23.4 %	11.7 %	17.4 %	6.4 %

*Notes:* These figures are averages over groups of grade/subject combinations, disaggregated according to whether the combination received textbooks from ICS. In both years, “textbooks provided” refers to English and math in grades 6 and 7 and science in grade 8, while “textbooks not provided” refers to science in grades 6 and 7 and English and math in grade 8. This information is available only for years 2 and 3, since the relevant student questionnaires were administered only in those years.

on being present. The only other pedagogical difference is a small increase in homework assigned in textbook schools.<sup>2</sup>

## II. Program Effect on Average Test Scores

This section discusses our econometric specification and presents two estimators of the effect of textbooks on average test scores: a level estimator and a difference-in-differences estimator. Both estimators suggest little effect of textbooks on average test scores. Glewwe, Kremer, and Moulin 2007 report results from a third estimator, a differences-in-differences, subject-based estimator that compares the difference in textbook schools between test scores in subject-grade combinations in which textbooks were and were not given to the same difference in comparison schools. This estimator suggests that one can rule out effects of textbooks on average test scores as low as 0.07 standard deviations. However, such an estimator would be subject to bias if there are learning spillovers from one subject to another. Robustness checks show that the results for the first two estimators are not driven by selection or attrition bias. We then show that retrospective estimates yield misleading results in the Kenyan context.

### A. Econometric Specification

For each grade and subject combination, we normalize test scores by subtracting the mean score, and dividing by the standard deviation in the 75 comparison schools. Thus, a student with a normalized score of 0.1 is 0.1 standard deviations above the mean. For reference, moving from 0.0 to 0.1 standard deviations in a normal distribution moves a student from the 50th to the 54th percentile.

Test scores may be correlated among students in the same class and school due to unobserved teacher and headmaster characteristics, so we use an error components

<sup>2</sup> The classroom observation data provide a third source of information on textbook availability, but it is difficult to compare textbook availability using these data because textbook per student ratios were collected only if textbooks were used when classrooms were observed, and textbooks were used more often in textbook schools than in comparison schools. Moreover, the data provide information only on whether pupils had textbooks at school on a given day, not overall textbook access. Some students may have left textbooks at home.



econometric model with school, grade, and subject random effects.<sup>3</sup> Assume that the test score of student  $i$  in grade  $j$ , subject  $k$ , and school  $s$ ,  $t_{ijks}$ , is given by

$$(1) \quad t_{ijks} = \alpha_{jk} + \beta_{jk}p_s + u_{jks} + e_{ijks} \quad j = 3, 4, \dots, 8 \quad k = \text{English, math, science.}$$

The dummy variable  $p_s$  indicates whether school  $s$  received textbooks. Random assignment of textbooks to schools ensures that  $E[p_s u_{jks}] = E[p_s e_{ijks}] = 0$ . All estimates of (1) use Generalized Least Squares (GLS) to account for within-school correlation of students' test scores for a given grade and subject without imposing a specific distribution (e.g., normality) on the residuals.

We also combine several grades in a subject to estimate the (weighted) average impact of textbooks on test scores in that subject. This specification decomposes  $u_{jks}$  in (1) into a school specific term,  $u_{ks}$ , and a grade specific term conditional on the school term,  $v_{jks}$ :

$$(2) \quad t_{ijks} = \alpha_{3k}D_{3i} + \alpha_{4k}D_{4i} + \dots + \alpha_{8k}D_{8i} + \beta_k p_s + u_{ks} + v_{jks} + e_{ijks}$$

$k = \text{English, math, science.}$

Equation (2) includes grade specific intercepts with corresponding dummy variables.

Our final estimates combine all grades and subjects for which data are available:<sup>4</sup>

$$(3) \quad t_{ijks} = \alpha_{3E}D_{3Ei} + \alpha_{3M}D_{3Mi} + \alpha_{3S}D_{3Si} + \dots + \alpha_{8E}D_{8Ei} + \alpha_{8M}D_{8Mi}$$

$$+ \alpha_{8S}D_{8Si} + \beta p_s + u_s + w_{js} + v_{jks} + e_{ijks}.$$

Equations (2) and (3) are also estimated by GLS to account for their additive error structure.

For each year, the sample comprises all students tested in October of that year who were enrolled in January of year 1 in the 25 textbook schools or the relevant comparison group. Pupils who changed schools after January of year 1 are always classified by their initial school, so the estimated program effect is the impact of being offered the treatment (intention to treat) not the impact of the treatment itself (selection and attrition bias are discussed below). The comparison schools are the

<sup>3</sup> An alternative specification is to allow for arbitrary correlation across students in the same schools (and classrooms) by using a "robust" variance-covariance matrix that allows for such correlation. We reestimated our main results using this specification and found very similar results. For example, in the first column of Table 4, using this approach gives the same point estimate (0.022) and a slightly smaller standard error (0.085). For the third column of Table 4, it gives a lower point estimate (-0.014) and a slightly smaller standard error (0.049).

<sup>4</sup> Equation (3) assumes that students have the same teacher for all subjects, so that  $w_{js}$  is a teacher-specific effect and  $v_{jks}$  is a subject-specific effect conditional on having that teacher. In upper grades, teachers specialize by subject, so the error term should be  $u_s + w_{ks} + v_{jks} + e_{ijks}$ , where  $w_{ks}$  is a teacher specific effect for the teacher teaching subject  $k$  in all grades, and  $v_{jks}$  is the grade-specific impact of that teacher. In practice, these two different error structures for equation (3) yield similar results. Also, adding an individual-level random effect when stacking across subjects for equation (3) had almost no effect on the estimates and only slightly reduced estimated standard errors.

75-school comparison group for year 1, the 50-school group for year 2, and the 25-school group for years 3 and 4. The smaller sample sizes for years 3 and 4, and the changes in curriculum and textbook availability discussed above lead us to focus on years 1 and 2.

### *B. Impact on Average Test Scores*

Aggregating over all grades and subjects, equation (3) gives a statistically insignificant impact of 0.02 standard deviations (Table 4, column 1). The standard error of 0.087 allows one to reject (at the 5 percent level) the hypothesis that the true (average) effect was 0.20 or higher. Separate estimates by sex (not shown in Table 4) revealed very little difference, by gender, in the impact of the program.

There is no evidence that the impact of textbooks rose over time. At the end of year 2, all textbook school students in grades 4–7 had had English textbooks for two years, and those in grades 4 and 6 had had math textbooks for two years. Aggregating over both subjects for these grades yields a statistically insignificant impact of 0.02 (Table 4, column 2), the same as the estimated effect after one year, ruling out an average impact of 0.23 or higher (at the 5 percent significance level). Estimates of program impacts for years 3 and 4, reported in the working paper version of this paper, are slightly negative and rule out impacts of 0.20 or higher.

Comparing differences in pretest and post-test scores across textbook and comparison schools yields more precise estimates (Table 4, columns 3 and 4). These differences-in-difference estimates for the schools that participated in the year 1 pretest, aggregated over subjects, yield estimates of the treatment effect of 0.02 and  $-0.05$  after one and two years, respectively. Their smaller standard errors rule out impacts of 0.13 and 0.11 or higher. Estimates after three and four years (not reported in Table 4) rule out impacts of 0.10 and 0.08 or higher, respectively.

Thus, both types of estimates show impacts of textbooks close to zero. Depending on the method used, one can reject effects as small as 0.20 or 0.13 standard deviations.

### *C. Robustness Checks*

The results do not appear to be due to selection or attrition bias. Dropout and transfer rates from year 0 to year 1 across textbook and comparison schools show small, statistically insignificant differences (Table 5, first two rows), suggesting that selection into textbook schools is not a serious concern.

Textbook schools may have promoted more pupils to grades 3, 5, and 7, hoping to receive more textbooks. From year 0 to year 1, 22 percent of textbook school students repeated a grade, while 26 percent did so in the comparison schools (Table 5, row 3). Yet, any bias from differential repetition is likely to be very small. To see why, consider the extreme assumption that the “extra” 4 percent of students promoted in the textbook schools are the weakest students, i.e., those who scored lowest on the year 1 pretests. “Demoting” these students one grade and reestimating the year 2 level regression in Table 4 yields a textbooks impact of only 0.044 standard deviations (standard error of 0.098). This is only slightly larger than the year 2 estimate in Table 4 despite the extreme assumption.

TABLE 4—IMPACT OF TEXTBOOK PROGRAM ON NORMALIZED TEST SCORES

Dependent variable	Normalized test score <sup>a b</sup> (1)	Normalized test score <sup>b</sup> (2)	Normalized test score minus pretest score <sup>c</sup> (3)	Normalized test score minus pretest score <sup>c</sup> (4)
Textbook school	0.023 (0.087)	0.020 (0.104)	0.018 (0.053)	-0.046 (0.071)
Received a textbook				
Region and sex dummies	Y	Y	Y	Y
Years exposed to textbooks	1	2	1	2
Grades	3-8	4-7	3-8	4-7
Observations	24,132	12,663	11,321	7,354

Notes: Standard errors in parentheses.

<sup>a</sup>Running the same regressions for individual subjects English, math, and science (not shown in this table), yields similar results, with the coefficients on textbooks never statistically significantly different from zero.

<sup>b</sup>Sample includes all children enrolled in January of year 1 who took the relevant October/November test

<sup>c</sup>Sample includes all children who were enrolled in January of year 1 and took the relevant October/November test as well as the pretest in January of year 1.

\*\*\* Significant at the 1 percent level.

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

Attrition bias is also unlikely to explain the results. In year 1, the difference in the percent of pupils not tested is very small, 25.4 percent in textbook schools and 25.6 percent in the 25-school comparison group (Table 5, row 5). In year 2, 33.3 percent of pupils in comparison schools and 31.0 percent in textbook schools were not tested. Yet, differences in the pretest scores of untested pupils in textbook and comparison schools are insignificant, with different signs for different subjects. Thus, the slightly higher rate at which pupils in textbook schools are tested in year 2 is unlikely to induce seriously biased estimates of textbooks' impact on test scores. Attrition in years 3 and 4 is higher, another reason we downplay those results, but there is no evidence that it differs across groups. Finally, the simplest selection and attrition stories cannot explain the failure to find effects in the difference-in-differences specification.

#### D. Retrospective Estimates

The failure to find an impact of textbooks contrasts sharply with many people's assumptions, but also with the positive results of most retrospective studies. For example, Heyneman, Farrell, and Sepulveda-Stuardo (1978) find positive effects of textbooks on test scores in 15 out of 18 studies, with some studies finding larger effects for disadvantaged students. Fuller (1986) reports significant effects of textbooks in 14 out of 22 studies, and Fuller and Clarke (1994) found the same in 19 out of 26 studies. The four papers examined in detail in Lockheed and Hanushek's (1988) review of developing country studies report that textbooks raised test scores by 0.34, 0.36, 0.30, and 0.06 standard deviations of individual test scores. Jamison et al. (1981) compared 48 grade 1 classrooms in Nicaragua that were randomly assigned to receive radio mathematics education with 20 that received math workbooks and

TABLE 5—SELECTION AND ATTRITION DURING YEAR ONE

	Textbook	Comparison	Difference
Drop outs and transfers from year 0 to start of year 1 (20 schools) <sup>a</sup>			
Drop outs (percent)	5.3	6.0	-0.7
Transfers out (percent)	5.2	3.6	1.6
Composition of students, beginning of year 1 (50 schools) <sup>b</sup>			
Repeaters (percent) <sup>c</sup>	21.9	26.0	-4.1***
Transfers in (percent)	11.2	10.3	0.9
Students present at start of year 1 but not tested at end of year (50 schools) <sup>d</sup>			
Year 1 (percent)	25.4	25.6	-0.2
Year 2 (percent)	31.0	33.3	-2.3***
Year 3 (percent)	38.6	39.9	-1.2
Year 4 (percent)	45.2	47.9	-2.7

*Notes:* Significance is based on probit regressions with school random effects. The regressors are a constant and dummy variables for textbook schools, sex, and each grade. Regressions with 50 schools add dummies for geographic regions.

<sup>a</sup> Data for year 0 is available for only 20 schools (10 in Group 1 and 10 in Group 4).

<sup>b</sup> The other 50 schools were first visited in October of year 1, and data were collected only for children being tested, not for children who may have dropped out or transferred out between January and October of year 1.

<sup>c</sup> The percentage of repeaters is underestimated slightly for both types of schools because there is no information on repetition for nearly one-half of students who transferred in (about 6 percent of all students).

<sup>d</sup> The year 1 results on whether students were tested include all students. The year 2 results exclude students in grade 8 in year 1, since most were no longer in school and could not be tested. Similarly, the year 3 results exclude children in grades 7 and 8 in year 1, and the year 4 results exclude children in grades 6–8 in year 1. All results are based on the 50 schools (groups 1 and 4) that were visited at the start of year 1.

\*\*\* Significant at the 1 percent level.

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

20 that served as controls.<sup>5</sup> After one year, pupils who received workbooks scored one-third of a standard deviation higher than control group pupils, a difference significant at the 1 percent level. No significant interaction was found between pretest scores and receipt of workbooks, although workbooks narrowed gaps between rural and urban students.

Retrospective estimates using Kenyan data could lead one to conclude that textbooks raise test scores. Cross-sectional (nonexperimental) variation in textbook availability in year 1 for the 75-school comparison group can be used to estimate the effect of textbooks on test scores. We estimated OLS regressions (with school-level random effects) of normalized test scores on (a) a dummy variable indicating students with privately owned textbooks; (b) the textbook to student ratio for school-owned textbooks (calculated for each grade-subject combination in each school); and (c) other school and family variables. Pupils with privately owned English textbooks scored 0.18 standard deviations higher on English exams, controlling for parental education and land owned, and math and science books raised those test scores by 0.09 and 0.05 standard deviations (Table 6, row 1). The English and math impacts are highly significant, but not the science impact. Aggregating over all subjects yields an impact of 0.12 with a tight standard error (0.016).

<sup>5</sup> As discussed above, Heyneman, Jamison, and Montenegro (1984) compared Philippine schools randomly selected to receive one textbook per one or two students, and found little difference in test scores. They also compared these two scenarios with no textbooks and find a substantial positive impact of textbooks on test scores. Yet, these estimates compare the same schools before and after receiving textbooks and so are not based on randomized trials.

TABLE 6—CROSS-SECTIONAL RETROSPECTIVE ESTIMATES OF IMPACT OF TEXTBOOKS ON TEST SCORES

	English	Math	Science	All
Student-owned books	0.178*** (0.024)	0.087*** (0.026)	0.054 (0.047)	0.116*** (0.016)
School-owned books	-0.010 (0.170)	-0.354*** (0.145)	-0.554 (0.442)	-0.190** (0.095)
Sample size	10,115	10,129	10,068	30,312

*Notes:* Asymptotic standard errors are in parentheses. All regression included school random effects, a constant term and dummy variables for grade. Other variables included are parental education (for students in grades 6–8, for whom data are available), land owned, parental participation in school, teacher education, and teacher training. The test scores used are the ICS tests of October of year 1 for grades 3 and 4, the district tests of October of year 1 for grades 5–7 and the KCPE tests of November of year 1 for grade 8.

\*\*\* Significant at the 1 percent level.

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

School-owned textbooks and test scores are negatively correlated in these estimates (Table 6, row 2) but this may be because donors gave more textbooks to the neediest schools. A difference-in-differences estimate of the impact of school-owned textbooks is possible using data on a 1994 World Bank textbook provision project conducted by the Jomo Kenyatta Foundation. It provided English and math textbooks to 95 of Busia and Teso's 333 primary schools at roughly one book per two pupils and one Swahili and one science book per four pupils, targeting the neediest schools. Difference-in-differences estimates of school-level average scores on government exams in grades 6, 7, and 8 for about 80 percent of the 334 schools suggest that in 1994 textbooks raised grade 7 test scores by 0.50 standard deviations and, in 1995, grade 7 and 8 scores were raised by about 0.65 standard deviations (Table 7).<sup>6</sup> No significant impact is seen for grade 6 or for grade 8 in 1994.

To summarize Section II, our randomized trial yields no evidence that textbooks raised test scores for the average student. Our two estimators of the impact on average test scores allow us to reject effects greater than 0.20 and 0.13 standard deviations in year 1, and data from later years provide no evidence that textbooks' impact on test scores accumulates over time. These results do not seem to be driven by selection or attrition bias. In contrast, retrospective estimates using cross-

TABLE 7—PANEL RETROSPECTIVE ESTIMATES—IMPACT OF JOMO KENYATTA TEXTBOOKS ON TEST SCORES

Grade	1994	1995
6	-0.157 (0.171)	-0.091 (0.266)
7	0.497** (0.252)	0.641** (0.291)
8	0.020 (0.172)	0.676*** (0.204)

*Notes:* Standard errors are in parentheses. Each cell represents a regression of normalized test scores in 1994 or 1995 minus normalized test scores in 1993 on geographic dummy variables and a dummy variable for whether the school received textbooks (the coefficient on the last dummy is the one reported). Sample sizes ranged from 255 to 274. Scores are normalized across students not across schools. The school mean scores have been standardized, in terms of the distribution across students, using a sample of students from the SAP schools in each year.

\*\*\* Significant at the 1 percent level.

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

<sup>6</sup> We convert standard deviations of school mean test scores into estimated standard deviations of individual test scores using a small sample of schools for which we have student level data.

sectional data on privately owned textbooks and longitudinal data on school-owned textbooks suggest a significant positive impact of textbooks on average test scores.<sup>7</sup>

### III. Interactions between Initial Test Scores and Program Impact

This section presents evidence that the provision of textbooks in rural Kenya primarily benefited the strongest students. This result contrasts sharply with findings in the literature that textbooks should most benefit the weakest students and holds despite the fact that weaker students are less likely to own a textbook initially.<sup>8</sup> Note, that if textbooks benefit the strongest students disproportionately, parents should be more likely to buy textbooks for such students, biasing upward retrospective estimates of the impact of textbooks on test scores. We first examine the impact of interactions between student pretest scores and assignment to a textbook school on post-test scores. Interacting program assignment and student characteristics should be done cautiously, given the potential for data mining, but conditioning on initial values of the dependent variable is a natural interaction to examine average effects. We also present evidence on other educational outcomes, showing that textbooks increased progression to secondary school for eighth graders but did not reduce grade repetition or raise attendance in lower grades. This supports the hypothesis that textbooks mostly benefited strong students since only those students reach grade 8 and have a hope of going to secondary school. We then present evidence that many students had difficulty reading the textbooks. Finally, we present results from the grants program, which also seems to disproportionately benefit initially high scoring students, and examine alternative explanations of our results, none of which garner evidence in their support.

#### A. Interaction Effects

In year 1, an interaction term between the program variable and the average pretest score is highly significant when aggregating across all subjects. The program increased scores by 0.057 standard deviations more for students with pretest scores one standard deviation above the mean (Table 8, column 1).<sup>9</sup> This is also true in year 2 when the interaction term is even slightly higher at 0.061 standard deviations (Table 8, column 2).

Most of the difference in program impacts seems to reflect interactions between the program and student characteristics, not interactions with school or teacher characteristics. To check this, we added school fixed effects to the regressions in Table 8 (not shown). This prevents estimation of the program effect but not estimation of a within-school interaction term. For English, math, and all subjects combined, within-school interactions are only slightly smaller than the overall interaction

<sup>7</sup> Of course, it is possible that more sophisticated methods such as regression discontinuity design methods would have been less subject to bias than simple OLS analyses.

<sup>8</sup> The correlation between mean pretest score and ownership of private textbooks in group 4 schools in year 1 ranged from 0.10 (science) to 0.15 (English).

<sup>9</sup> Similar results are found using only the English pretest score (recall that all textbooks are written in English). This is not surprising because English pretest scores and average pretest scores are highly correlated.

TABLE 8—NORMALIZED TEST SCORES AS A FUNCTION OF TREATMENT AND PRETEST SCORE

Dependent variable	Normalized test score (year 1)	Normalized test score (year 2)	Normalized test score minus pretest scores (year 1)
Textbook school	0.060 (0.061)	-0.016 (0.088)	0.021 (0.060)
Pretest score	0.430*** (0.013)	0.342*** (0.016)	-0.338*** (0.016)
Pretest × textbook school	0.057*** (0.018)	0.061*** (0.022)	0.042** (0.021)
Observations	11,342	7,393	11,321

*Notes:* The first two columns are regressions of October/November test scores for the relevant year on dummy variables for textbook school, sex, and region, the average of the three year 1 pretest scores, and an interaction term between the textbook school dummy and the average pretest score. The dependent variable in the third column is the October/November test score minus pretest score. Each regression includes all children enrolled in January of year 1 who took the January pretest and the relevant October/November test.

\*\*\* Significant at the 1 percent level.

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

effects in Table 8. For example, the coefficient is 0.048 for all subjects combined in year 1, only slightly smaller than the estimate of 0.057.

Pretest scores may measure initial achievement with substantial noise, so the coefficients on the pretest and on the interaction between the pretest and the program may have attenuation bias, underestimating the true impact of initial academic achievement, and its interaction with the program, on later achievement. To get a sense of the potential magnitude of such a correction, suppose the true coefficient on initial academic achievement were one, as in many difference-in-difference specifications. If so, attenuation bias appears to be large, since that coefficient is only 0.43 in year 1 (and lower in year 2). Applying a similar correction factor to the estimated year 1 interaction effect yields an interaction effect of 0.13 standard deviations.

Interactions are also significant in regressions of changes in scores. Table 8, column 3 adds an interaction term to the year 1 regression in Table 4, column 3, in which the dependent variable is the (post-test) score minus the pretest score. That term is significant at the 5 percent level.

Table 9 splits the sample into quintiles, based on average pretest scores, and re-estimates the level regressions in Table 4 by quintile, thus allowing treatment effects and school random effects to vary by quintile. Aggregating across all subject-grade combinations, the estimated effects of textbooks on test scores in year 1 by quintile, from lowest to highest, are -0.05, -0.02, 0.03, 0.14, and 0.22 standard deviations (Table 9). These effects are statistically insignificant for quintiles 1-3, but significant for quintiles 4 and 5 at the 10 percent and 5 percent levels, respectively.<sup>10</sup> Since pretests are noisy, some students in the top quintile were not necessarily in the top quintile of initial knowledge. Thus, the true effect of knowledge in the top quintile may be greater than 0.22 standard deviations. The pattern of effects across quintiles

<sup>10</sup> This pattern generates the positive interaction effect seen above and is consistent with not finding a significant average effect, so these interaction and average effects do not necessarily imply negative impacts on weak students.

TABLE 9—PROGRAM IMPACT ON NORMALIZED TEST SCORES, BY QUINTILE OF PRETEST SCORES

Years exposed	Quintile 1 (1)	Quintile 2 (2)	Quintile 3 (3)	Quintile 4 (4)	Quintile 5 (5)
1	-0.049 (0.064)	-0.021 (0.069)	0.032 (0.073)	0.142* (0.079)	0.218** (0.096)
2	-0.077 (0.081)	-0.109 (0.094)	-0.089 (0.104)	0.021 (0.101)	0.173 (0.131)

*Notes:* Each row represents five random effects regressions, one for each quintile (based on pretest scores from January of year 1), of post-test scores on a dummy variable indicating whether a child is in a textbook school and on dummy variables for region and sex. The sample consists of all children enrolled in January of year 1 who took both the pretest in year 1 and the relevant post-test. All results are aggregated over all grade/subject combinations that received textbooks.

\*\*\* Significant at the 1 percent level.

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

in year 2 is similar, except that the effects for the top two quintiles are less precisely estimated and statistically insignificant.

It is difficult to use our data to test for interaction effects after year 2, but the data suggest no interaction effects in later years. This is not surprising. First, textbook availability and usage converged in later years (Tables 2 and 3), so most academically strong students in comparison schools probably had at least a shared textbook. Second, textbook use in class converged over time for treatment and comparison schools, perhaps because new textbooks were issued in year 2, and teachers may not have wanted students to use the old books, even though they were quite similar.<sup>11</sup> Third, all interaction effects are based on pretests at the start of year 1, and, over time, they presumably measure current achievement and motivation less precisely. The fall from year 1 to year 2 in the pretest score coefficient in Table 8 supports this. Finally, the sample size falls each year as children drop out or finish primary school, reducing the precision of the estimates.

### B. Other Educational Outcomes

Evidence for other educational outcomes is consistent with the finding that textbooks primarily benefited academically strong students. Consider drop out rates and grade repetition. Pupils in grades 3–7 in year 1 had four possible outcomes for year 2: stay in school and be promoted, stay in school and repeat, drop out, or transfer out. The same holds true for grade 8 pupils, but promotion has two forms: finish primary school and leave school, or finish primary school and enter secondary school. After one year, grade 3–7 pupils show no significant differences in dropout or repetition rates across textbook and comparison schools (Table 10, columns 1 and 2).<sup>12</sup> Textbooks also did not affect pupil absences. These results imply no program effect

<sup>11</sup> Classroom observations in year 2 in grade-subject combinations given textbooks showed that textbook use in class was 16.4 percentage points higher in textbook schools but, by year 3, this gap was only 8.2 percentage points.

<sup>12</sup> Here, statistical significance is based on a probit regression with school random effects.



TABLE 10—PROMOTION, REPETITION, AND DROPPING OUT FROM YEAR 1 TO YEAR 2

	Lower grades (3–7)		Upper grade (8)	
	Textbook schools (1)	Comparison schools (2)	Textbook schools (3)	Comparison schools (4)
Stayed, promoted	0.53	0.53		
Finished primary, no secondary			0.32***	0.41
Entered secondary			0.43**	0.38
Stayed, repeated	0.21	0.21	0.16	0.14
Dropped out	0.17	0.17	0.01	0.03
Transferred out	0.08	0.09	0.06	0.04
Number of students	5,009	4,838	447	440

*Notes:* For promotion, repetition and dropping out, comparison schools are the 25-school comparison group. Lower and upper grades refer to grade of pupils in year 1. Tests for statistical significance are based on probits with school random effects.

\*\*\* Difference with comparison group significant at the 1 percent level.

\*\* Difference with comparison group significant at the 5 percent level.

\* Difference with comparison group significant at 10 percent.

on average and below average pupils, those most likely to repeat, drop out, or be absent.

In contrast, grade 8 students in the textbook schools in year 1 were more likely to enter secondary school in year 2 than comparison school students (43 percent versus 38 percent). This difference is significant at the 5 percent level (Table 10, columns 3 and 4) and is consistent with textbooks being most helpful to academically strong students, since grade 8 is de facto selective and only strong students go to secondary school.

### C. Why Do Textbooks Increase Scores Only for Strong Students?

One possible reason why students with high initial achievement may benefit more from textbooks is that those books are too difficult for other students to use effectively. Indeed, the median students in lower grades seem to have difficulty even reading the textbooks. Fifty of the 100 schools were randomly selected for visits in year 4, and the median student (by class rank) in grades 3–8 was asked to read the English textbook provided by the program. In grade 3, only 15 percent of the median students could read the grade 3 English textbook, and only 29 percent of the grade 4 median students could read their English textbooks (Table 11). This problem is less common in upper grades. The figures are 62 percent for grade 5, 85 percent for grade 6, and about 95 percent for grades 7 and 8. Yet, even students who can read the words in the textbooks may have difficulty effectively using a textbook in their third language (Table 11, columns 2–4).

Differences in whether students took textbooks home may explain part, but not most, of the differential impact of textbooks across weak and strong students. Students in grades 6–8 were allowed to take textbooks home, but younger students were not. Grades 6–8 students in textbook schools who took the textbooks home had higher average pretest scores (0.21) than those who did not (–0.03), a difference significant at the 1 percent level. Yet, the year 1 interaction effect for the level regression for grades 3–5, in which students could not take textbooks home, is 0.051, only

TABLE 11—UNDERSTANDING OF THE ENGLISH TEXTBOOK BY THE MEDIAN STUDENT

Grade	Can read the book (percent) (1)	Can answer questions in English, about the passage (percent) (2)	If unable to answer question in English, can answer when asked in Kiswahili (percent) (3)	Can answer written questions in English, from the book (percent) (4)
3	15	6	58	33
4	29	12	37	15
5	62	38	38	25
6	85	79	27	10
7	94	77	17	52
8	96	92	0	90

*Notes:* The data consist of the median student in each grade in a random sample of 50 schools. Thus, for each grade, there are data for 50 students, one from each school. The data on answering questions in English include all students in the sample. Those unable to read the passage had it read to them in English.

slightly smaller than the estimate of 0.072 for grades 6–8 (these are not reported in Table 8).

#### *D. Evidence from Another Program and Alternative Explanations*

Evidence from another source, the grant program conducted in the 25 group 2 schools, further supports the hypothesis that textbooks were best suited for the strongest students. The 25 schools given grants in year 2 spent almost half the funds on textbooks. Thus, we have a second, quasi-independent randomized trial to evaluate.<sup>13</sup> We cannot rule out that nontextbook grant expenditures may have also affected scores, but most of that expenditure was on classroom construction, which takes time, so estimated effects should mainly reflect textbook purchases.

Table 12 shows estimates of equation (3) for the schools given grants in year 2. Level results (column 1) show a point estimate of the effect of receiving a grant on average test scores of 0.13 standard deviations, but it is insignificant. Difference-in-differences results (column 2) yield marginally significant results of about 0.12 standard deviations. While the point estimates in columns 1 and 2 of the effect of textbooks purchased using grants are larger than those for the textbook program, the differences with the corresponding estimators in Table 4 (columns 1 and 3) are not statistically significant. If the higher point estimate reflects a real difference, rather than sampling variation, it may be due to schools that particularly need textbooks choosing to buy them, or it may reflect nontextbook expenditure from the grants.

Estimates based on the 25 schools given grants offer further support for the proposition that textbooks are most useful for students with strong preparation. First, an interaction term between pretest scores and the program variable (Table 12, column 3) was positive and statistically significant at the 1 percent level. Second, estimates

<sup>13</sup> Strictly speaking, this trial is not fully independent since group 2 schools are control schools for evaluating year 1 outcomes, and analyses of year 2 data for grants and for textbooks use the same comparison schools (the 50-school comparison group), but it is independent in the sense that the 25 textbook schools are not in the grants sample.

TABLE 12—IMPACT OF PROVIDING GRANTS TO GROUP 2 SCHOOLS IN YEAR 2  
(after one year of the program)

Dependent variable	Normalized test score	Normalized test score	Normalized test score	Normalized test score
	(Year 2)	(Year 2) minus Year 1 test score	score (Year 2)	(Year 2) minus Year 1 test score
	(1)	(2)	(3)	(4)
Textbook school	0.130 (0.123)	0.117* (0.068)	0.124 (0.079)	0.124 (0.079)
Pretest score			0.494*** (0.011)	−0.506*** (0.011)
Pretest score × textbook school			0.051*** (0.019)	0.051*** (0.019)
Observations	7,486	7,486	7,486	7,486

Notes: Standard errors in parentheses. The sample includes all children enrolled in group 2, 3, or 4 schools in January of Year 1 and who took the Year 2 October/November test.

\*\*\* Significant at the 1 percent level.

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

similar to those in Table 9 (not shown) show significant impacts for quintile 5 (5 percent level in level regressions and 10 percent in differenced regressions) but not for the other four quintiles.

One possible reason for the absence of an effect of textbooks on average scores is crowding out; other efforts to improve schools may have declined in response to textbook provision. Data from a school questionnaire (which collected information from the school headmaster) and a school committee questionnaire (school committees consist of selected teachers and parents), suggest that the program crowded out harambee fundraising, which focuses on classroom construction in small schools. Yet, even if ICS assistance reduced that fundraising, it would have little short-run impact on test scores because constructing new classrooms takes time and the flow of services from new classrooms extends over many years, so the resulting test score gain is probably small in any one year. Moreover, larger schools show no evidence of crowding out, yet the estimated average program effect (from level regressions) for year 1 over all grades and subjects is 0.02 with a standard error of 0.12, virtually identical to the estimate of 0.022 in Table 4.<sup>14</sup> Thus, crowding out is unlikely to explain why textbooks do not affect average test scores.

A final possibility is that the tests were too hard for most students. This would explain both little or no impact on most students and the significant impact on the best students (for whom the tests may have been appropriate). Indeed, the year 1 district tests were quite difficult for many students. In some subject-grade combinations, average scores were not much higher than random guessing, which implies little information in those tests. Yet, we find no evidence for textbook effects in subject-grade combinations with higher average test scores. Overall, the information content of the tests is higher for grades 6–8 than for grades 3–5, but neither

<sup>14</sup> The year 1 regression in Table 4 was re-estimated after adding an interaction term between school size and the textbook schools dummy variable. This interaction term was small and completely insignificant.

level nor difference regressions show significant impacts in grades 6–8, nor even of higher point estimates than in grades 3–5. Moreover, the ICS tests used in years 2, 3, and 4 were intentionally designed to be easier. The mean scores were much higher, conveying more information. Yet, estimates in year 2 were also close to zero. For a detailed discussion, see Glewwe, Kremer, and Moulin (2002).

#### IV. Conclusion

Providing textbooks to schools where few students have them seems to be an obvious way to improve learning in developing countries. Textbook provision is almost universally accepted, even by those who doubt the effectiveness of increased school spending. Yet, our results show that providing the official government textbooks in Kenya did not increase average test scores, although the scores of students with high initial achievement did increase. This suggests that the official government textbooks used were ill-suited for the typical student, which is supported by the finding that the median student in lower grades had difficulty reading the official English textbook. This may help explain the positive coefficient on textbooks in retrospective studies. Parents are more likely to obtain textbooks for academically strong students, leading to upward bias in regressions estimating the impact of textbooks on test scores. Indeed, our retrospective estimates using data from the comparison schools and from school level data for all schools in Busia and Teso, which could suffer from this bias, show positive impacts of textbooks that are statistically significant.

The finding that provision of Kenya's official government textbooks increased test scores only for above average students is arguably part of a much larger mismatch between the official curriculum and the needs of many students. Kenyan schools are judged mainly on average KCPE results, giving them little incentive to focus on students who will not make it through grade 8 or will bring down average scores if they get there. This mismatch may well play an important role in generating high repetition (21 percent) and dropout (17 percent) rates. In many developing countries—where students and teachers are frequently absent and scores on international standardized tests are very low—using curricula and textbooks that cover material at the same pace as in the developed world and assume that students know material from previous grade levels yields a serious mismatch. An evaluation of achievement grouping in Kenya (Esther Duflo, Pascaline Dupas, and Kremer 2007) further supports the idea that the Kenyan curriculum may leave many students behind. Students with low pretest scores fared much better when grouped with peers with similar pretest scores, perhaps because teachers could teach to their initial academic level.

Why is the system geared toward the strongest students? We suggest that the current system reflects three common characteristics of developing countries: (a) the adoption of a centralized, uniform national curriculum and education system; (b) substantial heterogeneity in the student population, due in part to rapid educational expansion; and (c) disproportionate political power of elites. The following paragraphs discuss each of these characteristics.

In contrast to the United States, many developing countries are characterized by a centralized educational system with a single national curriculum, competitive national exams to enter higher levels of education, and little local control over schools. This

is the case in Kenya and in many post-colonial countries. Such systems may reflect political goals, such as unifying countries with substantial ethnic diversity. While many European countries also have centralized education systems, those systems include opportunities to follow a vocational track. Such tracks are rare in developing countries. In most African countries less than 5 percent of secondary pupils attend vocational school. In Kenya, the rate is only 1.7 percent. The situation is similar in most of Asia, although China is an important exception. In developed countries, specialization typically occurs in secondary school, yet many primary students in developing countries are older than the ages at which European students begin tracking. Among pupils starting grade 8 (the last grade of primary school in Kenya) in our sample, the average age is 15, and 20 percent are 16 years old or older. Among pupils starting grades 6 or 7, the average age is 13.6 and 38 percent are age 14 or older. A typical US student who finishes 13 years of education has spent half of his or her time in a system with within-school tracking, while a typical Kenyan student who completes eight years of education spends no time in a tracked system.

As access to education expands, curricula are often adjusted. For example, at the start of the twentieth century, US secondary schools served only 10 percent of the population and were highly academic, geared towards preparing students for tertiary education. In 1910, for example, 50 percent of high school students took Latin (Diane Ravitch 2000). In the following decades, the United States rapidly expanded secondary education. By the mid-twentieth century, high school enrollment reached 65 percent, bringing in children from very different backgrounds. This educational expansion led to a transformed curriculum, including the incorporation of vocational education (Ravitch 1983).

Following independence, Kenya also expanded access to education. Its gross primary enrollment ratio almost doubled in 20 years, from 47 percent in 1960 (just before independence) to 90 percent in 1980 (Anil B. Deolalikar 1999). However, its curriculum changed very little. Kenya's colonial education system was designed to produce a small group of elite Africans to work for the colonial bureaucracy. The demanding curriculum had no remedial education for those who fell behind. Today's Kenyan students are extremely heterogeneous in family background, preparation for schooling, and economic status. Middle-class children in Nairobi and other cities grow up with constant exposure to English, good nutrition, and electricity, while the children of subsistence farmers hear very little English until they go to school, have poor health and nutrition, and no electricity, which substantially limits study time at home.

As enrollment increased, Kenya did make some changes to its curriculum, but the system remained primarily oriented toward the elites. References to Kenyan history and geography replaced references to English history and geography, and pictures in books are of Kenyan, not English, children. But the language of instruction remains English, most students' third language. Tanzania is an interesting contrast. At independence, Tanzania also had a centralized, uniform education system and considerable heterogeneity in pupils' backgrounds. Yet, unlike Kenya, Tanzania chose an education system arguably better suited to most of its population. It uses Swahili, not English, as the language of instruction, and invested more in primary education relative to secondary education (John B. Knight and Richard H. Sabot 1990).

How did Kenya end up with such a system? Historically, four factors led to a Kenyan curriculum that was not targeted toward the typical student. First, since independence, Kenya has been for the most part a de facto or de jure one party state, and elites have held disproportionate political power. These elites may have preferred an education system targeted to their children's needs. Ministry of Education officials may design curricula and textbooks in their children's interests and see efforts to design a curriculum more suitable for rural areas as lowering academic standards. Second, even parents of average students may favor an elite curriculum in order to secure more desirable peers in their children's schools. If an ambitious curriculum causes weak pupils to leave the system, average students may benefit from better than average peers, and from more resources per student. Third, teachers also have incentives to demand advanced textbooks and devote less effort to pupils who cannot read them. Primary schools are judged by students' scores on the KCPE exam, and students who drop out before taking the KCPE exam are excluded from calculations of school performance. Thus, teachers have incentives to use textbooks suited to the strongest students. Indeed, their work load decreases when a weak student drops out of school. Finally, even the market for textbooks does not cater to the typical student because it consists primarily of elite students and teachers since poorer students rarely buy textbooks.

Moving beyond the educational system itself, the broader structure of the economy and the political system did not lead to demands for increased vocational education. First, civil service jobs historically offered large rents, and access to these jobs depends on formal, academic qualifications, not vocational training. In this setting, no ethnic group had incentives to push for vocational education, or to alter local schools in ways that might improve learning for weaker pupils but could harm chances for strong pupils to obtain civil service jobs by winning the academic sweepstakes, since such a policy would reduce the representation of that ethnic group in the civil service. Second, without multi-party democracy, political competition was largely between elites from each region, so the system favored these elites. Third, in an economy with limited industry and where civil service jobs were the main route to economic security, vocational education had little constituency.

Recent political events in Kenya may be changing its political economy. When the Cold War ended, international donors gained more leverage over the government and imposed multi-party democracy, forcing politicians to appeal to poor voters and reducing the rents available to elites. A key element of politicians' recent efforts to appeal to these voters was a promise to abolish school fees, which greatly increased enrollment. Many urban elites consequently left for private schools. Perhaps this will induce a shift in the public curricula toward average students.

We suggest two policies that might help less prepared students in Kenya. The first is remedial education for children who have fallen behind the official curriculum. Abhijit V. Banerjee et al. (2005) examined such a program in India and found it very effective. The program offers remedial education to children who reach grade 3 or 4 without mastering basic skills. They leave the classroom and receive tutoring for two hours per day. A randomized evaluation of the program shows that test scores in treatment schools rose by 0.14 standard deviations in the first year and by 0.28

standard deviations after two years, and that those impacts were strongest for the weakest students.

A second possibility would be to allow different schools or different programs within schools to teach the curriculum at different speeds. For example, some schools could cover the primary school curriculum, which is currently designed for eight years, in seven years, while others could cover it in nine or ten years. All students would take the same curriculum, so students would not be irretrievably sorted into an academic track or a more basic vocational track at an early age. Students' opportunities for secondary school would be based solely on their performance on the KCPE, regardless of their years in primary school. Singapore has such a policy. Strong students are put into an "express" track and take the GCE O-level exam (a requirement to apply to junior colleges) after four years. The rest are placed in the "normal" track, taking an intermediate exam after four years. Those who do well, take the GCE O-level exam at the end of the fifth year (Ministry of Education of Singapore, <http://www.moe.gov.sg/education/secondary>). In one sense, such a system would formalize and rationalize what happens informally, and inefficiently, now. Officially, all Kenyan students face an eight-year curriculum, but many students require nine or ten years. It is inefficient to cover the same material two times, a year apart, in classes that mix pupils who have seen the material before with those who have not.

The findings in this paper have implications for other developing countries because the problem of a curriculum poorly matched to typical students' needs, leading to high repetition and dropout rates, is not peculiar to Kenya. It occurs in much of the developing world. In sub-Saharan Africa, 15.6 percent of primary school children repeated a grade in 2002–2003, and some countries have higher rates, such as 34.4 percent in Gabon and 25.8 percent in Cameroon. Only 68.6 percent of African students reach grade 5 (UNESCO 2006). The grade 5 survival rate in India is 61 percent, and pupils whose parents never attended school have great difficulty keeping up with the curriculum (Banerjee et al. 2007). Indeed, a World Bank (1997) report provides evidence that primary school textbooks in India are too difficult for many students. High drop out and repetition rates in many developing countries suggest that curricula are not well suited to the average student.

Many have argued that distortions in educational systems limit the efficacy of additional education spending. Filmer and Pritchett (1999) expand on this argument, claiming that although teacher-centered inputs may be ineffective, nonteacher inputs could be very effective. Our results suggest that political economy distortions in educational systems may limit even the impact of one of the most basic nonteacher inputs, textbooks.

Many developing countries share the underlying features that lead to a political economy of education in which many students are left behind: a centralized, uniform education system, the heterogeneous student population that comes with rapid educational expansion, and disproportionate elite influence. Future research should examine potential reforms that could broaden access to learning and schooling, including reorientation of curricula, decentralization, tracking, vocational education, and remedial education.

## REFERENCES

- Bamgbose, Ayo.** 2004. "Language of Instruction Policy and Practice in Africa." Paris: UNESCO. [http://www.unesco.org/education/languages\\_2004/languageinstruction\\_africa.pdf](http://www.unesco.org/education/languages_2004/languageinstruction_africa.pdf).
- Banerjee, Abhijit V., Shawn Cole, Esther Duflo, and Leigh Linden.** 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *Quarterly Journal of Economics*, 122(3): 1235–64.
- Deolalikar, Anil B.** 1999. "Primary and Secondary Education in Kenya: A Sector Review." Human Development Sector Unit, Eastern Africa Region, The World Bank. Revised February 1999.
- Duflo, Esther, Pascaline Dupas, and Michael Kremer.** 2007. "Peer Effects, Pupil-Teacher Ratios, and Teacher Incentives: Evidence from a Randomized Evaluation in Kenya." Unpublished.
- Fuller, Bruce.** 1986. "Raising School Quality in Developing Countries: What Investments Boost Learning?" World Bank Discussion Paper No. 2. The World Bank: Washington D.C.
- Fuller, Bruce, and Prema Clarke.** 1994. "Raising School Effects While Ignoring Culture? Local Conditions and the Influence of Classrooms, Tools, Rules and Pedagogy." *Review of Educational Research*, 64(1): 119–57.
- Glewwe, Paul, Michael Kremer, and Sylvie Moulin.** 2002. "Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya." Unpublished.
- Glewwe, Paul, Michael Kremer, and Sylvie Moulin.** 2007. "Many Children Left Behind? Textbooks and Test Scores in Kenya." National Bureau of Economic Research Working Paper 13300.
- Hanushek, Eric A.** 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature*, 24(3): 1141–77.
- Hanushek, Eric A.** 1995. "Interpreting Recent Research on Schooling in Developing Countries." *World Bank Research Observer*, 10(2), 227–46.
- Heyneman, Stephen P., Joseph P. Farrell, and Manuel A. Sepulveda-Stuardo.** 1978. "Textbooks and Achievement: What We Know." World Bank Staff Working Paper 298.
- Heyneman, Stephen P., Dean T. Jamison, and Xenia Montenegro.** 1984. "Textbooks in the Philippines: Evaluation of the Pedagogical Impact of Nationwide Investment." *Educational Evaluation and Policy Analysis*, 6(2): 139–50.
- Jamison, Dean, Barbara Searle, Klaus Galda, and Stephen Heyneman.** 1981. "Improving Elementary Mathematics Education in Nicaragua: An Experimental Study of the Impact of Textbooks and Radio on Achievement." *Journal of Educational Psychology*, 73(4): 556–67.
- Knight, John B., and Richard H. Sabot.** 1990. *Education, Productivity and Inequality: The East African Natural Experiment*. Oxford University Press.
- Lockheed, Marlaine E., and Eric Hanushek.** 1988. "Improving Educational Efficiency in Developing Countries: What Do We Know?" *Compare*, 18(1): 21–38.
- National Council of Educational Research and Training.** 2002. Seventh All India School Education Survey (Provisional Statistics). New Delhi: National Council of Educational Research and Training. <http://gov.ua.nic.in/aises/pdf>.
- Pritchett, Lant, and Deon Filmer.** 1999. "What Education Production Functions Really Show: A Positive Theory of Education Expenditures." *Economics of Education Review*, 18(2): 223–39.
- Ravitch, Diane.** 2000. *Left Back: A Century of Failed School Reforms*. New York: Simon & Schuster.
- Ravitch, Diane.** 1983. *The Troubled Crusade: American Education, 1945–1980*. New York: Basic Books.



**This article has been cited by:**

1. Sarah Baird, Ephraim Chirwa, Craig McIntosh, Berk Özler. 2010. The short-term impacts of a schooling conditional cash transfer program on the sexual behavior of young women. *Health Economics* **19**:S1, 55-68. [[CrossRef](#)]
2. Abhijit V. Banerjee, , Esther Duflo, . 2010. Giving Credit Where It Is DueGiving Credit Where It Is Due. *Journal of Economic Perspectives* **24**:3, 61-80. [[Abstract](#)] [[View PDF article](#)] [[PDF with links](#)]
3. Paul Glewwe, , Nauman Ilias, , Michael Kremer. 2010. Teacher IncentivesTeacher Incentives. *American Economic Journal: Applied Economics* **2**:3, 205-227. [[Abstract](#)] [[View PDF article](#)] [[PDF with links](#)]
4. G. Lassibille, J.-P. Tan, C. Jesse, T. Van Nguyen. 2010. Managing for Results in Primary Education in Madagascar: Evaluating the Impact of Selected Workflow Interventions. *The World Bank Economic Review* **24**:2, 303-329. [[CrossRef](#)]
5. Abhijit V. Banerjee, Esther Duflo. 2009. The Experimental Approach to Development Economics. *Annual Review of Economics* **1**:1, 151-178. [[CrossRef](#)]
6. Michael Kremer, Alaka Holla. 2009. Improving Education in the Developing World: What Have We Learned from Randomized Evaluations?. *Annual Review of Economics* **1**:1, 513-542. [[CrossRef](#)]