NBER WORKING PAPER SERIES

MANY CHILDREN LEFT BEHIND? TEXTBOOKS AND TEST SCORES IN KENYA

Paul Glewwe
Michael Kremer
Sylvie Moulin

Working Paper 13300
http://www.nber.org/papers/w13300

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2007

This project was a joint undertaking of 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 Lipeyah, 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 costs of this evaluation were 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, the countries they represent, or the National Bureau of Economic Research.

Many Children Left Behind? Textbooks and Test Scores in Kenya
Paul Glewwe, Michael Kremer, and Sylvie Moulin
NBER Working Paper No. 13300
August 2007
JEL No. C93,I20,O15,P16

ABSTRACT

A randomized evaluation suggests that a program which provided official textbooks to randomly selected rural Kenyan primary schools did not increase test scores for the average student. In contrast, the previous literature suggests that textbook provision has a large impact on test scores. Disaggregating the results by students? initial academic achievement suggests a potential explanation for the lack of an overall impact. Textbooks increased scores for students with high initial academic achievement and increased the probability that the students who had made it to the selective final year of primary school would go on to secondary school. However, students with weaker academic backgrounds did not benefit from the textbooks. Many pupils could not read the textbooks, which are written in English, most students? third language. The results are consistent with the hypothesis that the Kenyan education system and curricular materials are oriented to the academically strongest students rather than to typical students. More generally, many students may be left behind in societies that combine 1) a centralized, unified education system; 2) the heterogeneity in student preparation associated with rapid expansion of education; and 3) disproportionate elite power.

Paul Glewwe
Dept of Applied Economics, U of MN
1994 Buford Ave.
St. Paul MN 55108
pglewwe@umn.edu

Sylvie Moulin
World Bank
1818 H Street, N.W.
Washington, DC 20433

Michael Kremer
Harvard University
Department of Economics
Littauer Center M20
Cambridge, MA 02138
and NBER
mkremer@fas.harvard.edu

I. Introduction

Many economists argue that increasing educational expenditure will have a limited impact on learning in distorted educational systems (see, for example, Hanushek 1995 and Pritchett and Filmer 1999). However, even skeptics about the impact of education spending in such circumstances believe that providing textbooks in environments where they are scarce can substantially increase test scores (see literature reviews by Heyneman, Farrell, and Sepulveda-Stuardo 1978, Fuller 1986, Lockheed and Hanushek 1988, and Fuller and Clarke 1994). Indeed some political economy models of distortions in education expenditure suggest that non-teacher inputs will have a large impact on student performance relative to the impact of increased spending on teachers (Pritchett and Filmer, 1999). Policymakers have acted in accordance with this view. For example, when the World Bank increased loans to Kenya after the end of the Moi regime, one of the first loans was for a massive textbook supply program.

We report the results of a randomized evaluation of a program that provided textbooks to rural Kenyan primary schools. Unlike previous studies, we find that textbooks had little effect on the scores of a typical student, and can reject the hypothesis that textbooks raised average student test scores by 0.07 or more standard deviations.

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. Moreover, a randomized evaluation of a program that gave grants to similar schools, half of which were spent on textbooks, shows similar results.

Why did textbooks not raise scores? One clue is that textbooks appear to have improved the scores of students with higher pre-test scores. That is, an interaction term between pre-test scores and assignment to the textbook program has a highly significant positive correlation with

post-test scores. Consistent with the hypothesis that textbooks were useful only for the strongest students, students who made it to the selective final year of primary school (grade 8) were more likely to enter secondary school in the next year if they were in schools that received text-books. There is little evidence that textbooks reduced grade repetition or dropping out, consistent with the finding of no impact on weaker students. Analysis of the grant program yields similar results.

The hypothesis that the official Kenyan government textbooks were of limited use to many students is plausible. English is the medium of instruction in Kenyan schools, but it is the third language for many pupils, including those examined here. Moreover, pupil and teacher absence rates are high, so many pupils fall behind the official curriculum and are likely to have difficulty with the English textbooks used under that curriculum. Our data show that many students cannot read the textbooks.

This raises a larger issue. The Kenyan government's adoption of a centralized, uniform education system makes it very difficult to serve the entire population, given the great heterogeneity in the educational and economic background of students in a setting where education has expanded rapidly. The historical legacy of colonial education in Kenya, and the political economy of Kenya in the post-independence period, may have produced an education system that favors the most advantaged students instead of the typical Kenyan student.

Many other developing countries appear to have a similar mismatch in curricula. Many schools, particularly after the early primary years, 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 (Bamgbose, 2004). In India, most primary schools are taught in the most common regional mother tongue, but 87% of secondary schools teach in another language (NCERT, 2002).

Poor pupil performance on national and international standardized tests, high grade repetition, and high dropout rates suggest that many pupils in developing countries fall behind the curriculum. For example, in Sri Lanka, academic tests given to students at the start of grade 5 revealed that only 37% of students had mastered the reading skills (in Sinhala or Tamil) for fourth grade, and only 38% had mastered the math skills, (World Bank, 2004). In developing countries, 6.2% of primary school pupils repeat a grade each year, compared to 0.8% in developed countries, and 25.5% drop out of primary school, compared to 2.3% in developed countries (UNESCO, 2006). Arguably, the mismatch in curricula helps account for high dropout rates and low learning levels in much of the developing world.

Our results suggest that distortions in education systems due to political economy factors may go well beyond favoring spending on teachers relative to non-teacher inputs, and may be so severe as to compromise even policies with apparently clear benefits, such as textbook provision.

This paper is organized as follows: Section II provides background on primary education in Kenya, and explains the design of the textbook program. Section III presents evidence that the program had little impact on average scores. Section IV argues that the program did benefit the students with highest initial student achievement. Section V interprets these results within the larger dynamics of the political economy of education in Kenya and Section VI concludes.

II. Background: Primary Education in Kenya and the School Assistance Program

This section presents an overview of Kenya's primary education system, and explains how the textbooks program was implemented. It then discusses the selection of schools for the program, the test score data, and initial conditions in textbook and comparison schools. Finally, it examines the program's impact on textbook availability and pedagogy.

A. Primary Education in Kenya. In Kenyan schools, during the time period covered in this study, the Ministry of Education set the curriculum, administered national and district exams, and hired teachers. Local school committees, composed mostly of parents, were responsible for almost all other school costs. Fundraising for major capital expenses, such as construction, was done via *harambees*: large fundraising drives. Recurrent costs (minor repairs, chalk, books for teachers) were covered by fees set by each school. In practice, headmasters and parents often bargained over how much of the official fee the parent had to pay.

Almost all Kenyan children start primary school. Students in grades one, two, and three are taught in English, Kiswahili, and the local language (Kiluhya in two thirds of our sample and Ateso in the rest). From grade four on, all instruction is in English. At the end of grade eight, students take a national exam, the Kenya Certificate of Primary Education (KCPE) that determines which secondary schools they can attend. Some primary schools promote only strong students to grade eight in order to maintain high average scores on the KCPE exam. The students who are not admitted to grade 8 repeat grade seven or drop out. Many students also drop out in earlier grades; in our sample about 35% of the students who were in grade 3 when the program started had dropped out after three years. Physical facilities at rural schools are minimal; classrooms are often dilapidated, and sometimes non-existent. Schools usually had textbooks for teachers, but only a few for students. A Ministry of Education survey in 1990 found a pupil-textbook ratio of 17 in primary schools.

At the start of the project, the schools we examine owned very few textbooks; 80% of the students in the sample were in classrooms that had less than one English textbook for every 20 students and the analogous figures for math and science were 78% and 89%, respectively. In response, some parents purchased textbooks for their children; in our sample about 80-90% of

the textbooks that students had were purchased by their parents, rather than provided by the school. Yet even after including textbooks purchased by parents, in grades three, four and five, only one out of six students had English and math textbooks, while in grades six and seven one out of four had these textbooks. Very few students had textbooks in other subjects. Students who reached grade eight had more textbooks; about 40 percent had math and English textbooks. Access to textbooks was somewhat better than these statistics suggest, however, since typically two or three students shared a bench and if one had a textbook they would share it.

B. Textbook and Grant Provision. In late 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), which was funded by a Dutch non-profit organization, International Christelijk Steunfonds (ICS). Schools in Busia and Teso are typical of others in Kenya; in 1995, average KCPE scores in these districts were roughly at the median for Kenya as a whole. From these 100 schools, 25 were randomly selected to receive the official government textbooks in early 1996. (Kenya's school year begins in January and ends in November, with long breaks in the spring and summer.) Grades 3 - 7 received English textbooks, and grades 3, 5 and 7 received math textbooks. Grade 8 received science textbooks since many of those students already had math and English textbooks. In early 1997, math textbooks were given to grades 4 and 6, and agriculture textbooks to grade 8. Each grade and subject that was given textbooks also received one copy of the associated teacher's guide.

Textbooks were given at less than a one-to-one ratio, based on Heyneman, Jamison, and Montenegro's (1984) finding of little difference in test scores between Philippine schools randomly allocated one textbook for every two pupils and those given a textbook for each pupil. As noted above, sharing textbooks is common in rural Kenyan primary schools. A 60 percent

textbook per pupil ratio was provided in English and science, and a 50 percent ratio in math. Pupils in grades 3-5 could not take textbooks home, but pupils in grades 6-8 were put in pairs to share textbooks and were supposed to be able to take the textbook 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% of this money was spent on new textbooks, 46% was spent on construction, 10% on equipment, and the remainder on supplies. Any short run effect of grants is likely to be due primarily to textbook purchases, for two reasons. First, new classroom construction in Kenya often takes years to complete, as schools initially build what they can afford, and then stop construction until more funds become available. Second, the grants were too small to pay for an entire classroom.

C. School Selection. The 100 SAP schools were chosen because they were thought to need assistance and (with one exception) had not participated in a previous textbook distribution program (discussed in Section III). The median school average test score among SAP schools on the 1995 grade 6 and 7 district exams 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. Thus student academic performance in the 100 SAP schools was somewhat lower than in the average among 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, in alphabetical order by division names, into a single list. 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 that could be used to purchase

educational materials (including textbooks). Group 3 and 4 schools received similar grants in early 1998 and 2000 respectively. In the rest of this paper, 1995, 1996, 1997, 1998, 1999, and 2000 are referred to as years 0, 1, 2, 3, 4, and 5, respectively, of the program.

D. Description of Tests. The Ministry of Education administers district-wide exams to upper-grade primary school students to measure their understanding of all subjects in the official curriculum. Grade 8 students take the Kenya Certificate of Primary Education (KCPE) exam, which determines continuation into secondary school.[3] In all four years of the program, ICS administered additional tests in the 100 SAP schools; those tests were modeled closely on the Kenyan government tests and developed by Ministry of Education officials. At the start of year 1, ICS administered baseline tests in English, math, and science for grades 3 through 8 in the 25 group 1 schools (which received textbooks at the beginning 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 those grades that did not administer a district exam, i.e. grades 3 and 4 in year 1 and grades 3 through 7 in year 2. In years 3 and 4, ICS conducted exams in grades 3 through 8 even though there were district exams in most grades.

The 25 schools that received textbooks in year 1 can be compared to three sets of comparison schools. 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 them the 75-school comparison group. At the end of 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

[3] 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, and in year 3, they were given only in Busia district, not in Teso district.

results that use the year 1 (January 1996) pre-test scores must compare the textbook schools to the 25-school comparison group, since only those comparison schools have pretest scores.

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 be subject to common effects, such as headmaster quality.

Averaging across grades, pre-program differences between group 1 and group 4 schools in English and math scores are never more than 0.05 standard deviations and never statistically significant. This is also true for science when averaging over all grades, but the difference for grade 8 science (the only grade given science textbooks) is larger, 0.17, and statistically significant at the 10% level. Results in the last columns of Table 1 from regressions that combine all grade-subject combinations show small, statistically insignificant differences: 0.06 standard deviations for grades later 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 relative to the supply in comparison schools. Yet the impact declined 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 in which textbooks were provided 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 virtually identical for both groups of schools, 0.02 and 0.03. In grade-subject combinations that received textbooks, textbook schools had fewer privately-owned textbooks than comparison schools, 0.09 vs. 0.15, suggesting that the program crowded

out 0.06 private books per student in year 1. Yet such crowding out is small compared to the difference in school-owned books; combining school and private books, the ratio was 0.74 for textbook schools and 0.19 for comparison schools, a gap of 0.55. In contrast, in the grade-subject combinations not given textbooks the ratios were almost identical, 0.08 and 0.09.

Similar results hold in later years, but the gap between the textbook and comparison schools narrows over time in grade-subject combinations that received textbooks. In year 2, the average book per pupil ratio for school-owned books fell to 0.55, perhaps due to loss of (or damage to) school-owned textbooks. Crowding out appears slightly higher; privately-owned books per pupil were 0.09 and 0.17 in the textbook and comparison schools, respectively.

At the end of year 2, a new Dutch-funded program distributed textbooks to 21 of the 100 schools. (The 21 schools were spread evenly over the four groups of 25 schools.) Thus in year 3 the stock of school-owned books increased in comparison schools. It held steady in textbook schools because the new books offset depreciation. Overall, the textbook gap between the groups narrowed. Moreover, there were some curriculum changes, and new editions of some textbooks appeared at the start of year 2. The changes were modest, but education in Kenya is structured around the official curriculum, and some teachers who had new textbooks may not have wanted their 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. Finally, by year 4 the gap in textbooks had declined even further for grade-subject combinations that received textbooks.[4] For school-owned books, the ratio was 0.43 in textbook schools and 0.10 in comparison schools, a difference of 0.33. Including private books, the ratios are 0.48 and 0.21. The receipt of

[4] The year 4 figures in Table 4 are for English and math books only, and include only students in grades 6-8. Data are not available for science books in year 4, and no data were collected for pupils in lower grades.

additional textbooks by some schools at the end of year 2 and the changes in curriculum and textbooks in year 2 lead us to focus most of our analysis on the results of years 1 and 2.

The data on school-owned textbooks in Table 2 are from school questionnaires that include counts of those textbooks during visits to schools. To see whether students really used them, both in class and at home, a student questionnaire was administered in years 2 and 3 to students in grades 6-8 (students in grades 3-5 were considered too young to fill it out).

Students in textbook schools report having much more access to school-owned textbooks in grade-subject combinations that received textbooks; 62% of textbook school students in year 2 report having access to a school-owned book in class for the grade-subject combinations that received textbooks, compared to only 8% of students in the 50-school comparison group (Table 3, Columns 1 and 2). By year 3, the difference had narrowed; 72% of students in textbook schools were issued textbooks, versus 28% in comparison schools. More than half of students in textbooks schools report that they can take home school texts on subjects for which textbooks were given (Table 3, cols. 3 and 4), compared to less than 10% of comparison school students.

Finally, trained observers collected classroom observation data to see whether the program affected pedagogy. Overall, there were few noticeable effects. The main difference (not shown in Table 3) is that pupils use textbooks in class more often in textbook schools than in comparison schools, although differences were modest and dissipate by year 3. Specifically, in year 2 textbooks were used in 62.2% of the classroom observation sessions in the textbook schools, but only in 45.8% of the observation sessions in the comparison schools, a difference that is statistically significant at the 5% level. By year 3 this difference was smaller (45.6% in textbook schools vs. 37.4% in comparison schools) and not statistically significant. The difference in textbook use in year 2 is due both to an increase in teachers' presence in the

classroom and to an increase in teachers' use of textbooks conditional on being in the classroom. The only other significant difference in pedagogy is a small increase in homework assignment in the schools that received textbooks.[5]

III. Program Effect on Average Test Scores

After first discussing our econometric specifications, this section presents three different estimates of the effect of textbooks on average test scores: a levels estimator, a difference-in-differences estimator, and a differences-in-differences subject-based estimator that compares the difference between test scores in textbook schools in subject-grade combinations in which textbooks were and were not given to the same difference in comparison schools. All methods show little effect of textbooks on average test scores; depending on the method, they allow us to reject (at the 5% level) an effect greater than 0.2, 0.1, or 0.07 standard deviations respectively. Robustness checks show that these results are not driven by selection or attrition bias. We then show that simple retrospective estimates yield misleading results in the Kenyan context.

A. Econometric specification. Test scores units are arbitrary, so for each grade and subject combination we normalize test scores by subtracting the mean test score in the 75 comparison schools and dividing by the standard deviation. 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 percentile to the 54th.

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 econometric

[5] 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 textbooks schools than in

model with school, grade, and subject random effects. Assume that the test score of student i in grade j, subject k, and school s, t_{ijks}, is given by:

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

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. There is little difference between aggregating potentially disparate effects across grades and aggregating potentially disparate effects across students, as is routinely done. 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}:

$$t_{ijks} = \alpha_{3k}D_{3i} + \alpha_{4k}D_{4i} + \ldots + \alpha_{8k}D_{8i} + \beta_k p_s + u_{ks} + v_{jks} + e_{ijks} \quad k = \text{English, Math, Science.} \qquad (2)$$

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

Our final estimates combine all grades and subjects for which data are available:[6]

$$t_{ijks} = \alpha_{3E}D_{3Ei} + \alpha_{3M}D_{3Mi} + \alpha_{3S}D_{3Si} + \ldots + \alpha_{8E}D_{8Ei} + \alpha_{8M}D_{8Mi} + \alpha_{8S}D_{8Si} + \beta p_s + u_s + w_{js} + v_{jks} + e_{ijks}. \qquad (3)$$

---

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.

[6] 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.

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. Children 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 issues are discussed in detail below). The comparison schools are the 75-school comparison group for year 1, the 50-school comparison group for year 2, and the 25-school comparison 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 the first two years of results.

B. Impact on Average Test Scores. The program effects are expressed in standard deviations of test scores, i.e. the coefficient indicates by how much the average test scores increased in schools that received textbooks compared to schools that didn't. The estimated program effects for year 1 range from zero for English to 0.06 for math (not shown).[7] Aggregating over all subjects gives a statistically insignificant impact of 0.02 standard deviations (Table 4 column 1). The standard error of 0.086 allows one to reject (at the 5% significance 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

---

[7] All regressions include controls for sex and for Busia's seven geographic divisions. Regressions without such controls are also consistent and yield similar results, but adding them often improves the precision of the estimates.

grades yields a statistically insignificant impact of 0.02 (Table 4 column 2), the same as the estimated average effect after one year, which rules out an average impact of 0.23 or higher (at the 5% significance level). All estimates of program impacts for years 3 and 4 (not reported in Table 4) are slightly negative and rule out impacts of 0.20 or higher.

Comparing differences in pre-test and post-test scores across textbook and comparison schools yields more precise estimates (Table 4, Columns 3 and 4). These difference in difference estimates for the schools that participated in the year 1 pre-test, aggregating over subjects, yield point estimates for the textbook treatment effect after one and two years of 0.02 and -0.04, 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) are slightly negative, and rule out impacts of 0.10 and 0.08 or higher, respectively.

We also use a difference-in-difference subject-based estimator that regresses test scores on dummy variables for whether students were in textbook schools *and* whether they were in subject-grade combinations that received textbooks. This estimator requires another assumption, that textbooks in one subject do not affect scores in another, but it is more robust to selection or attrition bias. To see the intuition, first calculate the difference in normalized test scores between treatment and comparison schools in subject-grade combinations provided with textbooks. Then calculate the difference in scores between treatment and comparison schools in subject-grade combinations that did not get textbooks. The difference between these differences is the subject-based estimator. Comparing subjects within a single grade that vary in whether a textbook was received effectively compares test scores across subjects for the same student. This approach may offer greater precision by differencing out random variation at the school, class, and student level. If provision in one subject diverts student time from other subjects, this estimator will

overestimate the impact of textbooks, while positive spillovers across subjects will underestimate it. (The latter scenario implies a positive coefficient on the textbook-school dummy.)

This method works best for year 1 because that year has the most variation across grade-subject combinations in receipt of textbooks (grades 4 and 6 received math textbooks in year 2). Combining all grades and subjects into one regression (table 4, column 5) yields a slightly negative (–0.01 standard deviations) estimate of the direct effect of receiving a textbook. The small standard error allows us to reject (5% level) the hypothesis that textbooks raise the average test score by 0.07 or more standard deviations.

Thus, all three estimates show that the impact of textbooks is close to zero. Depending on the assumptions made, one can reject effects as large as 0.2, 0.1, or 0.07 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 students into grades 3, 5 and 7, hoping that they would receive more textbooks. Twenty-two percent of textbook school students repeated a grade from year zero to year 1, while 26% did so in the comparison schools (table 5, third row). Yet any bias due to differential repetition is likely very small. To see why, consider the extreme assumption that the "extra" 4% of students promoted in the textbook schools are the most marginal students: those who scored the lowest on the year 1 pre-tests. "Demoting" these students down one grade and re-estimating the level regressions in Table 4 for year 2 yields an estimated impact of textbooks of only 0.044 standard deviations, with a standard error of 0.098. This is only slightly larger than the year 2 estimate in Table 4, despite the extreme assumption.

Attrition bias is also unlikely to explain the results. In year 1, there is little difference in the percentage of students not tested: 26.0% in the textbook schools and 26.3% in the 25-school comparison group (Table 5, row 5). In year 2, 33.3% of students in comparison schools and 31.0% in the textbook schools were not tested. Yet differences in pre-test scores of students who were not tested in textbook and comparison schools are insignificant, with different signs for different subjects. Thus the slightly higher propensity of children in textbook schools to be tested in year 2 is unlikely to lead to substantially biased estimates of textbooks' impact on test scores. Attrition in years 3 and 4 is higher, another reason we do not focus on those results, but there is no evidence that it is asymmetric across groups. Note also that the simplest selection and attrition stories cannot account for the failure to find effects in the difference-in-differences specification or the differences-in-differences specification across subjects.

D. Retrospective Estimates. The failure to find an impact of textbooks contrasts sharply not only with many people's priors, 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 of 18 studies, noting that some studies find greater effects of textbooks for disadvantaged students. Fuller (1986) reports significant effects of textbooks in 14 of 22 studies, and Fuller and Clarke (1994) found significant effects in 19 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 first-grade classrooms in Nicaragua that were randomly assigned to receive radio mathematics education, with 20 that received math workbooks and 20 that served as controls.[8] After one year, pupils who received

[8] Heyneman, Jamison and Montenegro (1984) compare Philippine schools randomly selected to receive one textbook for either every one or two students, and find little difference in test scores. They also compare these two

workbooks scored one-third of a standard deviation higher than control group pupils, a difference significant at the 1% level. No significant interaction was found between pre-test scores and receipt of workbooks, although workbooks narrowed gaps between rural and urban students.

Simple retrospective estimates using Kenyan data also suggest that textbooks raise test scores. Cross-sectional (non-experimental) variation in textbook availability in the 75 schools that did not receive textbooks in year 1 can be used to estimate the impact of textbooks on test scores. We present OLS regressions (with school-level random effects) of normalized test scores on (i) a dummy variable indicating students with privately owned textbooks; (ii) the textbook to student ratio for school-owned textbooks (calculated separately for each grade-subject combination in each school); and (iii) other school and family characteristics. Students with privately owned English textbooks scored 0.18 standard deviations higher on English exams, controlling for parental education and land ownership, and those with math and science books scored 0.09 and 0.05 standard deviations higher on those exams (Table 6, row 1). The English and math impacts are highly significant, but the science impact is insignificant. Aggregating over all three subjects yields an impact of 0.12, with a tight standard error of 0.016.

School-owned textbooks and test scores are negatively correlated in these estimates (table 6, row 2) but this may be because the government or donors gave more textbooks to the neediest schools. A difference-in-difference estimate of the impact of school-owned textbooks can be obtained from data on a World Bank textbook provision project carried out in 1994 by the Jomo Kenyatta Foundation. It provided textbooks to 95 of Busia's and Teso's 334 primary schools, at ratios of about one book for every two pupils in English and math and one for every four in Swahili and science, targeting the neediest schools. We have school-level average scores on

---

scenarios with no textbooks, and they 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.

government exams in grades 6, 7, and 8 for about 80% of the 334 schools. A basic difference-in-differences analysis suggests that textbooks raised grade 7 test scores by 0.50 standard deviations in 1994 and raised grade 7 and grade 8 scores by about 0.65 standard deviations in 1995 (Table 7).[9] No significant impact is seen on the grade 6 scores, nor on the 1994 grade 8 scores.

Overall, retrospective estimates of textbooks' impacts on test scores based on cross-sectional data on privately-owned textbooks or longitudinal data on school-owned books suggest a stronger positive impact of textbooks on test scores than is found by a randomized evaluation.[10]

To summarize Section III, there is no evidence that the program increased test scores for the average student. Our three estimators for the impact on average test scores allow us to reject average effects greater than 0.20, 0.13 and 0.07 standard deviations in year 1, and data from later years provide no evidence that textbooks' impact on test scores accumulates over time. This result does not seem to be driven by selection or attrition bias. Moreover, it differs from results found by using standard OLS techniques to assess the impact of textbooks; simple retrospective estimates show sizeable and statistically significant impacts of textbooks on student test scores.

IV: Interactions between initial test scores and program impact

What explains the lack of an impact of textbooks on test scores in data from a randomized evaluation? Why are retrospective estimates positive, even controlling for parental education and wealth? To shed some light, it is useful to disaggregate the impact of textbooks by students' initial academic achievement. One might have expected to find the strongest impact for the weakest students, since they were less likely to own a text book initially. The correlation

[9] 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.

[10] 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.

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). Consistent with this fact, in grades 6-8 in the textbook schools, students with lower test scores are more likely to report receiving a textbook from their school (the average pre-test score was 0.22 higher for students not given a textbook). Moreover, as noted above, some retrospective studies suggest that textbooks are most helpful to the weakest students (though these estimates may be biased).

This section presents evidence that the provision of textbooks in rural Kenya primarily benefited the strongest students. We first examine interactions between initial student pre-test scores and the impact of textbooks on post-test scores. Interacting program impact and student characteristics should be done with caution, given the potential for data mining, but conditioning on initial values of the dependent variable is a natural interaction to examine to understand the main effects. We then 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 the program mostly benefited strong students, since only those students reach grade 8 and have a hope to progress to secondary school. Next, we present evidence that many students had difficulty even reading the textbooks, let alone using them effectively. Finally, we examine a few alternative explanations of our results and find little evidence to support any of those explanations. It is worth noting that if textbooks disproportionately benefit the strongest students, one would expect parents to be more likely to buy textbooks for such students, biasing upward retrospective estimates of the impact of textbooks on test scores.

A. Interaction Effects. In year 1, an interaction term between the program variable and the average pre-test score is highly significant when aggregating across all subjects:[11] the program increased scores by 0.057 standard deviations more for students with pre-test scores one standard deviation above the mean (Table 8, column 1). This is also true after two years; the interaction term is even slightly higher at 0.064 standard deviations (Table 8, column 2).

Most of the difference in the program impact seems to reflect interactions between the program and student characteristics, instead of interactions with school or teacher characteristics. To check this, we added school fixed effects to the regressions in Table 8 (not shown in that table); this prevents estimation of the program effect but allows 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 effects in Table 8; for example, the coefficient of 0.048 for all subjects combined in year 1 is only slightly smaller than the estimate of 0.057.

Pre-test scores may measure initial achievement with substantial noise, so the coefficients on the pre-test and on the interaction between the pre-test and the program may have attenuation bias, underestimating the true impact of initial academic achievement, and its interaction with the program, on later achievement. Suppose the true coefficient on initial academic achievement is one, as in many difference-in-difference specifications. If so, attenuation bias is 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.

One approach to correcting attenuation bias is to instrument pre-test scores with scores in other subjects. An important caveat is that instrumental variable (IV) estimates are consistent only if, conditional on the pre-test score in subject k, a student's pre-test score in another subject

[11] Similar results are found using only the English pre-test score (recall that all textbooks are written in English). This is not surprising because English pre-test scores and average pre-test scores are highly correlated.

has no effect on his or her post-test score in subject k. Thus if English skills help students learn math and science and math skills help in science the exclusion restriction will be violated. Yet it is unlikely that math and science skills help in English, so IV estimates of differential program effects in English should be consistent.

IV estimates of the textbook effect are about what one would expect under the hypothesis that there is an underlying coefficient of one on initial levels of learning, and that the coefficient on the pre-test is less than one because the test is a noisy signal of initial learning. The IV estimate of the interaction is 0.14 standard deviations in year 1 and 0.13 in the year 2, and both are significant at the 1% level (Table 8, column 3). Given potential problems with instrumenting math scores with English scores, we ran regressions in which math scores were instrumented only by science scores. The results were very similar, both in magnitude and significance.

Interactions are also significant in regressions of changes in scores and regressions that distinguish between being in a textbook school and receiving textbooks in particular subjects. Table 8 (column 4) adds an interaction term to the year 1 regression in column 3 of Table 4, in which the dependent variable is the (post-test) score minus the pre-test score. That term is significant at the 5% level. Lastly, column 5 of Table 8 shows results after adding a term that interacts receipt of textbooks with the (average) pre-test score. This interaction of 0.07 standard deviations is significant at the 1% level for all subjects combined.

The estimates in Table 8 require the interaction between the program and pre-test scores to be linear, and they constrain school random effects to be equal for students of different ability, ruling out scenarios where some schools are above average for high-ability students but below average for low-ability students. Table 9 relaxes these constraints by dividing the sample into

quintiles, based on average pre-test scores, and re-estimating the level regressions in Table 4 by quintile. This allows 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 for each 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 are significant for quintiles 4 and 5 at the 10% and 5% levels, respectively.[12] Since pre-tests are noisy, some students in the top quintile were not necessarily in the top quintile of initial learning. Thus the true effect in the top quintile of initial learning will be greater than 0.22 standard deviations. The pattern of effects across quintiles in year two is quite 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 beyond year 2 of the program, but the data do not suggest any interaction effects in later years. This is not surprising. First, textbook availability and usage converged in later years, as seen in Tables 2 and 3, so in later years most academically strong students in comparison schools probably had access to a shared textbook. Second, textbook use in class converged between treatment and comparison schools over time, 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. (Classroom observations in year 2 in grade-subject combinations that received textbooks showed that textbook use in class was 16.4 percentage points higher in textbooks schools; but by year 3 this difference was only 8.2 percentage points) Third, all interaction effects are based on pre-tests at the start of year 1, and as time passes they presumably measure current achievement and motivation less precisely. This is consistent with the fall from year 1 to year 2 in the pre-test

[12] 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.

score coefficient in Table 8. Finally, the sample size falls each year as children drop out or complete primary school, reducing the precision of the estimates.

B. Other Educational Outcomes. Evidence for other educational outcomes is consistent with the finding that textbook provision benefited academically strong students but not weaker ones. Consider drop out rates and grade repetition. Students enrolled in grades 3-7 in year 1 had four possible outcomes in year 2: stay in school and be promoted, stay in school and repeat, drop out, or transfer out. The same is true for grade 8 students, except that promotion has two forms: finish primary and leave school, or finish primary and enter secondary school. After one year, children in grades 3-7 show no significant differences in dropout, transfer or repetition rates across the textbook and comparison schools (table 10, columns 1 and 2). (Statistical significance is based on a probit regression with school random effects). Textbooks also had no impact on absence rates. These results are consistent with no program effect on average and below average students, who are most likely to be absent, repeat, and drop out.

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% vs. 38%). This difference is significant at the 5% 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 academically strong students go to secondary school.

C. Why Do Textbooks Only Increase the Scores of Strong Students? One possible reason why students with high initial academic achievement may benefit more from the official textbooks provided by the program is that those textbooks are too difficult for the other students. Indeed, all the textbooks are written in English, the third language for almost all students, and 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 16% of the median students could read the grade 3 English textbook, and only 28% of the grade 4 median students could read their English textbooks (Table 11). Difficulty literally reading the textbook is less common in upper grades – the figures are 67% for grade 5 and over 90% for grades 6-8. Yet even if students can read the words in the textbooks, many may have difficulty effectively using a text book in their third language (Table 11 columns 2-4).

Differences in whether students took textbooks home may explain part of the differential impact of textbooks across weak and strong students, but not most of it. Recall that students in grades 6-8 were allowed to take textbooks home, but not younger students. In grades 6-8, students in textbook schools who took the textbooks home had higher average pre-test scores (0.21) than students who did not (-0.03), a difference statistically significant at the 1% level. Yet the year 1 interaction effect for the level regression (without instruments) in grades 3-5, in which students could not take textbooks home, is 0.051, only slightly smaller than the estimate of 0.072 for grades 6-8 (these estimates are not reported in Table 8).

In developing countries, where both students and teachers are frequently absent (our data show absence rates of about 20% for both students and teachers), many students fall behind the official curriculum. Indeed, international comparisons reveal very low test scores for developing countries, even though most developing countries that participate are middle income countries. For example, only 3% to 6% of 15 year old students in developed countries had very low reading skills on the PISA standardized international reading test, while this share was 20% to 30% in most of the participating developing countries (OECD, 2004). In such settings, using a

curriculum and associated textbooks that assume that students know the material associated with previous grade levels will yield a serious mismatch.

D. Robustness Checks, Evidence from Other Programs, and Alternative Explanations. Evidence from another source, the grant program, conducted in the 25 Group 2 schools, lends further support to the hypothesis that textbooks were best suited for the strongest students. The 25 schools that were given grants in year 2 spent almost half of their grants on textbooks. Thus we have a second, quasi-independent randomized trial to evaluate. (Strictly speaking, this is not a fully independent trial since the Group 2 schools are control schools for evaluating year 1 outcomes, and analyses of the year 2 data for grants and for textbooks use the same comparison schools (50-school comparison group), but it is independent in the sense that the 25 textbooks schools are excluded from the grants sample.) We cannot rule out that grant expenditures for purposes other than buying textbooks may have also affected scores in these schools, but most of that expenditure was on classroom construction, which takes time, so effects likely largely reflect textbook purchases.

Estimates of equations (2) and (3) for the schools given grants in year 2 are shown in Table 12. Level results (columns 1) show an average point estimate of the effect of receiving the grant on average test scores of about 0.13 standard deviations, but it is not significant. Difference-in-differences results (columns 3), yield marginally significant results of about 0.12 standard deviations. Estimates based on the 25 schools that received grants are also consistent with the results in Tables 8 and 9. First, an interaction term between pre-test scores and the program dummy variable (not shown in Table 12) was positive and statistically significant (at the 5% level). Second, estimates similar to those in Table 9 (not shown) show significant impacts for quintile 5 (5% level in level regressions and 10% in differenced regressions) but no

significant impacts for the other four quintiles. While the point estimates in column 1 for the effect of textbooks purchased using grants is substantially larger than those for the textbook program, the difference is not statistically significant. If the higher point estimate reflects a real difference, rather than sampling variation, it may be due to schools choosing to buy textbooks that are particularly in need of them, or it may reflect non-textbook expenditure from the grants.

Another piece of evidence consistent with the conjecture that curricula may leave many students behind is from an evaluation of a program that hired extra teachers for some Kenyan schools, allowing grades to be split between two teachers (Duflo, Dupas, and Kremer, 2007). In some schools, students were assigned randomly to the new teachers, whereas in others students were grouped by initial academic preparedness based on initial test scores. Students with low pre-test scores fared considerably better when grouped with other students who had similar pre-test scores, presumably because teachers could teach to their initial academic background.

One possible reason for little 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 and a school committee questionnaire suggest that in small schools the program crowded out *harambee* fundraising, which focuses on classroom construction. However, even if ICS assistance reduced that fundraising, there would be 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, and the resulting test score gain is presumably 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 across 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.[13] Thus crowding out is unlikely to explain the lack of impact of textbooks on average test scores.

A final explanation for our results is that the tests were "too hard" for most students; this would explain both the lack of an impact on most students and the significant impact on the best students (for whom the tests may have been appropriate). Indeed, the district tests in year 1 were quite difficult for many students; in some subject-grade combinations, the average scores were not much higher than random guessing, which implies little information content in those tests. Yet there was no evidence for textbook effects in those subject-grade combinations with higher average test scores. In particular, the information content of the tests is higher for grades 6-8 than in grades 3-5, but in both level and difference regressions there is no evidence of a significant impact in grades 6-8 or even of higher point estimates than in grades 3-5. Moreover, the ICS tests used in year 2 and later years 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 more detailed information, see Glewwe, Kremer and Moulin (2002).

Section V: Political Economy of Textbook and Curriculum Mismatch

The fact that provision of Kenya's official government textbooks only led to test score increases for the most talented students is arguably part of a much larger mismatch between the curriculum and the needs of many students. Schools are judged mainly on average KCPE results, giving them little incentive to focus on students who will not make it through 8th grade, or will bring down average scores if they get there. This mismatch leads to high repetition rates (21%,

[13] 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.

see Table 10), since the typical student needs more time than is allotted to learn the required material, and high dropout rates (17%, see Table 10), as students fall behind and cannot follow the material presented in class.

The problem compounds over time. Since teachers are judged based on their students' performance on tests covering the material in their grade, teachers for the next grade may have incentives not to pick up where their predecessors ended, but instead start with the more advanced material in the official curriculum for that grade.

Why is the system geared toward the strongest students? This brings the discussion into the realm of political economy, a topic inherently more speculative than estimating the impact of textbooks on students' educational outcomes. We argue that the problem is due to the confluence of three factors common in many developing countries: 1) the adoption of a centralized, uniform national curriculum and education system; 2) substantial heterogeneity in the student population, due in part to rapid educational expansion; and 3) disproportionate political power of elites.

A. Centralized Uniform Curriculum. Kenya, like many post-colonial countries, adopted a centralized system with a single uniform national curriculum and competitive national exams to enter higher levels of education. This is not the only possible choice: some nations either decentralize authority or have multiple tracks within a single national system, instead of a single centralized uniform curriculum. The United States exemplifies the former alternative: there is no national exam, many decisions are decentralized to local school districts, and parents have some choice of school district. Historically, much of Europe pursued the latter alternative: a national education and exam system that includes different types of schools, so some students follow an academic and others a vocational track. In Germany, for example, students are placed into different types of schools at age 10: highly academic, commercial (high vocational) or

general (part-time vocational and apprenticeship training). In 2001, only 43% of secondary students were in highly academic schools (Federal Statistical Office, http://www.destatis.de/themen/e/thm_bildung1.htm).

In contrast, many developing countries have a single national system with little local control over schools. Such systems may reflect political goals, such as unifying countries left with substantial ethnic diversity at the end of colonial rule. Yet, unlike Europe, students have few opportunities to follow a vocational track. In most African countries, for example, less than 5% of secondary pupils attend a vocational school. In Kenya, only 1.7% of secondary students attend such schools. The situation is similar in Asia; in most Asian developing countries – China being a major exception –only a small percentage of secondary students are in vocational tracks.

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 eight (the last grade of primary school in Kenya) in our sample, the average age is 15, and 20% are 16 or older. Among pupils starting grades six or seven the average age is 13.6 and 38% are age 14 or older. A typical U.S. student who finishes 13 years of education has spent half of his or her time in a system that had tracking within schools. In contrast, a typical Kenyan student who completes eight years of education has spent no time in a tracked system.

B. Rapid Educational Expansion and a Heterogeneous Student Population. As access to education expands, curricula are often adjusted. For example, at the start of the 20th century, U.S. secondary schools served only 10% of the population and were highly academic, preparing students for tertiary education. In 1910, 50% of high school students took Latin, (Ravitch, 2000). In the following decades, the United States rapidly expanded secondary

education; by the mid 20th century, high school enrollment reached 65%, bringing in children from very different backgrounds. This educational expansion led to a transformed curriculum, including the incorporation of vocational education (Ravitch, 1983). Similar issues have arisen more recently in Europe, as access to tertiary education and secondary academic education has expanded. Changes in curriculum have been controversial in each case.

As in much of Africa, access to education expanded tremendously in Kenya following independence. However, changes in the curriculum did not match the changes in enrollment. Kenya's colonial education system was designed to produce a small group of elite Africans to work in the colonial bureaucracy. It had a demanding curriculum with no remedial education for those who fell behind. In 1960, just before independence, Kenya's gross primary enrollment ratio was only 47%. By 1980, it had almost doubled to 90% (Deolalikar, 1999). With the introduction of free primary education, the primary school completion rate is now 91% (World Bank, 2007). Kenyan students are extremely heterogeneous in their family background, preparation for schooling, and economic circumstances. 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 live without electricity, which substantially limits study time at home.

Kenya made some changes to its curriculum, but the system remained oriented towards elites. References to Kenyan history and geography replaced references to English history and geography. Pictures in books are of Kenyan, not English, children. But the language of instruction remains English, most students' third language. The curriculum is arguably better suited to students with educated parents than to typical pupils. The factors discussed above combine to create a system that is ill-suited for many students, so that providing basic inputs

such as textbooks benefits only the strongest pupils.

C. Disproportionate Elite Power. Even given Kenya's heterogeneity and its decision to adopt a uniform curriculum, Kenya could have adopted a curriculum and textbooks more suited to the average student. Tanzania is an interesting contrast; at independence it also had a centralized, uniform education system and considerable heterogeneity in pupils' backgrounds. Yet unlike Kenya, Tanzania chose an education system that is arguably well suited to most of the population (which reflected Nyerere's socialist views) but less suited to the elites' interests. For example, Tanzania uses Swahili, not English, as the language of instruction. This may have eased the attainment of basic literacy for the disadvantaged, but it restricted educational and economic opportunities for those who reached higher levels of education. Tanzania also invested heavily in primary education, but much less in higher levels of schooling, perhaps creating sizeable economic costs by neglecting secondary education (Knight and Sabot, 1990). Recent political developments in the US provide another example suggesting that national standards need not always suit the interests of elites. Reback (2006) argues that test-score based accountability systems that focus on minimum competency – such as the "No Child Left Behind Act" – generate incentives for schools to target efforts on low-performing students: those on the margin of passing the tests. Using data on Texas' 1990s school accountability program, he shows that students with low initial test scores had the largest increase in test scores when schools had strong incentives to raise the test scores of their low-achieving students.

Historically, four factors led to a system in Kenya in which curricula – and textbooks in particular – were not targeted towards the typical student. First, for most of its post-independence history, Kenya has been a de facto or de jure one party state, and elites have had disproportionate political power. These elites tend to prefer an education system targeted to their

children's needs. Ministry of Education officials who design curricula and textbooks tend to favor their children's interests, as well as to reflect the norms of their profession, seeing efforts to design a curriculum more suitable for rural areas as lowering academic standards. To the extent that elites in post-independence societies see things through the prism of international comparisons, they may resent efforts to lower the level of the curriculum and textbooks.

Second, even parents of average students may favor a curriculum designed for stronger students in order to secure more desirable peers in their children's schools, because a curriculum best suited for the typical student may cause elites to switch from government to private 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. Students who drop out before taking the KCPE exam are excluded from calculations of school performance, so teachers have incentives to use textbooks suited to the strongest students. More generally, in the current system teachers have little incentive to help weaker students, and 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; that market consists primarily of elite students and teachers, since poorer students rarely buy textbooks. Thus textbooks are usually designed for students at the top of the distribution of student performance rather than for typical or below average students.

D. Why Did Kenya Choose These Policies? The analysis above suggests that many children may be left behind when three elements are combined: 1) a centralized, uniform curriculum; 2) heterogeneity in the student population associated with rapid educational

expansion; and 3) disproportionate political influence of elites on curricula decisions.

The political economy of Kenya from independence in 1960 to the end of the Cold War helps explain why it chose these education policies. The focus on national unity at independence made a centralized education system attractive. As the political coalition in power represented the most educated ethnic groups, maintaining a centralized, meritocratic system allowed these groups to obtain jobs as teachers throughout the country, rather than only in their home areas.

Post-independence, civil service jobs offered large rents, and access to these jobs depended on formal, academic educational qualifications, not vocational training. In this situation, no ethnic group had incentives to push for vocational education in its area, or to alter the schools in its area in ways that might improve learning for weaker pupils but could harm chances for strong pupils to win the academic sweepstakes, since such a policy would reduce the representation of that ethnic group in the civil service.

Without multi-party democracy, political competition was largely between elites from each region, so the system favored these elites. In an economy with limited industry, where civil service jobs were the main route to economic security, vocational education had no constituency of industrialists needing skilled workers, nor of parents seeking a better life for their children.

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 the elite. A key element of politicians' recent efforts to appeal to these voters was a promise to abolish school fees. This led to a large jump in enrollment. Many urban elites felt that the influx of new pupils without any increase in the number of teachers led to a decline in school quality, and left for private schools. This could induce a shift in the public curricula toward the average

student. Yet elites are likely to retain a powerful hold on the bureaucracy for some time, and while abolishing school fees may be a salient issue that politicians can use to appeal to the poor during elections, reforming the curriculum may be less salient and more subject to civil servant control. Thus it is unclear whether curricula will be reformed to better suit typical students.

Two policies could help less prepared students. A first option is remedial education for children who have fallen behind the official curriculum. Banerjee, Cole, Duflo and Linden (2005) examined one such program in India and found it to be very effective. India's curriculum also appears to be too difficult for many students (who often are first-generation students), and many fall behind. The program offers remedial education to children who reach grade three or four 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 benefited weakest students the most.

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 may cover it in nine or ten years. All students would take the same curriculum, so students would not be irretrievably sorted at an early age into an academic track or a more basic, vocational track. 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, Web Page).

In one sense, such a system would formalize and rationalize what happens informally – and inefficiently – now. Officially, everyone in Kenya faces an eight-year curriculum, but many students require nine or ten years. Yet 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.

Another way in which some Kenyan schools already deliver heterogeneous education is by teachers not covering the entire curriculum in one year; it is routine in a given year for teachers not to finish that year's curriculum. Yet structuring education in this way is inefficient, since teachers in the next grade are unlikely to teach the unfinished curriculum of the previous grade, but rather start with the curriculum for that grade. This tendency is reinforced by the exams, which cover the assigned curriculum for that grade. Thus a grade 7 teacher who helps students complete the grade 6 curriculum may produce students who score lower on the grade 7 district exams. Such teachers are unlikely to be appreciated by headmasters, colleagues, or the community in which they teach. In contrast, if teachers cover only the grade 7 curriculum, at least some students will do well.

VI. Conclusion

Providing textbooks to schools where few students have them seems to be an obvious way to raise students' educational performance 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 textbooks in Kenya did not increase average test scores, although it did increase the scores of students with high initial achievement. The latter finding suggests that the government's official textbooks used were ill-suited for the typical

student. 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. Proximate reasons why textbooks are ill-suited to the typical student in this population may include the fact that textbooks are bought only by a small minority of students and that teachers focus effort on the best prepared students, who are likely to make it through the exam at the end of primary school. More fundamentally, this may reflect a larger dilemma with centralized uniform educational systems, a heterogeneous student population, and entrenched elite power.

The problem of a curriculum that does not match typical students' needs and thus leads 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% of primary school children repeated a grade in 2002-03, and some countries have higher rates, such as 34.4% in Gabon and 25.8% in Cameroon. Only 68.6% of African students reach grade 5 (UNESCO, 2006). The grade 5 survival rate in India is 61%, and first-generation pupils have great difficulty keeping up with the curriculum (Banerjee *et al.* 2005). Indeed, a World Bank (1997, p.168) report provides evidence that primary school textbooks in India are too difficult for many students. High drop out rates and repetition in many developing countries suggest that the curricula are not well suited to the average student.

Many have argued that distortions in education 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, non-teacher inputs are likely to be very effective. Our results suggest that political economy distortions in education systems may even limit the impact of one of the most basic non-teacher 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, disproportionate elite influence, and the heterogeneous student population that comes with rapid educational expansion is likely to leave many children behind. Future research should examine potential reforms that could broaden access to learning, as well as to schooling, including reorientation of curricula, decentralization, tracking and vocational education, and remedial education.

Table 1: Differences in Normalized Pre-Test Scores between Textbook Schools and 25-School Comparison Group

Subject	English		Math		Science		All subjects combined	
Grade	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 pre-test scores from January of year 1on 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 pre-test 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

		School-owned books		**Privately-owned books**		**Total**	
Program Year	*Subject/Grade given textbooks?*	*Textbook schools*	*Comparison schools*[a]	*Textbook schools*	*Comparison schools*	*Textbook schools*	*Comparison schools*
1	Yes	0.65	0.04	0.09	0.15	0.74	0.19
	No	0.02	0.03	0.05	0.07	0.08	0.09
2	Yes	0.55	0.04	0.09	0.17	0.64	0.21
	No	0.04	0.03	0.08	0.08	0.12	0.12
3	Yes	0.52	0.11	0.08	0.14	0.61	0.25
	No	0.11	0.09	0.09	0.09	0.20	0.19
4	Yes	0.43	0.10	0.05	0.11	0.48	0.21
	No	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.

[a] Comparison schools: 75-school group in year 1, 50-school group in year 2 and 25-school group in year 3 and 4.

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	*Comparison schools*	*Textbook schools*	*Comparison schools*
		(1)	*(2)*	*(3)*	*(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

Note: 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 where administered only in those two years.

Table 4:Impact of Textbook Program on Normalized Test Scores

Dependent Variable	Normalized test score [a b]	Normalized test score [b]	Normalized test score minus pretest score [c]	Normalized test score minus pretest score [c]	Relative normalized test score [d]
	(1)	(2)	(3)	(4)	(5)
Textbook school	0.022 (0.086)	0.023 (0.105)	0.019 (0.053)	-0.039 (0.070)	0.036 (0.083)
Received a textbook					-0.009 (0.040)
Region and sex dummies	YES	YES	YES	YES	YES
Years exposed to textbooks	1	2	1	2	1
Grades	3-8	4-7	3-8	4-7	3-8
Observations	24,132	12,487	11,321	7,377	47,116

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%. Standard errors in parentheses.

[a] 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.

[b] Sample includes all children enrolled in January of year 1 who took the relevant October/November test

[c] Sample includes all children who were enrolled in January of year 1 and took the relevant October/November test as well as the pre-test in January of year 1.

[d] Relative normalized test scores by whether subject grade combination received text book. Sample includes all pupils in 25 textbook schools and 75-school comparison group in January of year 1 who took the year 1 post-tests.

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)[a]			
Dropouts (%)	5.3	6.0	-0.7
Transfers out (%)	5.2	3.6	1.6
Composition of students, beginning of year 1 (50 schools)[b]			
Repeaters (%)[c]	21.9	26.0	-4.1***
Transfers in (%)	11.2	10.3	0.9
Students present at start of year 1 but not tested at end of year 1 (100 schools)[d]			
Year 1 (%)	26.0	26.3	-0.3
Year 2 (%)	31.0	33.3	-2.3***
Year 3 (%)	38.6	39.9	-1.2
Year 4 (%)	45.2	47.9	-2.7

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%

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.

[a] Data for year 0 is available for only 20 schools (10 in Group 1 and 10 in Group 4).

[b] 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.

[c] The percentage of repeaters is slightly underestimated for both types of schools because there is no information on repetition for nearly one half of students who transferred in (about 6% of all students).

[d] 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.

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

* significant at 10%, ** significant at 5%, *** significant at 1%.

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.

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

Each cell represents a regression of normalized test score in 1994 or 1995 minus normalized test score 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.

Table 8: Normalized Test Scores as a Function of Treatment and Pre-Test Score

Dependent Variable	Normalized test score (year 1)	Normalized test score (year 2)	Normalized test score (year 1, IV estimates)	Normalized test score minus pre-test scores (year 1)	Relative normalized test score (year 1)
Textbook school	0.060	-0.014	-0.035	0.021	0.065
	(0.061)	*(0.083)*	*(0.066)*	*(0.060)*	*(0.078)*
Received a textbook					-0.006
					(0.047)
Pre-test score	0.429	0.345	0.839	-0.338	0.384
	(0.013)	*(0.016)*	*(0.042)*	*(0.016)*	*(0.007)*
Pre-test*Textbook school	0.057***	0.064***	0.144***	0.042**	
	(0.018)	*(0.022)*	*(0.055)*	*(0.021)*	
Pre-test*Received a textbook					0.070***
					(0.014)
Number of observations	11,342	7393	11,211	11,321	22,130

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%

The first three columns are regressions of October/November test scores on the relevant year on dummy variables for textbook school, sex and region, the average of the three year 1 pretest scores, and an inter-action term between the textbook school dummy and the average pretest score. The interaction term for IV results is the pre-test score for the subject (not the average over subjects); the instruments were the scores on the two other tests. The dependent variable in the 4th column is the October/November test score minus pre-test score. Each regression includes all children enrolled in January of year 1 who took the January pre-test and the relevant October/November test.

Table 9: Normalized Test Scores by Quintile of Pre-Test Scores

Years exposed	*Quintile 1*	*Quintile 2*	*Quintile 3*	*Quintile 4*	*Quintile 5*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
1	-0.049 *(0.064)*	-0.021 *(0.069)*	0.032 *(0.073)*	0.142* *(0.079)*	0.218** *(0.096)*
2	-0.080 *(0.081)*	-0.096 *(0.094)*	-0.090 *(0.103)*	0.022 *(0.100)*	0.173 *(0.133)*

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%

Each row represents five random effects regressions, one for each quintile (based on pre-test 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 pre-test in year 1 and the relevant post-test. All results are aggregated over all grade/subject combinations that received textbooks.

Table 10: Promotion, Repetition, and Dropping Out

	Lower Grades (3 – 7)		***Upper Grade (8)***	
	Textbook schools	*Comparison schools*	*Textbook schools*	*Comparison schools*
	(1)	*(2)*	*(3)*	*(4)*
Stayed, promoted	.53	.53		
Finished primary, no secondary			.32***	.41
Entered secondary			.43**	.38
Stayed, repeated	.21	.21	.16	.14
Dropped out	.17	.17	.01	.03
Transferred out	.08	.09	.06	.04
Number of students	5009	4838	447	440

Notes: * difference with comparison group significant at 10%, ** significant at 5%, *** significant at 1%

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.

Table 11: Understanding of the English Textbook by the Median Student

Grade	*Can read the book (%)*	*Can answer questions, in English, about the passage (%)*	*If unable to answer question in English, can answer when asked in Kiswahili (%)*	*Can answer written questions, in English, from the book (%)*
	(1)	*(2)*	*(3)*	*(4)*
3	16	10	51	37
4	28	10	61	12
5	67	49	51	47
6	92	86	82	39
7	96	77	22	56
8	100	99	-	92

Note: 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.

Table 12: Impact of Providing Grants to Group 2 Schools in Year 2 (after one year of the program)

By subject (grades 3-8 combined)	*Test score levels in year 2*		*Test score differences (year 1 to 2)*	
	Coefficient on textbook schools	*Number of observations*	*Coefficient on textbook schools*	*Number of observations*
	(1)	*(2)*	*(3)*	*(4)*
English	0.129 *(0.112)*	12,323	0.111* *(0.067)*	7,563
Math	0.143 *(0.098)*	12,215	0.101* *(0.061)*	7,490
Science	0.124 *(0.101)*	12,168	0.137* *(0.078)*	7,482
All subjects	0.132 *(0.100)*	36,706	0.121** *(0.057)*	22,535

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%. 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.

References

Bamgbose, Ayo. 2004. "Language of Instruction Policy and Practice in Africa." Paris: UNESCO.

Banerjee, Abhijit, Shawn Cole, Esther Duflo and Leigh Linden. 2005. "Remedying Education: Evidence from Two Randomized Experiments in India." Department of Economics, MIT.

Deolalikar, Anil B. 1999. "Primary and Secondary Education in Kenya: A Sector Review." Revised February 1999.

Fuller, Bruce. 1986. "Raising School Quality in Developing Countries: What Investments Boost Learning?" *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), 119-157.

Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 2002. "Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya."

Hanushek, Eric A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24(3):1141-1177.

Hanushek, Eric A. 1995. "Interpreting Recent Research on Schooling in Developing Countries," *World Bank Research Observer*, 10 (August), 227-246.

Heyneman, Stephen P., Joseph P. Farrell and Manuel A. Sepulveda-Stuardo. 1978. "Textbooks and Achievement: What We Know." *World Bank Staff Working Paper* No. 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):139-150.

Iyer, Lakshmi, Paul Glewwe, and Michael Kremer. 2002. "School Funding and Student Performance: Evidence from Kenya." *Department of Economics, Harvard University*. Draft paper.

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, and Richard 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," New Delhi: National Council of Educational Research and Training.

Organization for Economic Co-operation and Development. 2004. "Learning for Tomorrow's World: First Results from PISA 2003." Paris: *OECD*.

Pritchett, Lant, and Deon Filmer. 1999. "What Education Production Functions Really Show: A Positive Theory of Education Expenditures," *Economics of Education Review*. 18(2):223-239.

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*

Reback, Randall (2006). "Teaching to the Rating: School Accountability and the Distribution of Student Achievement," Working paper (July).

UNDP. 1990. "Human Development Report." *Oxford University Press*. New York.

UNESCO. 1999. "Statistical Yearbook." *UNESCO*: Paris, France and Bernan: Lanham, Maryland.

UNESCO. 2006. "Education for All Global Monitoring Report." *UNESCO*: Paris, France.

World Bank. 1990. "World Development Report." *Oxford University Press.* New York.

World Bank. 1997. *Primary Education in India*. The World Bank. Washington, DC.

World Bank. 2004. "Treasures of the Education System in Sri Lanka." The World Bank. Washington, DC.