Featured Article

The Role of Classroom Resources and National Educational Context in Student Learning Gains: Comparing Botswana, Kenya, and South Africa

MARTIN CARNOY, MOSES NGWARE, AND MOSES OKETCH

We take an innovative approach to estimating student mathematics learning in the sixth grade of three African countries. The study reinforces the notion that beyond the quality of the teaching process in classrooms, national contextual factors are important in understanding the contribution that schooling makes to student performance. Our approach enhances more typical cross-sectional production function estimates in three ways: (1) to respond to critiques that production function estimates usually do not include classroom processes, we measure both teacher characteristics and teaching process variables and include them in the model; (2) to more clearly identify student learning with schooling processes, we estimate the gain in learning associated with a student's exposure to teaching characteristics and processes during the sixth-grade academic year in each country; and (3) to begin to address the issue of possible "national institutional factors" influencing student achievement, we use a comparative approach to approximate and initiate discussion of "country fixed effects."

Introduction

Improving student performance in school may have a high economic and social payoff (Hanushek et al. 2013), but policy analysts in developing countries have surprisingly little empirical data on which to base educational strategies for raising achievement (UNESCO Education for All 2005). A major problem in assembling evidence to guide such strategies is that student learning is a complex process subject to many family and community factors outside the control of school authorities (Rothstein 2004). Schools themselves are also complex institutions. Much of the technology of teaching and learning, including curriculum and time specifications for presenting subject matter, is set outside of classrooms and schools—the relationships between teachers, parents, and administrators are imbedded in each society's politi-

Received February 26, 2014; revised August 23, 2014; accepted November 10, 2014; electronically published March 25, 2015

Comparative Education Review, vol. 59, no. 2.

0010-4086/2015/5902-0002$10.00

cal history, and the development of educational expectations and standards are themselves products of that history.[1]

In the context of these constraints, educators and social scientists attempt to model and estimate empirically how schools in specific societies can improve students' learning. The vast majority of these studies focus on teachers and teaching, for good reason. Teachers are the key contact that students have with the schooling process. However, a number of studies have stressed the role of other factors that are likely to influence both teacher and student performance, such as school leadership (e.g., Spillane et al. 2001), national curricula (Schmidt et al. 2001), and particular institutional factors, such as decentralized administration of schooling (Woessmann 2004; Fuchs and Woessmann 2007).

In this article, we employ a production function model to estimate student mathematics learning in the sixth grade of three African countries. We enhance more typical cross-sectional production function estimates in three ways: (1) to respond to critiques that production function estimates usually do not include classroom processes (Levin 1980), we measure both teacher characteristics and teaching process variables and include them in the model; (2) to more clearly identify student learning with schooling processes, we estimate the gain in learning associated with a student's exposure to teaching characteristics and processes during the sixth-grade academic year in each country; and (3) to begin to address the issue of possible "national institutional factors" influencing student achievement, we use a comparative approach to approximate and initiate discussion of "country fixed effects."

The three countries we study are low-income Kenya, where grade 6 students score consistently much higher in mathematics than students in middle-income Botswana, whose grade 6 students, in turn, score higher than students in middle-income South Africa (for past test score comparisons, see SACMEQ [2010]). In Botswana and South Africa (only the North West Province), we randomly sampled 120 public, non-fee-paying schools and their sixth-grade mathematics classrooms on either side (50 km) of the Botswana–South Africa border. The samples were designed to exploit a "natural experiment" (Knight and Sabot 1990) in which colonial and postcolonial history caused Setswana-speaking people to be governed by two nation-states that pursued very different national education policies (Carnoy et al. 2012). Our Kenya survey was conducted in the same year (2009) as part of a different but coordinated study that randomly sampled nationally two sets of 36 schools (high and low scoring on the Kenya Certificate of Primary Education examination [KCPE]) and randomly chose and surveyed one of their sixth-grade mathematics classrooms. The Kenya survey used the same

[1] See Levin (1980), Carnoy and Levin (1985), Carnoy and Samoff (1989), and Carnoy et al. (2007).

questionnaires and student tests as the Botswana–South Africa study, the same teacher and principal questionnaires, a similar teacher test, and the same teacher videotaping and opportunity to learn assessment methodology (Ngware et al. 2010).

We used the data from the surveys in the three countries to estimate how family, classroom, and school inputs are related to differences in student learning gains in the samples in each country, with the Kenya sample divided into its two parts. Since Kenya is not contiguous to Botswana or South Africa, the Kenya samples could not be designed in the same way as the Botswana–South Africa study to test the possible effects on student achievement of national policy differences between Kenya and the other two countries. That said, we use the results from all three surveys to discuss the possible influence on student achievement of national differences in educational systems beyond usually measured classroom resources and processes. The Botswana–South Africa study was specifically designed with this end in mind, and, whereas the Kenya samples do not represent all schools in Kenya, the fact that students in both the "high-scoring" and "low-scoring" samples achieve at much higher levels than students in Botswana and South Africa can serve to initiate the discussion of national education differences that might explain higher student achievement in Kenya.

As in the rest of the world, the process of education in these three African countries' classrooms is couched in the context of their significantly different histories. Botswana, South Africa, and Kenya were all at one time English colonies, but their historical separation from English colonial legacies and their subsequent educational development varied enormously. Although we can only describe these historical sociopolitical contexts briefly in this article, we use our data to explore tentatively how much of the differences in student learning gains may reside in unobserved national/regional factors related to broader educational policy histories.

Research Background

In developed countries, much of the empirical research on school factors affecting student learning focuses on teachers. It shows that students with more effective teachers perform better on achievement tests.[2] The emphasis in identifying effective teachers has been on teacher characteristics associated with higher student outcomes. For example, some studies suggest that greater teacher experience contributes significantly to student achievement.[3] Others have shown that positive effects on student outcomes result from teacher subject matter knowledge (Hill et al. 2005) or proxies

[2] See, e.g., Nye et al. (2004), Rockoff (2004), Hanushek et al. (2005), and Boyd et al. (2006).

[3] See Ferguson and Ladd (1996), Rockoff (2004), Hanushek et al. (2005), and Clotfelter et al. (2007).

for subject matter knowledge, such as teacher scores on a test of literacy or verbal ability.[4] Rockoff et al. (2008) find small but significant effects on student achievement of a combination of teachers' cognitive and noncognitive skills measured at the time of hire.

Another theme in explaining better student test performance is opportunity to learn (OTL). In developed countries, OTL is mainly defined as the quality of the subject matter curriculum and curriculum coverage. The 1995 Trends in International Mathematics and Science Survey (TIMSS), for example, focused on comparing OTL in mathematics and science across countries. The conclusion of those studies was that exposure to mathematics and science driven by national or regional curricula was important in explaining average student performance (Schmidt et al. 1997, 2001). Most recently, the Programme in International Student Assessment (PISA) 2012 test included an analysis of such exposure to different types of mathematical concepts across national education systems (OECD 2013). In developing countries, OTL is defined more in terms of actual time spent teaching relevant subject matter (Fuller et al. 1994). A large study of teacher time use in Latin American classrooms suggests that more time on task could result in significant student learning gains (Bruns and Luque 2014).

There is also a long history of linking teaching practices and styles to student achievement (for a review, see Hill et al. 2005). More recently, researchers have used larger-scale samples to measure the effect of teaching practices on student test score gains. In the United States, this has culminated in an extensive study of teacher effectiveness—a sample of 3,000 teachers in a number of US urban areas. The data have been used to predict successfully teachers' effectiveness on the basis of past student achievement gains (value added) in a teacher's classroom, assessment of four videotapes of the teacher teaching, and student responses to a survey regarding the teacher's practices (Kane et al. 2013).

Considerably less has been written on system differences, and the main aspects of system differences that have been studied emerge from two rather different positions on what makes systems more effective. Woessmann (2004) and Fuchs and Woessmann (2007) argue for a market approach to education, claiming that the more public educational systems are characterized by market conditions (e.g., school autonomy, competition from private providers), the more effective the systems. Using cross-national international test data—TIMSS and PISA—they estimate positive significant coefficients for school autonomy (as reported by principals) and, in the case of the 2003 PISA test, a positive effect of a nation's proportion of private schools. Schmidt et al. (2001) take a diametrically different position, contending that strong national mathematics curricula (OTL) and their implementation

[4] See Rice (2003), Boyd et al. (2006), and Clotfelter et al. (2007, 2010).

(implicitly, through effective public interventions) are fundamental to understanding how effective schools and teachers are in producing student math learning. Both explanations implicitly argue that historically imbedded national educational policy differences and educational cultures (school autonomy and national curricula/the implementation of national curricula) may explain much of the difference in student achievement levels. In the three African countries we study, such policy differences were shaped by their colonial/postcolonial histories.

Conceptual Framework

Our analysis draws on a two-part conceptual framework: the first part is a production function model of how the knowledge students bring from their homes combines with teachers' knowledge, skills, and effort to produce student learning within different national contexts (Levin 1980). The second part is a conception of state social capital (Carnoy and Marshall 2005; Carnoy et al. 2007) that frames how this combination results in different levels of learning gains in different national political contexts. That conception of state social capital argues that the sociopolitical organizational context—often historical and not easily measurable—can raise or lower expectations for student achievement and thus raise the payoff to family and educational resources.

Our model of classroom learning is centered on teachers and the teaching process. It consists of five main components: (i) the teacher's capacity to teach the material (content knowledge and pedagogical content knowledge); (ii) the teacher's pedagogical skills; (iii) students' OTL, which may influence student learning gains directly and indirectly through their interaction with the quality of teaching; (iv) the students' socioeconomic background, which may influence student learning gains directly and indirectly, through its potential impact on teacher/school expectations of students and the process of teaching itself; and (v) the outcome of the process—student learning gains.

The notion that sociopolitical structures shape the process of student learning (whether in families or schools) is hardly new.[5] In this article, however, we are less concerned with a generalized theory of how sociopolitical structures shape the process of schooling and student learning than with the simpler idea that the institutional structures emerging from national political histories may be important to the level of student learning produced in schools. This is so because, among other things, these institutional structures influence family and school learning expectations, teacher preparation, state-teacher-family relations, national curriculum, and school language

[5] See Bowles and Gintis (1975), Foucault (1975), Carnoy and Levin (1985), and Carnoy et al. (2007).

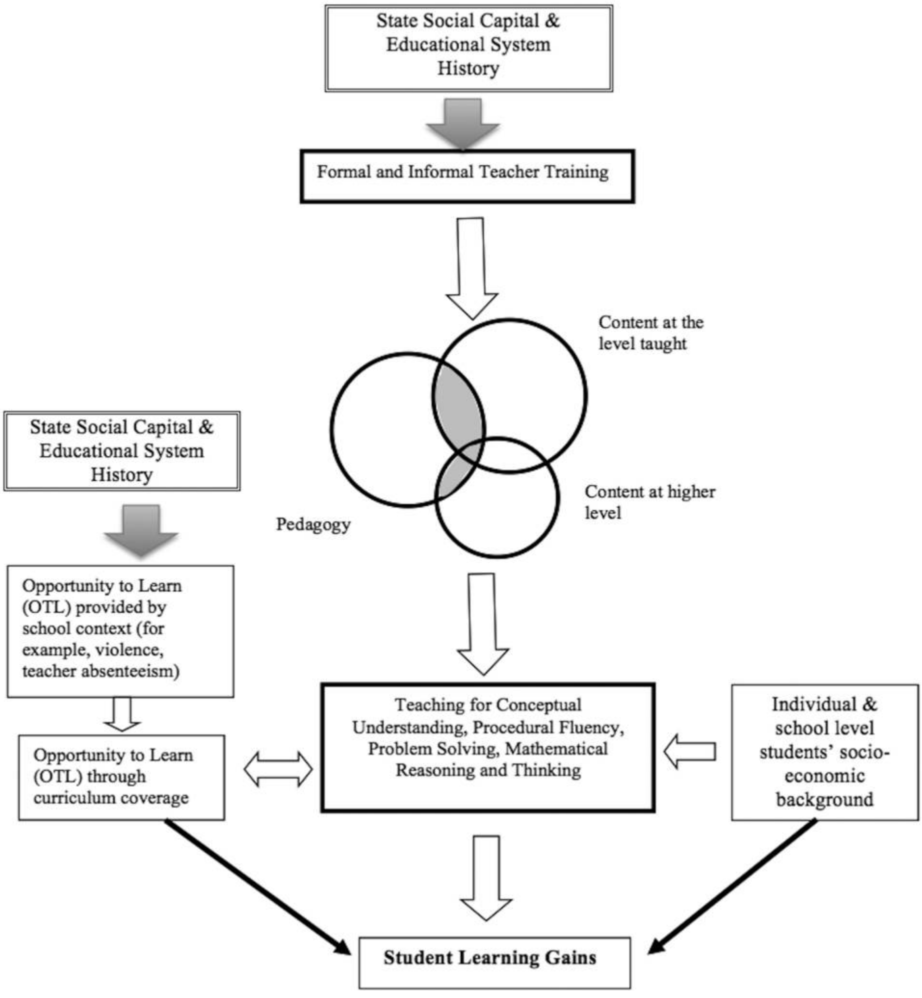

Fig. 1.—Estimating student mathematics achievement: conceptual framework

policy. The parts of our conceptual framework are described in the schematic shown in figure 1.

The Data

We collected two sets of data in 2009 as part of two separate, but coordinated, studies of student mathematics achievement in southern and eastern Africa.[6] The first of our studies, conducted under the auspices of the

[6] The Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) had conducted three large national-level surveys of sixth-grade classrooms in this region (1997, 2001,

Human Sciences Research Council and the University of Botswana drew random samples of 120 schools in Southeastern Province, Botswana, and North West Province, South Africa, on either side of the Botswana–South Africa border (of which 116 schools—58 in each country—supplied sufficient data) and 126 grade 6 classrooms from those schools (64 in Botswana and 62 in South Africa). The second study was conducted in Kenya by the African Population and Health Research Center (APHRC). The APHRC sought to compare high- and low-scoring schools in Kenya, so it drew random samples of 36 sixth-grade mathematics classrooms from 36 schools scoring in the lowest 20 percent on the KCPE and 36 classrooms from 36 schools scoring in the top 20 percent on the KCPE. The samples were drawn in five districts (now counties) plus Nairobi. The districts selected were representative of Kenya as a whole, but the schools sampled were not representative of all schools in Kenya (Ngware et al. 2010). In Botswana and South Africa, we made three visits to each school; in Kenya, two.

The sampled schools in Botswana are spread across four districts: Gaborone, South East, Southern, and Kgatleng districts. The schools are located close to the capital city of Gaborone and the town of Lobatse and are within a radius of about 80 km. Both Gaborone and Lobatse are near the South African border posts of Tlokweng and Mahikeng. Private schools and grade 6 classes with enrollments of less than 20 were excluded from the sample. A total population of 107 schools with enrollment of 6,835 grade 6 pupils constituted the sampling base. Sixty schools were sampled. Of these, 58 schools agreed to participate in the study.

In bordering North West Province, South Africa, we used a two-level stratified random sampling methodology in public schools with grade 6 learners. According to SACMEQ results, grade 6 students in North West Province score somewhat above the average for South Africa in mathematics (503 vs. 495; SACMEQ 2010). We sampled schools in the Mahikeng and Ramotshere Moiloa local municipalities, which border Botswana and encompass two of the major urban centers in the province, Mahikeng and Zeerust. A total of 155 schools with a learner population of 9,157 constituted the sampling frame, with 12 percent of schools urban, 72 percent Tribal Area and 16 percent "Rural Formal."[7]

The Kenya sample was stratified by the average school score on the KCPE (the eighth-grade primary-school-leaving exam) at both the district and the

2007), but these provided student test data at only one point in time and collected limited data on classroom processes, especially OTL and teaching quality. The difficulty of drawing policy-relevant inferences from SACMEQ data motivated us to gather and analyze our own more complete data.

[7] The categories Tribal Area and Rural Formal were combined into one "rural area." The Mahikeng and Ramotshere Moiloa local municipalities were not individually sampled but were combined into one stratum. All urban schools in the sampling frame were retained, constituting 20 percent of the sample. A simple random sample, without replacements, was drawn from the remaining (all rural) schools to make up the other 80 percent of the sample.

school level. The sample was taken in six districts (Baringo, Embu, Garissa, Muranga, Gucha, and Nairobi). Two of these districts consistently scored in the lowest 10 percent on the KCPE in 2005–8, two ranked in the middle 20 percent, and two in the top 10 percent. In each district, six schools were randomly selected from those scoring in the highest 20 percent and six schools from the lowest 20 percent on the KCPE exam. We will treat these two samples as representative nationally of high- and low-scoring schools. In that sense, they need to be treated separately and are not representative of all Kenya schools. As noted earlier, the samples were stratified and representative within these two populations of schools, but the schools selected in each stratified sample are not representative of all schools in Kenya. Table 1 describes the type of data we collected from students, teachers, and school principals. All the samples ended up being about 20 percent urban. Although Kenyan students, on average, came from a lower socioeconomic background than students in Botswana and South Africa, student socioeconomic background varied considerably in all three countries. Table 2 shows that another major difference between the Kenya and Botswana/South Africa samples is that the surveys in Botswana and South Africa purposefully focused on non-fee-paying schools, whereas the Kenya sample includes a relatively high percentage of private-run institutions (about 35 percent, including 7 percent operated by nongovernmental organizations).[8] To check the comparability of our production estimates, we include a "private school effect" in our two subsamples of Kenyan schools.

Student Achievement

In Botswana and South Africa, we gave students a mathematics test once near the beginning of the sixth grade (March) and once near the end (early November). Kenyan students were first tested in July in sixth grade rather than in March and again at the very beginning of seventh grade. The fact that the posttest was given in the first month of the new school year may bias the Kenya posttest scores downward because of knowledge loss over vacation, but we have no evidence for such bias. The student test was the same for students in Botswana and South Africa, and 35 of 40 items were the same in Kenya. The test was largely set at a fifth-grade level with a reasonable

[8] Fee-paying institutions in Botswana and South Africa are a much smaller proportion of schools than in Kenya (SACMEQ 2010). The proportion of private school students in 2009 in South Africa was 3 percent; in Botswana, 6 percent; and in Kenya, 11 percent. The 3 percent figure in South Africa does not include Model C schools (former white schools) that generally charge high fees and receive a reduced government subsidy. The proportion of private school students in our two Kenya samples is much higher than the proportion for Kenya as a whole. Further, the socioeconomic status (SES) of students attending fee-paying schools in Botswana and South Africa is distinctly higher than of those in non-fee-paying schools (SACMEQ 2010), whereas many fee-paying institutions in Kenya are geared to low SES pupils. We show below that private school students in the lower-scoring school subsample, but not in the high-scoring subsample, make somewhat higher gains than public school students in the same (high- or low-scoring) group of schools.

representation of questions related to the grade 6 curriculum, which is very similar in Botswana and South Africa but somewhat more advanced in Kenya.

A usual issue is student "loss" between the pretest and the posttest. This was not a problem in two cases—we were able to "match" more than 90 percent of the students on their pre- and posttests in Botswana and South Africa. But in Kenya, this figure dropped to 80 percent because a number of students who were not in the original sample were tested after vacation, and a number of students who were in the original sample were not captured in the second test (the number of observations on test scores versus other variables shown in table 2 reflect these differences). Since we had considerable information on the students in Kenya who only took the pretest compared to those who took both pre- and posttests (such as their individual characteristics, their family background, and their initial test score), we were able to test statistically for possible bias from nonrepresentativeness of students taking the posttest.[9] We found no such bias and used the reduced sample for our production function estimates.

Students in our Botswana and South Africa samples did poorly on the mainly fifth-grade-level mathematics test applied near the beginning of grade 6 (mean = 28.6 percent on the initial test in North West Province and 33.5 percent in Botswana) and made relatively small gains in the 7 months between the pre- and posttest (3 percentage points in North West and 4 percentage points in Botswana). These differences were statistically significant. Kenyan students scored much higher on the initial test we applied there (35 questions were identical; five questions were of somewhat greater difficulty)—47.7 percent overall, 39.9 percent in the sample of low-scoring schools—and made large gains in the 7 months between tests (10.5 points in the high-scoring schools, 9.8 points overall, and 9.6 points in the sample of low-scoring schools). The standard deviation of individual test scores was relatively high, about 12 points in the Botswana and North West Province samples, 12 points in the Kenya subsample of low-scoring schools, and 15 points in the Kenya high-scoring school subsample. There was also considerable variation in student gains, also about a standard deviation of 10–11 points (see table 2).

Figure 2 shows average pretest scores by item for the 35 items taken by students in all three countries (as well as items 36–40, which were the same in Botswana and North West Province but differed in Kenya). Students generally did well and poorly on the same items, yet Kenya students scored much

[9] Specifically, we used the repeated imputation method to impute posttest scores to those students who took the first test but not the second and then estimated our two Kenya production functions with and without including the "missing" observations with imputed scores. The coefficients of our independent variables changed little, implying no selection bias from not including students who did not take the second test.

TABLE 1

Source and Description of Key Variables Used in Estimates of Student Mathematics Achievement Gain

Variable	Instrument Source for Variable	Source of Instrument	Variable Description	Comment
Initial test score	Student test administered in March 2009 in BW and SA; July 2009 in Kenya	Test based on fifth- and sixth-grade curriculum in BW and SA	40-question mathematics test, with multiple parts to some questions; score = % of items correct, with multiple-part items given partial credit	35 common items between Kenya test and BW and SA test; APHRC (Kenya) changed five items to fit Kenyan curriculum but similar to BW and SA items
Second test score	Student test administered in November 2009 in BW and SA; February 2010 in Kenya	Same as initial test	Same as initial test with number changes in several items and several changes in placement order of items	Same; both initial and second tests had the same 35 common test items and similar test items 36–40
Individual student characteristics and family background, including articles in the home	Student questionnaire	Instrument developed by BW and SA researchers; modified and approved by APHRC	Standard definitions of student characteristics and family background; articles in the home appropriate for societies in question	
Opportunity to learn	Detailed protocols developed for student notebook analysis	For BW and SA, protocol based on subjects covered by items on student test; for Kenya, protocol based on proportion of curriculum covered	For BW and SA, number of lessons given on each item, test topics covered, number of classes held; for Kenya, percentage of standard Kenyan curriculum covered	Because measures are different in BW and SA from Kenya, standardized score for each country is used (mean = 0, standard deviation = 1)
Years teaching experience	Teacher questionnaire	Instrument developed by BW and SA researchers; modified and approved by APHRC	Years teaching as reported by teacher	

Teaching quality	Videotaped lesson(s) for each teacher analyzed on the basis of 12-item protocol	Protocol developed by two expert math educators on the basis of mathematics teaching literature	On the basis of protocol score, teacher's teaching rated 1 (low), 2 (better), or 3 (best)	
Teacher test score	Teacher test administered in all three countries in July 2009	24-item test developed by BW and SA researchers based on BW and SA curriculum; five items from student test	Percentage of items correct	APHRC changed several items on the BW and SA teacher test it administered to Kenyan teachers in order to fit sixth-grade Kenyan curriculum
Observed class size	Classroom videotapes		Number of students observed in videotape	
Average class SES	Student questionnaire	Average of articles in the home index for all individual students in each class	Index of variables in the home; variables are weighted inversely to percentage of students reporting article in home	
Violence index	Student, teacher, and principal questionnaires	All three questionnaires designed by BW and SA researchers, modified and approved by APHRC	In BW and SA, class average of sum of indexes of student, teacher, and principal views of scale of student-student violence and teacher-student violence; in Kenya, class average of scale of teacher assessment of student-student violence	Because of differences between measures of BW and SA index and Kenya index, class average violence index standardized for each class (mean = 0, SD = 1).

NOTE.—BW = Botswana; SA = South Africa; APHRC = African Population and Health Research Center; SES = socioeconomic status.

TABLE 2
LOWER- AND HIGHER-SCORING SCHOOLS, MEANS, AND STANDARD DEVIATIONS OF SCHOOL SAMPLE RESULTS FOR STUDENT AND TEACHER VARIABLES

Variable	South Africa			Botswana			Kenya Lower-Scoring Schools			Kenya High-Scoring Schools		
	N	Mean	SD	*N*	Mean	SD	*N*	Mean	SD	*N*	Mean	SD
Second test score	3,485	.316	.125	1,612	.384	.144	788	.485	.149	1,119	.637	.174
Initial test score	3,581	.295	.122	1,666	.346	.125	794	.399	.118	1,119	.532	.151
Student characteristics:												
Sex (female = 1)	3,581	.51	.50	1,666	.49	.50	1,047	.50	.50	1,389	.44	.50
Learner's age	3,581	12.49	1.21	1,666	12.03	1.14	1,047	12.4	3.57	1,389	12.2	3.04
English spoken at home (no = 1)	3,581	.95	.21	1,520	.84	.37	1,047	.07	.25	1,389	.06	.25
Preschool attendance (yes = 1)	3,581	.72	.45	1,641	.57	.50	1,047	.83	.38	1,389	.90	.29
Highest parents' education = primary	3,581	.13	.34	1,666	.16	.37	1,047	.45	.50	1,389	.32	.46
Highest parents' education = some secondary	3,581	.10	.30	1,666	.12	.33	1,047	.10	.31	1,389	.07	.26
Highest parents' education = secondary complete	3,581	.14	.34	1,666	.12	.33	1,047	.22	.41	1,389	.21	.41
Highest parents' education ≥ secondary	3,581	.25	.31	1,666	.25	.34	1,047	.12	.23	1,389	.28	.35
Missing parent education	3,581	.40	.49	1,666	.36	.48	1,047	.10	.30	1,389	.12	.32
No books at home	3,581	.05	.22	1,666	.13	.37	1,047	.22	.42	1,389	.16	.37
1–10 books at home	3,581	.27	.44	1,666	.35	.47	1,047	.57	.49	1,389	.57	.50
10–50 books at home	3,581	.50	.40	1,666	.36	.43	1,047	.12	.32	1,389	.16	.49
More than 50 books at home	3,581	.17	.37	1,666	.13	.34	1,047	.08	.20	1,389	.11	.20
Books missing	3,581	.01	.11	1,666	.03	.17	1,047	.00	.03	1,389	.00	.00
Never read at home	3,581	.07	.25	1,666	.07	.25	1,047	.02	.15	1,389	.01	.11
Receive daily news (no = 1)	3,581	.49	.50	1,645	.60	.49	1,047	.56	.50	1,389	.50	.50

Have magazines at home (no = 1)	3,581	.58	.49	1,647	.70	.46	1,047	.66	.47	1,389	.61	.49
Have fridge at home (no = 1)	3581	.33	.47	1,646	.57	.49	1,047	.88	.33	1,389	.78	.41
Have car at home (no = 1)	3,581	.49	.50	1,647	.42	.49	1,047	.86	.35	1,389	.77	.42
Teacher/class characteristics:												
OTL total lessons on test items*	3,530	60.83	22.41	1,666	50.76	21.32	...	...	...	...	...	...
OTL total test topics*	3,530	39.88	13.49	1,666	38.80	10.67	788	45.04	15.50	1,119	52.67	15.78
Years teaching experience	3,581	23.06	9.05	1,666	15.44	7.97	788	14.15	10.01	1,119	15.51	9.12
Better teaching quality	3,581	.54	.50	1,666	.49	.50	788	.53	.50	1,119	.47	.50
Best teaching quality	3,581	.18	.38	1,666	.15	.36	788	.12	.33	1,119	.15	.36
Teacher test score	3,581	46.70	11.2	1,666	53.50	.95	788	57.97	14.30	1,119	61.73	16.25
Observed class size	3,581	40.49	12.32	1,666	28.65	5.16	788	39.85	18.84	1,119	44.65	16.52
Standardized class average SES†	3,581	.17	.97	1,666	−.37	.96	1,047	−.50	.91	1,389	.38	1.56
Standardized violence index‡	3,581	−.32	.85	1,666	.71	.92	1,047	.08	1.14	1,389	−.09	.80
% students private school							1,047	.31		1,389	.27	
% students private "other" school							1,047	.10		1,389	.06	

SOURCE.—South Africa (SA), Botswana (BW), and Kenya school surveys, 2009.

* Measures of opportunity to learn (OTL) are the same in SA and BW (total lessons on test items in each classroom during 7 months of academic year or number of test mathematics topics covered); in Kenya, the measure is proportion of required math curriculum topics covered during school year.

† Socioeconomic status (SES) measure is an index constructed from articles in the home reported by students. Here, we show the standardized version of the variable that is used in the regressions.

‡ Violence indicator in BW and SA is index of student, teacher, and principal assessment of learner-to-learner and teacher-to-learner violence in school; in Kenya, violence indicator is teacher assessment of learner-to-learner violence. Since indexes of reported violence differ somewhat in Kenya survey, we estimate a standardized index within each country.

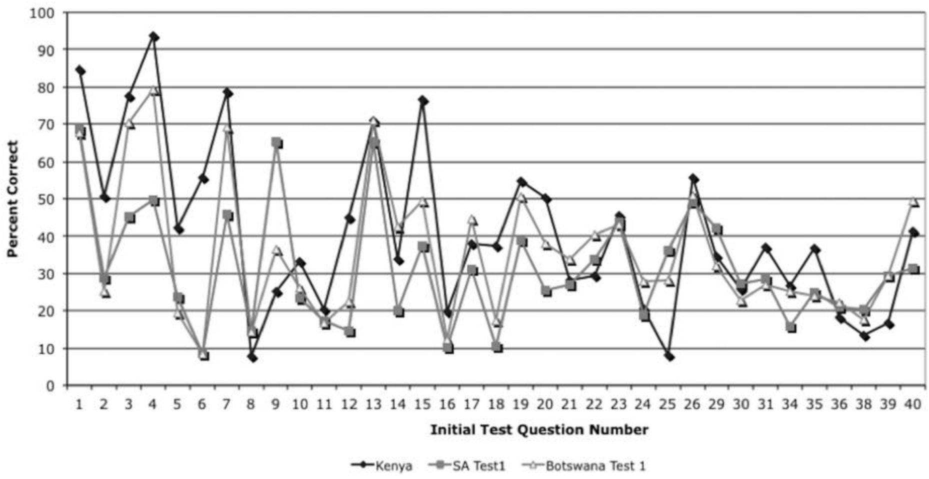

FIG. 2.—Botswana, Kenya, and South Africa: initial sixth-grade mathematics test score means, by country, 2009. SOURCE.—Botswana, North West Province, and Kenya Surveys, 2009. A color version of this figure is available online.

higher across items and Botswana students scored somewhat higher across items than South African students on the pre- and posttests. The gains on the posttest were also greatest in Kenya (table 2). Thus, one cannot argue that students in Kenya and Botswana are learning parts of the curriculum that North West students are not; Botswana and especially Kenyan students do better because they seem to be learning some parts of the curriculum significantly better than students in North West Province.

Teacher Experience, Education, and Knowledge

We also gave teachers a detailed questionnaire and tested their mathematics knowledge. Some of the mathematics questions were taken from the student test. The teachers in our South African sample averaged considerably more teaching experience (were older) than in Botswana and Kenya—23 years in North West versus 15 in Botswana and 14–16 in Kenya (table 2). Botswana and Kenya teachers were much more likely to have been in their current school a relatively shorter time (11 years in South Africa, 5 years in Botswana and Kenya).

We gave all teachers in our samples a mathematics test to measure their content knowledge of the sixth-grade curriculum. The Botswana and Kenya teachers in our samples scored significantly higher than the teachers in our South Africa sample (table 2). As on the student test, teachers in the three countries did well and poorly on the same items, and where there was overlap between items on the student and teacher tests, teachers did poorly on items such as fractions and so did the students.[10]

Teachers' Teaching Quality

We intended to film each teacher giving two lessons during the year, but this proved difficult (some teachers resisted having their lessons filmed repeatedly), so we ended up with 83 videotaped mathematics lessons for 64 classrooms in Botswana, 100 videotaped lessons for 62 classrooms in South Africa, and 70 videotaped lessons for 72 teachers in Kenya.[11] Evaluating teachers' teaching on the basis of one or two videotapes is a limitation of our study. That said, since teachers were given previous notice that their lesson would be filmed, it is likely that they gave a "good performance" and that the videotape represents the way they teach math on a "good day." Some countries, such as Chile, use single videotapes as an important part of a portfolio to evaluate public school teachers.[12]

The same two experts in mathematics pedagogy coded each videotape on three dimensions: (*a*) mathematical proficiency, based on the National Research Council's (2001) five subcategories, or strands (conceptual understanding, procedural fluency, strategic competence, adaptive reasoning, and productive disposition); (*b*) cognitive demand, based on Stein et al.'s (2000) classification of higher and lower cognitive demand in four subcategories (memorization, procedures without connections, procedures with connections, and doing mathematics); and (*c*) teacher's observed knowledge, based on Shulman's (1986) three subcategories (grade-level mathematics knowledge, general pedagogical knowledge, and mathematical knowledge in teaching).

For each of these 12 subcategories and for each videotaped lesson, coders indicated whether the teacher displayed the relevant element in her or his teaching. On the basis of how well the teacher did, coders gave each

[10] Although we did not include teacher preservice training in the variables we used to estimate student achievement gains, mainly because teacher preparation did not vary greatly within country, we gathered considerable information from the teacher questionnaire on teacher preparation. It helps explain differences across countries in teacher performance on the exam we gave them. All three education systems have raised the level of teacher training required for primary school certification in the past 20 years, but because of their age, almost all the teachers in our South Africa sample learned their mathematics at junior-secondary- (graduation at ninth grade) and their teacher training in upper-secondary-school level colleges. In contrast, because they were younger, most of our sampled Botswana and Kenya teachers had had to complete twelfth grade to pursue primary teacher education. The level of mathematics preparation is therefore greater, on average, in our Botswana and Kenya samples.

[11] In two schools (one in the high-scoring group and one in the low-scoring group), lessons were videotaped, but in science instead of math. For those teachers we imputed the average teaching score of the teachers in their 36-school group. In our regression results, reported below, we also made estimates without these two teachers included—the difference in estimated coefficients was negligible.

[12] Public school teachers in Chile have been evaluated since 2005 on the basis of a portfolio that includes only one videotaped lesson (Taut et al. 2010). Data from these evaluations have been used successfully to estimate "teacher quality effects" on student achievement (Carrasco 2014).

taught lesson an overall score of 1 (minimum level), 2 (better teaching), or 3 (best teaching). We used this overall score for our analysis of learner gains.

The findings suggest that teachers focused on procedural fluency in their lessons, in large part because this is what the curricula in each country emphasize, and that some teachers value conceptual understanding before learners move to the manipulation of symbols or computation. This is also consistent with the kinds of questions used by teachers in the classroom. However, not all teachers were able to teach conceptually in an efficient way. In Botswana and South Africa, only about 22–23 percent, and in Kenya 33 percent, of the lessons' mathematical tasks included aspects of reasoning (comprehending why algorithms and rules work and justifying steps or final answers) and strategic competence (or problem solving). Finally, less than half the lessons included teaching that helped learners see mathematics as sensible, useful, and worthwhile combined with a belief in their ability to do mathematics ("productive disposition"). This category was observed only during the lessons in which learners were involved in either the application or reasoning of mathematics. In those lessons, learners seemed to enjoy and value the logical thinking and problem-solving activities (table 3).

In addition, a high fraction of the lessons in both countries focused only on memorization or procedures without connections. Relatively few lessons in Botswana and South Africa (23–30 percent), but 43 percent in Kenya, required students to understand the meaning of operations or underlying concepts behind the procedures, and almost none required students to investigate or explore relationships between mathematical ideas ("doing mathematics"), yet 6 percent in Kenya versus 2 percent in Botswana and South Africa taught lessons that fell into this latter category. Teachers in Botswana and Kenya were much more likely to show knowledge of grade 6 mathe-

TABLE 3
Relative Frequency of Mathematics Teaching Characteristics in Videotaped Lessons and Overall Teachers' Rating (%)

Variable	Botswana	South Africa	Kenya
Strands of proficiency:			
Conceptual understanding	35	58	41
Procedural fluency	93	78	91
Problem solving	20	22	24
Reasoning	22	23	33
Productive disposition	40	47	41
Cognitive level of tasks:			
Memorization	72	92	76
Procedures without connections	84	59	80
Procedures with connections	23	30	43
Doing mathematics	2	2	6
Overall rating:			
Low	40	33	40
Better	43	56	47
Best	17	11	13

Source.—Botswana, South Africa, and Kenya school samples, 2009.

matics, and few teachers in the three countries (13–17 percent) displayed the necessary pedagogical content knowledge to teach that level of mathematics.

Given these teaching assessments, it is not surprising that few teachers in our two samples received an overall rating of 3. Of the 83 videotapes evaluated in Botswana, 33 were "low," 35 "better," and 15 "best." Of the 100 videotapes evaluated in South Africa, 33 were "low," 56 were "better," and only 11 were "best." In Kenya, of the 70 lessons we videotaped, 28 were "low," 33 were "better," and 9, "best." In relative terms, the Botswana and Kenya samples suggested that a higher fraction of teachers were teaching at a low level overall than in South Africa, but a higher fraction were teaching at the highest level.

Opportunity to Learn

In Botswana and South Africa, we measured classroom OTL by analyzing the contents of three (best) student notebooks in each classroom at two points during the school year. Because all teachers are required to have their students record lessons in notebooks, in Botswana and South Africa, we were able to measure the number of daily mathematics lessons taught up to the beginning of November (1 month short of the end of school), the number of mathematics topics covered by our student test items taught in that same period, and the number of times the teacher had touched on one of our student test items. In Kenya, researchers also used grade 6 exercise books to measure OTL as the number and proportion of required math subtopics taught during the academic year. Both measures indicate curriculum coverage, since the items on the test were based on the official curriculum.[13]

Exposure to mathematics is a major problem in the three countries, but especially so in South Africa. For example, the average number of math lessons in the 7 months covered by our student notebook observations was very low in South Africa (52 of the 140 scheduled number of math lessons in March through the end of October) and low in Botswana (78 of 140 lessons). Kenyan teachers only covered about one-half the sixth-grade math curriculum, and the difference between teachers in the sample of high-scoring KCPE schools and the lower-scoring KCPE schools was significant (53 vs. 45 percent, about 0.5 SD). In South Africa and Botswana, we learned from teacher and principal interviews that the reasons for such low coverage are multiple, including scheduled in-service training, departmental and union meetings during school hours, absences due to illness, and teachers "bumping" math lessons because they do not understand how to teach the math concepts in the curriculum. The student notebooks in South Africa

[13] Because the OTL measures differ in Botswana/South Africa and Kenya, for our analysis, we used the standardized form of these data, using a mean of 0 and a standard deviation of 1.

and (less so) in Botswana supported this last reason, since the pacing in many of the math lessons was extremely slow and teachers gave multiple lessons on the same (easier) topics and often never covered harder topics.

Since the Kenya samples represent two "slices" of the overall population of Kenya's schools, it is worth noting how similar the variation in the student test scores, student background variables, and teacher variables in each of these tranches is to the other slice and to the variation in the broader Botswana and South Africa samples.

Estimation Strategy

At the center of our analysis is a model of how the knowledge students bring from home interacts with school and classroom/teacher factors to produce student learning (Levin 1980). The model is centered on students' classroom experience and consists of six main components: (*a*) teachers' capacity to teach the curriculum material, as measured by the their mathematics content knowledge (teacher test score) and their teaching experience; (*b*) mathematics teaching quality, as measured by ratings from lesson videotapes, that influences student learning gains directly and indirectly through teachers' capacity to teach the material; (*c*) a measure of OTL (exposure to mathematics concepts) that influences student learning gains directly and indirectly through the capacity of teachers to teach these concepts; (*d*) student characteristics including gender, age, individual family academic resources, prior achievement (pretest score), and the average family academic resources of students in the class (all of these student characteristics may influence student learning gains directly and indirectly through their potential impact on teacher/school expectations of students, hence, the quality of the teaching they face and their OTL); (*e*) the number of students in the class and the reported level of student-to-student violence in the class, both of which can influence the student's learning; (*f*) the outcome of the process (students' posttest mathematics achievement).

Education production within the classroom takes place through a complex process. In particular, student inputs (such as family resources) and classroom inputs (such as teacher characteristics and OTL) are systematically related to each other and to student outcomes. To better understand the direct and indirect impacts of various inputs on student outcomes, it is helpful to model these complex relationships explicitly.

We use a series of equations to model and estimate the relationships in each country (in Kenya, we estimate separate production functions for high-scoring schools and low-scoring schools) between student achievement outcomes (posttest scores) and (*a*) teacher content knowledge, teacher experience, and teaching quality (how well teachers teach mathematics in the sixth grade); (*b*) OTL (how much exposure students have to mathematics in their classrooms during the sixth grade); (*c*) student characteristics at the individual and classroom levels; and (*d*) class size and reported class violence.

Based on the production function literature and earlier models (e.g., Schmidt et al. 2001), three assumptions underlie our model. First, we assume that the quality of mathematics teaching is related to teacher content knowledge and a quadratic function of teacher experience and possibly to the classroom average of students' socioeconomic background (eq. [1a]). The relationship between mathematics teaching quality and the classroom average of students' socioeconomic background reflects the notion that teacher resources may be allocated to students (and vice versa), partly on the basis of students' family academic resources.

$$\mathrm{TQ}_j = C_1 + \gamma_1 \mathrm{TMK}_j + \gamma_2 \mathrm{TExp}_j + \gamma_3 \left(\mathrm{TExp}_j\right)^2 + \gamma_4 \mathrm{AvgX}_j + e_j, \qquad (1a)$$

where TQ_j is mathematics teacher *j*'s teaching quality as rated from videotaped lessons, TMK_j is teacher *j*'s mathematics content knowledge as measured by the teacher's test score, TExp_j is teacher *j*'s years of teaching experience, AvgX_j is the average socioeconomic index of students in teacher *j*'s classroom, and e_j is an error term.

Second, we assume that OTL is also related to teacher mathematics content knowledge, a quadratic function of teacher experience, and to the classroom average of students' socioeconomic background (eq. [1b]). In this formulation, OTL acts as a complex mediator of capacity, in which teachers who have greater capacity to teach mathematics are more likely to expose students to more mathematics. TQ and OTL could be related to each other, but they are measures of rather different aspects of classroom processes. TQ measures the quality with which teachers teach mathematical concepts; OTL measures how much exposure students receive to the mathematics curriculum in grade 6. In all three countries OTL is influenced by many factors apart from TQ or even teacher mathematics knowledge and teacher experience.[14] Many mathematics lessons are not taught because of teacher illness, union meetings, and in-service training. The academic resources students bring to class may also influence teachers' coverage of the curriculum. Teachers may be less likely to expose students with low levels of family resources to as much mathematics as students with high levels of family resources—controlling for average class SES in equation (1b) gives us some indication whether this is the case in our three country samples. We use classroom-level data to estimate equations (1a) and (1b).

$$\mathrm{OTL}_j = C_2 + \beta_1 \mathrm{TMK}_j + \beta_2 \mathrm{TExp}_j + \beta_3 \left(\mathrm{TExp}_j\right)^2 + \beta_4 \mathrm{AvgX}_j + u_j, \qquad (1b)$$

[14] The correlation between TQ and OTL is very low in all three countries.

where OTL_j is exposure to mathematics in teacher j's classroom as measured by student notebook analysis, and u_j is an error term.

Third, we assume that student achievement is cumulative and is a function of previous achievement, students' family resources, teacher's capacity to teach mathematics, teaching quality, OTL, and classroom/school context, particularly the average family resources of students in the classroom, the class size, and the reported level of violence in the classroom. Typically, students' performance is estimated without controlling for students' previous achievement, so we, too, estimate such a model (eq. [2]). We call this the "posttest score only" model.

$$\begin{aligned} A_{ij\text{post}} = {} & C_3 + \Sigma b_k X_{ijk} + \delta_1 \mathrm{TMK}_j + \delta_2 \mathrm{TQ}_j + \delta_3 \mathrm{TExp}_j + \delta_4 \left(\mathrm{TExp}_j\right)^2 \\ & + \delta_5 \mathrm{OTL}_j + \delta_6 \mathrm{AvgX}_j + \delta_7 C_j + \delta_8 V_j + v_{ij}, \end{aligned} \tag{2}$$

where $A_{ij\text{post}}$ is the standardized (mean $= 0$, SD $= 1$) posttest mathematics score for student i in classroom j, X_{ijk} is a vector of family characteristics k of student i in classroom j, C_j is the number of students observed in class j, V_j is the average student-to-student violence index reported in class j, and v_{ij} is an error term.

A standard problem inherent in estimating the relation between classroom inputs and student mathematics achievement is that students accumulate mathematics knowledge before schooling and over many years in school. We attempt to address this problem in our model by controlling for students' pretest score as well as their family resources. Specifically, we estimate the following two equations:

$$\begin{aligned} A_{ij\text{post}} = {} & C_4 + a_1 A_{ij\text{pre}} + \Sigma b'_k X_{ijk} + \phi_1 \mathrm{TMK}_j + \phi_2 \mathrm{TQ}_j + \phi_3 \mathrm{TExp}_j + \phi_4 \left(\mathrm{TExp}_j\right)^2 \\ & + \phi_5 \mathrm{OTL}_j + \phi_6 \mathrm{AvgX}_j + \phi_7 C_j + \phi_8 V_j + v'_{ij}; \end{aligned} \tag{3a}$$

$$\begin{aligned} A_{ij\text{post}} - A_{ij\text{pre}} = {} & C_5 + \Sigma b''_k X_{ijk} + \theta_1 \mathrm{TMK}_j + \theta_2 \mathrm{TQ}_j + \theta_3 \mathrm{TExp}_j + \theta_4 \left(\mathrm{TExp}_j\right)^2 \\ & + \theta_5 \mathrm{OTL}_j + \theta_6 \mathrm{AvgX}_j + \theta_7 C_j + \theta_8 V_j + v''_{ij}. \end{aligned} \tag{3b}$$

Equations (3a) and (3b) control for students' accumulated achievement at the beginning of the "treatment year" (sixth grade). Equation (3a), which we call the "posttest controlling for initial test" model, estimates less biased relations between school resources and student academic achievement than the "posttest score only" model but does not address the issue that a student's baseline achievement is likely correlated with the error term (Ladd 2008). Equation (3b), which is a more restrictive value-added model than equation (3a), reduces the potential bias from this correlation by directly subtracting the standardized sixth-grade pretest score from the standardized sixth-grade posttest score. Specifically, equation (3a) assumes "decay" in what the student knew at the beginning of the sixth grade (the coefficient $a_1 < 1$, and the "decay" in student knowledge $= 1 - a_1$), so the teacher/school effect on gains includes this decay and therefore may be biased upward. In equation (3b), we explicitly make a_1 in equation (3a) equal to 1 and thus assume that the decay in student knowledge from the time of the pretest to the posttest is equal to zero. This reduces the amount of knowledge gain attributable to classroom and school inputs. The estimated coefficients for teacher quality and OTL in equation (3b) are therefore more conservative and potentially less biased.

Equations (2), (3a), and (3b) are two-level ordinary least squares (OLS) regressions in which learners are nested in classrooms, and we must account for within-classroom variation as well as across-classroom variation. The coefficients of classroom-level variables in such two-level OLS estimates are unbiased, but unless we account for the within-classroom variation, the standard errors of the estimated coefficients for the classroom variables are biased downward. We use a cluster correction to adjust the standard errors upward. The reported significance levels of our estimated coefficients in table 6 are based on adjusted standard errors.

Our model assumes that part of the influence of teacher mathematics knowledge and teacher experience on student achievement gain works through the quality of teaching, part through the exposure of students to OTL and part through other classroom processes (class management or individual interactions with students on mathematics-related issues). Thus, for example, in equation (3a), the total influence of teacher mathematics knowledge (TMK) on student post achievement (controlling for initial achievement) is $\phi_1 + \phi_2\gamma_1 + \phi_5\beta_1$, where ϕ_1 can be viewed as the direct, or "extra," classroom influence of TMK, $\phi_2\gamma_1$ is TMK's influence on achievement through teaching quality, and $\phi_5\beta_1$ is TMK's influence through OTL. Similar breakdowns of influence on student achievement would apply to teacher experience. When we discuss our results, we show an example of these breakdowns for each country through our estimates of equations (1a)–(3b).

In addition to these individual country production functions, we aggregate the data from our four samples to estimate whether there is a significant difference in students' mathematics performance between them, once we control for the student and classroom variables we specified in equations (3a) and (3b). Equations (4a) and (4b) are estimated for the combined sample with the addition of dummy variables for Botswana and Kenya low- and high-scoring schools. The reference dummy variable is South Africa. Thus, combining data from the three country samples allows us to estimate whether there is a significant country "fixed" effect on student mathematics gain for Kenya (divided in the samples of low- and high-scoring schools) and Botswana relative to South Africa. The estimated country coefficients for Botswana and for Kenya low- and high-scoring schools can represent many factors, including classroom resources and processes that we have not been able to account for, and broader societal variables (state social capital) that can be characterized as historical educational policies. We can only speculate what these are with three countries in our comparison, but it is worth discussing the possible size of such factors because they may play an important role in explaining how well students achieve in school.

$$A_{ij\mathrm{post}} = \text{Independent variables as in equation (3a)} + \Sigma\mathrm{Country}_{\mathrm{n}}; \qquad (4a)$$

$$A_{ij\mathrm{post}} - A_{ij\mathrm{pre}} = \text{Independent variables as in equation (3b)} + \Sigma\mathrm{Country}_{\mathrm{n}}, \qquad (4b)$$

where $\mathrm{Country}_n$ is a dummy variable for Botswana, Kenya low scoring, and Kenya high scoring (reference dummy = South Africa).

Results

Teaching Quality, Opportunity to Learn, and Teacher Characteristics

We hypothesized that in each of the three countries under study, teaching quality and the OTL provided to students by their mathematics teachers are related to teachers' mathematics knowledge and teaching experience. Table 4 shows the results of estimating equation (1a) with and without controls for class average students' SES. The sample sizes are relatively small, but we still get significant results.[15]

We find that in South Africa, teacher's mathematics knowledge (test score) is associated with a higher cumulative score given the teacher on her videotaped lesson. In Botswana and high-scoring Kenya schools, there is no significant relationship between the teacher's video score and either the teacher's math test score or years of teaching experience. In the sample of low-scoring schools in Kenya, the teaching rating is negatively related to teaching experience up to about 13 years of teaching experience and then rises. That is, "better" teaching in those lower-scoring schools is associated with fewer years of teaching experience. We hypothesized that classes where

[15] We show the OLS estimates in which we use a continuous value of the teaching evaluation (1, 2, or 3) as our measure of TQ. We also estimated an ordinal logit for TQ. The odds ratios in the ordinal logit estimates were statistically significant (or not) and the same sign for every variable as in the OLS estimates.

TABLE 4

ESTIMATED TEACHING QUALITY AND TEACHER CHARACTERISTICS, WITHOUT AND WITH CONTROLS FOR STUDENT AND CLASS SES, BY COUNTRY

	South Africa		Botswana		Kenya Lower Scoring		Kenya High Scoring	
Variable	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Standardized teacher test score	.340***	.346***	.021	.029	−.051	−.051	.191	.137
	(.084)	(.086)	(.153)	(.156)	(.112)	(.116)	(.134)	(.138)
Teacher experience	−.044	−.044	−.008	−.005	−.080**	−.079**	−.015	−.027
	(.038)	(.038)	(.037)	(.038)	(.038)	(.039)	(.044)	(.044)
Teacher experience^2	.001	.001	.000	.000	.003**	.003**	−.000	−.000
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
Standardized average classroom SES	...	−.019	...	−.030	...	.022	...	.107
		(.083)		(.094)		(.119)		(.078)
Intercept	2.240***	2.239***	1.788***	1.781***	1.996***	2.005***	1.994***	2.013***
Observations	62	62	64	64	36	36	36	36
Adjusted R^2	.274	.261	−.047	−.063	.140	.141	−.004	.022

SOURCE.—South Africa, Botswana, and Kenya school samples, 2009.
NOTE.—Standard errors in parentheses. SES = socioeconomic status.
** $P = .05$.
*** $P = .01$.

students have higher levels of family resources may have "better" mathematics teaching; however, we do not find evidence of such a relationship in our samples.

Teachers with more mathematics knowledge may feel more comfortable with the subject matter in the curriculum; hence, they may be more likely to spend time teaching the curriculum material. They may also cover more topics because they may be more familiar with a greater range of topics. As far as giving more lessons on the topics related to the test items, the relationship is less clear: teachers with more mathematics knowledge may like teaching mathematics more, so they may be more likely to give a larger number of lessons on the topics, but they may also be more "efficient" and could spend less time on each topic. They could also choose to devote more time to topics not covered by our student test—topics more challenging for students in their classes.

Table 5 shows the results of estimating equation (1b) for OTL. Teachers' mathematics knowledge is only positively related to OTL in South Africa, which also shows the largest effect size of teacher test score on OTL.[16] Teachers' experience is positively related to standardized OTL in South Africa and the high-scoring Kenya schools. This suggests that in those two cases, more experienced teachers provided more mathematics coverage during the academic year. To the contrary, in Botswana, the estimates suggest that less experienced teachers provided more mathematics coverage.

As in the case of teaching quality (table 4), the average SES of students in the class does not generally appear to be significantly related to OTL. However, for the sample of low-scoring Kenya schools, classrooms with higher average SES students apparently get significantly less curriculum coverage (about 0.5 SD less). We should also note that in South Africa, teacher test score is positively and rather highly correlated (0.30) with average class SES, suggesting that, at least in South Africa, teachers with a higher test score may be selecting on higher-SES classrooms.

Thus, our results indicate that higher levels of teacher mathematics knowledge contribute to better mathematics teaching only in South African schools and that, if anything, more teaching experience is negatively associated with better teaching. Our estimates also suggest that teachers with more mathematics knowledge are more likely to provide more mathematics coverage in all three countries, but experienced teachers may or may not provide more mathematics coverage. The relation between teacher test score and both the quality of teaching and OTL is strongest in South Africa. This is important, since it suggests that improvement in teacher mathematics knowledge in South Africa could have a bigger impact on teaching qual-

[16] Because our measure of OTL is different in Kenya from the measure used in Botswana/South Africa, we standardized observed OTL (mean = 0, SD = 1) for each of the four samples.

TABLE 5

Opportunity to Learn (Standardized) and Teacher Characteristics without and with Controls for Student and Class SES, by Country

Variable	South Africa		Botswana		Kenya Lower Scoring		Kenya High Scoring	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Standardized teacher test score	.278**	.233*	.170	.200	.129	.139	.104	.130
	(.132)	(.138)	(.197)	(.200)	(.191)	(.167)	(.176)	(.186)
Teacher experience	.116*	.118*	−.162***	−.154***	−.034	−.075	.102*	.108*
	(.061)	(.061)	(.048)	(.048)	(.063)	(.057)	(.057)	(.059)
Teacher experience^2	−.003**	−.003**	.005***	.005***	.001	.003	−.004*	−.004**
	(.001)	(.001)	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
Standardized average classroom SES	...	.147	...	−.118	...	−.573***	...	−.051
		(.132)		(.120)		(.171)		(.105)
Intercept	−1.103	−1.091	.989***	.960***	.144	−.074	−.436	−.445
Observations	62	62	64	64	36	36	36	36
Adjusted R^2	.063	.067	.148	.148	−.070	.189	.031	.007

Source.—South Africa, Botswana, and Kenya school samples, 2009.
Note.—Standard errors in parentheses. SES = socioeconomic status.
* $P = .10$.
** $P = .05$.
*** $P = .01$.

ity and OTL than in Botswana and Kenya, where teachers scored higher on the teacher test.

The Impact of Teacher Skills on Learner Mathematics Achievement Gains

The results in table 6 present our estimates in each country for posttest score (eq. [2]) and our two models of achievement gains (eqq. [3a] and [3b]). Overall, the estimates show that teacher capacity to teach mathematics in these three countries is generally important in influencing student achievement and achievement gains. In South Africa, Botswana, and low-scoring Kenya schools samples, better classroom teaching, as evaluated from the videotapes, has a positive and large impact on the learner achievement gain (models 2 and 3). Teacher mathematics knowledge has a positive direct effect on learner mathematics gains in our Botswana and, especially, our Kenya low-scoring-school sample. The size of the positive effect is about .05 of a standard deviation in Botswana and 0.09–0.11 standard deviations in low-scoring Kenya schools. In South Africa, however, our structural model equations (eqq. [1a], [1b], [3a], and [3b]) come into play. The positive effects of teacher mathematics knowledge on student achievement gains are indirect, through teachers' teaching quality (table 4) and OTL (table 5). The indirect effect through improving teacher quality by 1 standard deviation is equal to $\theta_2 \times \gamma_1$, or 0.34 (table 4) $\times$ 0.28 (table 6, model 2) $=$ 0.10. Similarly, the indirect effect of increased teacher knowledge through improved OTL equals 0.27 (table 5) $\times$ 0.07 (table 6, model 2) $=$ 0.02.[17]

Greater OTL also has a significant positive effect on achievement gains in the South Africa sample. Teacher experience is not significantly related to achievement gains in any of the three samples, although, as we showed in tables 4 and 5, this may be partly the result of correlations between teacher experience and two of the other treatment variables, teaching quality and OTL.[18] Reported violence in the schools is negatively associated with test score gains in both lower-scoring and higher-scoring Kenya schools.[19]

[17] The direct effect on student achievement gain of increasing teacher knowledge by 1 standard deviation in South Africa is negative, equal to -0.08 (table 6, model 2). Thus, the total effect in South Africa on student achievement gain of increasing teacher knowledge may be $-0.08 + 0.10 + 0.02 = 0.04$ standard deviations—about the same as the direct (and only) effect of increasing teacher knowledge on student achievement gain in Botswana.

[18] A traditionally important problem in identifying the effect of teacher skills on student learning gains is selection bias: more skilled teachers may get assigned to classrooms/schools with higher student mathematics gains, and students (or parents of students) who are more motivated learners may choose more skilled teachers or schools with more skilled teachers. As discussed, we find little evidence of this type of selection bias in any of our samples (see tables 4 and 5), except in South Africa (through the correlation of teacher test scores and average classroom social class) and in Kenya's low-scoring schools. Yet, we still take the issue seriously and control for several classroom/school variables that could be proxies for "better" classroom/school conditions, which teachers and students/parents might choose if they were selecting themselves into high-learning-gains situations. These control variables are average classroom student socioeconomic background, class size, private/public school (in Kenya only), and a violence index based on student and teacher reports of frequency of student-to-student violence.

[19] The differences in results between the more restricted version of the value-added specification (eq. [3b]) and the specification in eq. (3a) are small, except for the impact on the coefficients of

Students' family resources and classroom factors that we were able to measure account for about 20 percent of the posttest score we gave students and, including students' initial test score, 41 (South Africa) to 65 percent (Kenya's high-scoring schools) of students' second test score (table 6). It also appears that except in Kenya's high-scoring schools, our measures of teacher quality and OTL are related to student mathematics learning in grade 6.

It could be argued that because the Kenya samples do not represent the broad spectrum of Kenya schools, but only two slices of the distribution, the coefficients of classroom variables such as teaching quality and OTL are biased downward, particularly for the Kenya high-scoring sample. There are several reasons to believe that this is not the source of the small effect of teacher variables on value added in that subsample. First, table 2 suggests that the variation in our teacher variables is as great in the high-scoring subsample as in the low-scoring subsample; second, the adjusted R^2s of the table 6 value-added regressions in the high-scoring sample are as high or higher than in the lower-scoring samples in South Africa and Botswana; and third, the fact that teacher variables are as statistically significant in Kenya's low-scoring schools as in South Africa and Botswana suggests that the lack of significant coefficients for teacher skill variables in high-scoring Kenyan schools reflects unobserved sources of variation across high-scoring (also higher social class) Kenyan schools rather than the effect of sample truncation.

Estimates of Country Fixed Effects

In addition to these estimates of within-country achievement gains, we estimate the "fixed effect" on grade 6 gains of being a student attending school in each of these countries. This fixed effect is due to achievement gains over and above the gains due to country differences in students' background characteristics and classroom resources.[20]

Estimates combining the samples (eqq. [4a] and [4b]) are shown in table 7. These show that even when we control for student socioeconomic background and teacher/class variables—teaching skills (as measured by teaching video analysis), teacher mathematics test score, OTL coverage,

class SES (decrease in 3b) and private school (increase in 3b) in the low-scoring Kenya sample, suggesting that generally our classroom treatment variables are not highly correlated with students' initial score.

[20] There are alternative methods of approximating the unexplained country effect using production functions (see, e.g., Carnoy and Marshall 2005). A common method is to compare the predicted value of student outcomes in the lower-scoring country X, using the estimated production function for that country and the values of classroom resources in the higher-scoring country Y as well as the predicted value of student outcomes in country X using its own values of classroom resources (Oaxaca 1977). Our regression results with country fixed effects give similar results to the Oaxaca method. The advantage of our method is that we can compare student achievement differences across countries, controlling for student and classroom/school resource variation. Its disadvantage is that the coefficients of classroom/school resources are averaged across countries.

TABLE 6

Estimated Student Mathematics Achievement (Z-score), by Country, 2009

Variable	South Africa			Botswana			Lower-Scoring Kenya Schools			High-Scoring Kenya Schools		
	Posttest Score Only (Model 1)	Posttest Controlling for Initial Test (Model 2)	Posttest Minus Pretest (Model 3)	Posttest Score Only (Model 1)	Posttest Controlling for Initial Test (Model 2)	Posttest Minus Pretest (Model 3)	Posttest Score Only (Model 1)	Posttest Controlling for Initial Test (Model 2)	Posttest Minus Pretest (Model 3)	Posttest Score Only (Model 1)	Posttest Controlling for Initial Test (Model 2)	Posttest Minus Pretest (Model 3)
Initial test score		.454***			.702***			.704***			.668***	
		(.06)			(.02)			(.03)			(.03)	
Teacher experience	−.01	−.01	−.00	−.01	.01	.02	.03*	.01	.00	−.01	.01	.02
	(.02)	(.01)	(.02)	(.02)	(.18)	(.01)	(.02)	(.01)	(.01)	(.04)	(.02)	(.03)
Teacher experience^2	.00	.00	.00	.00	−.00	−.00	−.00*	−.00	.00	.00	−.00	−.00
	(.00)	(.1)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.03)	(.00)	(.00)
Standardized teacher test score[a]	−.06	−.08*	−.11**	.08	.05*	.05	.18***	.09***	.11**	−.15	−.01	.04
	(.07)	(.04)	(.05)	(.05)	(.03)	(.03)	(.05)	(.03)	(.05)	(.09)	(.04)	(.06)
Better teaching quality	.29**	.28***	.34**	.26**	.14**	.13**	.42***	.22***	.30***	.05	.06	.09
	(.12)	(.10)	(.14)	(.10)	(.05)	(.06)	(.11)	(.06)	(.09)	(.23)	(.11)	(.16)
Best teaching quality	.29*	.26**	.28	.24	.08	.04	.31*	.19**	.28*	.50*	.25	.30
	(.29)	(.10)	(.18)	(.14)	(.08)	(.08)	(.17)	(.10)	(.17)	(.28)	(.16)	(.22)
Standardized OTL[b]	.07	.07**	.08***	.01	.01	.01	.13**	.04	.03	.15	.07	.08
	(.05)	(.03)	(.03)	(.04)	(.03)	(.03)	(.06)	(.03)	(.04)	(.12)	(.06)	(.07)

Observed class size	−.003	−.003	−.003	−.00	−.01	−.01	−.01***	−.004***	−.005***	.01	.01	.01
	(.00)	(.00)	(.00)	(.01)	(.00)	(.00)	(.00)	(.001)	(.002)	(.01)	(.00)	(.01)
Standardized class SES[c]	.38***	.20***	.03	.13**	.06**	.05	.15**	−.02	.01	.25***	.09***	.07
	(.06)	(.04)	(.04)	(.05)	(.03)	(.03)	(.06)	(.03)	(.04)	(.06)	(.03)	(.05)
Standardized violence index[d]	−.01***	−.00	.02	.03	.02	.02	−.12***	−.05***	−.06*	−.32***	−.12*	−.10
	(.05)	(.03)	(.05)	(.05)	(.03)	(.03)	(.03)	(.02)	(.03)	(.11)	(.06)	(.09)
Private school							.04	.07	.13*	.03	.00	−.00
							(.09)	(.05)	(.07)	(.25)	(.11)	(.1)
Constant	5.16***	3.57***	2.33**	5.54***	2.19***	1.36	−.71**	−.29**	−.32	−.69*	−.43*	−.60
Learner characteristics included?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,485	3,485	3,485	1,612	1,612	1,612	788	788	788	1,119	1,119	1,119
Adjusted R^2	.235	.409	.033	.188	.561	.039	.211	.554	.067	.305	.647	.068

SOURCE.—South Africa (SA), Botswana (BW), and Kenya data samples, 2009.
NOTE.—Standard errors in parentheses.
[a] Teacher test in Kenya is somewhat different and has a different metric, so we use a standardized metric for each country.
[b] Opportunity to learn (OTL) variable in BW and SA is total lessons per item on student test items; in Kenya, percent of curriculum covered.
[c] Average class socioeconomic status (SES) is the same composite indicator in the three countries; we use the standardized index score.
[d] Violence indicator in BW and SA is index of student, teacher, and principal assessment of learner-to-learner and teacher-to-learner violence in school; in Kenya, violence indicator is teacher assessment of learner-to-learner violence.
* P = .10.
** P = .05.
*** P = .01.

TABLE 7
ESTIMATES OF LEARNER MATHEMATICS ACHIEVEMENT AND ACHIEVEMENT GAIN AS A FUNCTION OF TEACHER CHARACTERISTICS, SCHOOL CHARACTERISTICS, AND COUNTRY EFFECTS, 2009

Variable	Posttest Score (Model 1)	Posttest Controlling for Initial Test Score (Model 2)	Test Score Gain (Posttest Minus Pretest Score; Model 3)
Standardized initial test score		.571***	
		(.033)	
Years of teaching experience	.003	.007	.014**
	(.008)	(.005)	(.006)
Teaching experience^2	−.000	−.000**	−.000***
	(.000)	(.000)	(.000)
Better teaching quality	.185***	.133***	.191**
	(.060)	(.043)	(.074)
Best teaching quality	.128	.101*	.142
	(.081)	(.053)	(.092)
Standardized teacher test score[a]	−.108	−.084*	−.118*
	(.103)	(.052)	(.068)
Standardized teacher test score^{2}[a]	.144	.103*	.136*
	(.130)	(.062)	(.071)
Standardized OTL[a]	.073***	.047***	.056***
	(.028)	(.015)	(.020)
Observed class size	−.002	−.002	−.002
	(.003)	(.002)	(.003)
Average class SES[a]	.186***	.069***	.024
	(.030)	(.017)	(.023)
Standardized violence index[a]	−.060**	−.026*	−.011
	(.028)	(.016)	(.028)
Kenya high-scoring schools	1.515***	.749***	.632***
	(.204)	(.119)	(.140)
Kenya lower-scoring schools	.957***	.537***	.550**
	(.165)	(.093)	(.118)
Botswana	.561***	.247***	.164**
	(.084)	(.046)	(.057)
Intercept	−.580***	−.399***	−.508***
Student characteristics included	Yes	Yes	Yes
Adjusted R^2	.498	.692	.096

SOURCE.—South Africa, Botswana, and Kenya school samples, 2009.
NOTE.—Unweighted student achievement and achievement gain standardized across all three countries with mean = 0 and standard deviation = 1. Standard errors in parentheses. OTL = opportunity to learn; SES = socioeconomic status; N = 7,004.
[a] Standardized variable (mean = 0, standard deviation = 1).
* P = .10.
** P = .05.
*** P = .01.

class size, classroom average socioeconomic background (peer effect), and the school violence index—students in Botswana and in Kenya, especially, have significantly larger gains on the mathematics test than students in South Africa. Thus, even given the seeming lack of direct comparability in the students sampled in Kenya with our Botswana/South Africa samples, there appears to be a large "residual effect" that cannot be explained by differences in the socioeconomic background of the students or the variables we use to measure teacher quality or classroom conditions. The Botswana residual effect on scores relative to South Africa is 0.16–0.25 standard deviations in test score gains (depending which method is used to measure gains); the Kenya lower-scoring school effect on test score gains is about 0.54–0.55 standard deviations; and the Kenya high-scoring school effect on test score gains is 0.63–0.75 standard deviations. These coefficients are all statistically significant at the 5 percent or 1 percent level. Although the two Kenya samples together do not represent all Kenya schools, they are representative of two groups of schools, and students in both groups of Kenya schools make considerably larger mathematics achievement gains, when controlling for student background and classroom variables, than the students in either Botswana or South Africa. There are certainly other school and classroom variables we have not observed in Botswana and Kenya that could influence student performance, but the very size of the country fixed effects, particularly for the two Kenya samples, suggests that the educational system as a whole in that country is able to achieve much higher gains than in South Africa or Botswana, for students of similar social class backgrounds and with similar educational resources (at least those we have been able to measure).

Conclusion

Our analysis suggests that school resources make major contributions to student learning gains in three African countries. The fact that increasing teacher skills and teaching quality can influence student outcomes is important. The great advantage of our analysis is that we can connect the mathematics gains made by learners in our sample with specific teachers, specific classroom conditions, and specific classroom processes. Even though one can question whether our student test truly measured how much mathematics learners knew near the beginning and toward the end of the school year, we based the test on what they were supposed to have learned before entering grade 6 and how much they should have learned during that academic year. We also carefully measured their exposure to the mathematics curriculum and the quality of teacher knowledge and teaching along multiple dimensions. Yet, even when we control for differences in teachers' mathematics knowledge, teaching skills, and curriculum coverage, Kenyan grade 6 students in both lower- and higher-scoring schools tested considerably higher on our mathematics test than Botswana students and even higher than students in the North West Province of South Africa.

We are acutely aware of the limitations of our data for explaining country-level student achievement differences with a sample of only three countries at one point in time and in comparing our Kenya samples with our Botswana/South Africa samples, which were specifically designed for such an intercountry analysis. Nevertheless, our data also have important advantages in beginning to discuss such explanations. Our main outcome variable is student achievement gains in a single grade. Thus, our estimates of country differences are more likely than those using cross-sectional international test score data to be related to unmeasured school/classroom effectiveness rather than a number of other effects that result from outside school influences. Our measures of classroom resources are also more detailed than in typical intercountry production function estimates.

What might these residual country fixed effects represent? As we discussed in our review of the literature, some have argued that greater school autonomy and competition among schools is an important explainer of country differences in student achievement. If the relatively high fraction of private schools in Kenya represents greater autonomy and competition in the school system, this might conceivably explain the large fixed effects for each of the two sets of Kenyan schools. An argument against this is that in our Kenya estimates the effect on student achievement gains of private education is significant (but relatively small) only for the lower-scoring Kenya sample (table 6). However, since we sampled neither private schools in Botswana nor private and Model C schools in South Africa, we could not assess the effect of private/more autonomous schooling across countries.

There are other possible (and, based on our school/classroom observations, more likely) explanations for these country fixed effects. For example, they could reflect different expectations educational systems and the broader society have internalized regarding how demanding schooling should be or the level of mathematics children should be expected to learn in primary school classrooms. As observed in the Kenya teacher videotapes, for example, the technical quality of mathematics teaching was about the same as in South Africa and Botswana, but the level of mathematics being taught by the teachers in Kenya was higher (table 3). At some level of society (central government, parents), higher levels of math exposure reflect higher expectations of student learning.[21]

Kenya's examination system, originally inherited from the British, has likely contributed to shaping these higher expectations. The KCPE, now given at the end of eighth grade, has traditionally set a standard for teachers that is high stakes for both students (entry into a good secondary school) and schools competing for prestige. Bishop (1997) has argued convincingly that such high-stakes end of school/entrance tests significantly raise teacher and student performance.

The other side of the expectations coin is South Africa, where apartheid deprecated school quality for Africans and consciously and systematically

[21] There is some evidence, for example, that British colonial (missionary) education in east Africa and Rhodesia (Zimbabwe) established relatively high standards for Africans' education despite their commitment to preparing Africans for subjugated roles in colonial society (see, e.g., Whitehead [1993] commenting on Mwiria [1991]). Kenyan parents have had a tradition of demanding relatively high levels of attainment for their children and were willing to pay for low-quality, privately financed Harambee schools when the public sector was unwilling to provide sufficient primary and, later, secondary schooling (Mwiria 1985).

lowered academic expectations and shortchanged African education financially (Fedderke and Luiz 2002). Further, because of the central role that schooling played under apartheid in the reproduction of racial/social inequality, schools became places of organized political resistance to black oppression rather than places focused on academic learning. The legacy of apartheid-era institutional structures was evident in the schools we sampled. Veteran teachers continue to view themselves primarily as political actors. In part, but only in part, this legacy is reflected in the very low numbers of mathematics lessons teachers in our sample offered students and the low level of teachers' mathematics knowledge. Beyond that, most teachers and school directors are held to an extremely low standard of how much students could and should learn in school (Carnoy et al. 2012).

These observations are suggestive, but whatever the reasons for them may be, the corresponding effects appear to be large and demand explanation. Although most such country effects may not be exportable, if we recognize their importance, others may serve to help policy makers move beyond marginal fixes toward more productive changes in educational contexts.

References

Bishop, John. 1997. "The Effect of National Standards and Curriculum-Based Exams on Achievement." *American Economic Review* 87 (2): 260–64.

Bowles, Samuel, and Herbert Gintis. 1975. *Schooling in Capitalist America*. New York: Basic.

Boyd, Donald, Pamela Grossman, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2006. "How Changes in Entry Requirements Alter the Teacher Workforce and Affect Student Achievement." *Education Finance and Policy* 1 (2): 176–216.

Bruns, Barbara, and Javier Luque. 2014. *Better Teachers for Latin America and the Caribbean: New Evidence on How to Raise Teacher Quality and Student Learning*. Washington, DC: World Bank.

Carnoy, Martin, Linda Chisholm, and Bagele Chilisa, eds. 2012. *The Low Achievement Trap*. Pretoria: Human Sciences Research Council Press.

Carnoy, Martin, Amber Gove, and Jeffery Marshall. 2007. *Cuba's Academic Advantage*. Stanford, CA: Stanford University Press.

Carnoy, Martin, and Henry Levin. 1985. *Schooling and Work in the Democratic State*. Stanford, CA: Stanford University Press.

Carnoy, Martin, and Jeffery Marshall. 2005. "Cuba's Academic Performance in Comparative Perspective." *Comparative Education Review* 49 (2): 230–61.

Carnoy, Martin, and Joel Samoff. 1989. *Education and Social Transition in the Third World*. Princeton, NJ: Princeton University Press.

Carrasco, Rafael. 2014. "Leveling the Playing Field: How Can We Address Educational Inequalities?" PhD diss., Stanford University.

Clotfelter, Charles T., Helen Ladd, and Jacob Vigdor. 2007. "Teacher Credentials and Student Achievement: Longitudinal Analysis with Student Fixed Effects." *Economics of Education Review* 26 (6): 673–82.

Clotfelter, Charles T., Helen Ladd, and Jacob Vigdor. 2010. "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects." *Journal of Human Resources* 45 (3): 655–82.

Fedderke, Johannes, and John Luiz. 2002. "Production of Educational Output: Time Series Evidence from Socioeconomically Heterogeneous Populations; The Case of South Africa, 1927–1993." *Economic Development and Cultural Change* 51 (1): 161–87.

Ferguson, Ron, and Helen Ladd. 1996. "How and Why Money Matters: An Analysis of Alabama Schools." In *Holding School Accountable*, ed. Helen Ladd. Washington, DC: Brookings.

Foucault, Michel. 1975. *Discipline and Punish*. New York: Pantheon.

Fuchs, Thomas, and Ludger Woessmann. 2007. "What Accounts for International Differences in Student Performance: A Reexamination Using PISA Data." *Empirical Economics* 32 (2–3): 433–64.

Fuller, Bruce, Haiyan Hua, and Conrad Snyder. 1994. "When Girls Learn More than Boys: The Influence of Time in School and Pedagogy in Botswana." *Comparative Education Review* 38 (3): 347–76.

Hanushek, Eric, Paul Peterson, and Ludger Woessmann. 2013. *Endangering Prosperity*. Washington, DC: Brookings.

Hanushek, Eric A., John Kain, and Steven Rivkin. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73:417–58.

Hill, Heather, Brian Rowan, and Deborah Ball. 2005. "Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement." *American Educational Research Journal* 42 (2): 371–406.

Kane, Thomas, Daniel F. McCaffrey, Trey Miller, and Douglas O. Staiger. 2013. "Have We Identified Effective Teachers?" Bill and Melinda Gates Foundation, Seattle. http://metproject.org/downloads/MET_Validating_Using_Random_Assignment_Research_Paper.pdf.

Knight, John, and Richard Sabot. 1990. *Education, Productivity, and Inequality: The East African Natural Experiment*. Washington, DC: World Bank.

Ladd, Helen. 2008. "Teacher Effects: What Do We Know?" In *Teacher Quality*, ed. Greg Duncan and James Spillane. Evanston, IL: Northwestern University Press.

Levin, Henry M. 1980. "Educational Production Theory and Teacher Inputs." In *The Analysis of Educational Productivity*, vol. 2 of *Issues in Macro Analysis*, ed. Charles Bidwell and Douglas Windham. Cambridge, MA: Ballinger.

Mwiria, Kilemi. 1985. "The Kenya Harambee School Movement: An Historical Perspective." PhD diss., Stanford University.

Mwiria, Kilemi. 1991. "Education for Subordination: African Education in Colonial Kenya." *History of Education* 20 (3): 261–73.

National Research Council. 2001. *Adding It Up: Helping Children Learn Mathematics*. Washington, DC: National Academy.

Ngware, Moses, Moses Oketch, Maurice Mutisya, and Benta Abuya. 2010. *Classroom Observation Study: A Report on the Quality and Learning in Primary Schools in Kenya*. Nairobi: African Population and Health Research Center.

Nye, Barbara, Spyros Konstantopoulos, and Larry V. Hedges. 2004. "How Large Are Teacher Effects?" *Educational Evaluation and Policy Analysis* 26 (3): 237–57.

Oaxaca, Ronald. 1977. “The Persistence of Male-Female Earnings Differences.” In *Distribution of Economic Well-Being*, ed. F. Thomas Juster. Cambridge, MA: National Bureau of Economic Research.

OECD (Organization for Economic Cooperation and Development). 2013. *PISA 2012 Results: What Students Know and Can Do*. Paris: OECD.

Rice, J. 2003. *Teacher Quality: Understanding the Effectiveness of Teacher Attributes*. Washington, DC: Economic Policy Institute.

Rockoff, Jonah E. 2004. “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data.” *American Economic Review* 94 (2): 247–52.

Rockoff, Jonah E., Brian Jacob, Thomas Kane, and Douglas O. Staiger. 2008. “Can You Recognize an Effective Teacher When You Recruit One?” NBER Working Paper no. 14485, National Bureau of Economic Research, Cambridge, MA.

Rothstein, Richard. 2004. *Class and Schools*. New York: Teachers College Press.

SACMEQ (Southern African Consortium for Monitoring Education Quality). 2010. “What Are the Levels and Trends in Reading and Mathematics Achievement?” *SACMEQ Policy Issues*, no. 2 (September).

Schmidt, William H., Curtis C. McKnight, Richard Houang, Hsingchi Wang, David Wiley, Leland Cogan, and Richard Wolfe. 2001. *Why Schools Matter: A Cross-National Comparison of Curriculum and Learning*. San Francisco: Jossey-Bass.

Schmidt, William H., Curtis C. McKnight, and S. A. Raizen. 1997. *A Splintered Vision: An Investigation of U.S. Science and Mathematics Education*. Dordrecht: Kluwer.

Shulman, Lee. 1986. “Those Who Understand: Knowledge Growth in Teaching.” *Educational Researcher* 15 (2): 4–14.

Spillane, James, Richard Halverson, and John Diamond. 2001. “Investing in School Leadership Practice: A Distributed Perspective.” *Educational Researcher* 30 (3): 23–28.

Stein, Mary Kay, Margaret Smith, Marjorie Henningsen, and Edward Silver. 2000. *Implementing Standards-Based Mathematics Instruction: A Casebook for Professional Development*. New York: Teachers College Press.

Taut, Sandy, Veronica Santelices, Carolina Araya, and Jorge Manzi. 2010. “Theory Underlying a National Teacher Evaluation Program.” *Evaluation and Program Planning* 33 (4): 477–86.

UNESCO Education for All. 2005. *Global Monitoring Report*. Paris: UNESCO.

Whitehead, Clive. 1993. “Education for Subordination: Some Reflections on Kilimi Mwiria’s Account of African Education in Colonial Kenya.” *History of Education* 22 (1): 85–93.

Woessmann, Ludger. 2004. “How Equal Are Educational Opportunities? Family Background and Student Achievement in Europe and the United States.” CESifo Working Paper no. 1162, CESifo Group, Munich.